



**HAL**  
open science

# A minimality property of the value function in optimal control over the Wasserstein space

Cristopher Hermosilla, Averil Prost

► **To cite this version:**

Cristopher Hermosilla, Averil Prost. A minimality property of the value function in optimal control over the Wasserstein space. 2024. hal-04427139

**HAL Id: hal-04427139**

**<https://hal.science/hal-04427139v1>**

Preprint submitted on 30 Jan 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A MINIMALITY PROPERTY OF THE VALUE FUNCTION IN OPTIMAL CONTROL OVER THE WASSERSTEIN SPACE \*

CRISTOPHER HERMOSILLA<sup>†</sup> AND AVERIL PROST<sup>‡</sup>

**Abstract.** An optimal control problem with (possibly) unbounded terminal cost is considered in  $\mathcal{P}_2(\mathbb{R}^d)$ , the space of Borel probability measures with finite second moment. We consider the metric geometry associated with the Wasserstein distance, and a suitable weak topology rendering  $\mathcal{P}_2(\mathbb{R}^d)$  locally compact. In this setting, we show that the value function of a control problem is the minimal viscosity supersolution of an appropriate Hamilton-Jacobi-Bellman (HJB) equation. Additionally, if the terminal cost is bounded and continuous, we show that the value function is the unique viscosity solution of the HJB equation.

**Key words.** Wasserstein space, Optimal control problems, viscosity solutions, weak topology

**MSC codes.** 35F21, 35R06, 49Lxx

**1. Introduction.** We consider a Hamilton-Jacobi-Bellman (HJB) equation arising from an optimal control problem whose state space is a set of measures. More precisely, the unknown of our equation belongs  $\mathcal{P}_2(\mathbb{R}^d)$ , the space of Borel probability measures with finite second moment. It is well-known that this setting is suitable for the modelling of optimal control of population dynamics in crowd motion [29, 13] or biology [11]. In these approaches, the configuration at time  $s \in [0, T]$  of a population is represented by a measure  $\mu_s \in \mathcal{P}_2(\mathbb{R}^d)$ , and the evolution in time is assumed to satisfy a continuity equation of the form

$$(1.1) \quad \partial_s \mu_s + \operatorname{div}(f \# \mu_s) = 0, \quad s \in [0, T], \quad \mu_0 = \nu.$$

The equation (1.1) is understood in the sense of distributions, and is the measure counterpart of an ODE with initial term  $\nu \in \mathcal{P}_2(\mathbb{R}^d)$ , and dynamic  $f \in \mathcal{C}(\mathbb{R}^d; \mathbb{T} \mathbb{R}^d)$ . In this work, we study a controlled version of the continuity equation. Given some compact control set  $U \subset \mathbb{R}^k$  and a function  $u \in L^0([0, T]; U)$ , we consider a controlled dynamic  $f : U \rightarrow \mathcal{C}(\mathbb{R}^d; \mathbb{T} \mathbb{R}^d)$  and the associated *controlled continuity equation*

$$(1.2) \quad \partial_s \mu_s + \operatorname{div}(f[u(s)] \# \mu_s) = 0, \quad s \in [0, T], \quad \mu_0 = \nu.$$

Let  $(\mu_s^{0, \nu, u})_{s \in [0, T]}$  be the solution of (1.2), whose meaning and well-posedness will be discussed in Section 3. We are concerned with a Mayer type problem associated with a terminal cost  $\mathfrak{J} : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R} \cup \{+\infty\}$ , which may enforce terminal state constraints when taking the value  $+\infty$ . To compute the optimal control, a general approach is to study the value function  $V : [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R} \cup \{\pm\infty\}$  of the problem

$$(1.3) \quad V(t, \nu) := \inf \{ \mathfrak{J}(\mu_T^{t, \nu, u}) \mid u \in L^0([t, T]; U) \}.$$

From the theory in finite-dimensional and Hilbert spaces, it is expected that (1.3) is linked to an HJB equation of the form

$$(1.4) \quad -\partial_t V(t, \mu) + H(\mu, D_\mu V(t, \mu)) = 0, \quad V(T, \mu) = \mathfrak{J}(\mu).$$

---

\*Submitted to the editors 26/01/2024.

**Funding:** This work was supported by the Center for Mathematical Modeling (CMM) and ANID-Chile under BASAL funds for Center of Excellence FB210005 and Fondecyt Regular 1231049.

<sup>†</sup>Departamento de Matemática, Universidad Técnica Federico Santa María, Valparaíso, Chile. (cristopher.hermosill@usm.cl).

<sup>‡</sup>Laboratoire de Mathématiques, INSA Rouen Normandie, Saint-Étienne-du-Rouvray, France. (averil.prost@insa-rouen.fr).

38 More precisely, a classical result of the HJB theory in Euclidian spaces is that whenever  
 39  $\mathfrak{J}$  is real-valued and uniformly continuous, the value function  $V$  is the unique solution  
 40 of (1.4) in the sense of viscosity solutions [14, 15]. In the case where  $\mathfrak{J}$  may take infinite  
 41 values, the value function may still be characterized as the smallest supersolution in  
 42 the classical (Euclidean) sense. The purpose of this work is to transpose these results  
 43 to problems where the state evolves in the space of Wasserstein measures, that is,  
 44 our main results is concerned with showing that the value function  $V$  is the smallest  
 45 supersolution of (1.4) in an ad-hoc sense for the space  $\mathcal{P}_2(\mathbb{R}^d)$ .

46 The HJB equation we are interested in is a particular instance of a PDE on the  
 47 Wasserstein space, which has attracted a lot of interest since the seminal work of  
 48 Otto [28] on the porous media equation. The corpus of results of the theory concerns  
 49 gradient flows in the space of measures [4, 1, 12], the master equation in the theory  
 50 of mean-field games [9, 10, 7], Hamiltonian systems and flow equations [3, 2] and  
 51 optimal control problems [27, 24, 16]. As far as classical solutions are concerned, the  
 52 dominating theory is the *Lions differentiability*, that gives a proper definition of the  
 53 gradient of an application  $\varphi : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$  [25, 6]. The strength of this idea is to  
 54 embed measures into an external Hilbert space  $L^2_{\mathbb{P}}$ , and use the Hilbertian structure  
 55 therein. This strategy proved successful to study the master equation whenever the  
 56 data is sufficiently smooth to expect a  $C^1$  solution [8, Theorem 1.5].

57 In the case when such regularity is not achievable, the most effective strategy so  
 58 far uses *semidifferentials* defined in an appropriate tangent space. Indeed, viscosity  
 59 solutions may be defined by imposing some sign conditions on the sub and superdiffer-  
 60 ential of the solution at any point, thus accounting for the non-existence of a gradient.  
 61 A standard choice in the literature is the *regular tangent space*, defined as

$$62 \quad \text{Tan}_{\mu} \mathcal{P}_2(\mathbb{R}^d) := \overline{\{\nabla\varphi \mid \varphi \in C_c^1(\mathbb{R}^d; \mathbb{R})\}}^{L^2_{\mu}(\mathbb{R}^d; \mathbb{T}\mathbb{R}^d)}.$$

63 This space comes from the study of continuity equations, and may be used to define  
 64 viscosity solutions [17, 23]. However, it appears that  $\text{Tan}_{\mu}$  is too small to contain  
 65 all the directions issued from  $\mu$ , since it does not allow the splitting of mass. One  
 66 could instead consider a *general tangent space*  $\mathbf{Tan}_{\mu}$  build from the geodesics, whose  
 67 definition and properties are delayed to Subsection 2.2. However, using  $\mathbf{Tan}_{\mu}$  does  
 68 not bring any additional smoothness, but significantly complexifies the manipulation.  
 69 Hence some authors adopted the strategy to restrict by penalization to regular mea-  
 70 sures  $\mu$  where  $\text{Tan}_{\mu}$  and  $\mathbf{Tan}_{\mu}$  coincide, to obtain comparison principles; see [16].

71 In this work, we consider *directional derivatives* as our available infinitesimal  
 72 information over the variations of a function. This corresponds to a step back in  
 73 the theory of partial differential equations: instead of considering equations over  
 74 a gradient  $\nabla\varphi$  in some appropriate dual space, we consider an equation over the  
 75 application  $q \mapsto \langle \nabla\varphi, q \rangle$  defined over the tangent directions  $q$ . This point of view allow  
 76 us to avoid altogether gradients and semidifferentials, since we only need to manipulate  
 77 functions  $\varphi$  that are directionally differentiable. The notation  $D_{\mu}V(t, \mu)$  in (1.4) refers  
 78 to the application that to each  $\xi \in \mathbf{Tan}_{\mu} \mathcal{P}_2(\mathbb{R}^d)$ , associates the directional derivative  
 79 of  $V$  at  $(t, \mu)$  in the direction  $\xi$ . The *control Hamiltonian* can be defined over the  
 80 maps  $p : \mathbf{Tan}_{\mu} \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$  as follows

$$81 \quad H(\mu, p) := \sup_{u \in U} -p(\pi^{\mu}(f[u] \# \mu)).$$

82 Here  $\pi^{\mu}$  denotes the projection over  $\mathbf{Tan}_{\mu}$ . Note that whenever  $p(q) = \langle \nabla\varphi, q \rangle$ , the  
 83 Hamiltonian  $H$  coincides with the classical control Hamiltonian. A similar definition

84 of  $D_\mu V$  and  $H$  is used in [21] to study viscosity solutions of general HJ equations in  
 85 CAT(0) spaces, and in [20, 22] to treat the Wasserstein space by using its curvature.

86 The results of [22] include a strong comparison principle in the case of measure-  
 87 dependant dynamics and locally uniformly continuous terminal costs  $\mathfrak{J} : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ .  
 88 To do so, a restrictive notion of upper semicontinuity is introduced, that corresponds  
 89 to the upper semicontinuity of the function  $U : B \mapsto \sup_{\mu \in B} u(\mu)$  in the space of  
 90 nonempty bounded and closed subsets of  $\mathcal{P}_2(\mathbb{R}^d)$  endowed with the Hausdorff dis-  
 91 tance. This semicontinuity is not equivalent to upper semicontinuity in Wasserstein  
 92 nor narrow topology. Although it is clear that semicontinuity in the Wasserstein  
 93 topology is not sufficient to provide a good notion, the question stays open in the case  
 94 of the narrow topology. Indeed, the Wasserstein space is narrowly locally compact,  
 95 and many of the technicalities of [22] could be avoided using this property. This ques-  
 96 tion is the first aim of the present paper: we consider a particularly interesting weak  
 97 topology  $\tau$ , built as the inductive limits of the narrow topology on each Wasserstein  
 98 ball (see Subsection 2.4). Additionally, we consider the case of state constraints at the  
 99 final time, and use ideas from [26] to show, *mutatis mutandis*, that the (discontinuous)  
 100 value function can be characterized by the HJB equation in this case.

101 **1.1. Main contributions and standing assumptions.** The main results of  
 102 the paper are the following. First, assume that the dynamic  $f$  is Lipschitz, that  
 103 the set  $f[U] \subset \mathcal{C}(\mathbb{R}^d; \mathbb{T}\mathbb{R}^d)$  is convex, and that  $\mathfrak{J} : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R} \cup \{+\infty\}$  is proper,  
 104 lower bounded and  $\tau$ -lower semicontinuous. Then the value function  $V$  is the minimal  
 105 supersolution of the HJB equation (1.4) in the sense of Definition 4.1; see Theorem 5.5.  
 106 Second, in the case where  $\mathfrak{J}$  is additionally bounded and  $\tau$ -continuous, we are able to  
 107 prove that  $V$  is actually the unique viscosity solution of (1.4).

108 We furthermore provide a strong comparison principle by a rather direct gener-  
 109 alization of the arguments of [14]. The difficulty then lies in proving that the value  
 110 function is itself  $\tau$ -lower semicontinuous. In particular, we have to restrict to measure-  
 111 independent dynamics: this may be understood with the analogy of the weak topology  
 112 over  $L^2$  spaces, where in general, the composition of a convex function  $\mathfrak{J}$  with the flow  
 113 of an ODE stays convex only if the flow is linear.

114 In the sequel, we make the following standing assumptions.

115 *Hypothesis 1.1* (on the dynamic). The set  $f[U] \subset \mathcal{C}(\mathbb{R}^d; \mathbb{T}\mathbb{R}^d)$  is nonempty,  
 116 convex as a set of functions, and closed in the topology of local uniform convergence.  
 117 Moreover, there exists  $C_f \geq 0$  such that  $|f[u](0)| + \text{Lip}(f[u]) \leq C_f$  for each  $u \in U$ .

118 *Hypothesis 1.2* (on the terminal cost). The function  $\mathfrak{J} : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R} \cup \{+\infty\}$   
 119 is proper, lower bounded and  $\tau$ -lower semicontinuous.

120 The rest of the paper is organized as follows. Section 2 gathers the definition of the  
 121 Wasserstein space and the metric differential structure over it, as well as the topologies  
 122 in use over the dynamic and the state space. In Section 3, we study the trajectories  
 123 of (1.2) and the continuity properties of the reachable sets in the topology  $\tau$ . The  
 124 HJB equation (1.4) is revisited in Section 4, where we define viscosity solutions and  
 125 prove a strong comparison principle. Section 5 is devoted to the link between the  
 126 value function and the HJB equation, and contains our main results.

127 **2. Preliminary material.** If  $X$  and  $Y$  are two measure spaces, the symbol  
 128  $\#$  is used to denote the pushforward operator, which to any probability measure  
 129  $\mu \in \mathcal{P}(X)$  and measurable application  $g : X \rightarrow Y$ , associates another probability  
 130 measure  $g\#\mu \in \mathcal{P}(Y)$  given by  $(g\#\mu)(A) = \mu(g^{-1}(A))$  for any measurable  $A \subset Y$ .

131 **2.1. The Wasserstein space.** Let  $\pi_x, \pi_y : (\mathbb{R}^d)^2 \rightarrow \mathbb{R}^d$  denote the canonical  
 132 projections, i.e.  $\pi_x(a, b) = a$ . Given  $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ , let  $\Gamma(\mu, \nu) \subset \mathcal{P}((\mathbb{R}^d)^2)$   
 133 denote the set of *transport plans*  $\eta = \eta(x, y)$  with first marginal  $\pi_x \# \eta$  equal to  $\mu$ ,  
 134 and second marginal  $\pi_y \# \eta$  equal to  $\nu$ . We say that  $\mu$  has finite second moment  
 135 if  $\int_{x \in \mathbb{R}^d} |x|^2 d\mu(x) < \infty$ , and denote  $\mathcal{P}_2(\mathbb{R}^d)$  the set of such measures. This set is  
 136 endowed with so-called *Wasserstein distance*, defined by

$$137 \quad d_{\mathcal{W}}^2(\mu, \nu) := \inf_{\eta \in \Gamma(\mu, \nu)} \int_{(x, y) \in (\mathbb{R}^d)^2} |x - y|^2 d\eta(x, y).$$

138 The set of optimal transport plans is denoted  $\Gamma_o(\mu, \nu)$ . Notice that the curves  $t \mapsto$   
 139  $\eta_t := ((1-t)\pi_x + t\pi_y) \# \eta$  parametrized by  $\eta \in \Gamma_o(\mu, \nu)$  exactly describe the geodesics  
 140 linking  $\mu$  to  $\nu$ . The squared Wasserstein distance happens to be semiconcave along  
 141 geodesics, i.e. for all  $\sigma \in \mathcal{P}_2(\mathbb{R}^d)$  and  $\eta \in \Gamma_o(\mu, \nu)$ , it follows that

$$142 \quad (2.1) \quad d_{\mathcal{W}}^2(\eta_t, \sigma) \geq (1-t)d_{\mathcal{W}}^2(\mu, \sigma) + td_{\mathcal{W}}^2(\nu, \sigma) - t(1-t)d_{\mathcal{W}}^2(\mu, \nu), \quad \forall t \in [0, 1].$$

143 This curvature property implies the existence of directional derivatives of  $d_{\mathcal{W}}^2(\cdot, \sigma)$   
 144 along geodesics. However geodesics are parameterized over  $[0, 1]$ , and may not be  
 145 extended over  $[0, \infty)$ . This poses a conceptual problem, since the positive multiples  
 146 of directional derivatives may not always be represented as directional derivatives  
 147 over "scaled" directions. To overcome this issue, the definition of directions is rather  
 148 understood through the concept of a tangent cone.

149 **2.2. Tangent and cotangent bundles.** We refer the reader to [18, Chap. 4]  
 150 for details in this section. We denote  $T\mathbb{R}^d := \bigcup_{x \in \mathbb{R}^d} \{x\} \times T_x \mathbb{R}^d$  the tangent bundle of  
 151  $\mathbb{R}^d$ , endowed with the distance  $|(x, v) - (y, w)|^2 = |x - y|^2 + |v - w|^2$ . For the sake of  
 152 notation, when it is clear from the context, we will identify applications  $f : \mathbb{R}^d \rightarrow T\mathbb{R}^d$   
 153 with their second-coordinate applications  $\bar{f}$  defined by  $f(x) = (x, \bar{f}(x))$ .

154 To manipulate tangent directions instead of transport plans, we perform an equiv-  
 155 alent of the change of variable  $(x, y) \rightarrow (x, y - x)$ . Namely, for any  $\eta \in \Gamma(\mu, \nu)$ , let  
 156  $\xi = \xi(x, v) \in \mathcal{P}_2(T\mathbb{R}^d)$  be given by  $\xi := (\pi_x, \pi_y - \pi_x) \# \eta$ .

157 For a given  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ , we write  $\mathcal{P}_2(T\mathbb{R}^d)_\mu := \{\xi \in \mathcal{P}_2(T\mathbb{R}^d) \mid \pi_x \# \xi = \mu\}$ .  
 158 This set can be understood as the largest set of velocities issued from  $\mu$ , that can be  
 159 scaled with the operation  $\lambda \cdot \xi := (\pi_x, \lambda\pi_v) \# \xi$ . Then the curve  $t \mapsto ((1-t)\pi_x + t\pi_y) \# \eta$   
 160 coincides with  $t \mapsto \exp_\mu(t \cdot \xi) := (\pi_x + t\pi_v) \# \xi$ . This curve generalizes the applications  
 161  $t \mapsto x + tv$ , by sending the mass that  $\xi$  puts over  $(x, v)$  to the point  $x + tv$ . The  
 162 exponential map  $\exp_\mu$  admits a partial inverse

$$163 \quad \exp_\mu^{-1}(\nu) := \left\{ \xi \in \mathcal{P}_2(T\mathbb{R}^d) \mid \exp_\mu(\xi) = \nu, \int_{(x, v) \in T\mathbb{R}^d} |v|^2 d\xi = d_{\mathcal{W}}^2(\mu, \nu) \right\}.$$

164 To measure the distance between  $\xi, \zeta \in \mathcal{P}_2(T\mathbb{R}^d)_\mu$ , one introduces a set of 3-plans

$$165 \quad \Gamma_\mu(\xi, \zeta) := \{\alpha = \alpha(x, v, w) \in \mathcal{P}(T^2\mathbb{R}^d) \mid (\pi_x, \pi_v) \# \alpha = \xi, (\pi_x, \pi_w) \# \alpha = \zeta\},$$

166 where  $T^2\mathbb{R}^d := \{(x, v, w) \mid x \in \mathbb{R}^d, v, w \in T_x \mathbb{R}^d\}$ , and the application

$$167 \quad W_\mu : (\mathcal{P}_2(T\mathbb{R}^d)_\mu)^2 \rightarrow \mathbb{R}^+, \quad W_\mu^2(\xi, \zeta) := \inf_{\alpha \in \Gamma_\mu(\xi, \zeta)} \int_{(x, v, w) \in T^2\mathbb{R}^d} |v - w|^2 d\alpha(x, v, w),$$

168 which turns out to be a distance in  $\mathcal{P}_2(\mathbb{T}\mathbb{R}^d)_\mu$ . Let  $0\#\mu \in \mathcal{P}_2(\mathbb{T}\mathbb{R}^d)_\mu$  be the  
169 probability over the tangent space concentrated on the null velocity. We denote

$$170 \quad |\cdot|_\mu : \mathcal{P}_2(\mathbb{T}\mathbb{R}^d)_\mu \rightarrow \mathbb{R}^+, \quad |\xi|_\mu := W_\mu(\xi, 0\#\mu).$$

171 In particular, if  $\xi = f\#\mu$  and  $\zeta = g\#\mu$  for some fields  $f, g \in L^2_\mu(\mathbb{R}^d; \mathbb{T}\mathbb{R}^d)$ , then  
172  $W_\mu(f\#\mu, g\#\mu) = \|f - g\|_{L^2_\mu}$ . The general tangent cone to a measure  $\mu$  is defined by

$$173 \quad \mathbf{Tan}_\mu \mathcal{P}_2(\mathbb{T}\mathbb{R}^d) := \overline{\{\alpha \cdot \xi \in \mathcal{P}_2(\mathbb{T}\mathbb{R}^d) \mid \alpha \in \mathbb{R}^+ \text{ and } \xi \in \exp_\mu^{-1}(\mathcal{P}_2(\mathbb{R}^d))\}}^{W_\mu}.$$

174 It is stable by scaling, and  $(\pi_x, (1-t)\pi_v + t\pi_w)\#\alpha \in \mathbf{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$  for  $t \in [0, 1]$  and  
175  $\alpha \in \Gamma_\mu(\xi, \zeta)$  whenever  $\xi, \zeta \in \mathbf{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$ . The orthogonal projection is well-defined  
176 as the unique application

$$177 \quad \pi^\mu : \mathcal{P}_2(\mathbb{T}\mathbb{R}^d)_\mu \rightarrow \mathbf{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d) \quad \text{such that} \quad W_\mu(\xi, \pi^\mu \xi) = \min_{\zeta \in \mathbf{Tan}_\mu} W_\mu(\xi, \zeta).$$

178 Intuitively, the set  $\mathbf{Tan}_\mu$  represents the set of available directions issued from  $\mu$ , and  
179 lays the path for the metric counterpart of the dual space.

180 **DEFINITION 2.1.** For a function  $p : \mathbf{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ , set

$$181 \quad \|p\|_\mu := \sup \{|p(\xi)| \mid \xi \in \mathbf{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d), |\xi|_\mu = 1\}.$$

182 We define the metric cotangent bundle  $\mathbb{T}$  as  $\mathbb{T} := \bigcup_\mu \{\mu\} \times \mathbb{T}_\mu$ , where

$$183 \quad \mathbb{T}_\mu := \left\{ p : \mathbf{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R} \mid \begin{array}{l} \|p\|_\mu < \infty, \quad p(\lambda\xi) = \lambda p(\xi) \quad \forall \lambda \geq 0, \\ \text{and } p \text{ is Lipschitz w.r.t. } W_\mu. \end{array} \right\}.$$

184 We then recover  $|p(\xi)| \leq \|p\|_\mu |\xi|_\mu$ , for any  $\xi \in \mathbf{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$  and  $p \in \mathbb{T}_\mu$ . El-  
185 ements of  $\mathbb{T}_\mu$  may be built from directional derivatives of sufficiently smooth maps.  
186 As an important example, let  $\mu, \sigma \in \mathcal{P}_2(\mathbb{R}^d)$ . Due to the semiconcavity (2.1), the  
187 application  $[0, 1] \ni h \mapsto \frac{1}{h} (d_{\mathcal{W}}^2(\exp_\mu(h \cdot \xi), \sigma) - d_{\mathcal{W}}^2(\mu, \sigma))$  is bounded and monotone,  
188 so that  $d_{\mathcal{W}}^2(\cdot, \sigma)$  is directionally differentiable along all elements  $\xi \in \mathbf{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$ ,  
189 and there holds

$$190 \quad D_\mu d_{\mathcal{W}}^2(\cdot, \sigma)(\xi) = \inf_{\eta \in \exp_\mu^{-1}(\sigma)} \inf_{\alpha \in \Gamma_\mu(\xi, \eta)} \int_{(x,v,w) \in \mathbb{T}^2 \mathbb{R}^d} \langle v, w \rangle d\alpha(x, v, w).$$

191 It turns out that the squared distance is directionally differentiable along any  $\xi$  in  
192  $\mathcal{P}_2(\mathbb{T}\mathbb{R}^d)_\mu$ , and by [20, Theorem 3.8], there holds  $D_\mu d_{\mathcal{W}}^2(\cdot, \sigma)(\xi) = D_\mu d_{\mathcal{W}}^2(\cdot, \sigma)(\pi^\mu \xi)$   
193 for all  $\xi \in \mathcal{P}_2(\mathbb{T}\mathbb{R}^d)_\mu$ . Moreover,  $\|D_\mu d_{\mathcal{W}}^2(\cdot, \sigma)\|_\mu = 2d_{\mathcal{W}}(\mu, \sigma)$ .

194 **2.3. The topology over the dynamics.** For convenience, denote

$$195 \quad (2.2) \quad |b|_{\text{ucc}} := |b(0)| + \sum_{n \in \mathbb{N}} 2^{-n} \sup_{x \in \mathbb{R}^d} \{|b(x)| \mid \|x\| \leq n\}, \quad \forall b \in \mathcal{C}(\mathbb{R}^d; \mathbb{T}\mathbb{R}^d).$$

196 The topology induced by  $|\cdot|_{\text{ucc}}$  on  $\mathcal{C}(\mathbb{R}^d; \mathbb{T}\mathbb{R}^d)$  is that of the uniform convergence over  
197 compact sets, and  $(\mathcal{C}(\mathbb{R}^d; \mathbb{T}\mathbb{R}^d), |\cdot|_{\text{ucc}})$  is a Banach space. As  $f[U]$  is a set of equiLip-  
198 schitz and equibounded maps (Hypothesis 1.1), it is compact in  $(\mathcal{C}(\mathbb{R}^d; \mathbb{T}\mathbb{R}^d), |\cdot|_{\text{ucc}})$ .  
199 Indeed, by Arzelà-Ascoli, the set  $\{b|_K \mid b \in f[U]\}$  is compact in  $(\mathcal{C}(K; \mathbb{T}\mathbb{R}^d); |\cdot|_{\text{ucc}})$   
200 for any compact  $K$ , and then a diagonal argument proves the claim.

201 LEMMA 2.2 (Weak compactness). *For each nontrivial compact  $I \subset \mathbb{R}$ , the set*  
 202  *$L^1(I; f[U])$  is weakly compact in  $L^1(I; (\mathcal{C}(\mathbb{R}^d; \mathbb{T}\mathbb{R}^d), |\cdot|_{ucc}))$ .*

203 *Proof.* Let  $X := (\mathcal{C}(\mathbb{R}^d; \mathbb{T}\mathbb{R}^d), |\cdot|_{ucc})$ . As  $f[U]$  is nonempty, convex and closed  
 204 in  $X$ , it is weakly closed by Hahn-Banach. As it is weakly closed and compact, it is  
 205 weakly compact by James' Theorem [19, Theorem 5]. Consequently,  $L^1(I; f[U])$  is relatively  
 206 weakly compact by Diestel's Theorem [30, Proposition 7]. Finally,  $L^1(I; f[U])$   
 207 is closed and convex, hence weakly closed, thus weakly compact in  $L^1(I; X)$ .  $\square$

208 **2.4. The topology  $\tau$  over  $\mathcal{P}_2(\mathbb{R}^d)$ .** The set  $\mathcal{P}_2(\mathbb{R}^d)$  may be endowed with  
 209 several topologies, for example the narrow topology and the topology induced by  $d_W$ .  
 210 The main advantage of the narrow topology is that closed balls for the Wasserstein  
 211 distance are compact. However, it does not hold that any narrowly converging sequence  
 212 is bounded w.r.t. the Wasserstein distance. To circumvent this issue, we  
 213 consider another intermediate topology on  $\mathcal{P}_2(\mathbb{R}^d)$ , obtained as the inductive limit  
 214 of the narrow topology induced on each closed ball of radius  $n \in \mathbb{N}$ .

215 DEFINITION 2.3 (Topology  $\tau$  ([18, Definition 2.16])). *For each  $n \in \mathbb{N}$ , let  $K_n :=$*   
 216  *$\overline{\mathcal{B}_W(\delta_0, n)}$  the Wasserstein closed ball centered in  $\delta_0$  of radius  $n$ , and denote  $\tau_n$  the*  
 217 *topology on  $K_n$  induced by the narrow topology. Let  $\iota_n : K_n \rightarrow \mathcal{P}_2(\mathbb{R}^d)$  be the*  
 218 *canonical injection. The topology  $\tau$  is the finest topology on  $\mathcal{P}_2(\mathbb{R}^d)$  that lets each  $\iota_n$*   
 219 *be continuous from  $(K_n, \tau_n)$  to  $(\mathcal{P}_2(\mathbb{R}^d), \tau)$ .*

220 In other words,  $\tau$  is the strict inductive limit of the topologies  $\tau_n$ . Let us collect  
 221 the principal characteristics of  $\tau$  from [18, Definition 2.16].

222 LEMMA 2.4 (Properties of  $(\mathcal{P}_2(\mathbb{R}^d), \tau)$ ). *Let  $\tau$  be given by Definition 2.3.*

- 223 1. *A set  $A \subset \mathcal{P}_2(\mathbb{R}^d)$  is closed in  $\tau$  if and only if each  $A \cap K_n$  is closed in  $\tau_n$ .*
- 224 2. *A sequence  $(\mu_n)_n \subset \mathcal{P}_2(\mathbb{R}^d)$  converges in  $\tau$  towards some  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$  if and*  
 225 *only if  $\mu_n \xrightarrow{n} \mu$  and  $\sup_{n \in \mathbb{N}} d_W(\mu_n, \delta_0) < \infty$ . We then denote  $\mu_n \xrightarrow{\tau} \mu$ .*
- 226 3. *Wasserstein closed balls are compact and sequentially compact in  $\tau$ .*
- 227 4.  *$(\mathcal{P}_2(\mathbb{R}^d), \tau)$  is not first-countable.*
- 228 5. *The squared Wasserstein distance is sequentially  $\tau$ -lower semicontinuous.*

229 Observe that from Item 4, the topology  $\tau$  is not metrizable, and we do not di-  
 230 rectly have that sequential lower semicontinuity is equivalent to lower semicontinuity  
 231 in this topology. However, the class of spaces where these two properties coincide  
 232 (the *Fréchet-Urysohn spaces*) is larger than first-countable spaces, and  $(\mathcal{P}_2(\mathbb{R}^d), \tau)$   
 233 happens to be one of these. In particular, it implies that the squared Wasserstein  
 234 distance is also  $\tau$ -lower semicontinuous.

235 LEMMA 2.5. *Assume that each  $(K_n, \tau_n)$  is Fréchet-Urysohn, and let  $K = \bigcup_n K_n$ .*  
 236 *Then the inductive limit  $(K, \tau)$  is also Fréchet-Urysohn.*

237 *Proof.* Since each closed set is also sequentially closed, it is enough to prove the  
 238 converse in  $(K, \mathcal{P}_2(\mathbb{R}^d))$ . Let  $A \subset K$  be sequentially closed. For each  $n \in \mathbb{N}$ , consider  
 239  $(x_m)_m \subset A_n := A \cap K_n$  a sequence converging in  $\tau_n$  to some  $x \in K_n$ . By the continuity  
 240 of  $\iota_n$ , the sequence  $(\iota_n(x_m))_m \subset A$  converges towards  $\iota_n(x)$ , and as  $A$  is sequentially  
 241 closed,  $\iota_n(x) \in A$ . Then  $A_n$  is sequentially closed, and as  $(K_n, \tau_n)$  is Fréchet-Urysohn,  
 242  $A_n$  is closed in  $\tau_n$ . But Item 1 in Lemma 2.4 implies that  $A$  is closed.  $\square$

243 Consequently, in the sequel, we make no distinction between lower (resp. upper)  
 244 semicontinuity and sequential lower (resp. upper) semicontinuity for the topology  $\tau$ .

### 245 3. Trajectories in the Wasserstein space.

246 **3.1. Existence and properties of the trajectories.** Let  $0 \leq t \leq T < \infty$ ,  
 247  $\nu \in \mathcal{P}_2(\mathbb{R}^d)$  and consider the continuity inclusion

$$248 \quad (3.1) \quad \partial_s \mu_s \in -\operatorname{div}(f[U] \# \mu_s), \quad s \in [t, T], \quad \mu_t = \nu.$$

249 A curve  $(\mu_s^{t,\nu})_{s \in [t, T]}$  is a solution of (3.1) if it is absolutely continuous in  $\mathcal{P}_2(\mathbb{R}^d)$ , and  
 250 if there exists a measurable map  $b \in L^1([t, T]; f[U])$  such that  $\partial_s \mu_s = -\operatorname{div}(b_s \# \mu_s)$   
 251 in the sense of distributions, that is,

$$252 \quad \int_{s \in [0, T]} \int_{x \in \mathbb{R}^d} [\partial_s \varphi(s, x) + \langle \nabla \varphi(s, x), b_s(x) \rangle] d\mu_s(x) = 0 \quad \forall \varphi \in C_c^\infty((t, T) \times \mathbb{R}^d).$$

253 The following result stems from the combination of [5, Theorems 4.2 and 4.5].

254 **PROPOSITION 3.1** (Existence, uniqueness and representation). *Assume Hypoth-*  
 255 *esis 1.1. For each  $(t, \nu) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^d)$ , the set  $\mathcal{S}^{t,\nu} \subset AC([t, T]; \mathcal{P}_2(\mathbb{R}^d))$  of*  
 256 *solutions of (3.1) is nonempty and compact in  $\mathcal{C}([t, T]; \mathcal{P}_2(\mathbb{R}^d))$ . Moreover, it holds*

$$257 \quad (3.2) \quad \mathcal{S}^{t,\nu} = \left\{ (S_s^{t,b} \# \nu)_{s \in [t, T]} \mid b \in L^1([t, T]; f[U]) \right\},$$

258 where for each  $b \in L^1([t, T]; f[U])$ , the semigroup  $S^{t,b} : [t, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  is the unique  
 259 solution of the underlying Cauchy problem

$$260 \quad (3.3) \quad \frac{d}{ds} S_s^{t,b}(x) = b_s(S_s^{t,b}(x)), \quad S_t^{t,b}(x) = x.$$

261 **Remark 3.2.** Denote again  $C_f$  an upper bound over  $|b(0)| + \operatorname{Lip}(b)$  for each  $b \in$   
 262  $f[U]$ . By a Grönwall Lemma, for any  $b \in L^1([t, T]; f[U])$ , the solution of (3.3) satisfies

$$263 \quad (3.4) \quad |S_s^{t,b}(x) - x| \leq (1 + |x|) \left( e^{C_f(s-t)} - 1 \right), \quad \forall s \in [t, T].$$

264 Consequently, for each  $\mu \in \mathcal{S}^{t,\nu}$ , we have  $d_W(\mu_s, \nu) \leq (1 + d_W(\delta_0, \nu)) (e^{C_f(s-t)} - 1)$ .

265 We now turn to qualitative properties of the trajectories in (3.2). First, as the set  
 266 of dynamics  $f[U]$  does not depend on the measure variable, the trajectories enjoy a  
 267 linear structure. More precisely, let  $S^{t,b} : [t, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  be the semigroup solution of  
 268 (3.3) for some  $b \in L^1([0, T]; f[U])$ . Then for each  $\nu \in \mathcal{P}_2(\mathbb{R}^d)$ , the curve  $s \mapsto S_s^{t,b} \# \nu$   
 269 is a solution of (3.1), and for all  $\nu_0, \nu_1 \in \mathcal{P}_2(\mathbb{R}^d)$  and  $t \in [0, 1]$ , there holds

$$270 \quad S_s^{t,b} \# ((1-t)\nu_0 + t\nu_1) = (1-t)S_s^{t,b} \# \nu_0 + tS_s^{t,b} \# \nu_1.$$

271 Here addition and product are understood in the Banach space of measures, that  
 272 is,  $[(1-t)\nu_0 + t\nu_1](A) = (1-t)\nu_0(A) + t\nu_1(A)$  for each measurable  $A \subset \mathbb{R}^d$ . This  
 273 linearity property is the key point to prove the lower semicontinuity of the value  
 274 function in the topology  $\tau$ .

275 **3.2. Continuity properties of reachable sets.** For each  $0 \leq t \leq s \leq T$  and  
 276  $\nu \in \mathcal{P}_2(\mathbb{R}^d)$ , define the *reachable set* of the continuity inclusion via the formula

$$277 \quad (3.5) \quad R_s^{t,\nu} := \{ \mu_s \mid \mu \in \mathcal{S}^{t,\nu} \}.$$

278 We now investigate the behavior of the reachable sets under convergence in  $\tau$ .



279 LEMMA 3.3 (Sequential  $\tau$ -lower semicontinuity). *Assume Hypothesis 1.1. Let*  
 280  *$(t_n, \nu_n)_n \subset [0, T] \times \mathcal{P}_2(\mathbb{R}^d)$  such that  $t_n \rightarrow t \in [0, T]$  and  $\nu_n \xrightarrow{\tau} \nu \in \mathcal{P}_2(\mathbb{R}^d)$ . For any*  
 281  *$\mu \in R_T^{t, \nu}$ , there is  $(\mu_n)_n \subset \mathcal{P}_2(\mathbb{R}^d)$  such that  $\mu_n \in R_T^{t_n, \nu_n}$  and  $\mu_n \xrightarrow{\tau} \mu$ .*

282 *Proof.* Let  $b \in L^1([t, T]; f[U])$  such that  $\mu = S_T^{t, b} \# \nu$ . Define a sequence  $(b_n)_n$  by

$$283 \quad b_n \in L^1([t_n, T]; f[U]), \quad b_n(s) := b_{\max(s, t)} \quad \forall s \in [t_n, T].$$

284 Consider  $\mu_n \in R_T^{t_n, \nu_n}$  given by  $S_T^{t_n, b_n} \# \nu_n$ . Using (3.4), one has

$$285 \quad d_{\mathcal{W}}(\mu_n, \nu_n) \leq \sqrt{\int_{x \in \mathbb{R}^d} \left| S_T^{t_n, b_n}(x) - x \right|^2 d\nu_n} \leq (T - t_n) C_f (1 + d_{\mathcal{W}}(\nu_n, \delta_0)) e^{C_f(T-t_n)}.$$

286 As  $(\nu_n)_n$  is bounded in  $d_{\mathcal{W}}$  by the definition of  $\tau$ -convergence, so is the sequence  
 287  $(\mu_n)_n$ . Moreover, for any  $\varphi \in \mathcal{C}_b(\mathbb{R}^d; \mathbb{R})$ , we have

$$288 \quad |\langle \varphi, \mu_n \rangle - \langle \varphi, \mu \rangle| = \left| \int_{x \in \mathbb{R}^d} \varphi \left( S_T^{t_n, b_n}(x) \right) d\nu_n - \int_{x \in \mathbb{R}^d} \varphi \left( S_T^{t, b}(x) \right) d\nu \right| \leq \mathcal{A}_1^n + \mathcal{A}_2^n,$$

289 where  $\mathcal{A}_1^n := \left| \langle \varphi \circ S_T^{t_n, b_n} - \varphi \circ S_T^{t, b}, \nu_n \rangle \right|$  and  $\mathcal{A}_2^n := \left| \langle \varphi \circ S_T^{t, b}, \nu_n - \nu \rangle \right|$ . On the one  
 290 hand, the composition  $\varphi \circ S_T^{t, b}$  is continuous, so that  $\mathcal{A}_2^n$  goes to 0 when  $n \rightarrow \infty$ . On  
 291 the other hand, for each  $x \in \mathbb{R}^d$  and  $s \in [\max(t, t_n), T]$  we have

$$292 \quad \left| S_s^{t_n, b_n}(x) - S_s^{t, b}(x) \right| = \left| \int_{r=t_n}^s b_{\max(r, t)} \left( S_r^{t_n, b_n}(x) \right) dr - \int_{r=t}^s b_r \left( S_r^{t, b}(x) \right) dr \right|$$

$$293 \quad \leq \int_{r=\min(t_n, t)}^t |b_t \left( S_r^{t_n, b_n}(x) \right)| dr + \int_{r=t}^s C_f |S_r^{t_n, b_n}(x) - S_r^{t, b}(x)| dr$$

$$294 \quad \leq C_f \beta_n(x) |t - t_n| + \int_{r=t}^s C_f |S_r^{t_n, b_n}(x) - S_r^{t, b}(x)| dr,$$

$$295$$

296 where the last inequality follows (3.4), with  $\beta_n(x) := e^{C_f(|t-t_n|)}(1 + |x|)$ . Applying a  
 297 Grönwall Lemma, one has  $\left| S_T^{t_n, b_n}(x) - S_T^{t, b}(x) \right| \leq C_f \beta_n(x) |t - t_n| e^{C_f(T-t)}$ . Therefore,  
 298 if  $\varphi$  is in addition Lipschitz continuous we get

$$299 \quad |\langle \varphi, \mu_n \rangle - \langle \varphi, \mu \rangle| \leq \text{Lip}(\varphi) |t - t_n| C_f e^{C_f(|t-t_n|+(T-t))} (1 + d_{\mathcal{W}}(\nu_n, \delta_0)) + \mathcal{A}_2^n \xrightarrow[n \rightarrow \infty]{} 0.$$

300 By a density argument (e.g. [4, Section 5.1]) we can conclude that  $\langle \varphi, \mu_n \rangle \rightarrow \langle \varphi, \mu \rangle$  for  
 301 any  $\varphi \in \mathcal{C}_b(\mathbb{R}^d; \mathbb{R})$ . Furthermore, since  $(\mu_n)_n$  bounded in  $d_{\mathcal{W}}$ , there holds  $\mu_n \xrightarrow{\tau} \mu$ .  $\square$

302 LEMMA 3.4 (Sequential  $\tau$ -upper semicontinuity). *Assume Hypothesis 1.1. Let*  
 303  *$(t_n, \nu_n)_n \subset [0, T] \times \mathcal{P}_2(\mathbb{R}^d)$  such that  $t_n \rightarrow t \in [0, T]$  and  $\nu_n \xrightarrow{\tau} \nu \in \mathcal{P}_2(\mathbb{R}^d)$ . For*  
 304 *each  $n$ , let  $\omega_n \in R_T^{t_n, \nu_n}$ . Then there is  $\omega \in R_T^{t, \nu}$  and a subsequence so that  $\omega_{n_k} \xrightarrow{\tau_k} \omega$ .*

305 *Proof.* For each  $n$ , let  $\mu^n \in S^{t_n, \nu_n}$  such that  $\mu_T^n = \omega_n$ , and denote  $b^{0, n} \in$   
 306  $L^1([t_n, T]; f[U])$  a driving field for  $\mu^n$ . Using (3.4), there holds

$$307 \quad d_{\mathcal{W}}(\omega_n, \nu_n) \leq \sqrt{\int_{x \in \mathbb{R}^d} \left| S_T^{t_n, b_n}(x) - x \right|^2 d\nu_n} \leq (T - t_n) C_f \sqrt{1 + d_{\mathcal{W}}^2(\nu_n, \delta_0)} e^{C_f(T-t_n)}.$$

308 As the convergence of  $(\nu_n)_n$  in the topology  $\tau$  implies that  $d_{\mathcal{W}}^2(\nu_n, \delta_0)$  is bounded in-  
 309 dependently of  $n$ , we deduce that  $(\omega_n)_n$  stays in a Wasserstein ball. From Lemma 2.4,

310 Wasserstein balls are sequentially compact in  $\tau$ , so that up to a non relabeled subse-  
 311 quence,  $\omega_n \xrightarrow{\tau} \omega$  for some  $\omega \in \mathcal{P}_2(\mathbb{R}^d)$ . Stays to prove that  $\omega \in R_T^{t,\nu}$ . We divide the  
 312 rest of the proof into several parts.

313 **Extracting a dynamic** Let  $b^n \in L^1([t, T]; f[U])$  defined by  $b_s^n = b_{\max(s, t_n)}^{0, n}$  for any  
 314  $s \in [t, T]$ . Denote  $(S_s^{t, b^n})_{s \in [t, T]}$  the semigroup related to the dynamic  $b^n$ . Setting  
 315  $\beta_n(x) := e^{C_f(|t-t_n|)}(1 + |x|)$ , one has by (3.4) that

$$316 \quad |S_s^{t, b^n}(x) - S_s^{t_n, b^n}(x)| \leq C_f \beta_n(x) |t - t_n| + C_f \int_{\max(t, t_n)}^s |S_r^{t, b^n}(x) - S_r^{t_n, b^n}(x)| dr,$$

317 so that by a Grönwall Lemma, for each  $s \in [\max(t, t_n), T]$ , there holds

$$318 \quad (3.6) \quad \left| S_s^{t, b^n}(x) - S_s^{t_n, b^n}(x) \right| \leq C_f \beta_n(x) |t - t_n| e^{C_f(s - \max(t, t_n))}.$$

319 Recall that  $f[U]$  is endowed with the topology of uniform convergence on compact  
 320 sets. By Lemma 2.2,  $L^1([t, T]; f[U])$  is weakly compact in  $L^1(I; X)$ , where  $X$  equals  
 321  $(\mathcal{C}(\mathbb{R}^d; \mathbb{T}\mathbb{R}^d), |\cdot|_{\text{ucc}})$  and  $|\cdot|_{\text{ucc}}$  is defined in (2.2). Then, up to a further (non relabeled)  
 322 subsequence,  $b^n \xrightarrow{\tau} b$  for some  $b \in L^1(I; f[U])$ . Let us show that  $\omega = S_T^{t, b} \# \nu$ , or  
 323 equivalently, that

$$324 \quad (3.7) \quad \langle \varphi, S_T^{t, b} \# \nu \rangle = \langle \varphi, \omega \rangle, \quad \forall \varphi \in C_c^\infty(\mathbb{R}^d; \mathbb{R}).$$

325 **Estimates** Let  $\varphi \in C_c^\infty(\mathbb{R}^d; \mathbb{R})$ , and denote  $\text{Lip}(\varphi)$  its Lipschitz constant. As the  
 326 convergence in  $\tau$  implies weak convergence, we have that  $\langle \varphi, \omega \rangle = \lim_{n \rightarrow \infty} \langle \varphi, \omega_n \rangle =$   
 327  $\lim_{n \rightarrow \infty} \langle \varphi, S_T^{t_n, b^n} \# \nu_n \rangle$ . For each fix  $n$ , decompose

$$328 \quad (3.8) \quad \left| \langle \varphi, S_T^{t, b} \# \nu \rangle - \langle \varphi, S_T^{t_n, b^n} \# \nu_n \rangle \right| \leq \left| \langle \varphi \circ S_T^{t, b}, \nu - \nu_n \rangle \right| + \mathcal{A}_1^n + \mathcal{A}_2^n,$$

329 where

$$330 \quad \mathcal{A}_1^n := \left| \langle \varphi, S_T^{t, b} \# \nu_n - S_T^{t_n, b^n} \# \nu_n \rangle \right| \quad \text{and} \quad \mathcal{A}_2^n := \left| \langle \varphi, S_T^{t, b^n} \# \nu_n - S_T^{t_n, b^n} \# \nu_n \rangle \right|.$$

331 As  $x \mapsto S_T^{t, b}(x)$  is continuous, then  $\varphi \circ S_T^{t, b}$  belongs to  $\mathcal{C}_b(\mathbb{R}^d; \mathbb{R})$ , and the convergence  
 332  $\nu_n \xrightarrow{\tau} \nu$  ensures that  $\langle \varphi \circ S_T^{t, b}, \nu - \nu_n \rangle \rightarrow 0$ . Moreover, using (3.6) we get

$$333 \quad \mathcal{A}_2^n \leq \text{Lip}(\varphi) |t - t_n| C_f e^{C_f(|t-t_n| + (T - \max(t, t_n)))} (1 + d_{\mathcal{W}}(\nu_n, \delta_0)) \xrightarrow{n \rightarrow \infty} 0.$$

334 We turn to  $\mathcal{A}_1^n$ . Let  $\Omega \subset \mathbb{R}^d$  be the fattened compact

$$335 \quad \Omega := \left\{ x \in \mathbb{R}^d \mid \exists y \in \text{supp}(\varphi), |x - y| \leq TC_f \left( 1 + \sup_{z \in \text{supp}(\varphi)} |z| \right) e^{C_f T} \right\}.$$

336 By (3.4),  $\Omega$  contains every trajectory  $s \mapsto S_s^{t, \beta}(x)$  such that  $S_T^{t, \beta}(x) \in \text{supp}(\varphi)$  for  
 337 some  $\beta \in L^1([t, T]; f[U])$ . As  $\varphi$  vanishes outside its support,

$$338 \quad \mathcal{A}_1^n \leq \text{Lip}(\varphi) \int_{x \in \Omega} \left| S_T^{t, b}(x) - S_T^{t_n, b^n}(x) \right| d\nu_n(x).$$

339 The application  $\psi : s \mapsto |S_s^{t,b}(x) - S_s^{t,b^n}(x)|$  satisfies

$$340 \quad (3.9) \quad \psi(T) \leq C_f \int_{s=t}^T \psi(s) ds + \left| \int_{s=t}^T b(S_s^{t,b}(x)) ds - \int_{s=t}^T b^n(S_s^{t,b}(x)) ds \right|.$$

341 Since  $S_s^{t,b}(x) \in \Omega$  for all  $s \in [t, T]$ , the linear map  $\beta \mapsto \int_{s=t}^T \beta(S_s^{t,b}(x)) ds$  is continuous  
 342 in  $L^1([t, T]; f[U])$ . By weak convergence,  $\left| \int_{s=t}^T b(S_s^{t,b}(x)) ds - \int_{s=t}^T b^n(S_s^{t,b}(x)) ds \right| =:$   
 343  $\varepsilon_n(x) \xrightarrow[n \rightarrow \infty]{} 0$ , and using a Grönwall Lemma on (3.9), we obtain that

$$344 \quad \mathcal{A}_1^n \leq \text{Lip}(\varphi) \int_{x \in \Omega} \varepsilon_n(x) e^{C_f T} d\nu_n(x).$$

345 Using the compactness of  $\Omega$ , the application  $\varepsilon_n(x)$  is uniformly bounded and converges  
 346 pointwise to 0, so that by Lebesgue dominated convergence theorem,  $\mathcal{A}_1^n \xrightarrow[n \rightarrow \infty]{} 0$ . From  
 347 here, we conclude that the right hand-side of (3.8) vanishes with  $n$ , proving the claim.  $\square$

348 **3.3. Approximation along a subsequence.** Recall that  $\exp_\mu(h \cdot b \# \mu) =$   
 349  $(\pi_x + hb \circ (\pi_x)) \# \mu$  for each  $b \in f[U]$ . Since the trajectories of the controlled system  
 350 may lack  $C^1$  regularity in time, we are not allowed to linearize them. However,  
 351 we can still approximate a trajectory issued from  $\mu$  by a "linear" curve  $h \mapsto \exp_\mu(h \cdot b)$   
 352 along some given subsequence.

353 **LEMMA 3.5 (Sequential approximation).** *Assume Hypothesis 1.1. Let  $(\mu_s)_{s \in [t, T]}$*   
 354 *be a solution of (3.1). Then there exists  $\bar{b} \in f[U]$  and a sequence  $(h_n)_{n \in \mathbb{N}}$  such that*

$$355 \quad \lim_{n \rightarrow \infty} \frac{d_{\mathcal{W}}(\mu_{t+h_n}, \exp_{\mu_t}(h_n \cdot \bar{b} \# \mu_t))}{h_n} = 0.$$

356 *Proof.* Let  $b \in L^1([t, T]; f[U])$  and  $(\mu_s)_{s \in [t, T]}$  such that  $\partial_s \mu_s + \text{div}(b_s \# \mu_s) = 0$ .  
 357 For  $h > 0$  such that  $t + h \leq T$ , consider

$$358 \quad \bar{b}^h(x) := \int_{s=t}^{t+h} b_s(x) ds = \frac{1}{h} \int_{s=t}^{t+h} b_s(x) ds, \quad \forall x \in \mathbb{R}^d.$$

359 Here  $\bar{b}^h \in f[U]$  by convexity. By compactness of  $f[U]$  in the topology of uniform  
 360 convergence over compact sets, there exists a sequence  $(h_n)_n \searrow 0$  and some  $\bar{b} \in f[U]$   
 361 such that  $\bar{b}^{h_n}$  converges uniformly over compact sets towards  $\bar{b}$ . From (3.4) we get

$$362 \quad \frac{1}{h_n} d_{\mathcal{W}}(\mu_{t+h_n}, \exp_{\mu_t}(h_n \cdot \bar{b} \# \mu_t)) \leq \frac{1}{h_n} \sqrt{\int_{x \in \mathbb{R}^d} |S_{t+h_n}^{t,b}(x) - (x + h_n \bar{b}(x))|^2 d\mu_t}$$

$$363 \quad = \sqrt{\int_x \left| \int_{s=t}^{t+h_n} b_s(S_s^{t,b}(x)) ds - \bar{b}(x) \right|^2 d\mu_t}$$

$$364 \quad \leq \sqrt{\int_{x \in \mathbb{R}^d} \left| \int_{s=t}^{t+h_n} b_s(x) ds - \bar{b}(x) \right|^2 d\mu_t} + C_f \sqrt{\int_{x \in \mathbb{R}^d} \left| \int_{s=t}^{t+h_n} S_s^{t,b}(x) ds - x \right|^2 d\mu_t}$$

$$365 \quad \leq \sqrt{\int_{x \in \mathbb{R}^d} |\bar{b}^{h_n}(x) - \bar{b}(x)|^2 d\mu_t} + C_f (e^{h_n} - 1) (1 + d_{\mathcal{W}}(\mu_t, \delta_0)).$$

366

367 Let  $\varepsilon > 0$ . Since  $\mu_t \in \mathcal{P}_2(\mathbb{R}^d)$ , there exists  $R \geq 0$  large enough so that

$$368 \quad \int_{|x|>R} \left| \bar{b}^{h_n}(x) - \bar{b}(x) \right|^2 d\mu_t(x) \leq \int_{|x|>R} (2C_f(1+|x|))^2 d\mu_t(x) \leq \varepsilon^2.$$

369

370 On the compact  $\overline{\mathcal{B}(0, R)}$ , the convergence  $\bar{b}^{h_n} \rightarrow \bar{b}$  is uniform with a modulus denoted  
 371  $\omega_R$ . Summarizing the above, we have

$$372 \quad \frac{d_{\mathcal{W}}\left(\mu_{t+h_n}, \exp_{\mu_t}(h_n \cdot \bar{b} \# \mu_t)\right)}{h_n} \leq (\varepsilon^2 + \omega_R^2(h_n))^{1/2} + C_f(e^{h_n} - 1)(1 + d_{\mathcal{W}}(\mu_t, \delta_0)).$$

373 Taking the limsup in  $n \rightarrow \infty$ , we get that

$$374 \quad \overline{\lim}_{n \rightarrow \infty} \frac{d_{\mathcal{W}}\left(\mu_{t+h_n}, \exp_{\mu_t}(h_n \cdot \bar{b} \# \mu_t)\right)}{h_n} \leq \varepsilon,$$

375 and  $\varepsilon > 0$  being arbitrary, we conclude.  $\square$

376 **4. The Hamilton-Jacobi equation.** Let  $H : \mathbb{T} \rightarrow \mathbb{R}$ , and consider the HJ  
 377 equation

$$378 \quad (4.1) \quad -\partial_t v(t, \mu) + H(\mu, D_\mu v(t, \mu)) = 0 \quad t \in (0, T), \quad v(T, \cdot) = \mathfrak{J}.$$

379 This section is devoted to the notion of solution adapted to (4.1). We first introduce  
 380 a definition of viscosity solutions using test functions, and then prove a comparison  
 381 principle that implies the uniqueness of the viscosity solution of (4.1).

382 **4.1. Definition of viscosity solutions.** We employ two distinct sets of test  
 383 functions for the sub and supersolutions. Denote

$$384 \quad \mathcal{T}_\pm := \left\{ (t, \mu) \mapsto \psi(t) \pm \sum_{i=1}^N \alpha_i d_{\mathcal{W}}^2(\mu, \nu_i) \left| \begin{array}{l} \psi \in \mathcal{C}^1((0, T); \mathbb{R}), N \in \mathbb{N}, \\ \text{and } (\alpha_i, \nu_i)_{i \in [1, N]} \subset \mathbb{R}^+ \times \mathcal{P}_2(\mathbb{R}^d). \end{array} \right. \right\}$$

385 In particular, test functions in  $\mathcal{T}_+$  are  $\tau$ -lower semicontinuous, locally Lipschitz and  
 386 directionally differentiable everywhere, and  $\mathcal{T}_- = -\mathcal{T}_+$ . As each term of the finite  
 387 sum of the measure component is directionally differentiable, so is each  $\varphi(t, \cdot)$  for  
 388  $\varphi \in \mathcal{T}_\pm$ , and there holds  $D_\mu \varphi(t, \cdot)(\xi) = \pm \sum_{i=1}^N \alpha_i D_\mu d_{\mathcal{W}}^2(\cdot, \nu_i)(\xi)$ .

389 We consider the following definition.

390 **DEFINITION 4.1 (Viscosity solution).**  $v : [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R} \cup \{\pm\infty\}$  is called  
 391 – a viscosity subsolution of (4.1) if it is  $\tau$ -upper semicontinuous, does not take  
 392 the value  $+\infty$ ,  $v(T, \mu) \leq \mathfrak{J}(\mu)$ , and for each  $\varphi \in \mathcal{T}_+$  such that  $v - \varphi$  reaches  
 393 a finite maximum in  $(t, \mu) \in (0, T) \times \mathcal{P}_2(\mathbb{R}^d)$ , there holds

$$394 \quad (4.2) \quad -\partial_t \varphi(t, \mu) + H(\mu, D_\mu \varphi(t, \mu)) \leq 0.$$

395 – a viscosity supersolution of (4.1) if it is  $\tau$ -lower semicontinuous, does not  
 396 take the value  $-\infty$ ,  $v(T, \mu) \geq \mathfrak{J}(\mu)$ , and for each  $\varphi \in \mathcal{T}_-$  such that  $v - \varphi$   
 397 reaches a finite minimum in  $(t, \mu) \in (0, T) \times \mathcal{P}_2(\mathbb{R}^d)$ , there holds

$$398 \quad (4.3) \quad -\partial_t \varphi(t, \mu) + H(\mu, D_\mu \varphi(t, \mu)) \geq 0.$$

399 – a viscosity solution of (4.1) if it is both a sub and supersolution.

400 **4.2. Comparison principle.** The comparison principle is the key result in the  
 401 viscosity theory. It essentially says that subsolutions are always smaller than supersol-  
 402 utions in the pointwise sense. This gives the uniqueness of the viscosity solution, and  
 403 in the classical theory, also allows to obtain existence for general nonconvex Hamil-  
 404 tonians. Owing to the local compactness of  $(\mathcal{P}_2(\mathbb{R}^d), \tau)$ , our strategy to obtain a  
 405 comparison principle is quite close to that of [14]. We begin by the adaptation of [14,  
 406 Proposition 3.7] in our case.

407 **LEMMA 4.2** (Penalization lemma). *Let  $(X, \Theta)$  be a topological space,  $\mathcal{O} \subset X$*   
 408 *be nonempty,  $\Phi : \mathcal{O} \rightarrow \mathbb{R} \cup \{-\infty\}$  be  $\Theta$ -upper semicontinuous and proper in  $\mathcal{O}$ ,*  
 409  *$\Psi : \mathcal{O} \rightarrow \mathbb{R}$  be  $\Theta$ -lower semicontinuous and nonnegative. For any  $a > 0$ , set  $\Gamma_a :=$*   
 410  *$\sup_{x \in \mathcal{O}} [\Phi(x) - a\Psi(x)]$ . Assume that  $-\infty < \lim_{a \rightarrow \infty} \Gamma_a < \infty$ , and let  $x_a \in \mathcal{O}$  be chosen such*  
 411 *that  $\lim_{a \rightarrow \infty} (\Gamma_a - (\Phi(x_a) - a\Psi(x_a))) = 0$ . Then the following holds:*

- 412 1.  $\lim_{a \rightarrow \infty} a\Psi(x_a) = 0$ ,  
 413 2. whenever  $\hat{x} \in \mathcal{O}$  is a limit point of  $(x_a)_a$  in  $(X, \Theta)$ , then  $\Psi(\hat{x}) = 0$  and  
 414  $\lim_{a \rightarrow \infty} \Gamma_a = \Phi(\hat{x}) = \sup_{\Psi(x)=0} \Phi(x)$ .

415 *Proof.* Let

$$416 \quad \varepsilon_a := \Gamma_a - (\Phi(x_a) - a\Psi(x_a)),$$

417 so that  $\lim_{a \rightarrow \infty} \varepsilon_a = 0$ . Since  $\Psi \geq 0$ , the map  $a \mapsto \Gamma_a$  decreases when  $a$  increases,  
 418 and  $\lim_{a \rightarrow \infty} \Gamma_a$  exists and is finite. Furthermore,

$$419 \quad \Gamma_{a/2} \geq \Phi(x_a) - \frac{a}{2}\Psi(x_a) \geq \Phi(x_a) - a\Psi(x_a) + \frac{a}{2}\Psi(x_a) = \Gamma_a - \varepsilon_a + \frac{a}{2}\Psi(x_a),$$

420 which implies that  $a\Psi(x_a) \leq 2(\varepsilon_a + \Gamma_{a/2} - \Gamma_a)$ , hence  $\lim_{a \rightarrow \infty} a\Psi(x_a) = 0$ .

421 Suppose now that  $a_n \rightarrow \infty$  and  $x_{a_n} \rightarrow_n \hat{x} \in \mathcal{O}$ . Then  $\lim_{a_n \rightarrow \infty} \Psi(x_{a_n}) = 0$ , and  
 422 by lower semicontinuity,  $\Psi(\hat{x}) = 0$ . Moreover, since

$$423 \quad \Phi(x_{a_n}) - a_n\Psi(x_{a_n}) = \Gamma_{a_n} - \varepsilon_{a_n} \geq \sup_{\Psi(x)=0} \Phi(x) - \varepsilon_{a_n},$$

424 and  $\Phi$  is upper semicontinuous, the result holds.  $\square$

425 The comparison principle will rely on the next assumptions on the Hamiltonian.

426 **Hypothesis 4.3** (Structure of the Hamiltonian). Assume that there exists a con-  
 427 stant  $C_H \geq 0$  such that for all  $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$  and  $p, q \in \mathbb{T}_\mu$ ,

$$428 \quad (4.4) \quad |H(\mu, p + q) - H(\mu, p)| \leq C_H (1 + d_{\mathcal{W}}(\mu, \delta_0)) \|p\|_\mu$$

429 and for all  $a \geq 0$ ,

$$430 \quad (4.5) \quad H(\mu, -aD_\mu d_{\mathcal{W}}^2(\cdot, \nu)) - H(\nu, aD_\mu d_{\mathcal{W}}^2(\mu, \cdot)) \leq aC_H d_{\mathcal{W}}^2(\mu, \nu).$$

431 The fact that the Hamiltonian issued from the control problem (1.3) satisfies  
 432 Hypothesis 4.3 is proved in [22, Section 6, Lemmata 6 and 7]. We are now in a  
 433 position to state our comparison principle.

434 **PROPOSITION 4.4** (Comparison principle). *Assume Hypotheses 1.1, 1.2, and 4.3.*  
 435 *Let  $v : [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R} \cup \{-\infty\}$  be a subsolution of (4.1) bounded from above,*  
 436 *and  $w : [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R} \cup \{+\infty\}$  be a supersolution of (4.1) bounded from below.*  
 437 *Assume that there exists  $\sigma \in \mathcal{P}_2(\mathbb{R}^d)$  such that  $v(T, \sigma) - w(T, \sigma) \in \mathbb{R}$ . Then*

$$438 \quad \Gamma := \sup_{(t, \mu) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^d)} (v(t, \mu) - w(t, \mu)) \leq \sup_{\mu \in \mathcal{P}_2(\mathbb{R}^d)} (v(T, \mu) - w(T, \mu)) =: \Gamma_T.$$

439 *Proof.* By assumption,  $\Gamma_T$  and  $\Gamma$  are finite. Denote  $\lceil v \rceil, \lceil -w \rceil$  upper bounds on  
 440  $v$  and  $-w$ . Up to replacing  $v$  by  $v - \Gamma_T$ , we may assume that  $\Gamma_T = 0$ . Assume by  
 441 contradiction that  $\Gamma > 0$ . Consider

$$442 \quad v^\alpha(t, \mu) := v(t, \mu) + \alpha(t - T)$$

443 for some  $\alpha > 0$  small enough so that

$$444 \quad \Gamma^\alpha := \sup_{(t, \mu) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^d)} (v^\alpha(t, \mu) - w(t, \mu)) > 0.$$

445 Let now  $\sigma \in \mathcal{P}_2(\mathbb{R}^d)$  be as in the statement, and  $\varepsilon_0 > 0$  so that for all  $0 < \varepsilon < \varepsilon_0$ ,

$$446 \quad \Gamma_\varepsilon^\alpha := \sup_{(t, \mu) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^d)} \left( v^\alpha(t, \mu) - w(t, \mu) - 2\varepsilon \left( d_{\mathcal{W}}^2(\sigma, \mu) + \frac{1}{t} \right) \right) > 0.$$

447 The sequence  $(\Gamma_\varepsilon^\alpha)_\varepsilon$  is uniformly bounded, nondecreasing when  $\varepsilon \searrow 0$  and converges  
 448 towards  $\Gamma^\alpha$ . For each  $\varepsilon, a > 0$ , let

$$449 \quad \Phi_{\varepsilon, a}((t, \mu), (s, \nu)) := v^\alpha(t, \mu) - w(s, \nu) - a \frac{d_{\mathcal{W}}^2(\mu, \nu) + |t - s|^2}{2}$$

$$450 \quad - \varepsilon \left( d_{\mathcal{W}}^2(\sigma, \mu) + d_{\mathcal{W}}^2(\sigma, \nu) + \frac{1}{t} + \frac{1}{s} \right).$$

451 The proof involves taking subsequences and diagonal sequences in  $\varepsilon$  and  $a$ . In  
 452 order to lighten the notation, let  $I_\varepsilon := \mathbb{R}^+$  for all  $\varepsilon > 0$ , and  $I := \bigcup_{0 < \varepsilon < \varepsilon_0} I_\varepsilon$  a set of  
 453 indexes that will be refined further on. For a fixed  $\varepsilon$ , we denote  $z_{\varepsilon, a} \xrightarrow{a \in I_\varepsilon} z_\varepsilon$  if  $z_\varepsilon$  is  
 454 the limit of the family  $(z_{\varepsilon, a})_{a \in I_\varepsilon}$  when  $a \rightarrow \infty$ . We divide the rest of the proof into  
 455 several parts.

456 **Point of maximum** Notice that if  $\Phi_{\varepsilon, a}((t, \mu), (s, \nu)) \geq \Phi_{\varepsilon, a}((T, \sigma), (T, \sigma))$ , then

$$457 \quad \varepsilon \left( d_{\mathcal{W}}^2(\sigma, \mu) + d_{\mathcal{W}}^2(\sigma, \nu) \right) \leq \lceil v^\alpha \rceil + \lceil -w \rceil + \frac{2\varepsilon}{T} - (v^\alpha(T, \sigma) - w(T, \sigma)) < \infty.$$

458 Then there exists  $R_\varepsilon > 0$  such that  $\{\Phi_{\varepsilon, a} \geq \Phi_{\varepsilon, a}((T, \sigma), (T, \sigma))\} \subset \overline{\mathcal{B}}((T, \sigma), R_\varepsilon)^2$ .  
 459 As balls of  $[0, T] \times \mathcal{P}_2(\mathbb{R}^d) \times [0, T] \times \mathcal{P}_2(\mathbb{R}^d)$  are compact in the product topology  
 460  $\mathcal{B}_{[0, T]} \times \tau \times \mathcal{B}_{[0, T]} \times \tau$ , and  $\Phi_{\varepsilon, a}$  is proper, upper bounded and upper semicontinuous in  
 461 this topology, there exists a maximum point  $z_{\varepsilon, a} := (t_{\varepsilon, a}, \mu_{\varepsilon, a}, s_{\varepsilon, a}, \nu_{\varepsilon, a})$  of  $\Phi_{\varepsilon, a}$  over  
 462 its domain. As  $R_\varepsilon$  is independant of  $a$ , we may extract a subsequence of  $a$  such that  
 463  $z_{\varepsilon, a}$  converges to some  $z_\varepsilon \in ([0, T] \times \mathcal{P}_2(\mathbb{R}^d))^2$ . Redefining each  $I_\varepsilon$  to only keep the  
 464 indexes of the said subsequence, we may assume that

$$465 \quad z_{\varepsilon, a} \xrightarrow{a \in I_\varepsilon} z_\varepsilon \quad \text{in } (\mathcal{B}_{[0, T]} \times \tau)^2, \quad \text{and} \quad \lim_{a \in I_\varepsilon} \Phi_{\varepsilon, a}(z_{\varepsilon, a}) \quad \text{exists.}$$

466 Applying Lemma 4.2, we get that possibly along a further refinement of  $I$ ,

$$467 \quad (4.6) \quad \lim_{a \in I_\varepsilon} a \left( d_{\mathcal{W}}^2(\mu_{\varepsilon, a}, \nu_{\varepsilon, a}) + |t_{\varepsilon, a} - s_{\varepsilon, a}|^2 \right) = 0 \quad \text{and} \quad \lim_{a \in I_\varepsilon} \Phi_{\varepsilon, a}(z_{\varepsilon, a}) = \Gamma_\varepsilon^\alpha.$$

468 **Staying away from the boundary** By construction,  $t_{\varepsilon, a} > 0$  and  $s_{\varepsilon, a} > 0$  for  
 469 each  $(\varepsilon, a) \in I$ . On the other hand, for each  $\varepsilon$ , there exists  $a_\varepsilon$  large enough so that

472  $t_{\varepsilon,a}, s_{\varepsilon,a} < T$  for all  $a \in I_\varepsilon \cap [a_\varepsilon, \infty)$ . Indeed, if it was not the case, then there would  
 473 exist  $(a_\varepsilon^n)_n \subset I_\varepsilon$  going to  $+\infty$  with  $n \rightarrow \infty$  such that  $T \in \{t_{\varepsilon,a_\varepsilon^n}, s_{\varepsilon,a_\varepsilon^n}\}$ . Since by  
 474 (4.6),  $|t_{\varepsilon,a_\varepsilon^n} - s_{\varepsilon,a_\varepsilon^n}|$  vanishes when  $n \rightarrow \infty$ , using the upper semicontinuity of the  
 475 semisolutions, we would have

$$476 \Gamma_\varepsilon^\alpha = \overline{\lim}_{n \rightarrow \infty} \Phi_{\varepsilon,a_\varepsilon^n}(z_{\varepsilon,a_\varepsilon^n}) \leq \overline{\lim}_{n \rightarrow \infty} v^\alpha(t_{\varepsilon,a_\varepsilon^n}, \mu_{\varepsilon,a_\varepsilon^n}) - w(s_{\varepsilon,a_\varepsilon^n}, \nu_{\varepsilon,a_\varepsilon^n}) \leq v^\alpha(T, z_\varepsilon) - w(T, z_\varepsilon).$$

477 This is absurd because  $\Gamma_\varepsilon^\alpha > 0$  and  $v^\alpha(T, z_\varepsilon) - w(T, z_\varepsilon) \leq 0$ , and we get that  
 478  $(t_{\varepsilon,a}, s_{\varepsilon,a}) \in (0, T)^2$  for  $a$  large enough. Up to refining the index set, we may as-  
 479 sume that this holds for all  $(\varepsilon, a) \in I$ .

480 **Applying the definition of semisolutions** For each  $(\varepsilon, a) \in I$ , define

$$481 \varphi(t, \mu) := \frac{a}{2} \left( d_{\mathcal{W}}^2(\mu, \nu_{\varepsilon,a}) + |t - s_{\varepsilon,a}|^2 \right) + \varepsilon \left( d_{\mathcal{W}}^2(\sigma, \mu) + \frac{1}{t} \right) - \alpha t,$$

$$482 \psi(s, \nu) := -\frac{a}{2} \left( d_{\mathcal{W}}^2(\mu_{\varepsilon,a}, \nu) + |t_{\varepsilon,a} - s|^2 \right) - \varepsilon \left( d_{\mathcal{W}}^2(\sigma, \nu) + \frac{1}{s} \right).$$

484 Since the squared Wasserstein distance is semiconcave,  $\varphi \in \mathcal{T}_+$  and  $\psi \in \mathcal{T}_-$ . As  
 485  $u - \varphi$  reaches a maximum in  $(t_{\varepsilon,a}, \mu_{\varepsilon,a}) \in (0, T) \times \mathcal{P}_2(\mathbb{R}^d)$ , applying the definition of  
 486 subsolution, we get

$$487 \alpha + \frac{\varepsilon}{t_{\varepsilon,a}^2} - a(t_{\varepsilon,a} - s_{\varepsilon,a}) + H\left(\mu_{\varepsilon,a}, \frac{a}{2} D_{\mu_{\varepsilon,a}} d_{\mathcal{W}}^2(\cdot, \nu_{\varepsilon,a}) + \varepsilon D_{\mu_{\varepsilon,a}} d_{\mathcal{W}}^2(\sigma, \cdot)\right) \leq 0.$$

488 Using that  $\frac{\varepsilon}{t_{\varepsilon,a}^2} \geq 0$ , the assumption (4.4) on the Hamiltonian and the estimate  
 489  $\|D_\mu d_{\mathcal{W}}^2(\cdot, \nu)\|_\mu \leq 2d_{\mathcal{W}}(\mu, \nu)$ , we get

$$490 (4.7) \quad \alpha - a(t_{\varepsilon,a} - s_{\varepsilon,a}) + H\left(\mu_{\varepsilon,a}, \frac{a}{2} D_{\mu_{\varepsilon,a}} d_{\mathcal{W}}^2(\cdot, \nu_{\varepsilon,a})\right) \\ 491 \quad - 2\varepsilon d_{\mathcal{W}}(\sigma, \mu_{\varepsilon,a}) C_H(1 + d_{\mathcal{W}}(\delta_0, \mu_{\varepsilon,a})) \leq 0.$$

492 Similarly,  $w - \psi$  reaches a minimum in  $(s_{\varepsilon,a}, \nu_{\varepsilon,a})$ . Using the same reasoning as above,

$$493 (4.8) \quad a(s_{\varepsilon,a} - t_{\varepsilon,a}) + H\left(\nu_{\varepsilon,a}, -\frac{a}{2} D_{\nu_{\varepsilon,a}} d_{\mathcal{W}}^2(\mu_{\varepsilon,a}, \cdot)\right) \\ 494 \quad + 2\varepsilon d_{\mathcal{W}}(\sigma, \nu_{\varepsilon,a}) C_H(1 + d_{\mathcal{W}}(\delta_0, \nu_{\varepsilon,a})) \geq 0.$$

495 Combining (4.7) and (4.8) and using the assumption (4.5), there holds  $\forall (\varepsilon, a) \in I$

$$496 (4.9) \quad \alpha \leq a C_H d_{\mathcal{W}}^2(\mu_{\varepsilon,a}, \nu_{\varepsilon,a}) + 2\varepsilon C_H \sum_{\varpi \in \{\mu_{\varepsilon,a}, \nu_{\varepsilon,a}\}} d_{\mathcal{W}}(\sigma, \varpi)(1 + d_{\mathcal{W}}(\delta_0, \varpi)).$$

497 **Vanishing perturbation** Recall that  $z_\varepsilon = \lim_{a \in I_\varepsilon} z_{\varepsilon,a}$ , where the convergence is  
 498 understood in  $\tau$  for the measure coordinates. Passing to the limit in  $I_\varepsilon \ni a \rightarrow \infty$  in  
 499 (4.9) will not give useful information, since the squared Wasserstein distance is only  
 500  $\tau$ -lower semicontinuous, and we will not obtain an inequality on  $z_\varepsilon$ . Therefore, we  
 501 extract a diagonal sequence of  $I$ . Let  $n_0$  be large enough so that  $2^{-n_0} \leq \varepsilon_0$ , and  
 502 denote  $\varepsilon_n := 2^{-n}$  for  $n \geq n_0$ . Proceeding by induction and using (4.6), we may build  
 503 a sequence  $(\varepsilon_n, a_n)_n \subset I$  such that  $a_n < a_{n+1}$  for which

$$504 a_n d_{\mathcal{W}}^2(\mu_{\varepsilon_n, a_n}, \nu_{\varepsilon_n, a_n}) \leq \frac{1}{n}, \quad \sup(\Phi_{\varepsilon_n, a_n}) \geq \Gamma_{\varepsilon_n}^\alpha - \frac{1}{n}, \quad \sup(\Phi_{\varepsilon_{n+1}, a_{n+1}}) \geq \sup(\Phi_{\varepsilon_n, a_n}).$$

505 The sequence  $(\sup(\Phi_{\varepsilon_n, a_n}))_n$  is nondecreasing and upper bounded by  $\Gamma^\alpha$ , thus con-  
 506 verges. On the other hand,

$$507 \quad 0 \leq \frac{\varepsilon_n}{2} (d_{\mathcal{W}}^2(\sigma, \mu_{\varepsilon_n, a_n}) + d_{\mathcal{W}}^2(\sigma, \nu_{\varepsilon_n, a_n}) + 0) \leq \Phi_{\varepsilon_{n+1}, a_{n+1}}(z_n) - \Phi_{\varepsilon_n, a_n}(z_n)$$

$$508 \quad \leq \sup(\Phi_{\varepsilon_{n+1}, a_{n+1}}) - \sup(\Phi_{\varepsilon_n, a_n}) \xrightarrow[n \rightarrow \infty]{} 0.$$

510 Evaluating (4.9) along the subsequence  $(\varepsilon_n, a_n)_n \subset I$  and passing to the limit in  
 511  $n \rightarrow \infty$ , we obtain  $\alpha \leq 0$ , which is absurd. Consequently,  $\Gamma \leq 0$ .  $\square$

512 **5. Characterisation of the solution in the case of control problems.** We  
 513 now study the properties of the value function  $V : [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R} \cup \{\infty\}$   
 514 associated to the control problem (1.3), given by

$$515 \quad V(t, \nu) := \inf_{\omega \in R_T^{t, \nu}} \mathfrak{J}(\omega).$$

516 Let us illustrate our setting with an example. Let  $\mathfrak{J} : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R} \cup \{+\infty\}$  be  
 517 given by

$$518 \quad \mathfrak{J}(\mu) := d_{\mathcal{W}, 4}^4(\mu, \delta_0) = \int_{x \in \mathbb{R}^d} |x|^4 d\mu(x).$$

519 The domain  $\text{dom } \mathfrak{J} = \mathcal{P}_4(\mathbb{R}^d)$  is closed in  $\tau$ , since the 4–Wasserstein distance  
 520  $d_{\mathcal{W}, 4}(\cdot, \delta_0)$  is narrowly lower semicontinuous. Take the dynamic  $f : U \rightarrow \mathcal{C}(\mathbb{R}^d; \mathbb{T} \mathbb{R}^d)$   
 521 parametrized by  $U = [0, 1]$  as

$$522 \quad f[u](x) := -ux.$$

523 Then  $f[U]$  is convex, compact in the topology of uniform convergence on compact sets,  
 524 and each  $f[u]$  satisfies  $|f[u](0) + \text{Lip}(f[u])| \leq 1$ . For each  $u(\cdot) \in L^0([0, T]; U)$ , the flow  
 525 of the underlying ODE is given by  $S_s^{t, f[u]}(x) = \exp(-\int_{r=t}^s u(r) dr) x$ . Consequently,

$$526 \quad \mathfrak{J}(S_s^{t, u(\cdot)}) = \int_{x \in \mathbb{R}^d} \exp\left(-4 \int_{r=t}^s u(r) dr\right) |x|^4 d\mu(x) = \exp\left(-4 \int_{r=t}^s u(r) dr\right) \mathfrak{J}(\mu),$$

527 and minimizing over  $u(\cdot) \in L^0([t, T]; U)$ , the value function is given by

$$528 \quad V(t, \mu) = \exp(-4(T-t)) \mathfrak{J}(\mu) \in \mathbb{R} \cup \{\infty\}.$$

529 Gathering intuition from the available theory in finite dimension and Hilbert  
 530 spaces, we may expect  $V$  to be a viscosity supersolution of the HJB equation (4.1)  
 531 for the Hamiltonian

$$532 \quad (5.1) \quad H : \mathbb{T} \rightarrow \mathbb{R}, \quad H(\mu, p) := \sup_{u \in U} -p(\pi^\mu(f[u] \# \mu)),$$

533 and a solution whenever  $\mathfrak{J}$  is real-valued and  $\tau$ -continuous. From this point on-  
 534 ward, unless otherwise stated, we assume that the Hamiltonian of the HJB equation  
 535 (4.1) is given by (5.1).

536 Let us verify that it is indeed the case in our example. If  $(t, \mu) \in \text{dom } V =$   
 537  $[0, T] \times \mathcal{P}_4(\mathbb{R}^d)$ , then the map  $V(t, \cdot)$  is directionally differentiable along trajectories  
 538 of the form  $s \mapsto \exp_\mu(\pi^\mu(s \cdot f[u] \# \mu))$  and its derivative satisfies

$$539 \quad D_\mu V(t, \cdot)(\pi^\mu(s \cdot f[u] \# \mu)) = \exp(-4(T-t))(-4u) \mathfrak{J}(\mu).$$



540 Hence in this case, we may compute the Hamiltonian and see that  $\forall(t, \mu) \in \text{dom } V$ ,

$$541 \quad -\partial_t V(t, \mu) + \sup_{u \in U} -D_\mu V(t, \cdot)(\pi^\mu(f[u] \# \mu)) \\ 542 \quad = -4 \exp(-4(T-t)) \mathfrak{J}(\mu) + \sup_{u \in [0,1]} 4u \exp(-4(T-t)) \mathfrak{J}(\mu) = 0. \\ 543$$

544 This directly implies that  $V$  is a viscosity supersolution in the sense of Definition 4.1.  
545 Indeed, if  $\varphi \in \mathcal{T}_-$  is such that  $V - \varphi$  reaches a finite minimum in  $(t, \mu) \in (0, T) \times$   
546  $\mathcal{P}_2(\mathbb{R}^d)$ , then  $\partial_t \varphi(t, \mu) \leq \partial_t V(t, \mu)$  and  $D_\mu \varphi(t, \cdot)(\xi) \leq D_\mu V(t, \cdot)(\xi)$  along each  $\xi \in$   
547  $\text{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$  such that  $D_\mu V(t, \cdot)(\xi)$  exists. Hence the supersolution inequality

$$548 \quad -\partial_t \varphi(t, \mu) + \sup_{u \in U} -D_\mu \varphi(t, \cdot)(\pi^\mu(f[u] \# \mu)) \geq 0.$$

549 This section shows that this situation is generic in our setting. We begin by the  
550 general case where  $\mathfrak{J}$  may be unbounded, and then restrict to a more regular case.

### 551 5.1. General case.

552 LEMMA 5.1 (Regularity of the value function). *Assume Hypotheses 1.1 and 1.2.*  
553 *Then each  $V(t, \cdot)$  is proper, and  $V$  is lower bounded and  $\tau$ -lower semicontinuous.*  
554 *Moreover, if  $\mathfrak{J}$  is bounded and  $\tau$ -continuous, then so is  $V$ .*

555 *Proof.* Lower boundedness of  $V$  follows from that of  $\mathfrak{J}$  and by its definition. Let  
556  $\omega \in \text{dom } \mathfrak{J} \subset \mathcal{P}_2(\mathbb{R}^d)$ , and let  $b \in f[U]$  be fixed. As  $-b$  is Lipschitz-continuous, the  
557 reversed continuity equation

$$558 \quad \partial_s \mu_s + \text{div}(-b \# \mu_s) = 0, \quad \mu_0 = \sigma$$

559 admits an unique solution  $(\mu_s)_{s \in [0, T]}$  such that  $\bar{\mu}_s := \mu_{T-s}$  is a solution of the forward  
560 equation  $\partial_s \bar{\mu}_t + \text{div}(b \# \bar{\mu}_t) = 0$ , and  $\bar{\mu}_T = \sigma$ . Hence  $V(t, \bar{\mu}_t) \leq \mathfrak{J}(\sigma) < \infty$ . Thus,  
561  $V(t, \cdot)$  is proper.

562 Since closedness and sequential closedness coincide in  $(\mathcal{P}_2(\mathbb{R}^d), \tau)$  (see Defini-  
563 tion 2.3), we only have to show that  $V$  is sequentially lower semicontinuous. Let  
564  $(t_n, \nu_n)_n \subset [0, T] \times \mathcal{P}_2(\mathbb{R}^d)$  such that  $t_n \rightarrow t \in [0, T]$  and  $\nu_n \xrightarrow{\tau} \nu \in \mathcal{P}_2(\mathbb{R}^d)$ . For  
565 each  $n$ , let  $\omega_n \in R_T^{t_n, \nu_n}$  such that  $V(t_n, \nu_n) \geq \mathfrak{J}(\omega_n) - 1/n$ . Using Lemma 3.4, pos-  
566 sibly along a subsequence, there exists  $\bar{\omega} \in R_T^{t, \nu}$  such that  $\omega_n \xrightarrow{\tau} \bar{\omega}$ . Then, by lower  
567 semicontinuity of  $\mathfrak{J}$  in  $\tau$ ,

$$568 \quad \liminf_{n \rightarrow \infty} V(t_n, \nu_n) \geq \liminf_{n \rightarrow \infty} \mathfrak{J}(\omega_n) - \frac{1}{n} \geq \mathfrak{J}(\bar{\omega}) \geq \inf_{\omega \in R_T^{t, \nu}} \mathfrak{J}(\omega) = V(t, \nu).$$

569 Assume now that  $\mathfrak{J}$  is bounded and  $\tau$ -continuous. Then  $V$  shares the same bound  
570 by definition. To prove that  $V$  is  $\tau$ -upper semicontinuous, it is enough to show that  
571 for any  $(t, \nu) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^d)$ , any  $[0, T] \ni t_n \rightarrow t$  and  $\nu_n \xrightarrow{\tau} \nu$ ,

$$572 \quad \overline{\lim}_{n \rightarrow \infty} V(t_n, \nu_n) \leq V(t, \nu).$$

573 Up to extraction, we may assume that  $\overline{\lim}_{n \rightarrow \infty} V(t_n, \nu_n) = \lim_{n \rightarrow \infty} V(t_n, \nu_n)$ . For each  
574  $\varepsilon > 0$ , let  $\mu_\varepsilon \in R_T^{t, \nu}$  such that  $V(t, \nu) \geq \mathfrak{J}(\mu_\varepsilon) - \varepsilon$ . By Lemma 3.3, there exists  
575  $\mu_{\varepsilon, n} \in R_T^{t_n, \nu_n}$  such that  $\mu_{\varepsilon, n} \xrightarrow{\tau} \mu_\varepsilon$ . Then, since  $\mathfrak{J}$  is  $\tau$ -continuous,

$$576 \quad \lim_{n \rightarrow \infty} V(t_n, \nu_n) \leq \overline{\lim}_{t \rightarrow \infty} \mathfrak{J}(\mu_{\varepsilon, n}) = \mathfrak{J}(\mu_\varepsilon) \leq V(t, \nu) + \varepsilon.$$

577 Letting  $\varepsilon \searrow 0$ , we conclude that  $V$  is  $\tau$ -continuous.  $\square$

578 **THEOREM 5.2 (Supersolution).** *The value function is a supersolution of (4.1)*  
 579 *in the sense of Definition 4.1.*

580 *Proof.* By Lemma 5.1,  $V$  is lower bounded and  $\tau$ -lower semicontinuous, and by  
 581 definition,  $V(T, \cdot) = \mathfrak{J} > -\infty$ . Let  $\varphi \in \mathcal{T}_-$  such that  $V - \varphi$  reaches a minimum in  
 582  $(t, \nu) \in (0, T) \times \mathcal{P}_2(\mathbb{R}^d)$ . By Proposition 3.1, the set of trajectories  $\mathcal{S}^{t, \nu}$  is compact in  
 583  $\mathcal{C}([t, T]; \mathcal{P}_2(\mathbb{R}^d))$  endowed with the topology of uniform convergence. Therefore,  $R_T^{t, \nu}$   
 584 is compact in  $(\mathcal{P}_2(\mathbb{R}^d), d_{\mathcal{W}})$ , hence in  $(\mathcal{P}_2(\mathbb{R}^d), \tau)$ . Since  $\mathfrak{J}$  is  $\tau$ -lower semicontinuous,  
 585 there exists  $(\boldsymbol{\mu}_s^{t, \nu})_{s \in [t, T]}$  such that

$$586 \quad V(t, \nu) = V(t + h, \boldsymbol{\mu}_{t+h}^{t, \nu}) \quad \forall h \in [0, T - t].$$

587 Recall that  $\phi(t, \mu) = \psi(t) + g(\mu)$ , where  $\psi \in \mathcal{C}^1((0, T); \mathbb{R})$  and  $g$  is locally Lipschitz,  
 588 directionally differentiable and  $\tau$ -lower semicontinuous. Thus, for any  $h \in [0, T - t]$  it  
 589 follows that

$$590 \quad \psi(t + h) - \psi(t) + g(\boldsymbol{\mu}_{t+h}^{t, \nu}) - g(\nu) \leq V(t + h, \boldsymbol{\mu}_{t+h}^{t, \nu}) - V(t, \nu) = 0.$$

591 Using Lemma 3.5, there exists a subsequence  $(h_n)_n \subset (0, T - t]$  with  $h_n \searrow_n 0$ , and  
 592 some  $b \in f[U]$  such that  $d_{\mathcal{W}}(\boldsymbol{\mu}_{t+h_n}^{t, \nu}, \exp_{\nu}(h_n \cdot b \# \nu)) = o(h_n)$ . Dividing the above by  
 593  $h_n > 0$ , and denoting  $\text{Lip}(g)$  a local Lipschitz constant of  $g$  in a ball centered in  $\nu$   
 594 and containing all  $\boldsymbol{\mu}_{t+h_n}^{t, \nu}$  and  $\exp_{\nu}(h_n \cdot b \# \nu)$ , we have

$$595 \quad \frac{\psi(t + h_n) - \psi(t)}{h_n} + \frac{g(\exp_{\nu}(h_n \cdot b \# \nu)) - g(\nu)}{h_n} \leq \text{Lip}(g) \frac{d_{\mathcal{W}}(\exp_{\nu}(h_n \cdot b \# \nu), \boldsymbol{\mu}_{t+h_n}^{t, \nu})}{h_n}.$$

596 Taking the limit in  $n \rightarrow \infty$  and using the respective differentiability of  $\psi$  and  $g$ ,

$$597 \quad \partial_t \psi(t) + D_{\mu} g(\nu)(b \# \nu) \leq 0.$$

598 By the construction of test functions,  $D_{\mu} g(\nu)(\xi) = D_{\mu} g(\pi^{\mu} \xi)$  for all  $\xi \in \mathcal{P}_2(\mathbb{T} \mathbb{R}^d)_{\nu}$ .  
 599 multiplying by  $-1$  and taking the maximum over all  $b \in f[U]$ , we obtain that  
 600  $-\partial_t \varphi(t, \nu) + H(\nu, D_{\mu} \varphi(t, \nu)) \geq 0$ , which is the desired property.  $\square$

601 **5.2. Case of continuous and bounded terminal cost.** We show that in the  
 602 case where  $\mathfrak{J}$  is bounded and  $\tau$ -continuous, the value function is also a subsolution of  
 603 (4.1). Owing to the comparison principle, it will then be the unique solution.

604 **THEOREM 5.3 (Subsolution property).** *Assume Hypotheses 1.1 and 1.2, and that*  
 605  *$\mathfrak{J}$  is bounded and  $\tau$ -continuous. Then the value function is a subsolution of (4.1).*

606 *Proof.* By Lemma 5.1, the value function is bounded and  $\tau$ -upper semicontinuous.  
 607 As  $V(T, \cdot) = \mathfrak{J}$ , we only have to prove the viscosity inequality (4.2). Let  $\varphi = \psi \ominus g \in$   
 608  $\mathcal{T}_+$  and  $(t, \mu) \in (0, T) \times \mathcal{P}_2(\mathbb{R}^d)$  such that  $V - \varphi$  reaches a maximum in  $(t, \mu)$ . Let  
 609 any  $b \in f[U]$  be fixed. Then the flow  $S_{t+}^{t, b}$  of the autonomous ODE  $\frac{d}{ds} y_s = b(y_s)$  is of  
 610 class  $\mathcal{C}^1$ , and there holds

$$611 \quad \lim_{h \searrow 0} \frac{d_{\mathcal{W}}(S_{t+h}^{t, b} \# \mu, \exp_{\mu}(h \cdot b \# \mu))}{h} = 0.$$

612 Denote  $\boldsymbol{\mu}_s := S_s^{t, b} \# \mu$ . Using the dynamic programming principle,

$$613 \quad 0 \leq V(s, \boldsymbol{\mu}_s) - V(t, \nu) \leq \varphi(s, \boldsymbol{\mu}_s) - \varphi(t, \nu) = \psi(s) - \psi(t) - [g(\boldsymbol{\mu}_s) - g(\nu)].$$

614 As  $g$  is locally Lipschitz, dividing by  $s - t$  and sending  $s \rightarrow t$ , we get

$$615 \quad \partial_t \psi(t) - D_\mu g(b\#\mu) \geq 0.$$

616 Since  $D_\mu g(\nu)(\xi) = D_\mu g(\pi^\mu \xi)$  if  $\xi \in \mathcal{P}_2(\mathbb{T}\mathbb{R}^d)_\mu$  by definition of  $\mathcal{T}_+$ , multiplying by  
617  $-1$  and taking the supremum over  $b \in f[U]$ , we get that  $V$  is a subsolution of (4.1).  $\square$

618 In the general case,  $V$  may take the value  $+\infty$  and has no chance to be a viscosity  
619 subsolution of the HJB equation (4.1). However, we may still prove that it is the  
620 smallest supersolution in the pointwise sense. The argument proceeds by truncature  
621 and regularization, relying on the following result.

622 **LEMMA 5.4** (Inf-convolution in  $\tau$ ). *Let  $\mathfrak{J} : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R} \cup \{\infty\}$  be lower bounded  
623 and  $\tau$ -lower semicontinuous. Then for each  $B > 0$ , there is a nondecreasing sequence  
624 of bounded  $\tau$ -continuous maps  $\mathfrak{J}_n : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$  that converge pointwise towards  $\mathfrak{J}$   
625 over  $\overline{\mathcal{B}}_{\mathcal{W}}(\delta_0, B)$ .*

626 *Proof.* Denote  $\mathbb{I}_B : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R} \cup \{\infty\}$  the characteristic function of the closed  
627 ball  $\overline{\mathcal{B}}_{\mathcal{W}}(\delta_0, B)$ , i.e.  $\mathbb{I}_B(\nu) = 0$  if  $d_{\mathcal{W}}(\delta_0, \nu) \leq B$ , and  $\mathbb{I}_B(\nu) = +\infty$  otherwise. Since  
628 closed Wasserstein balls are  $\tau$ -compact,  $\mathbb{I}_B$  is  $\tau$ -lower semicontinuous. Moreover,  
629 the function  $\nu \mapsto \mathfrak{J}(\nu) + \mathbb{I}_B(\nu)$  is narrowly lower semicontinuous. Indeed, this is due  
630 to the fact that the topology  $\tau$  coincides with the narrow topology on closed balls. Let  
631  $\mathfrak{d} : \mathcal{P}_2(\mathbb{R}^d) \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}^+$  be a metric inducing the topology of narrow convergence  
632 over  $\mathcal{P}_2(\mathbb{R}^d)$  (e.g. [4, Section 5.1]), and

$$633 \quad \mathfrak{J}_n(\mu) := \min \left( n, \inf_{\nu \in \mathcal{P}_2(\mathbb{R}^d)} (\mathfrak{J} + \mathbb{I}_B)(\nu) + n\mathfrak{d}(\mu, \nu) \right).$$

634 We directly have  $\mathfrak{J}_n(\mu) \leq \min(n, \mathfrak{J}(\mu)) \leq \mathfrak{J}(\mu)$  for all  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ . Moreover, for  
635 each  $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^d)$ ,

$$636 \quad \mathfrak{J}_n(\mu_0) - \mathfrak{J}_n(\mu_1) \leq \max \left( 0, \sup_{\nu \in \mathcal{P}_2(\mathbb{R}^d)} n(\mathfrak{d}(\mu_0, \nu) - \mathfrak{d}(\mu_1, \nu)) \right) \leq n\mathfrak{d}(\mu_0, \mu_1).$$

637 By symmetry,  $\mathfrak{J}_n$  is  $n$ -Lipschitz with respect to  $\mathfrak{d}$ , thus  $\tau$ -continuous. It is moreover  
638 bounded with values in  $[\min(0, \inf(\mathfrak{J})), n]$ . To prove pointwise convergence, let  $\mu \in$   
639  $\overline{\mathcal{B}}_{\mathcal{W}}(\delta_0, B)$  be fixed. Assume by contradiction that there exists  $M < \mathfrak{J}(\mu)$  such that  
640  $\mathfrak{J}_n(\mu) \leq M$  for all  $n$ . Let  $\varepsilon := \min(1, \mathfrak{J}(\mu) - M) > 0$ . Since  $\mathfrak{J} + \mathbb{I}_B$  is narrowly lower  
641 semicontinuous, there exists  $r > 0$  such that  $\mathfrak{d}(\mu, \nu) < r$  implies  $(\mathfrak{J} + \mathbb{I}_B)(\nu) \geq M + \varepsilon/2$ .  
642 Taking  $n$  large enough so that  $nr \geq M + \frac{\varepsilon}{2} - \inf(\mathfrak{J})$  and  $n \geq M + \varepsilon/2$ , we get

$$643 \quad \mathfrak{J}_n(\mu) \geq \min \left( n, \inf_{\mathfrak{d}(\mu, \nu) < r} (\mathfrak{J} + \mathbb{I}_B)(\nu) + n\mathfrak{d}(\mu, \nu), \inf_{\mathfrak{d}(\mu, \nu) \geq r} (\mathfrak{J} + \mathbb{I}_B)(\nu) + n\mathfrak{d}(\mu, \nu) \right)$$

$$644 \quad \geq \min \left( M + \frac{\varepsilon}{2}, \inf_{\mathfrak{d}(\mu, \nu) \geq r} \mathfrak{J}(\nu) + 0 + M + \frac{\varepsilon}{2} - \inf(\mathfrak{J}) \right) \geq M + \frac{\varepsilon}{2},$$

$$645$$

646 which is absurd. Thus the claim.  $\square$

647 Using this regularization, we obtain the following characterization.

648 **THEOREM 5.5** (Minimality property in the general case). *Assume Hypotheses 1.1  
649 and 1.2. Then for any supersolution  $v : [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R} \cup \{\infty\}$  of (4.1) such  
650 that  $v(T, \cdot)$  is proper, there holds*

$$651 \quad (5.2) \quad v(t, \nu) \geq V(t, \nu), \quad \forall (t, \nu) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^d).$$

652 *Consequently, the value function  $V$  is the smallest viscosity supersolution of (4.1).*

653 *Proof.* Let  $(\bar{t}, \bar{\nu}) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^d)$ . From Remark 3.2, the reachable set  $R_T^{\bar{t}, \bar{\nu}}$  is  
 654 contained in  $\overline{\mathcal{B}_{\mathcal{W}}(\delta_0, B)}$  for some sufficiently large  $B$ . If  $v(\bar{t}, \bar{\nu}) = \infty$ , the inequality  
 655 (5.2) is trivially satisfied. Assume now that  $v(\bar{t}, \bar{\nu}) < \infty$ . Let  $(\mathfrak{J}_n)_n$  be given by  
 656 Lemma 5.4. By Theorems 5.2 and 5.3, the HJB equation

$$657 \quad (5.3) \quad -\partial_t \vartheta_n(t, \mu) + H(\mu, D_\mu \vartheta_n(t, \mu)) = 0, \quad \vartheta_n(T, \mu) = \mathfrak{J}_n(\mu)$$

658 admits a unique solution given by

$$659 \quad V_n(t, \nu) = \inf_{\mu \in R_T^{t, \nu}} \mathfrak{J}_n(\mu), \quad \forall (t, \nu) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^d).$$

660 Since  $v(t, \nu) \geq \mathfrak{J}(\nu) \geq \mathfrak{J}_n(\nu)$ , the map  $v$  is a supersolution of each regularized problem  
 661 (5.3). Let  $\sigma \in \mathcal{P}_2(\mathbb{R}^d)$  such that  $v(T, \sigma) \in \mathbb{R}$ : since  $v(T, \sigma) \geq \mathfrak{J}(\sigma) \geq \mathfrak{J}_n(\sigma) =$   
 662  $V_n(T, \sigma)$ , we have  $-\infty < V_n(T, \sigma) - v(T, \sigma) \leq 0$ . In consequence, we can apply  
 663 Proposition 4.4, and deduce that  $v(t, \nu) \geq V_n(t, \nu)$  for any  $(t, \nu) \in \mathcal{P}_2(\mathbb{R}^d)$ .

664 By Proposition 4.4, the solutions  $V_n$  are ordered in the sense that  $V_{n+1}(t, \nu) \geq$   
 665  $V_n(t, \nu)$  for all  $n$ . Moreover,  $\mathfrak{J}_n \leq \mathfrak{J}$  implies that the subsolutions  $V_n$  are smaller  
 666 than the supersolution  $V$ . Hence the sequence  $(V_n(\bar{t}, \bar{\nu}))_n$  is nondecreasing and upper  
 667 bounded by  $v(\bar{t}, \bar{\nu}) < \infty$ , and converges. For each  $n$ , let  $\mu_n \in R_T^{\bar{t}, \bar{\nu}}$  such that  $V_n(\bar{t}, \bar{\nu}) \geq$   
 668  $\mathfrak{J}_n(\mu_n) - \frac{1}{n}$ . Using Lemma 3.4, some (non relabeled) subsequence converges in  $\tau$   
 669 towards some  $\bar{\mu} \in R_T^{\bar{t}, \bar{\nu}}$ . Using the monotonicity of the family  $(\mathfrak{J}_n)_n$  and the continuity  
 670 in  $\tau$  of each  $\mathfrak{J}_m$  for a fixed  $m$ ,

$$671 \quad \lim_{n \rightarrow \infty} V_n(\bar{t}, \bar{\nu}) \geq \lim_{n \rightarrow \infty} \mathfrak{J}_n(\mu_n) - \frac{1}{n} \geq \lim_{n \rightarrow \infty, n \geq m} \mathfrak{J}_m(\mu_n) - \frac{1}{n} = \mathfrak{J}_m(\bar{\mu}).$$

672 As  $\bar{\mu} \in \overline{\mathcal{B}_{\mathcal{W}}(\delta_0, B)}$ , the conclusion follows from taking the limit in  $m \rightarrow \infty$  to obtain

$$673 \quad v(\bar{t}, \bar{\nu}) \geq \lim_{n \rightarrow \infty} V_n(t, \nu) \geq \mathfrak{J}(\bar{\mu}) \geq V(\bar{t}, \bar{\nu}). \quad \square$$

674 **Acknowledgments.** The second author would like to thank the Center for  
 675 Mathematical Modeling, in particular to the CMM-PhD Visiting Program 2022 for the  
 676 opportunity to conduct this research at Universidad Técnica Federico Santa María.

677

#### REFERENCES

- 678 [1] L. AMBROSIO, E. BRUÉ, AND D. SEMOLA, *Lectures on Optimal Transport*, vol. 130 of UNITEXT,  
 679 Springer International Publishing, 2021, <https://doi.org/10.1007/978-3-030-72162-6>.  
 680 [2] L. AMBROSIO AND J. FENG, *On a class of first order Hamilton–Jacobi equations in metric*  
 681 *spaces*, Journal of Differential Equations, 256 (2014), pp. 2194–2245, [https://doi.org/10.](https://doi.org/10.1016/j.jde.2013.12.018)  
 682 [1016/j.jde.2013.12.018](https://doi.org/10.1016/j.jde.2013.12.018).  
 683 [3] L. AMBROSIO AND W. GANGBO, *Hamiltonian ODEs in the Wasserstein space of probability*  
 684 *measures*, Communications on Pure and Applied Mathematics, 61 (2008), pp. 18–53, <https://doi.org/10.1002/cpa.20188>.  
 685 [4] L. AMBROSIO, N. GIGLI, AND G. SAVARÉ, *Gradient Flows*, Lectures in Mathematics ETH Zürich,  
 686 Birkhäuser-Verlag, Basel, 2005, <https://doi.org/10.1007/b137080>.  
 687 [5] B. BONNET AND H. FRANKOWSKA, *Caratheodory theory and a priori estimates for continuity*  
 688 *inclusions in the space of probability measures*, May 2023, [https://arxiv.org/abs/2302.](https://arxiv.org/abs/2302.00963)  
 689 [00963](https://arxiv.org/abs/2302.00963).  
 690 [6] P. CARDALIAGUET, *Notes on Mean Field Games*, (2013), p. 59, [https://www.ceremade.](https://www.ceremade.dauphine.fr/~cardaliaguet/MFG20130420.pdf)  
 691 [dauphine.fr/~cardaliaguet/MFG20130420.pdf](https://www.ceremade.dauphine.fr/~cardaliaguet/MFG20130420.pdf).  
 692 [7] P. CARDALIAGUET, F. DELARUE, J.-M. LASRY, AND P.-L. LIONS, *The Master Equation and the*  
 693 *Convergence Problem in Mean Field Games*, no. 201 in Annals of Mathematics Studies,  
 694 Princeton University Press, 2019, <https://doi.org/10.48550/arXiv.1509.02505>.  
 695

- 696 [8] P. CARDALIAGUET AND A. PORRETTA, *An Introduction to Mean Field Game Theory*, in Mean  
697 Field Games: Cetraro, Italy 2019, Y. Achdou et al., eds., Springer International Publishing,  
698 2020, pp. 1–158, [https://doi.org/10.1007/978-3-030-59837-2\\_1](https://doi.org/10.1007/978-3-030-59837-2_1).
- 699 [9] R. CARMONA AND F. DELARUE, *Probabilistic Theory of Mean Field Games with Applications I*,  
700 vol. 83 of Probability Theory and Stochastic Modelling, Springer International Publishing,  
701 Cham, 2018, <https://doi.org/10.1007/978-3-319-58920-6>.
- 702 [10] R. CARMONA AND F. DELARUE, *Probabilistic Theory of Mean Field Games with Applications II*,  
703 vol. 84 of Probability Theory and Stochastic Modelling, Springer International Publishing,  
704 Cham, 2018, <https://doi.org/10.1007/978-3-319-56436-4>.
- 705 [11] J. A. CARRILLO, Y.-P. CHOI, AND M. HAURAY, *The derivation of swarming models: Mean-*  
706 *field limit and Wasserstein distances*, in Collective Dynamics from Bacteria to Crowds,  
707 A. Muntean and F. Toschi, eds., CISM International Centre for Mechanical Sciences,  
708 Springer, Vienna, 2014, pp. 1–46, [https://doi.org/10.1007/978-3-7091-1785-9\\_1](https://doi.org/10.1007/978-3-7091-1785-9_1).
- 709 [12] G. CAVAGNARI, G. SAVARÉ, AND G. SODINI, *A Lagrangian approach to totally dissipative evo-*  
710 *lutions in Wasserstein spaces*, (2023), <https://doi.org/10.48550/arXiv.2305.05211>.
- 711 [13] A. CORBETTA, *Multiscale Crowd Dynamics: Physical Analysis, Modeling and Applications*,  
712 PhD thesis, Eindhoven University of Technology, 2016.
- 713 [14] M. G. CRANDALL, H. ISHII, AND P.-L. LIONS, *User’s guide to viscosity solutions of second order*  
714 *partial differential equations*, June 1992, <https://arxiv.org/abs/math/9207212>.
- 715 [15] M. G. CRANDALL AND P.-L. LIONS, *Hamilton-Jacobi equations in infinite dimensions I.*  
716 *uniqueness of viscosity solutions*, Journal of Functional Analysis, 62 (1985), pp. 379–396,  
717 [https://doi.org/https://doi.org/10.1016/0022-1236\(85\)90011-4](https://doi.org/https://doi.org/10.1016/0022-1236(85)90011-4).
- 718 [16] S. DAUDIN, J. JACKSON, AND B. SEEGER, *Well-posedness of Hamilton-Jacobi equations in*  
719 *the Wasserstein space: Non-convex Hamiltonians and common noise*, Dec. 2023, <https://arxiv.org/abs/2312.02324>.
- 720 [17] W. GANGBO AND A. TUDORASCU, *On differentiability in the Wasserstein space and well-*  
721 *posedness for Hamilton–Jacobi equations*, Journal de Mathématiques Pures et Appliquées,  
722 125 (2019), pp. 119–174, <https://doi.org/10.1016/j.matpur.2018.09.003>.
- 723 [18] N. GIGLI, *On the Geometry of the Space of Probability Measures Endowed with the Quadratic*  
724 *Optimal Transport Distance*, PhD thesis, Scuola Normale Superiore di Pisa, Pisa, 2008.
- 725 [19] R. C. JAMES, *Weakly compact sets*, Transactions of the American Mathematical Society, 113  
726 (1964), pp. 129–140, <https://doi.org/10.1090/S0002-9947-1964-0165344-2>.
- 727 [20] F. JEAN, O. JERHAOUI, AND H. ZIDANI, *Deterministic optimal control on Riemannian mani-*  
728 *folds under probability knowledge of the initial condition*, SIAM Journal on Mathematical  
729 Analysis, Accepted (2024), <https://ensta-paris.hal.science/hal-03564787/>.
- 730 [21] O. JERHAOUI, *Viscosity Theory of First Order Hamilton Jacobi Equations in Some Metric*  
731 *Spaces*, PhD thesis, Institut Polytechnique de Paris, Paris, 2022.
- 732 [22] O. JERHAOUI, A. PROST, AND H. ZIDANI, *Viscosity solutions of centralized control problems in*  
733 *measure spaces*, <https://hal.science/hal-04335852>, (2023).
- 734 [23] C. JIMENEZ, *Equivalence between strict viscosity solution and viscosity solution in the space of*  
735 *Wasserstein and regular extension of the Hamiltonian in  $L^2$ -IP*. [https://hal.science/hal-](https://hal.science/hal-04136329)  
736 [04136329](https://hal.science/hal-04136329), 2023.
- 737 [24] C. JIMENEZ, A. MARIGONDA, AND M. QUINCAMPOIX, *Optimal control of multiagent systems in*  
738 *the Wasserstein space*, Calculus of Variations and Partial Differential Equations, 59 (2020),  
739 <https://doi.org/10.1007/s00526-020-1718-6>.
- 740 [25] P.-L. LIONS, *Jeux à champ moyen*, 2006.
- 741 [26] P.-L. LIONS AND P. SOUGANIDIS, *Differential Games, Optimal Control and Directional Deriva-*  
742 *tives of Viscosity Solutions of Bellman’s and Isaacs’ Equations.*, SIAM Journal on Control  
743 and Optimization, 23 (1984), pp. 566–583, <https://doi.org/10.1137/0323036>.
- 744 [27] A. MARIGONDA AND M. QUINCAMPOIX, *Mayer control problem with probabilistic uncertainty*  
745 *on initial positions*, Journal of Differential Equations, 264 (2018), pp. 3212–3252, <https://doi.org/10.1016/j.jde.2017.11.014>.
- 746 [28] F. OTTO, *The Geometry of Dissipative Evolution Equations: The Porous Medium Equation*,  
747 *Communications in Partial Differential Equations*, 26 (2001), pp. 101–174, <https://doi.org/10.1081/PDE-100002243>.
- 748 [29] B. PICCOLI AND A. TOSIN, *Pedestrian flows in bounded domains with obstacles*, Continuum  
749 *Mechanics and Thermodynamics*, 21 (2009), pp. 85–107, <https://doi.org/10.1007/s00161-009-0100-x>.
- 750 [30] A. ÜLGER, *Weak compactness in  $L^1(\mu, X)$* , Proceedings of The American Mathematical So-  
751 *ciety*, 113 (1991), pp. 143–143, <https://doi.org/10.1090/S0002-9939-1991-1070533-0>.