



# Infinite horizon value functions in the Wasserstein spaces

Ryan Hynd <sup>\*,1</sup>, Hwa Kil Kim <sup>2</sup>

Received 24 June 2014; revised 8 October 2014

Available online 10 December 2014

## Abstract

We perform a systematic study of optimization problems in the Wasserstein spaces that are analogs of infinite horizon, deterministic control problems. We derive necessary conditions on action minimizing paths and present a sufficient condition for their existence. We also verify that the corresponding generalized value functions are a type of viscosity solution of a time independent, Hamilton–Jacobi equation in the space of probability measures. Finally, we prove a special case of a conjecture involving the subdifferential of generalized value functions and their relation to action minimizing paths.

© 2014 Elsevier Inc. All rights reserved.

## 1. Introduction

In this paper, we consider analogs of the classical value functions

$$u(x) = \inf \left\{ \int_0^\infty e^{-\delta t} \left( \frac{1}{p} |\dot{\gamma}(t)|^p - V(\gamma(t)) \right) dt : \gamma \in AC_{p,\delta}(\mathbb{R}^d), \gamma(0) = x \right\} \quad (1.1)$$

in the space of probability measures. In the above formula,  $x \in \mathbb{R}^d$ ,  $p \in (1, \infty)$ ,  $\delta \in (0, \infty)$ ,  $V \in C^1(\mathbb{R}^d)$ , and  $AC_{p,\delta}(\mathbb{R}^d)$  is the collection of locally absolutely continuous paths  $\gamma : [0, \infty) \rightarrow \mathbb{R}^d$

\* Corresponding author.

E-mail address: [rhynd@math.upenn.edu](mailto:rhynd@math.upenn.edu) (R. Hynd).

<sup>1</sup> Partially supported by NSF grants DMS-1004733 and DMS-1301628.

<sup>2</sup> Partially supported by the BK21 Plus SNU Mathematical Sciences Division.

satisfying  $\int_0^\infty e^{-\delta t} |\dot{\gamma}(t)|^p dt < \infty$ . Under appropriate growth assumptions on  $V$ , minimizers of  $u(x)$  exist and satisfy the Euler–Lagrange ODE

$$\frac{d}{dt} (|\dot{\gamma}(t)|^{p-2} \dot{\gamma}(t)) = -\nabla V(\gamma(t)) + \delta |\dot{\gamma}(t)|^{p-2} \dot{\gamma}(t), \quad t > 0. \tag{1.2}$$

Another important fact is that  $u$  in (1.1) can be characterized as a viscosity solution of the Hamilton–Jacobi equation (HJE)

$$\delta u + \frac{1}{q} |\nabla u|^q + V(x) = 0, \quad x \in \mathbb{R}^d. \tag{1.3}$$

Here and throughout,  $1/q + 1/p = 1$ . Moreover,

$$|\dot{\gamma}(t)|^{p-2} \dot{\gamma}(t) = -\nabla u(\gamma(t)), \quad t > 0, \tag{1.4}$$

provided that  $u$  is differentiable at each  $\gamma(t)$ . The goal of this work is to extend some of these ideas to the space of probability measures.

In our prior study [18], we considered value functions on a finite horizon in the Wasserstein spaces. This current work addresses the analogous class of infinite horizon, *generalized value functions*,

$$\mathcal{U}(\mu) = \inf \left\{ \int_0^\infty e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}(t)\|^p - \mathcal{V}(\sigma(t)) \right) dt : \sigma \in AC_{p,\delta}(\mathcal{M}_p), \sigma(0) = \mu \right\}. \tag{1.5}$$

In Eq. (1.5), the  $p$ th Wasserstein space  $\mathcal{M}_p$  is

$$\mathcal{P}_p(\mathbb{R}^d) := \left\{ \text{Borel probability measures } \mu \text{ on } \mathbb{R}^d : \int_{\mathbb{R}^d} |x|^p d\mu(x) < \infty \right\}$$

equipped with the metric

$$W_p(\mu, \nu) := \inf \left\{ \left( \iint_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p d\pi(x, y) \right)^{1/p} : \pi \in \Gamma(\mu, \nu) \right\}. \tag{1.6}$$

The infimum in (1.6) is taken over  $\Gamma(\mu, \nu)$ , the subcollection of  $\mathcal{P}_p(\mathbb{R}^d \times \mathbb{R}^d)$  having first marginal  $\mu$  and second marginal  $\nu$ . A wealth of information about  $\mathcal{M}_p$  can be found in the references [2,20].

In (1.5),  $t \mapsto \|\dot{\sigma}(t)\|$  is the usual metric derivative of a locally  $p$ -absolutely continuous path  $\sigma : [0, \infty) \rightarrow \mathcal{M}_p$  (Definition 1.1.1 in [2]). By Theorem 8.3.1 [2], there is a Borel vector field  $v : \mathbb{R}^d \times [0, \infty) \rightarrow \mathbb{R}^d$  such that the continuity equation holds in the sense of distributions

$$\partial_t \sigma + \nabla \cdot (\sigma v) = 0, \quad \mathbb{R}^d \times [0, \infty).$$

Moreover,

$$\|v(t)\|_{L^p(\sigma(t))} = \|\dot{\sigma}(t)\| \quad \text{a.e. } t > 0.$$

The mapping  $v$  is known as the *minimal velocity* of  $\sigma$  as  $t \mapsto v(\cdot, t)$  is essentially uniquely determined and satisfies  $\|v(t)\|_{L^p(\sigma(t))} \leq \|w(t)\|_{L^p(\sigma(t))}$  (for a.e.  $t \geq 0$ ) for any other Borel field  $w$  satisfying the continuity equation with  $\sigma$ . For this work, an important subset of the space of locally  $p$ -absolutely continuous paths is  $AC_{p,\delta}(\mathcal{M}_p)$  which additionally requires

$$\int_0^\infty \int_{\mathbb{R}^d} e^{-\delta t} |v(x, t)|^p d\sigma_t(x) dt < \infty. \tag{1.7}$$

We show that minimizing paths for  $\mathcal{U}$  and their corresponding minimal velocities satisfy the *Euler–Poisson system*

$$\begin{cases} \partial_t(\sigma |v|^{p-2}v) + \nabla \cdot (\sigma |v|^{p-2}v \otimes v) = -\sigma \nabla \mathcal{V}(\sigma) + \sigma \delta |v|^{p-2}v, \\ \partial_t \sigma + \nabla \cdot (\sigma v) = 0 \end{cases} \tag{1.8}$$

in the sense of distributions on  $\mathbb{R}^d \times (0, \infty)$ . We also present a result on the existence of minimizing paths provided the potential  $\mathcal{V}$  satisfies an appropriate growth condition and is continuous with respect to the narrow topology. To this end, we make specific use of the fact that the narrow topology on  $\mathcal{P}_p(\mathbb{R}^d)$  can be metrized by the Lévy–Prokhorov metric  $\Lambda$  (defined in (3.6)).

**Theorem 1.1.** *Assume that  $\mathcal{V} \in C((\mathcal{P}_p(\mathbb{R}^d), \Lambda))$  and*

$$|\mathcal{V}(\mu)| \leq \alpha \Lambda(\mu, \rho)^{2r} + \beta, \quad \mu \in \mathcal{M}_p, \tag{1.9}$$

for some  $\alpha, \beta \in \mathbb{R}$ ,  $1 \leq r < p$  and  $\rho \in \mathcal{M}_p$ . Then  $\mathcal{U}(\mu)$  has a minimizing path for each  $\mu \in \mathcal{M}_p$ .

As in our previous paper [18], the main result of this work is that each generalized value function is a type of viscosity solution of an appropriate Hamilton–Jacobi equation.

**Theorem 1.2.** *Assume*

$$\mathcal{V}(\mu) \leq \alpha W_p(\mu, \varrho)^p + \beta, \quad \mu \in \mathcal{M}_p, \tag{1.10}$$

for some measure  $\varrho \in \mathcal{M}_p$ ,  $\alpha, \beta \in \mathbb{R}$ . Then there is  $\delta_0 = \delta_0(p, \alpha) > 0$  such that the functional  $\mathcal{U}$  (1.5) is a viscosity solution (Definition 5.3) of the HJE

$$\delta \mathcal{U} + \frac{1}{q} \|\nabla \mathcal{U}\|_{L^q(\mu)}^q + \mathcal{V}(\mu) = 0, \quad \mu \in \mathcal{M}_p, \tag{1.11}$$

provided  $\delta > \delta_0$ .

An important example occurs when  $\mathcal{V}$  is a *simple potential*:

$$\mathcal{V}(\mu) = \int_{\mathbb{R}^d} V(x) d\mu(x) \tag{1.12}$$

for some  $V \in C^1(\mathbb{R}^d)$ . We note that if  $V$  is not uniformly bounded,  $\mathcal{V}$  will not in general be narrowly continuous (Remark 7.1.11 in [2]). So it is not immediate that the corresponding generalized value function  $\mathcal{U}$  will have minimizing paths. However, we show that this is generally the case and provide a formula for the generalized value function (1.5) in terms of the classical one (1.1).

**Proposition 1.3.** *Assume  $\mathcal{V}$  is a simple potential and that  $V(x) = O(|x|^r)$  as  $|x| \rightarrow \infty$  for some  $1 \leq r < p$ . Then*

$$\mathcal{U}(\mu) = \int_{\mathbb{R}^d} u(x) d\mu(x), \quad \mu \in \mathcal{M}_p. \tag{1.13}$$

Moreover, there is also a Borel map  $\Psi : \mathbb{R}^d \times [0, \infty) \rightarrow \mathbb{R}^d$  such that for each  $x \in \mathbb{R}^d$ ,  $t \mapsto \Psi(x, t)$  is a minimizer for  $u(x)$ , and for each  $\mu \in \mathcal{M}_p$ ,

$$\sigma(t) := \Psi(\cdot, t) \# \mu, \quad t \in [0, \infty),$$

is a minimizing path for  $\mathcal{U}(\mu)$ .

A corollary of the above proposition provides a connection between the classical HJE (1.3) and the Euler–Poisson system (1.8). We show in Corollary 4.2 that if  $\sigma$  is any minimizing trajectory for  $\mathcal{U}(\mu)$  with minimal velocity  $v$ , then

$$|v(x, t)|^{p-2} v(x, t) = -\nabla u(x), \quad \sigma(t) \text{ a.e. } x \in \mathbb{R}^d, \tag{1.14}$$

for Lebesgue almost every  $t > 0$ . This identity can be seen at a heuristic level by differentiating the classical HJE (1.3) and comparing the result to the Euler–Poisson system (1.8), when  $\mathcal{V}$  satisfies (1.12).

There is a growing literature on value functions in the space of probability measures [9, 11–13, 18] and more generally in metric spaces [3, 5, 16, 19, 14, 15]. However, the bulk of these efforts have been restricted to optimization problems on a *finite time horizon*; as mentioned, we consider model infinite horizon problems in this paper. Regarding the Wasserstein space, the main technical differences with our work on finite horizon required a nonstandard, weighted Poincaré inequality (Lemma 2.1) and a new probabilistic representation of  $AC_{p,\delta}(\mathcal{M}_p)$  paths (Theorem A.3).

One aspiration we have is to further develop the theory of action minimizing paths. We conjecture that a necessary and sufficient condition for minimizing trajectories is that they satisfy a gradient flow condition analogous to (1.4); see Conjecture 3.3. Unfortunately, we are only able to verify a particular case in (3.5). Another open problem is to verify whether generalized value functions are unique as viscosity solutions of the HJE (1.11). These are both areas of ongoing

research. We also remark that uniqueness has been verified when  $p = 2$  with a different notion of solution [3].

This paper is organized as follows. In Section 2, we deduce several important properties of generalized value functions including various continuity assertions. In Section 3, we study minimizers of  $\mathcal{U}$ , derive the Euler–Poisson system and verify Theorem 1.1. We study simple potentials and prove Proposition 1.3 in Section 4; and in Section 5, we verify Theorem 1.2 which asserts the viscosity solution property of  $\mathcal{U}$ . We thank the Universitat Politècnica de Catalunya, the Math Science Research Institute and Nathan and Angela George for their hospitality during the writing of this paper.

## 2. Various properties

In this section, we will deduce various characteristics of generalized value functions  $\mathcal{U}$  (1.5). In particular, we will show that under the appropriate hypotheses that  $\mathcal{U}$  obeys a dynamic programming principle and is continuous. First, we will start by showing that  $\mathcal{U}$  is well defined and satisfies simple pointwise bounds provided  $\delta$  is large enough. To this end, we will need a type of weighted Poincaré inequality.

**Lemma 2.1.** *Assume  $p \in [1, \infty)$ . Then*

$$\left( \int_0^\infty e^{-\delta t} |u(t) - u(0)|^p dt \right)^{1/p} \leq \frac{p}{\delta} \left( \int_0^\infty e^{-\delta t} |\dot{u}(t)|^p dt \right)^{1/p} \tag{2.1}$$

for all  $\delta > 0$  and  $u \in AC_{loc}([0, \infty); \mathbb{R})$ .

**Proof.** Assume  $u(0) = 0$ ,  $p \in (1, \infty)$  and fix  $T > 0$ . Integrating by parts and using Hölder’s inequality gives

$$\begin{aligned} \int_0^T e^{-\delta t} |u(t)|^p dt &= -\frac{1}{\delta} |u(t)|^p e^{-\delta t} \Big|_0^T + \frac{p}{\delta} \int_0^T |u(t)|^{p-2} u(t) \dot{u}(t) e^{-\delta t} dt \\ &= -\frac{1}{\delta} |u(T)|^p e^{-\delta T} + \frac{p}{\delta} \int_0^T |u(t)|^{p-2} u(t) e^{-\frac{\delta}{q}t} \dot{u}(t) e^{-\frac{\delta}{p}t} dt \\ &\leq \frac{p}{\delta} \left( \int_0^T e^{-\delta t} |u(t)|^p dt \right)^{1-1/p} \left( \int_0^T e^{-\delta t} |\dot{u}(t)|^p dt \right)^{1/p}. \end{aligned}$$

Thus, (2.1) holds by sending  $T \rightarrow \infty$ . Likewise, for  $p = 1$

$$\int_0^T e^{-\delta t} |u(t) - u(0)| dt \leq \int_0^T e^{-\delta t} \left( \int_0^t |\dot{u}(s)| ds \right) dt$$

$$\begin{aligned}
 &= -\frac{1}{\delta} \left( \int_0^t |\dot{u}(s)| ds \right) e^{-\delta t} \Big|_0^T + \frac{1}{\delta} \int_0^T e^{-\delta t} |\dot{u}(t)| dt \\
 &\leq \frac{1}{\delta} \int_0^T e^{-\delta t} |\dot{u}(t)| dt.
 \end{aligned}$$

Again we conclude by sending  $T \rightarrow \infty$ .  $\square$

An immediate corollary of Lemma 2.1 is as follows.

**Corollary 2.2.** Assume  $p \in [1, \infty)$  and  $\delta > 0$ . Then

$$\left( \int_0^\infty e^{-\delta t} W_p(\sigma(t), \sigma(0))^p dt \right)^{1/p} \leq \frac{p}{\delta} \left( \int_0^\infty e^{-\delta t} \|\dot{\sigma}(t)\|^p dt \right)^{1/p} \tag{2.2}$$

for each locally  $p$ -absolutely continuous path  $\sigma : [0, \infty) \rightarrow \mathcal{M}_p$ .

**Proof.** Define the function  $u(t) := W_p(\sigma(t), u(0))$ , for  $t \in [0, \infty)$ . The triangle inequality implies  $|u(t_1) - u(t_2)| \leq W_p(\sigma(t_1), \sigma(t_2))$ ; in particular,  $u \in AC_{loc}([0, \infty); \mathbb{R})$ , and  $|\dot{u}(t)| \leq \|\dot{\sigma}(t)\|$ . Repeating the proof of (2.1), we easily conclude (2.2).  $\square$

**Lemma 2.3.** Assume the growth condition (1.10) and

$$p(2p/\delta)^p \alpha < 1. \tag{2.3}$$

Then

$$-\frac{1}{\delta} (\beta + 2^p \alpha W_p(\mu, \varrho)^p) \leq \mathcal{U}(\mu) \leq -\frac{1}{\delta} \mathcal{V}(\mu). \tag{2.4}$$

In particular,  $\mathcal{U}(\mu) > -\infty$  for all large values of  $\delta$ .

**Proof.** Let  $\mu \in \mathcal{M}_p$  and  $\sigma$  be admissible for  $\mathcal{U}(\mu)$ . Using the growth assumption on  $\mathcal{V}$  and the weighted Poincaré inequality

$$\begin{aligned}
 \int_0^\infty e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}(t)\|^p - \mathcal{V}(\sigma(t)) \right) dt &\geq \int_0^s e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}(t)\|^p - (\beta + \alpha W_p(\rho, \sigma(t))^p) \right) dt \\
 &\geq -\frac{1}{\delta} (\beta + 2^p \alpha W_p(\mu, \varrho)^p) \\
 &\quad + \left( \frac{1}{p} - \alpha \left( \frac{2p}{\delta} \right)^p \right) \int_0^\infty e^{-\delta t} \|\dot{\sigma}(t)\|^p dt \\
 &\geq -\frac{1}{\delta} (\beta + 2^p \alpha W_p(\mu, \varrho)^p).
 \end{aligned}$$

The lower bound in (2.4) is now immediate. Choosing  $\sigma(t) = \mu$  for all  $t$  in (1.5) gives the upper bound in (2.4).  $\square$

Next we derive the all important *Dynamic Programming Principle*. The proof here is not so different from the well known arguments used to prove dynamic programming for (1.1) (see Lemma 7.1 of [10]), but we include it for completeness. We also remark that the HJE (1.11) is an infinitesimal version of the dynamic programming principle (2.5).

**Proposition 2.4.** For each  $T \in [0, \infty)$  and  $\mu \in \mathcal{M}_p$ ,

$$\mathcal{U}(\mu) = \inf \left\{ e^{-\delta T} \mathcal{U}(\sigma(T)) + \int_0^T e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}(t)\|^p - \mathcal{V}(\sigma(t)) \right) dt \right\} \tag{2.5}$$

where the infimum is taken over paths  $\sigma \in AC_{p,\delta}(\mathcal{M}_p)$  with  $\sigma(0) = \mu$ .

**Proof.** Let  $\sigma$  be admissible for  $\mathcal{U}(\mu)$ . Then  $[0, \infty) \ni t \mapsto \sigma(t + T)$  is admissible for  $\mathcal{U}(\sigma(T))$  and

$$\begin{aligned} & e^{-\delta T} \mathcal{U}(\sigma(T)) + \int_0^T e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}(t)\|^p - \mathcal{V}(\sigma(t)) \right) dt \\ & \leq \int_0^\infty e^{-\delta(t+T)} \left( \frac{1}{p} \|\dot{\sigma}(t+T)\|^p - \mathcal{V}(\sigma(t+T)) \right) dt + \int_0^T e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}(t)\|^p - \mathcal{V}(\sigma(t)) \right) dt \\ & = \int_0^\infty e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}(t)\|^p - \mathcal{V}(\sigma(t)) \right) dt. \end{aligned}$$

Consequently, the left hand side of (2.5) is at least as large as the right hand side.

Again let  $\sigma$  be admissible for  $\mathcal{U}(\mu)$  and  $\epsilon > 0$ . Choose a path  $\bar{\sigma} \in AC_{p,\delta}(\mathcal{M}_p)$  such that  $s \mapsto \bar{\sigma}(s + T)$  is  $\epsilon$  optimal for  $\mathcal{U}(\sigma(T))$ . That is

$$\mathcal{U}(\sigma(T)) + \epsilon > \int_0^\infty e^{-\delta t} \left( \frac{1}{p} \|\dot{\bar{\sigma}}(s+T)\|^p - \mathcal{V}(\bar{\sigma}(s+T)) \right) ds.$$

Changing variables in the above integral  $t = s + T$  and manipulating similar to how we did above gives

$$e^{-\delta T} \mathcal{U}(\sigma(T)) + \int_0^T e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}(t)\|^p - \mathcal{V}(\sigma(t)) \right) dt > -e^{-\delta T} \epsilon + \mathcal{U}(\mu).$$

It now follows that the right hand side of (2.5) is no less than the left hand side.  $\square$

**Remark 2.5.** It is clear that the infimum in (2.5) can be taken over paths  $\sigma \in AC_p([0, T], \mathcal{M}_p)$  with  $\sigma(0) = \mu$ .

**Remark 2.6.** Suppose that  $\sigma$  is a minimizer for  $\mathcal{U}(\mu)$ , then

$$\begin{aligned} \mathcal{U}(\mu) &= \int_0^\infty e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}(t)\|^p - \mathcal{V}(\sigma(t)) \right) dt \\ &= \int_0^T e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}(t)\|^p - \mathcal{V}(\sigma(t)) \right) dt + \int_T^\infty e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}(t)\|^p - \mathcal{V}(\sigma(t)) \right) dt \\ &= \int_0^T e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}(t)\|^p - \mathcal{V}(\sigma(t)) \right) dt \\ &\quad + e^{-\delta T} \int_0^\infty e^{-\delta s} \left( \frac{1}{p} \|\dot{\sigma}(T+s)\|^p - \mathcal{V}(\sigma(T+s)) \right) ds \\ &\geq \int_0^T e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}(t)\|^p - \mathcal{V}(\sigma(t)) \right) dt + e^{-\delta T} \mathcal{U}(\sigma(T)) \\ &\geq \mathcal{U}(\mu). \end{aligned}$$

The last inequality follows from (2.5). Hence  $\sigma$  is also a minimizer of

$$AC_{p,\delta}(\mathcal{M}_p) \ni \rho \mapsto e^{-\delta T} \mathcal{U}(\rho(T)) + \int_0^T e^{-\delta t} \left( \frac{1}{p} \|\dot{\rho}(t)\|^p - \mathcal{V}(\rho(t)) \right) dt \tag{2.6}$$

for each  $T > 0$ .

Next we derive continuity properties of value functions. First we show that  $\mathcal{U}$  is in general continuous and then refine this statement in terms of the modulus of continuity of  $\mathcal{V}$ .

**Proposition 2.7.**  $\mathcal{U} \in C(\mathcal{M}_p)$ .

**Proof.** Assume  $\mu_n \rightarrow \mu$  in  $\mathcal{M}_p$  and let  $\sigma$  be admissible for  $\mathcal{U}(\mu)$ . Define for  $\eta > 0$

$$\sigma_n(s) := \begin{cases} \underline{\sigma}_n(s/\eta), & 0 \leq s \leq \eta, \\ \sigma(s - \eta), & \eta \leq s < \infty, \end{cases}$$

where  $\underline{\sigma}_n : [0, 1] \rightarrow \mathcal{M}_p$  is a constant speed geodesic joining  $\mu_n$  to  $\mu$ . As  $\sigma_n$  is admissible for  $\mathcal{U}(\mu_n)$ ,



$$\begin{aligned}
 \mathcal{U}(\mu_n) &\leq \int_0^\infty e^{-\delta s} \left( \frac{1}{p} \|\dot{\sigma}_n(s)\|^p - \mathcal{V}(\sigma_n(s)) \right) ds \\
 &= \int_0^\eta e^{-\delta s} \left( \frac{1}{p} \|\dot{\sigma}_n(s)\|^p - \mathcal{V}(\sigma_n(s)) \right) ds + \int_\eta^\infty e^{-\delta s} \left( \frac{1}{p} \|\dot{\sigma}_n(s)\|^p - \mathcal{V}(\sigma_n(s)) \right) ds \\
 &= \int_0^\eta e^{-\delta s} \left( \frac{1}{p} \left( \frac{W_p(\mu, \mu_n)}{\eta} \right)^p - \mathcal{V}(\underline{\sigma}_n(s/\eta)) \right) ds \\
 &\quad + \int_0^\infty e^{-\delta(t+\eta)} \left( \frac{1}{p} \|\dot{\sigma}(t)\|^p - \mathcal{V}(\sigma(t)) \right) dt.
 \end{aligned}$$

Since  $W_p(\underline{\sigma}_n(s/\eta), \mu) \leq C$  for some universal constant  $C$ , we can pass to the limit above using the continuity of  $\mathcal{V}$  to obtain

$$\limsup_{n \rightarrow \infty} \mathcal{U}(\mu_n) \leq -\frac{1 - e^{-\delta\eta}}{\delta} \mathcal{V}(\mu) + e^{-\delta\eta} \int_0^\infty e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}(t)\|^p - \mathcal{V}(\sigma(t)) \right) dt.$$

As  $\eta$  and  $\sigma$  were arbitrary, we conclude that  $\limsup_{n \rightarrow \infty} \mathcal{U}(\mu_n) \leq \mathcal{U}(\mu)$ .

Now choose any  $\epsilon_n > 0$  tending to 0 as  $n \rightarrow \infty$ , and  $\sigma_n$  admissible for  $\mathcal{U}(\mu_n)$  for which

$$\mathcal{U}(\mu_n) > -\epsilon_n + \int_0^\infty e^{-\delta s} \left( \frac{1}{p} \|\dot{\sigma}_n(s)\|^p - \mathcal{V}(\sigma_n(s)) \right) ds.$$

Define

$$\sigma_{n,\eta}(t) := \begin{cases} \underline{\sigma}_n(1 - t/\eta), & 0 \leq t \leq \eta, \\ \sigma_n(t - \eta), & \eta \leq t < \infty, \end{cases}$$

and observe that  $\sigma_{n,\eta}$  is admissible for  $\mathcal{U}(\mu)$ . Similar to our computations above, we find

$$\begin{aligned}
 \mathcal{U}(\mu) &\leq \int_0^\eta e^{-\delta t} \left\{ \frac{1}{p} \left( \frac{W_p(\mu, \mu_n)}{\eta} \right)^p - \mathcal{V}(\underline{\sigma}_n(1 - t/\eta)) \right\} dt \\
 &\quad + \int_0^\infty e^{-\delta(\eta+s)} \left( \frac{1}{p} \|\dot{\sigma}_n(s)\|^p - \mathcal{V}(\sigma_n(s)) \right) ds \\
 &< \int_0^\eta e^{-\delta t} \left\{ \frac{1}{p} \left( \frac{W_p(\mu, \mu_n)}{\eta} \right)^p - \mathcal{V}(\underline{\sigma}_n(1 - t/\eta)) \right\} dt + (\epsilon_n + \mathcal{U}(\mu_n)) e^{-\delta\eta}.
 \end{aligned}$$

We easily compute  $\mathcal{U}(\mu) \leq \liminf_{n \rightarrow \infty} \mathcal{U}(\mu_n) e^{-\delta n} - \frac{1 - e^{-\delta \eta}}{\delta} \mathcal{V}(\mu)$ . Since  $\eta > 0$  was arbitrary,  $\mathcal{U}$  is lower semicontinuous.  $\square$

**Remark 2.8.** It is worth noting that the proof of continuity of the generalized value functions on infinite horizons requires less hypotheses than the proof for finite horizons (as the value function is continuous with no constraints on  $\delta$ ). See Proposition 2.5 of [18].

**Proposition 2.9.** *If  $V$  is uniformly continuous with modulus  $\omega$ , then  $\mathcal{U}$  is uniformly continuous with modulus  $\omega/\delta$ .*

**Proof.** Assume first that  $\mu^1$  and  $\mu^2$  are absolutely continuous with respect to Lebesgue measure. Let  $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$  be the unique Borel mapping that pushes  $\mu^1$  onto  $\mu^2$  and

$$W_p(\mu^1, \mu^2)^p = \int_{\mathbb{R}^d} |x - T(x)|^p d\mu^1(x) = \int_{\mathbb{R}^d} |y - T^{-1}(y)|^p d\mu^2(y).$$

Let  $\epsilon > 0$  and choose  $\sigma^2$  admissible for  $\mathcal{U}(\mu^2)$  such that

$$\mathcal{U}(\mu^2) > -\epsilon + \int_0^\infty e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}^2(t)\|^p - \mathcal{V}(\sigma^2(t)) \right) dt.$$

By the probabilistic representation of  $AC_{p,\delta}(\mathcal{M}_p)$  paths detailed in Theorem A.3, there is a Borel probability measure  $\eta$  on  $\mathbb{R}^d \times \Gamma$  such that

$$\sigma^2(t) := e(t)_{\#}\eta, \quad t \geq 0.$$

Here  $\Gamma$  is the space  $C([0, \infty); \mathbb{R}^d)$  with the topology of local uniform convergence,

$$e(t) : \mathbb{R}^d \times \Gamma \rightarrow \mathbb{R}^d; \quad (x, \gamma) \mapsto \gamma(t) \tag{2.7}$$

is the *evaluation map* and  $\eta$  is concentrated on  $(x, \gamma)$  satisfying  $\gamma(0) = x$  and  $\dot{\gamma}(t) = v(\gamma(t), t)$ , a.e.  $t \geq 0$ . The field  $v$  is the minimal velocity of  $\sigma^2$ .

Define a map

$$S(x, \gamma) := (T^{-1}(x), \gamma + T^{-1}(x) - x)$$

and a measure on  $\mathbb{R}^d \times \Gamma$

$$\lambda := S_{\#}\eta.$$

Also set

$$\sigma^1(t) := e(t)_{\#}\lambda, \quad t \geq 0.$$

As in Proposition 2.6 of [18], one checks that  $\sigma^1$  is admissible for  $\mathcal{U}(\mu^2)$ , that  $\|\dot{\sigma}^1(t)\| \leq \|\dot{\sigma}^2(t)\|$  for almost every  $t \geq 0$ , and

$$W_p(\sigma^1(t), \sigma^2(t)) \leq W_p(\mu^1, \mu^2).$$

Therefore, it follows that

$$\begin{aligned} \mathcal{U}(\mu^1) - \mathcal{U}(\mu^2) &\leq \epsilon + \int_0^\infty e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}^1(t)\|^p - \frac{1}{p} \|\dot{\sigma}^2(t)\|^p + \mathcal{V}(\sigma^2(t)) - \mathcal{V}(\sigma^1(t)) \right) dt \\ &\leq \epsilon + \int_0^\infty e^{-\delta t} \omega(W_p(\sigma^1(t), \sigma^2(t))) dt \\ &\leq \epsilon + \int_0^\infty e^{-\delta t} \omega(W_p(\mu^1, \mu^2)) dt \\ &= \epsilon + \frac{1}{\delta} \omega(W_p(\mu^1, \mu^2)). \end{aligned}$$

Interchanging,  $\mu^1$  and  $\mu^2$  yields the claim in the case where  $\mu^1$  and  $\mu^2$  are absolutely continuous. The general assertion now follows from the density of absolutely continuous measures in  $\mathcal{M}_p$  (Lemma 7.1.10 [2]) and the previous proposition.  $\square$

### 3. Minimizing trajectories

In this section, we prove our main existence result, [Theorem 1.1](#). Our main tool is a compactness lemma, which involves the weak or narrow convergence of sequences of paths in  $AC_{p,\delta}(\mathcal{M}_p)$ . Recall that a sequence  $\{\mu^k\}_{k \in \mathbb{N}} \subset \mathcal{P}_p(\mathbb{R}^d)$  converges to  $\mu \in \mathcal{P}_p(\mathbb{R}^d)$  narrowly if

$$\lim_{k \rightarrow \infty} \int_{\mathbb{R}^d} f(x) d\mu^k(x) = \int_{\mathbb{R}^d} f(x) d\mu(x)$$

for each bounded, continuous function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ ; see Chapter 5 of [2] for more on the narrow convergence of probability measures. We begin our study of minimizing paths by deriving the Euler–Poisson system (1.8) as a necessary condition.

**Theorem 3.1.** *Assume  $\mathcal{V}$  is Lipschitz continuous, and that for each  $\mu \in \mathcal{M}_p$ , there is a mapping  $\nabla \mathcal{V}(\mu) \in L^p(\mu)$  for which*

$$\lim_{\epsilon \rightarrow 0} \frac{\mathcal{V}((\text{id}_{\mathbb{R}^d} + \epsilon \eta) \# \mu) - \mathcal{V}(\mu)}{\epsilon} = \int_{\mathbb{R}^d} \nabla \mathcal{V}(\mu) \cdot \eta d\mu \tag{3.1}$$

for each  $\eta \in C_c^\infty(\mathbb{R}^d; \mathbb{R}^d)$ . Then for any minimizer  $\sigma$  of  $\mathcal{U}(\mu)$ , the equations

$$\begin{cases} \partial_t(\sigma |v|^{p-2} v) + \nabla \cdot (\sigma |v|^{p-2} v \otimes v) = -\sigma \nabla \mathcal{V}(\sigma) + \sigma \delta |v|^{p-2} v, \\ \partial_t \sigma + \nabla \cdot (\sigma v) = 0 \end{cases}$$

hold in the sense of distributions on  $\mathbb{R}^d \times (0, \infty)$ ; here  $v$  is the minimal velocity for  $\sigma$ .

**Proof.** Let  $\Psi \in C_c^\infty(\mathbb{R}^d \times (0, \infty); \mathbb{R}^d)$ , and set

$$\begin{cases} \sigma^\epsilon(t) := (\text{id}_{\mathbb{R}^d} + \epsilon\Psi(\cdot, t))_\# \sigma(t), \\ v^\epsilon(\cdot, t) := (v(t) + \epsilon(\partial_t\Psi(\cdot, t) + (v(\cdot, t) \cdot \nabla)\Psi(\cdot, t))) \circ (\text{id}_{\mathbb{R}^d} + \epsilon\Psi(\cdot, t))^{-1} \end{cases}$$

for  $\epsilon$  so small that  $x \mapsto x + \epsilon\Psi(x, t)$  is invertible for each  $t > 0$ . One checks that

$$\partial_t \sigma^\epsilon + \nabla \cdot (\sigma^\epsilon v^\epsilon) = 0, \quad \mathbb{R}^d \times (0, \infty),$$

and that  $\sigma^\epsilon$  is admissible for  $\mathcal{U}(\mu)$ . Thus

$$\begin{aligned} \mathcal{U}(\mu) &= \int_0^\infty e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}(t)\|^p - \mathcal{V}(\sigma(t)) \right) dt \\ &= \int_0^\infty e^{-\delta t} \left( \frac{1}{p} \int_{\mathbb{R}^d} |v(x, t)|^p d\sigma_t(x) - \mathcal{V}(\sigma(t)) \right) dt \\ &\leq \int_0^\infty e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}^\epsilon(t)\|^p - \mathcal{V}(\sigma^\epsilon(t)) \right) dt \\ &\leq \int_0^\infty e^{-\delta t} \left( \frac{1}{p} \int_{\mathbb{R}^d} |v^\epsilon(x, t)|^p d\sigma_t^\epsilon(x) - \mathcal{V}(\sigma^\epsilon(t)) \right) dt \\ &= \int_0^\infty e^{-\delta t} \left( \frac{1}{p} \int_{\mathbb{R}^d} |v + \epsilon(\partial_t\Psi + (v \cdot \nabla)\Psi)|^p d\sigma_t(x) - \mathcal{V}((\text{id}_{\mathbb{R}^d} + \epsilon\Psi(\cdot, t))_\# \sigma(t)) \right) dt. \end{aligned}$$

Consequently, the derivative on the right hand side expression above must be zero when taken at  $\epsilon = 0$ . Employing (3.1) and our assumption that  $\mathcal{V}$  is Lipschitz, we use limit theorems from measure theory to compute the derivative in question and find

$$0 = \int_0^\infty \int_{\mathbb{R}^d} e^{-\delta t} \{ (\partial_t\Psi + (v \cdot \nabla)\Psi) \cdot |v|^{p-2}v - \nabla\mathcal{V}(\sigma) \cdot \Psi \} d\sigma_t(x) dt. \tag{3.2}$$

Choosing  $\Psi(x, t) = e^{\delta t} \Phi(x, t)$ , where  $\Phi \in C_c^\infty(\mathbb{R}^d \times (0, \infty); \mathbb{R}^d)$  gives (1.8).  $\square$

Note that (3.2) also holds for  $\Psi \in C_c^1(\mathbb{R}^d \times (0, \infty); \mathbb{R}^d)$ . Also observe that by selecting  $\Psi(x, t) = \eta(x)f(t)$  in (3.2), where  $f \in C_c^\infty(0, \infty)$  and  $\eta \in C_c^1(\mathbb{R}^d)$ , gives

$$\frac{d}{dt} \int_{\mathbb{R}^d} e^{-\delta t} |v|^{p-2}v \cdot \eta d\sigma_t = \int_{\mathbb{R}^d} e^{-\delta t} \{ |v|^{p-2}v \cdot ((v \cdot \nabla)\eta) - \nabla\mathcal{V}(\sigma) \cdot \eta \} d\sigma_t. \tag{3.3}$$

Eq. (3.3) holds in the sense of distributions on  $(0, \infty)$ .

It now follows that the  $L^1[0, \infty)$  function

$$[0, \infty) \ni t \mapsto \int_{\mathbb{R}^d} e^{-\delta t} |v(x, t)|^{p-2} v(x, t) \cdot \eta(x) d\sigma_t(x)$$

can be identified with an absolutely continuous function belonging to  $W^{1,1}[0, \infty)$ . Moreover, standard arguments can be used to prove that  $[0, \infty) \ni t \mapsto \sigma(t)|v(\cdot, t)|^{p-2}v(\cdot, t)$  has a continuous representative with values in the dual space of the closure of  $C_c^1(\mathbb{R}^d; \mathbb{R}^d)$  with the norm

$$\|\eta\| := \sup_{\mathbb{R}^d} |\eta| + \sup_{\mathbb{R}^d} |\nabla \eta|.$$

See the proof of Lemma 8.1.2 in [2] for more details on this technical point.

**Corollary 3.2.** *Assume the hypotheses of Theorem 3.1 and identify  $t \mapsto \sigma(t)|v(\cdot, t)|^{p-2}v(\cdot, t) \in (C_c^1(\mathbb{R}^d; \mathbb{R}^d))'$  with its continuous representative. Then for all  $\eta \in C_c^1(\mathbb{R}^d; \mathbb{R}^d)$*

$$\lim_{t \rightarrow \infty} e^{-\delta t} \int_{\mathbb{R}^d} |v(t, x)|^{p-2} v(t, x) \cdot \eta(x) d\sigma_t(x) = 0 \tag{3.4}$$

and

$$\liminf_{\epsilon \rightarrow 0^+} \frac{\mathcal{U}((\text{id}_{\mathbb{R}^d} + \epsilon \eta) \# \sigma(t)) - \mathcal{U}(\sigma(t))}{\epsilon} \geq - \int_{\mathbb{R}^d} |v(x, t)|^{p-2} v(x, t) \cdot \eta(x) d\sigma_t(x) \tag{3.5}$$

for  $t > 0$ .

**Proof.** It is also not difficult to see that if  $g \in C^\infty([0, \infty))$ ,  $g(0) = 0$ , and  $g(t) = 1$  for all  $t$  large, then  $\Psi(t, x) = g(t)\eta(x)$  is a valid test function in (3.2). Substituting this test function yields

$$\begin{aligned} 0 &= \int_0^\infty \int_{\mathbb{R}^d} e^{-\delta t} \{(\partial_t \Psi + (v \cdot \nabla) \Psi) \cdot |v|^{p-2} v - \nabla \mathcal{V}(\sigma(t)) \cdot \Psi(t)\} d\sigma_t(x) dt \\ &= \lim_{T \rightarrow \infty} \int_0^T \int_{\mathbb{R}^d} e^{-\delta t} \{(\partial_t \Psi + (v \cdot \nabla) \Psi) \cdot |v|^{p-2} v - \nabla \mathcal{V}(\sigma(t)) \cdot \Psi(t)\} d\sigma_t(x) dt \\ &= \lim_{T \rightarrow \infty} \int_0^T \int_{\mathbb{R}^d} e^{-\delta t} \{(g'(t)\eta + (v \cdot \nabla)g(t)\eta) \cdot |v|^{p-2} v - \nabla \mathcal{V}(\sigma(t)) \cdot g(t)\eta\} d\sigma_t(x) dt \\ &= \lim_{T \rightarrow \infty} \left\{ g(t) \int_{\mathbb{R}^d} e^{-\delta t} |v|^{p-2} v \cdot \eta d\sigma_t \right\} \Big|_0^T \end{aligned}$$

$$\begin{aligned}
 & + \int_0^T g(t) \left[ -\frac{d}{dt} \int_{\mathbb{R}^d} e^{-\delta t} |v|^{p-2} v \cdot \eta d\sigma_t + \int_{\mathbb{R}^d} e^{-\delta t} \{ |v|^{p-2} v \cdot ((v \cdot \nabla)\eta) - \nabla V(\sigma) \cdot \eta \} d\sigma_t \right] \Bigg\} \\
 & = \lim_{T \rightarrow \infty} \int_{\mathbb{R}^d} e^{-\delta T} |v(x, T)|^{p-2} v(x, T) \cdot \eta(x) d\sigma_T(x)
 \end{aligned}$$

which is (3.4).

Now let  $f \in C_c^\infty(0, \infty)$  with  $f(T) = 1$ , and set

$$\Psi(x, t) := \eta(x) f(t) \in C_c^\infty(\mathbb{R}^d \times (0, \infty); \mathbb{R}^d).$$

Also denote  $\sigma^\epsilon$  and  $v^\epsilon$  as in the proof of Theorem 3.1. As  $\sigma$  is a minimizer of (2.6),

$$\begin{aligned}
 \mathcal{U}(\mu) & = e^{-\delta T} \mathcal{U}(\sigma(T)) + \int_0^T e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}(t)\|^p - \mathcal{V}(\sigma(t)) \right) dt \\
 & = e^{-\delta T} \mathcal{U}(\sigma(s)) + \int_0^T e^{-\delta t} \left( \frac{1}{p} \int_{\mathbb{R}^d} |v(x, t)|^p d\sigma_t(x) - \mathcal{V}(\sigma(t)) \right) dt \\
 & \leq e^{-\delta T} \mathcal{U}(\sigma^\epsilon(T)) + \int_0^T e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}^\epsilon(t)\|^p - \mathcal{V}(\sigma^\epsilon(t)) \right) dt \\
 & \leq e^{-\delta T} \mathcal{U}(\sigma^\epsilon(T)) + \int_0^T e^{-\delta t} \left( \frac{1}{p} \int_{\mathbb{R}^d} |v^\epsilon(x, t)|^p d\sigma_t^\epsilon(x) - \mathcal{V}(\sigma^\epsilon(t)) \right) dt \\
 & \leq e^{-\delta T} \mathcal{U}((\text{id}_{\mathbb{R}^d} + \epsilon\eta)_\# \sigma(T)) \\
 & \quad + \int_0^T e^{-\delta t} \left( \frac{1}{p} \int_{\mathbb{R}^d} |v + \epsilon(\partial_t \Psi + v \cdot \nabla \Psi)|^p d\sigma_t - \mathcal{V}((\text{id}_{\mathbb{R}^d} + \epsilon\Psi(\cdot, t))_\# \sigma(t)) \right) dt
 \end{aligned}$$

for each  $\epsilon > 0$  small enough. Similar to computations performed in the proof of Theorem 3.1

$$\begin{aligned}
 & \liminf_{\epsilon \rightarrow 0^+} \frac{\mathcal{U}((\text{id}_{\mathbb{R}^d} + \epsilon\eta)_\# \sigma(T)) - \mathcal{U}(\sigma(T))}{\epsilon} \\
 & \geq -e^{\delta T} \int_0^T \int_{\mathbb{R}^d} e^{-\delta t} \{ (\partial_t \Psi + (v \cdot \nabla)\Psi) \cdot |v|^{p-2} v - \nabla \mathcal{V}(\sigma(t)) \cdot \Psi(\cdot, t) \} d\sigma_t dt
 \end{aligned}$$

$$\begin{aligned}
 &= -e^{\delta T} \int_0^T \left[ f'(t) \left\{ \int_{\mathbb{R}^d} e^{-\delta t} |v|^{p-2} v \cdot \eta d\sigma_t \right\} \right. \\
 &\quad \left. + f(t) \left\{ \int_{\mathbb{R}^d} e^{-\delta t} |v|^{p-2} v \cdot ((v \cdot \nabla)\eta) - \nabla \mathcal{V}(\sigma) \cdot \eta \right\} d\sigma_t \right] dt \\
 &= - \int_{\mathbb{R}^d} |v(x, T)|^{p-2} v(x, T) \cdot \eta(x) d\sigma_T(x).
 \end{aligned}$$

The last inequality follows from an integration by parts and the use of the identity (3.3).  $\square$

We believe that more is true. We prove a special case of the below conjecture in the case of simple potentials; see Corollary 5.7.

**Conjecture 3.3.** *Under the assumptions of Theorem 3.1,*

$$-|v(\cdot, t)|^{p-2} v(\cdot, t) \in \nabla^{-\mathcal{U}}(\sigma(t)), \quad a.e. t \in (0, \infty).$$

$\nabla^{-\mathcal{U}}$  is specified in Definition 5.1.

We now will address some issues related to the existence of minimizing paths by using ideas from the calculus of variations. It will be necessary for us to employ the narrow convergence of measures. Recall that the Lévy–Prokhorov metric

$$\Lambda(\mu, \nu) := \inf\{\epsilon > 0: \mu(A) \leq \nu(A^\epsilon) + \epsilon, \nu(A) \leq \mu(A^\epsilon) + \epsilon, \text{ for Borel } A \subset \mathbb{R}^d\} \quad (3.6)$$

completely metrizes  $\mathcal{P}_p(\mathbb{R}^d)$  (see Chapter 6 of [6]); here  $A_\epsilon := \bigcup_{z \in A} B_\epsilon(z)$ . Moreover, the following inequality

$$\Lambda^2 \leq W_1 \quad (3.7)$$

holds (Corollary 2.18 of [17]). The inequality (3.7) is critical in the following compactness lemma.

**Lemma 3.4.** *Assume  $\{\sigma^k\}_{k \in \mathbb{N}} \subset AC_{p,\delta}(\mathcal{M}_p)$  for some  $p \in (1, \infty)$  and  $\delta > 0$ . Further suppose*

$$\sigma^k(0) = \mu$$

and

$$\sup_k \int_0^\infty e^{-\delta t} \|\dot{\sigma}^k(t)\|^p dt < \infty.$$

*Then there is a subsequence  $\{\sigma^{k_j}\}_{j \in \mathbb{N}}$  and  $\sigma \in AC_{p,\delta}(\mathcal{M}_p)$  such that*

(i)  $\sigma^{k_j} \rightarrow \sigma$  locally uniformly in  $(\mathcal{P}_p(\mathbb{R}^d), \Lambda)$ ,

(ii)

$$\liminf_{j \rightarrow \infty} \int_0^\infty e^{-\delta t} \|\dot{\sigma}^{k_j}(t)\|^p dt \geq \int_0^\infty e^{-\delta t} \|\dot{\sigma}(t)\|^p dt,$$

(iii)

$$\lim_{j \rightarrow \infty} \int_0^\infty e^{-\delta t} \Lambda(\sigma^{k_j}(t), \sigma(t))^{2r} dt = 0$$

for all  $1 \leq r < p$ .

**Proof.** Note for each  $T > 0$ ,

$$\int_0^T \|\dot{\sigma}^k(t)\|^p dt \leq e^{\delta T} \int_0^T e^{-\delta t} \|\dot{\sigma}^k(t)\|^p dt \leq e^{\delta T} \int_0^\infty e^{-\delta t} \|\dot{\sigma}^k(t)\|^p dt \leq C e^{\delta T}.$$

By Proposition 4.1 of our previous work [18], there is a subsequence of  $\sigma^k$  convergent in  $C([0, T]; (\mathcal{P}_p(\mathbb{R}^d), \Lambda))$ . Using a routine diagonalization argument, we obtain a sequence  $\{\sigma^{k_j}\}_{j \in \mathbb{N}}$  and  $\sigma \in C([0, \infty); (\mathcal{P}_p(\mathbb{R}^d), \Lambda))$  such that  $\sigma = \lim_{j \rightarrow \infty} \sigma^{k_j}$  locally uniformly in the narrow topology.

Also note that  $t \mapsto \|\dot{\sigma}^{k_j}(t)\|$  is bounded in  $L^p((0, \infty); e^{-\delta t} dt)$  and so has a further subsequence (not relabeled here) that converges weakly to some  $g$ . By the lower semicontinuity properties of  $W_p$

$$\begin{aligned} W_p(\sigma(t_1), \sigma(t_2)) &\leq \liminf_{j \rightarrow \infty} W_p(\sigma^{k_j}(t_1), \sigma^{k_j}(t_2)) \\ &\leq \liminf_{j \rightarrow \infty} \int_{t_1}^{t_2} \|\dot{\sigma}^{k_j}(t)\| dt \\ &= \liminf_{j \rightarrow \infty} \int_0^\infty e^{-\delta t} (e^{\delta t} \chi_{[t_1, t_2]}(t)) \|\dot{\sigma}^{k_j}(t)\| dt \\ &= \int_0^\infty e^{-\delta t} (e^{\delta t} \chi_{[t_1, t_2]}(t)) g(t) dt \\ &= \int_{t_1}^{t_2} g(t) dt \end{aligned}$$



for  $0 \leq t_1 \leq t_2 < \infty$ . As  $g \in L^p_{loc}(0, \infty)$ ,  $\|\dot{\sigma}(t)\| \leq g(t)$  for almost every  $t \geq 0$ . Moreover, weak convergence implies

$$\int_0^\infty e^{-\delta t} \|\dot{\sigma}(t)\|^p dt \leq \int_0^\infty e^{-\delta t} g(t)^p dt \leq \liminf_{j \rightarrow \infty} \int_0^\infty e^{-\delta t} \|\dot{\sigma}^{k_j}(t)\|^p dt.$$

By Egorov’s theorem, for every  $\epsilon > 0$ , there is a Borel measurable  $A_\epsilon \subset [0, \infty)$  such that  $\int_{A_\epsilon} e^{-\delta t} dt \leq \epsilon$  and  $\sigma^{k_j} \rightarrow \sigma$  on  $[0, \infty) \setminus A_\epsilon$  uniformly in the narrow topology. Consequently for  $1 \leq r < p$ ,

$$\begin{aligned} & \int_0^\infty e^{-\delta t} \Lambda(\sigma^{k_j}(t), \sigma(t))^{2r} dt \\ &= \int_{[0, \infty) \setminus A_\epsilon} e^{-\delta t} \Lambda(\sigma^{k_j}(t), \sigma(t))^{2r} dt + \int_{A_\epsilon} e^{-\delta t} \Lambda(\sigma^{k_j}(t), \sigma(t))^{2r} dt \\ &\leq \frac{1}{\delta} \left[ \sup_{t \in [0, \infty) \setminus A_\epsilon} \Lambda(\sigma^{k_j}(t), \sigma(t)) \right]^{2r} + \left( \int_0^\infty e^{-\delta t} \Lambda(\sigma^{k_j}(t), \sigma(t))^{2p} dt \right)^{r/p} \epsilon^{1-r/p} \\ &\leq \frac{1}{\delta} \left[ \sup_{t \in [0, \infty) \setminus A_\epsilon} \Lambda(\sigma^{k_j}(t), \sigma(t)) \right]^{2r} + \left( \int_0^\infty e^{-\delta t} W_p(\sigma^{k_j}(t), \sigma(t))^p dt \right)^{r/p} \epsilon^{1-r/p} \\ &\leq \frac{1}{\delta} \left[ \sup_{t \in [0, \infty) \setminus A_\epsilon} \Lambda(\sigma^{k_j}(t), \sigma(t)) \right]^{2r} + C\epsilon^{1-r/p}. \end{aligned}$$

The final estimate follows from the weighted Poincaré inequality and the hypotheses of this theorem. Thus,

$$\limsup_{j \rightarrow \infty} \int_0^\infty e^{-\delta t} \Lambda(\sigma^{k_j}(t), \sigma(t))^{2r} dt \leq C\epsilon^{1-r/p}.$$

The claim now follows from sending  $\epsilon \rightarrow 0^+$ .  $\square$

We are finally in a position to prove [Theorem 1.1](#).

**Proof of Theorem 1.1.** Let  $\mu \in \mathcal{M}_p$  and  $\epsilon_k$  be a sequence of positive numbers tending to 0 as  $k \rightarrow \infty$ . Choose paths  $\sigma^k$  admissible for  $\mathcal{U}(\mu)$  such that

$$\mathcal{U}(\mu) > -\epsilon_k + \int_0^\infty e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}^k(t)\|^p - \mathcal{V}(\sigma^k(t)) \right) dt \tag{3.8}$$

for  $k \in \mathbb{N}$ . By assumption (2.3), we manipulate (3.8) as in Lemma 2.3 and conclude  $\int_0^\infty e^{-\delta t} \|\dot{\sigma}^k(t)\|^p dt \leq C$ . As a result, the sequence  $\{\sigma^k\}_{k \in \mathbb{N}}$  satisfies the hypotheses of Lemma 3.4. Hence, there is a subsequence  $\{\sigma^{k_j}\}_{j \in \mathbb{N}}$  and  $\sigma$  admissible for  $\mathcal{U}(\mu)$  for which  $\sigma^{k_j}$  converges to  $\sigma$  as described in the previous assertion. Employing hypothesis (1.9) and the assumption that  $\mathcal{V}$  is continuous with respect to narrow convergence, we apply the dominated convergence theorem to find

$$\lim_{j \rightarrow \infty} \int_0^\infty \mathcal{V}(\sigma^{k_j}(t)) e^{-\delta t} dt = \int_0^\infty \mathcal{V}(\sigma(t)) e^{-\delta t} dt.$$

The claim is now immediate from passing to the limit  $k = k_j \rightarrow \infty$  in (3.8).  $\square$

### 4. Simple potentials

In this section, we will focus on value functions in the case of simple potentials (1.12)

$$\mathcal{V}(\mu) = \int_{\mathbb{R}^d} V(x) d\mu(x).$$

We shall further assume  $V \in C^1(\mathbb{R}^d)$  and satisfies

$$|V(x)| \leq a|x|^r + b \tag{4.1}$$

for some  $a, b \in \mathbb{R}$  and  $1 \leq r < p$ . Under this assumption, classical value functions  $u = u(x)$  (1.1) can be shown to be well defined, continuous and have minimizing paths for each  $x \in \mathbb{R}^d$  (see Lemma A.1 in Appendix A). Moreover, using the compactness built into this classical optimization problem, we obtain a measurable flow map associated with minimizing paths. This will be crucial to our proof of Proposition 1.3.

**Proposition 4.1.** *Assume (4.1), assume  $u$  is the classical value function (1.1), and define the set valued mapping*

$$F(x) := \left\{ \gamma \in AC_{p,\delta}(\mathbb{R}^d) : u(x) = \int_0^\infty e^{-\delta t} \left( \frac{1}{p} |\dot{\gamma}(t)|^p - V(\gamma(t)) \right) dt, \gamma(0) = x \right\}$$

for  $x \in \mathbb{R}^d$ . Then there is a measurable map  $\Phi : \mathbb{R}^d \rightarrow AC_{p,\delta}(\mathbb{R}^d)$  such that for each  $x \in \mathbb{R}^d$ ,  $\Phi(x) \in F(x)$ .

**Proof.** For  $\gamma, \xi \in AC_{p,\delta}(\mathbb{R}^d)$  define

$$\Sigma(\gamma, \xi) := \Pi(\gamma, \xi) + \left[ \int_0^\infty e^{-\delta t} |\dot{\gamma}(t) - \dot{\xi}(t)|^p dt \right]^{1/p}$$

where

$$\Pi(\gamma, \xi) = \sum_{k=0}^{\infty} \frac{1}{2^k} \frac{\max_{0 \leq t \leq k} |\gamma(t) - \xi(t)|}{1 + \max_{0 \leq t \leq k} |\gamma(t) - \xi(t)|}. \tag{4.2}$$

In Proposition A.2 of Appendix A, we verify that  $\Pi$  makes  $C([0, \infty); \mathbb{R}^d)$  into a complete, separable metric space. Moreover, convergence under  $\Pi$  is equivalent to local uniform convergence of  $\mathbb{R}^d$  valued paths on  $[0, \infty)$ . Employing these facts about  $\Pi$ , it is straightforward to check that  $AC_{p,\delta}(\mathbb{R}^d)$  is a complete, separable metric space under the distance  $\Sigma$ .

According to Lemma A.1 in Appendix A,  $F(x) \neq \emptyset$  for each  $x \in \mathbb{R}^d$ . It is also routine to verify that  $F(x) \subset AC_{p,\delta}(\mathbb{R}^d)$  is closed. By Theorem 8.3.1 of [1], it suffices to show that for each  $\eta \in AC_{p,\delta}(\mathbb{R}^d)$ ,

$$x \mapsto \text{dist}(\eta, F(x)) \text{ is Borel measurable}$$

to conclude the assertion; here  $\text{dist}(\eta, S) := \inf_{\xi \in S} \Sigma(\eta, \xi)$ ,  $S \subset AC_{p,\delta}(\mathbb{R}^d)$ . We show in fact that this function is lower-semicontinuous on  $\mathbb{R}^d$ .

To this end, assume  $x_n \rightarrow x \in \mathbb{R}^d$  and choose  $x_{n_j}$  such that

$$\liminf_{n \rightarrow \infty} \text{dist}(\eta, F(x_n)) = \lim_{j \rightarrow \infty} \text{dist}(\eta, F(x_{n_j}))$$

Employing Lemma A.1, we may select  $\gamma_j \in F(x_{n_j})$  such that  $\text{dist}(\eta, F(x_{n_j})) = \Sigma(\eta, \gamma_j)$ . Notice

$$\int_0^{\infty} e^{-\delta t} \left( \frac{1}{p} |\dot{\gamma}_j(t)|^p - V(\gamma_j(t)) \right) dt = u(x_{n_j}) \leq C$$

as  $u \in C(\mathbb{R}^d)$ . It follows now from (4.1) and the weighted Poincaré inequality that  $\int_0^{\infty} e^{-\delta t} \times |\dot{\gamma}_j(t)|^p dt$  is bounded independently of  $j \in \mathbb{N}$ . Moreover,  $|\gamma_j(0)| = |x_j| \leq C$ . By Lemma A.1, there is a  $\gamma \in AC_{p,\delta}(\mathbb{R}^d)$  such that a subsequence of  $\{\gamma_j\}_{j \in \mathbb{N}}$  (not relabeled) converges  $\gamma$  in  $(C([0, \infty); \mathbb{R}^d), \Pi)$  and  $\dot{\gamma}_j \rightarrow \dot{\gamma}$  weakly in  $L^p([0, \infty); e^{-\delta t} dt)$ . In particular,  $\gamma \in F(x)$ . This convergence implies

$$\begin{aligned} \text{dist}(\eta, F(x)) &\leq \Sigma(\eta, \gamma) \\ &\leq \liminf_{j \rightarrow \infty} \Sigma(\eta, \gamma_j) \\ &= \liminf_{j \rightarrow \infty} \text{dist}(\eta, F(x_{n_j})) \\ &\leq \liminf_{n \rightarrow \infty} \text{dist}(\eta, F(x_n)). \quad \square \end{aligned}$$

Proposition 4.1 establishes that for each  $x$ ,  $\Phi(x)$  is a minimizing path for  $u(x)$ . Define a new map

$$\Psi : \mathbb{R}^d \times [0, \infty) \rightarrow \mathbb{R}^d; \quad (x, t) \mapsto e(t) \circ (x, \Phi(x)) \tag{4.3}$$

which is measurable, since it is the composition of measurable mappings (recall the evaluation map  $e(t)$  defined in (2.7)). When  $u$  is differentiable along its minimizing trajectories,  $t \mapsto \Psi(x, t)$  satisfies the optimality equations (1.4)

$$\begin{cases} |\partial_t \Psi(x, t)|^{p-2} \partial_t \Psi(x, t) = -\nabla u(\Psi(x, t)), & t > 0, \\ \Psi(x, 0) = x. \end{cases}$$

Thus,  $\Psi$  is a type of measurable flow map associated with the ODE  $|\dot{\gamma}|^{p-2} \dot{\gamma} = -\nabla u(\gamma)$ .

**Proof of Proposition 1.3.** 1. Let  $\sigma$  be an admissible path for  $\mathcal{U}(\mu)$  and employ Theorem A.3 to write

$$\sigma(t) = e(t)_{\#} \eta, \quad t \geq 0.$$

Note that

$$\begin{aligned} \int_0^\infty e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}(t)\|^p - \mathcal{V}(\sigma(t)) \right) dt &= \int_0^\infty e^{-\delta t} \left( \frac{1}{p} \int_{\mathbb{R}^d} |v(x, t)|^p d\sigma_t(x) - \int_{\mathbb{R}^d} V(x) d\sigma_t(x) \right) dt \\ &= \int_0^\infty e^{-\delta t} \int_{\mathbb{R}^d} \left\{ \frac{1}{p} |v(x, t)|^p - V(x) \right\} d\sigma_t(x) dt \\ &= \int_0^\infty e^{-\delta t} \int_{\mathbb{R}^d \times \Gamma} \left\{ \frac{1}{p} |v(\gamma(t), t)|^p - V(\gamma(t)) \right\} d\eta(x, \gamma) dt \\ &= \int_{\mathbb{R}^d \times \Gamma} \int_0^\infty e^{-\delta t} \left\{ \frac{1}{p} |v(\gamma(t), t)|^p - V(\gamma(t)) \right\} dt d\eta(x, \gamma) \\ &= \int_{\mathbb{R}^d \times \Gamma} \int_0^\infty e^{-\delta t} \left\{ \frac{1}{p} |\dot{\gamma}(t)|^p - V(\gamma(t)) \right\} dt d\eta(x, \gamma) \\ &\geq \int_{\mathbb{R}^d \times \Gamma} u(x) d\eta(x, \gamma) \\ &= \int_{\mathbb{R}^d} u(x) d\mu(x). \end{aligned}$$

The interchange of integrals follows from Remark A.4, assumption (4.1) along with (2.1), and a routine application of Fubini’s theorem. We leave the details to the reader.

2. Now define the path  $\sigma(t) := \Psi(\cdot, t)_{\#} \mu$  for  $t \geq 0$ , where  $\Psi$  is defined in (4.3). Since  $t \mapsto \Psi(x, t)$  is a minimizer for  $u(x)$ , (2.3) implies

$$\int_{\mathbb{R}^d} \int_0^\infty e^{-\delta t} |\partial_t \Psi(t, x)|^p dt d\mu(x) < \infty$$

for  $\mu \in \mathcal{M}_p$ ; this bound follows closely to the proof of Lemma 2.3. Also notice for  $0 \leq s < t < \infty$ ,

$$\begin{aligned} W_p(\sigma(t), \sigma(s))^p &\leq \int_{\mathbb{R}^d} |\Psi(t, x) - \Psi(s, x)|^p d\mu(x) \\ &\leq \int_{\mathbb{R}^d} \left( \int_s^t |\partial_\tau \Psi(\tau, x)| d\tau \right)^p d\mu(x) \\ &\leq (t - s)^{p-1} \int_{\mathbb{R}^d} \int_s^t |\partial_\tau \Psi(\tau, x)|^p d\tau d\mu(x) \end{aligned}$$

which implies

$$\left( \frac{W_p(\sigma(t), \sigma(s))}{t - s} \right)^p \leq \frac{1}{t - s} \int_s^t \left( \int_{\mathbb{R}^d} |\partial_\tau \Psi(\tau, x)|^p d\mu(x) \right) d\tau.$$

As a result,

$$\|\dot{\sigma}(t)\|^p \leq \int_{\mathbb{R}^d} |\partial_t \Psi(t, x)|^p d\mu(x)$$

for a.e.  $t \geq 0$ . Therefore,

$$\begin{aligned} \mathcal{U}(\mu) &\leq \int_0^\infty e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}(t)\|^p - \mathcal{V}(\sigma(t)) \right) dt \\ &\leq \int_0^\infty e^{-\delta t} \left( \int_{\mathbb{R}^d} \left( \frac{1}{p} |\partial_t \Psi(t, x)|^p - V(\Psi(t, x)) \right) d\mu(x) \right) dt \\ &\leq \int_{\mathbb{R}^d} \left( \int_0^\infty e^{-\delta t} \left( \frac{1}{p} |\partial_t \Psi(t, x)|^p - V(\Psi(t, x)) \right) dt \right) d\mu(x) \\ &= \int_{\mathbb{R}^d} u(x) d\mu(x). \end{aligned}$$

In particular,  $\sigma$  is optimal for  $\mathcal{U}(\mu)$ .  $\square$

**Corollary 4.2.** *Let  $\sigma$  be a minimizing path for  $\mathcal{U}(\mu)$  with minimal velocity  $v$ . If  $u$  is differentiable along each of its minimizing trajectories, then*

$$|v(x, t)|^{p-2}v(x, t) = -\nabla u(x), \quad \sigma(t) \text{ a.e. } x \in \mathbb{R}^d, \tag{4.4}$$

for Lebesgue almost every  $t > 0$ .

**Proof.** Let  $\sigma$  be a minimizing path and use the probabilistic representation to write  $\sigma(t) = e(t)\#\eta$ . From part 1 of the proof of Proposition 1.3,  $\eta$  is concentrated on  $(x, \gamma)$  where  $\gamma$  is a minimizer for  $u(x)$ . By (1.4), we conclude for every  $t > 0$  and  $\eta$  almost every  $(x, \gamma)$

$$|\dot{\gamma}(t)|^{p-2}\dot{\gamma}(t) = -\nabla u(\gamma(t)).$$

From the canonical uniqueness of minimal velocities (Proposition 8.4.5 of [2]), we have for Lebesgue almost every  $t > 0$  and  $\eta$  almost every  $(x, \gamma)$

$$|v(\gamma(t), t)|^{p-2}v(\gamma(t), t) = |\dot{\gamma}(t)|^{p-2}\dot{\gamma}(t).$$

In particular, we have for Lebesgue almost every  $t > 0$

$$|v(e_t(x, \gamma), t)|^{p-2}v(e_t(x, \gamma), t) = -\nabla u(e_t(x, \gamma))$$

for  $\eta$  almost every  $(x, \gamma)$ . Recalling  $\sigma(t) = e(t)\#\eta$ , we conclude (4.4).  $\square$

**Example 4.3.** Assume  $V(x) = w \cdot x + c$ . Here  $w \in \mathbb{R}^d$  and  $c \in \mathbb{R}$  are fixed. The associated classical value function is

$$u(x) = -\frac{1}{\delta} \left( \frac{|w|^q}{q} + w \cdot x + c \right), \quad x \in \mathbb{R}^d,$$

with flow map  $\Psi(x, t) = x + t|\frac{w}{\delta}|^{q-2}\frac{w}{\delta}$ . By Proposition 1.3

$$\mathcal{U}(\mu) = -\frac{1}{\delta} \left( \frac{|w|^q}{q} + w \cdot \int_{\mathbb{R}^d} x d\mu(x) + c \right).$$

**Example 4.4.** Although Proposition 1.3 requires that  $V$  grow no more than  $|x|^r$  ( $1 \leq r < p$ ) for  $|x|$  large, the assertion is still valid in the case  $V(x) = -\frac{|x|^p}{p}$ . This follows because the associated action

$$u(x) = |a|^{q-2}a \frac{|x|^p}{p}$$

is  $C^1$  with smooth flow map  $\Psi(x, t) = xe^{-(|a|^{q-2}a)t}$ ; here  $a$  is the unique positive solution of the equation

$$\delta a + (p - 1)|a|^q - 1 = 0.$$

It can be checked in this case that the associated generalized value function is given by

$$\mathcal{U}(\mu) = \frac{|a|^{q-2}a}{p} \int_{\mathbb{R}^d} |x|^p d\mu(x).$$

### 5. Hamilton–Jacobi equations

This section is dedicated to the proof of the main result of this paper, [Theorem 1.2](#), and to the proof of a special case of [Conjecture 3.3](#). Our proof of [Theorem 1.2](#) requires us to define solutions of the abstract HJE [\(1.11\)](#)

$$\delta\mathcal{U} + \frac{1}{q} \|\nabla\mathcal{U}\|_{L^q(\mu)}^q + \mathcal{V}(\mu) = 0, \quad \mu \in \mathcal{M}_p.$$

Naturally, this will involve the *tangent space*

$$\text{Tan}_\mu \mathcal{M}_p := \overline{\{|\nabla\psi|^{q-2}\nabla\psi: \psi \in C_c^\infty(\mathbb{R}^d)\}}^{L^p(\mu)}$$

and the *cotangent space*

$$\text{CoTan}_\mu \mathcal{M}_p := \overline{\{\nabla\psi: \psi \in C_c^\infty(\mathbb{R}^d)\}}^{L^q(\mu)}$$

of  $\mathcal{M}_p$  at a given measure  $\mu$ . We will also make use of the following characterization

$$\text{Tan}_\mu \mathcal{M}_p = \overline{\{\lambda(r - \text{id}_{\mathbb{R}^d}): \lambda > 0, (\text{id}_{\mathbb{R}^d} \times r)_\# \mu \in \Gamma_0(\mu, r_\# \mu)\}}^{L^p(\mu)} \tag{5.1}$$

which is proved in Theorem 8.5.1 in [\[2\]](#); in [\(5.1\)](#)  $r: \mathbb{R}^d \rightarrow \mathbb{R}^d$  is a Borel measurable mapping. In the spirit of our previous work, we present a notion of sub- and super differential of functionals on  $\mathcal{M}_p$ . This notion is inspired by Definition 10.1.1 of [\[2\]](#) and Definition 3.1 of [\[11\]](#).

**Definition 5.1.**  $\xi \in \text{CoTan}_{\mu_0} \mathcal{M}_p$  belongs to the *superdifferential* of  $\mathcal{U}$  at  $\mu_0$  if

$$\mathcal{U}(\mu) \leq \mathcal{U}(\mu_0) + \inf_{\pi \in \Gamma_0(\mu_0, \mu)} \int \xi(x) \cdot (y - x) d\pi(x, y) + o(W_p(\mu_0, \mu))$$

as  $\mu \rightarrow \mu_0$ . Here  $\Gamma_0(\mu_0, \mu)$  is the set of optimal measures for  $W_p(\mu_0, \mu)$ . In this case, we write  $\xi \in \nabla^+ \mathcal{U}(\mu_0)$ . Likewise,  $\xi \in \text{CoTan}_{\mu_0} \mathcal{M}_p$  belongs to the *subdifferential* of  $\mathcal{U}$  at  $\mu_0$  if

$$\mathcal{U}(\mu) \geq \mathcal{U}(\mu_0) + \sup_{\pi \in \Gamma_0(\mu_0, \mu)} \int \xi(x) \cdot (y - x) d\pi(x, y) + o(W_p(\mu_0, \mu))$$

as  $\mu \rightarrow \mu_0$ . In this case, we write  $\xi \in \nabla^- \mathcal{U}(\mu_0)$ .

**Remark 5.2.** It is easy to verify that if both  $\nabla^+\mathcal{U}(\mu)$  and  $\nabla^-\mathcal{U}(\mu)$  are nonempty, then they must contain a common, single element which we denote  $\nabla\mathcal{U}(\mu)$  and call the derivative of  $\mathcal{U}$  at  $\mu$ . See Remark 3.2 of [11] for more on this comment. We are now ready to define an appropriate type of solution to the HJE (1.11). The following definition originates in the work of Gangbo, Nguyen, and Tudorascu [11].

**Definition 5.3.**  $\mathcal{U} \in USC(\mathcal{M}_p)$  is a viscosity subsolution of (1.11) if for all  $\xi \in \nabla^+\mathcal{U}(\mu_0)$ ,

$$\delta\mathcal{U}(\mu_0) + \frac{1}{q} \|\xi\|_{L^q(\mu_0)}^q + \mathcal{V}(\mu_0) \leq 0. \tag{5.2}$$

$\mathcal{U} \in LSC(\mathcal{M}_p)$  is a viscosity supersolution of (1.11) if for all  $\xi \in \nabla^-\mathcal{U}(\mu_0)$ ,

$$\delta\mathcal{U}(\mu_0) + \frac{1}{q} \|\xi\|_{L^q(\mu_0)}^q + \mathcal{V}(\mu_0) \geq 0. \tag{5.3}$$

Finally,  $\mathcal{U} \in C(\mathcal{M}_p)$  is a viscosity solution if it is both a sub- and supersolution.

**Proof of Theorem 1.2.** 1. Suppose  $\xi \in \nabla^+\mathcal{U}(\mu_0)$ . For  $\lambda > 0$  and  $r : \mathbb{R}^d \rightarrow \mathbb{R}^d$  such that  $(\text{id}_{\mathbb{R}^d} \times r)_{\#}\mu_0 \in \Gamma_0(\mu_0, r_{\#}\mu_0)$ , set

$$v := \lambda(r - \text{id}_{\mathbb{R}^d}).$$

We know such a  $v$  is  $L^p(\mu_0)$  dense in  $\text{Tan}_{\mu_0} \mathcal{M}_p$  by (5.1).

Define the path  $\sigma(t) := (\text{id}_{\mathbb{R}^d} + tv)_{\#}\mu_0$ , and notice

$$\sigma(t) = ((1 - t\lambda)\text{id}_{\mathbb{R}^d} + t\lambda r)_{\#}\mu_0 = ((1 - t\lambda)\pi^1 + t\lambda\pi^2)_{\#}[(\text{id}_{\mathbb{R}^d} \times r)_{\#}\mu_0].$$

Hence for  $t \in [0, 1/\lambda]$ ,  $\sigma$  is a constant speed geodesic joining  $\mu_0$  to  $r_{\#}\mu_0$  and

$$\|\dot{\sigma}(t)\| = \|v\|_{L^p(\mu_0)}, \quad t \in (0, 1/\lambda).$$

By dynamic programming (Proposition 2.4), for each  $0 < h < 1/\lambda$

$$\begin{aligned} \mathcal{U}(\mu_0) &\leq e^{-\delta h} \mathcal{U}(\sigma(h)) + \int_0^h e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}(t)\|^p - \mathcal{V}(\sigma(t)) \right) dt \\ &= e^{-\delta h} \mathcal{U}((\text{id}_{\mathbb{R}^d} + hv)_{\#}\mu_0) + \left( \frac{e^{-\delta h} - 1}{\delta} \right) \frac{\|v\|_{L^p(\mu_0)}^p}{p} - \int_0^h e^{-\delta t} \mathcal{V}(\sigma(t)) dt. \end{aligned}$$

Moreover, as

$$\pi(t) = (\text{id}_{\mathbb{R}^d} \times (\text{id}_{\mathbb{R}^d} + tv))_{\#}\mu_0 \in \Gamma_0(\mu_0, \sigma(t))$$



for all sufficiently small  $t > 0$ ,

$$\mathcal{U}(\mu_0) \leq e^{-\delta h} \left\{ \mathcal{U}(\mu_0) + h \int_{\mathbb{R}^d} v \cdot \xi d\mu_0 + o(h) \right\} + \left( \frac{e^{-\delta h} - 1}{\delta} \right) \frac{\|v\|_{L^p(\mu_0)}^p}{p} - \int_0^h e^{-\delta t} \mathcal{V}(\sigma(t)) dt.$$

Hence,

$$\delta \mathcal{U}(\mu_0) - \int_{\mathbb{R}^d} v \cdot \xi d\mu_0 - \frac{\|v\|_{L^p(\mu_0)}^p}{p} + \mathcal{V}(\mu_0) \leq o(1)$$

as  $h \rightarrow 0^+$ . Sending  $h$  to zero and taking the supremum over  $v$  gives the desired inequality (5.2).

2. Now suppose that  $\xi \in \nabla^- \mathcal{U}(\mu_0)$ . Fix  $\eta > 0$  and observe that for each  $h > 0$  there is  $\sigma_h$  admissible for  $\mathcal{U}(\mu_0)$  such that

$$\mathcal{U}(\mu_0) > -\eta h + \int_0^\infty e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}_h(t)\|^p - \mathcal{V}(\sigma_h(t)) \right) dt.$$

A routine computation shows

$$\mathcal{U}(\mu_0) > -\eta h + e^{-\delta h} \mathcal{U}(\sigma_h(h)) + \int_0^h e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}_h(t)\|^p - \mathcal{V}(\sigma_h(t)) \right) dt$$

and employing (1.10) and (2.3), we deduce

$$\int_0^h e^{-\delta t} \|\dot{\sigma}_h(t)\|^p dt \leq C$$

for all  $h > 0$ . Note  $C$  is a universal constant independent of  $h > 0$ . The following uniform estimate

$$W_p(\sigma_h(h), \mu) \leq Ch^{1-1/p} \tag{5.4}$$

is now immediate.

Let us now improve upon the estimate (5.4). By our computations above and the assumption that  $\xi$  belongs to the subdifferential of  $\mathcal{U}$  at  $\mu_0$ , for any  $\pi_h \in \Gamma_0(\mu_0, \sigma_h(h))$

$$\begin{aligned} \mathcal{U}(\mu_0) > -\eta h + e^{-\delta h} \left\{ \mathcal{U}(\mu_0) + \int \xi(x) \cdot (y - x) d\pi_h(x, y) + o(W_p(\mu_0, \sigma_h(h))) \right\} \\ + \int_0^h e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}_h(t)\|^p - \mathcal{V}(\sigma_h(t)) \right) dt. \end{aligned} \tag{5.5}$$

Further observe

$$\int_0^h e^{-\delta t} \frac{1}{p} \|\dot{\sigma}_h(t)\|^p dt \geq e^{-\delta h} \frac{W_p(\sigma_h(h), \mu_0)^p}{ph^{p-1}}$$

and by a version of Young’s inequality

$$\iint \xi(x) \cdot (y - x) d\pi_h(x, y) \geq -2^{q/p} \frac{\|\xi\|_{L^q(\mu_0)}^q}{q} h - \frac{W_p(\sigma_h(h), \mu_0)^p}{2ph^{p-1}}.$$

Combining these bounds with (5.5) gives

$$\begin{aligned} 0 > -\eta + \left(\frac{e^{-\delta h} - 1}{h}\right) \mathcal{U}(\mu_0) - o(1) \left[\frac{W_p(\sigma_h(h), \mu_0)}{h}\right] \\ + \frac{1}{p} \left(e^{-\delta h} - \frac{1}{2}\right) \left[\frac{W_p(\sigma_h(h), \mu_0)}{h}\right]^p - \frac{1}{h} \int_0^h e^{-\delta t} \mathcal{V}(\sigma_h(t)) dt - 2^{q/p} \frac{\|\xi\|_{L^q(\mu_0)}^q}{q} \end{aligned}$$

as  $h \rightarrow 0^+$ . Furthermore, the uniform estimate (5.4) implies  $\frac{1}{h} \int_0^h e^{-\delta t} \mathcal{V}(\sigma_h(t)) dt = \mathcal{V}(\mu_0) + o(1)$  as  $h \rightarrow 0^+$ , so we are now able to conclude

$$W_p(\sigma_h(h), \mu_0) \leq Ch. \tag{5.6}$$

This estimate is valid for all  $h > 0$  small enough for some constant  $C$ ; note the improvement over our previous estimate (5.4).

We return to (5.5), make use of (5.6), and again apply Young’s inequality to find

$$\begin{aligned} \mathcal{U}(\mu_0) > -\eta h + e^{-\delta h} \left[ \mathcal{U}(\mu_0) - \frac{\|\xi\|_{L^q(\mu_0)}^q}{q} h - \frac{W_p(\sigma(h), \mu_0)^p}{ph^{p-1}} + o(h) \right] \\ + \int_0^h e^{-\delta t} \left( \frac{1}{p} \|\dot{\sigma}(t)\|^p - \mathcal{V}(\sigma(t)) \right) dt \\ \geq -\eta h + e^{-\delta h} \mathcal{U}(\mu_0) - e^{-\delta h} \frac{\|\xi\|_{L^q(\mu_0)}^q}{q} h + o(h) - \mathcal{V}(\mu_0) h \end{aligned}$$

or

$$\delta \mathcal{U}(\mu_0) + \frac{\|\xi\|_{L^q(\mu_0)}^q}{q} + \mathcal{V}(\mu_0) > -\eta + o(1).$$

Send  $h$  and then  $\eta \rightarrow 0^+$  verifies the sought after inequality (5.3).  $\square$

**Remark 5.4.** We can extend Definition 5.1 to measures as follows. A measure  $\gamma \in \mathcal{P}_p(\mathbb{R}^d \times \mathbb{R}^d)$  belongs to the extended subdifferential of  $\mathcal{H}$  at  $\mu_0$  provided  $\pi_{\#}^1 \gamma = \mu_0$  and

$$\mathcal{H}(\mu) \leq \mathcal{H}(\mu_0) + \sup_{\rho} \iiint_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} z \cdot (y - x) d\rho(x, y, z) + o(W_p(\mu, \mu_0)), \tag{5.7}$$

as  $W_p(\mu, \mu_0) \rightarrow 0$ . Above, the supremum is taken over  $\rho \in \mathcal{P}_p(\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d)$  satisfying  $\pi_{\#}^{1,3} \rho = \gamma$  and  $\pi_{\#}^{1,2} \rho \in \Gamma_0(\mu_0, \mu)$ . For  $p = 2$ , Ambrosio and Feng used the extended subdifferential to successfully compare the corresponding viscosity sub- and supersolutions of Hamilton–Jacobi equations such as (1.11) [3]. It is an open problem to deduce whether viscosity sub- and supersolutions as defined in this paper admit a comparison principle.

**Remark 5.5.** Hamilton–Jacobi equations on Hilbert spaces have been well studied [4,8]. In particular, the recent work by Gomes and Nurbekyan [14] uncovers various properties of the corresponding value functions. These value functions are related to value functions  $\mathcal{U}$  defined in (1.5) when  $p = 2$  by identifying  $\mathcal{M}_2$  with a subset of  $L^2(\Omega)$ , for an appropriate set  $\Omega$ . This allows us to consider  $U : L^2(\Omega) \rightarrow \mathbb{R}$ , an extension of  $\mathcal{U}$ , which in turn satisfies a Hamilton–Jacobi equation related to Eq. (1.11) in the Hilbert space  $L^2(\Omega)$ . This approach is discussed in detail in Section 6 of the notes by Cardaliaguet [7].

We will conclude this paper with a proof of Conjecture 3.3 in the case of simple potentials (1.12). To this end, we shall need a lemma.

**Lemma 5.6.** Let  $g : \mathbb{R}^d \rightarrow \mathbb{R}$  be a Lipschitz continuous function and set

$$\mathcal{G}(\mu) := \int_{\mathbb{R}^d} g(x) d\mu(x), \quad \mu \in \mathcal{M}_p.$$

If  $g$  is differentiable for  $\mu$  almost every  $x \in \mathbb{R}^d$ , then  $\mathcal{G}$  is differentiable at  $\mu$  (in the sense of Remark 5.2) and

$$\nabla \mathcal{G}(\mu) = \nabla g, \quad \mu \text{ a.e. } x \in \mathbb{R}^d.$$

**Proof.** First we choose a Borel measurable mapping  $\zeta : \mathbb{R}^d \rightarrow \mathbb{R}^d$  that equals  $\nabla g$  for  $\mu$  almost every  $x \in \mathbb{R}^d$ . For instance, we may select  $\zeta := (\zeta^1, \dots, \zeta^n)$  as follows:

$$\zeta^i(x) := \limsup_{n \rightarrow \infty} \frac{g(x + \frac{e_i}{n}) - g(x)}{\frac{1}{n}}, \quad x \in \mathbb{R}^d,$$

for  $i = 1, \dots, n$ . And note that as  $g$  is Lipschitz continuous, this choice gives  $|\zeta(x)| \leq \sqrt{d} \text{Lip}(g)$  for all  $x \in \mathbb{R}^d$ .

Next, define

$$\omega(x, y) := \begin{cases} (g(y) - g(x) - \zeta(x) \cdot (y - x))/|y - x|, & y \neq x, \\ 0, & x = y. \end{cases}$$

Clearly,  $\omega$  is a Borel measurable on  $\mathbb{R}^d \times \mathbb{R}^d$  and  $|\omega(x, y)| \leq (1 + \sqrt{d}) \text{Lip}(g)$  for all  $x, y \in \mathbb{R}^d$ . And, by hypothesis,

$$\limsup_{y \rightarrow x} w(x, y) := \lim_{\epsilon \rightarrow 0^+} \sup_{y \in B_\epsilon(x)} w(x, y) = 0 \tag{5.8}$$

for  $\mu$  almost every  $x \in \mathbb{R}^d$ . For each  $\gamma \in \Gamma_o(\mu, \nu)$

$$\begin{aligned} \int_{\mathbb{R}^d} g(y) d\nu(y) &= \int_{\mathbb{R}^d} g(x) d\mu(x) + \iint_{\mathbb{R}^d \times \mathbb{R}^d} \zeta(x) \cdot (y - x) d\gamma(x, y) \\ &\quad + \iint_{\mathbb{R}^d \times \mathbb{R}^d} |x - y| w(x, y) d\gamma(x, y). \end{aligned}$$

Therefore, it suffices to show

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d} w(x, y) |x - y| d\gamma(x, y) = o(W_p(\mu, \nu))$$

as  $\nu \rightarrow \mu$ .

Assume  $\{\nu_n\}_{n \in \mathbb{N}}$  is a sequence converging  $\mu$  as  $n \rightarrow \infty$ , with  $W_p(\nu_n, \mu) > 0$  for each  $n$ . We know from Remark 7.1.6 of [2] that for any  $\gamma_n \in \Gamma_o(\nu_n, \mu)$ ,

$$\gamma_n \rightarrow (\text{id}_{\mathbb{R}^d} \times \text{id}_{\mathbb{R}^d}) \# \mu \tag{5.9}$$

narrowly in  $\mathcal{P}_p(\mathbb{R}^d \times \mathbb{R}^d)$ . We also have by Hölder’s inequality,

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d} w(x, y) |x - y| d\gamma_n(x, y) \leq W_p(\nu_n, \mu) \left( \iint_{\mathbb{R}^d \times \mathbb{R}^d} |w(x, y)|^q d\gamma_n(x, y) \right)^{1/q}. \tag{5.10}$$

Now let  $\delta > 0$  and choose  $\epsilon > 0$  so that

$$\int_{\mathbb{R}^d} \sup_{y \in B_\epsilon(x)} |w(x, y)|^q d\mu(x) \leq \delta.$$

Such an  $\epsilon > 0$  exists by a simple application of Lebesgue’s dominated convergence theorem. Also notice

$$\begin{aligned} &\iint_{\mathbb{R}^d \times \mathbb{R}^d} |w(x, y)|^q d\gamma_n(x, y) \\ &= \iint_{|x-y| < \epsilon} |w(x, y)|^q d\gamma_n(x, y) + \iint_{|x-y| \geq \epsilon} |w(x, y)|^q d\gamma_n(x, y) \end{aligned}$$

$$\begin{aligned} &\leq \int_{\mathbb{R}^d} \sup_{y \in B_\epsilon(x)} |w(x, y)|^q d\mu(x) + ((1 + \sqrt{d}) \text{Lip}(g))^q \gamma_n(\{(x, y) \in \mathbb{R}^d : |x - y| \geq \epsilon\}) \\ &\leq \delta + ((1 + \sqrt{d}) \text{Lip}(g))^q \gamma_n(\{(x, y) \in \mathbb{R}^d : |x - y| \geq \epsilon\}). \end{aligned}$$

Thus,  $\limsup_{n \rightarrow \infty} \iint_{\mathbb{R}^d \times \mathbb{R}^d} |w(x, y)|^q d\gamma_n(x, y) \leq \delta$  and by (5.9) and (5.10),

$$\limsup_{n \rightarrow \infty} \frac{\iint_{\mathbb{R}^d \times \mathbb{R}^d} w(x, y) |x - y| d\gamma_n(x, y)}{W_p(v_n, \mu)} \leq \delta^{1/q}.$$

This concludes the proof as  $\delta$  and the sequence  $\{v_n\}_{n \in \mathbb{N}}$  were chosen arbitrarily.  $\square$

**Corollary 5.7.** *Assume the hypotheses of Corollary 4.2 and suppose additionally that  $V$  is Lipschitz continuous. Then for any minimizing trajectory  $\sigma$  of the corresponding generalized value function  $\mathcal{U}(\mu)$ , with minimal velocity  $v$ ,*

$$-|v(\cdot, t)|^{p-2} v(\cdot, t) = \nabla \mathcal{U}(\sigma(t))$$

for Lebesgue almost every  $t > 0$ .

**Proof.** By our assumptions, the corresponding classical value function  $u$  (1.1) with potential  $V$  is also Lipschitz continuous. By our assumptions, the proof of Corollary 4.2 yields that  $u$  is differentiable at  $\sigma(t)$  almost every  $x \in \mathbb{R}^d$  with  $\nabla u = -|v(\cdot, t)|^{p-2} v(\cdot, t)$  for Lebesgue almost every  $t > 0$ . The desired conclusion now follows from Lemma 5.6 and Proposition 1.3.  $\square$

### Appendix A. Probabilistic representation

This appendix is dedicated to proving a probabilistic representation of  $AC_{p,\delta}(\mathcal{M}_p)$  paths. The foundation of this assertion is the following compactness result for paths in  $AC_{p,\delta}(\mathbb{R}^d)$ . An immediate consequence of the following lemma is that when  $V$  satisfies the bound (4.1), the classical value function  $u = u(x)$  (1.1) has minimizing paths for every  $x \in \mathbb{R}^d$ .

**Lemma A.1.** *Assume the sequence  $\{\gamma_k\}_{k \in \mathbb{N}} \subset AC_{p,\delta}(\mathbb{R}^d)$  satisfies*

$$\sup_{k \in \mathbb{N}} |\gamma_k(0)| < \infty$$

and

$$\sup_{k \in \mathbb{N}} \int_0^\infty e^{-\delta t} |\dot{\gamma}_k(t)|^p dt < \infty.$$

Then there is a subsequence  $\{\gamma_{k_j}\}_{j \in \mathbb{N}}$  and  $\gamma \in AC_{p,\delta}(\mathbb{R}^d)$  such that  $\gamma_{k_j} \rightarrow \gamma$  locally uniformly on  $[0, \infty)$  and

$$\lim_{j \rightarrow \infty} \int_0^\infty |\gamma_{k_j}(t) - \gamma(t)|^r e^{-\delta t} dt = 0$$

for each  $1 \leq r < p$ .

**Proof.** The proof follows closely with the argument given in the proof of Lemma 3.4, so we omit the details.  $\square$

Recall that for  $\gamma, \xi \in C([0, \infty); \mathbb{R}^d)$ , we define

$$\Pi(\gamma, \xi) = \sum_{k=0}^\infty \frac{1}{2^k} \frac{\max_{0 \leq t \leq k} |\gamma(t) - \xi(t)|}{1 + \max_{0 \leq t \leq k} |\gamma(t) - \xi(t)|}$$

in (4.2). It is immediate that  $\Pi$  is a metric for  $C([0, \infty); \mathbb{R}^d)$  and that  $\lim_{m \rightarrow 0} \Pi(\gamma_m, \gamma) = 0$  if and only if  $\gamma_m \rightarrow \gamma$  locally uniformly on  $[0, \infty)$ . We argue below that  $C([0, \infty); \mathbb{R}^d)$  is in fact a complete, separable metric space under this metric; this claim is of interest because  $C([0, \infty); \mathbb{R}^d)$  is not separable under the maximum norm.

**Proposition A.2.** *The metric space*

$$\Gamma := (C([0, \infty); \mathbb{R}^d), \Pi)$$

*is complete and separable.*

**Proof.** Let  $\{\gamma_m\} \subset \Gamma$  be a Cauchy sequence, assume  $k \in \mathbb{N}$  and fix  $\epsilon \in (0, 1)$ . Choose an integer  $N$  so large that

$$\Pi(\gamma_m, \gamma_n) < \frac{\epsilon}{2^k}, \quad m, n \geq N.$$

Notice that for  $t \in [0, k]$

$$\frac{1}{2^k} \frac{|\gamma_m(t) - \gamma_n(t)|}{1 + |\gamma_m(t) - \gamma_n(t)|} \leq \frac{1}{2^k} \frac{\max_{0 \leq s \leq k} |\gamma_m(s) - \gamma_n(s)|}{1 + \max_{0 \leq s \leq k} |\gamma_m(s) - \gamma_n(s)|} < \frac{\epsilon}{2^k}$$

from which we deduce

$$|\gamma_m(t) - \gamma_n(t)| < \epsilon, \quad t \in [0, k], \quad m, n \geq N.$$

Hence  $\{\gamma_m\}_{m \in \mathbb{N}} \subset C([0, k], \mathbb{R}^d)$  is Cauchy for each  $k$  and so converges locally uniformly to some function  $\gamma \in \Gamma$ . Thus,  $\Gamma$  is complete.

By the Weierstrass Approximation Theorem,  $C([0, k], \mathbb{R}^d)$  equipped with maximum norm is separable for each  $k \in \mathbb{N}$ . Let  $S_k := \{h_1^k, h_2^k, h_3^k, \dots\} \subset C([0, k], \mathbb{R}^d)$  be dense and extend each function in this family continuously on  $[0, \infty)$

$$f_j^k(t) := \begin{cases} h_j^k(t), & 0 \leq t \leq k, \\ h_j^k(k), & k \leq t < \infty, \end{cases}$$

$j \in \mathbb{N}$ . We'll now show that  $\bigcup_{k \in \mathbb{N}} S_k$  is dense in  $\Gamma$ .

Assume  $\gamma \in \Gamma$  and  $\epsilon > 0$ . Choose  $m \in \mathbb{N}$  so large that

$$\frac{1}{2^m} < \frac{\epsilon}{2},$$

and  $f_j^m \in \bigcup_{k \in \mathbb{N}} S_k$  such that

$$\max_{0 \leq t \leq m} |\gamma(t) - f_j^m(t)| < \epsilon/2.$$

By the definition of  $\Pi$ ,

$$\begin{aligned} \Pi(\gamma, f_j^m) &= \sum_{k=0}^m \frac{1}{2^k} \frac{\max_{0 \leq t \leq k} |\gamma(t) - f_j^m(t)|}{1 + \max_{0 \leq t \leq k} |\gamma(t) - f_j^m(t)|} + \sum_{k=m+1}^{\infty} \frac{1}{2^k} \frac{\max_{0 \leq t \leq k} |\gamma(t) - f_j^m(t)|}{1 + \max_{0 \leq t \leq k} |\gamma(t) - f_j^m(t)|} \\ &\leq \max_{0 \leq t \leq m} |\gamma(t) - f_j^m(t)| + \frac{1}{2^m} \\ &< \epsilon. \end{aligned}$$

Hence,  $\Gamma$  is separable.  $\square$

**Theorem A.3.** *Let  $\sigma \in AC_{p,\delta}(\mathcal{M}_p)$  for some  $1 < p < \infty$  and  $\delta > 0$ , and  $v$  be its minimal velocity. Then there is a Borel probability measure  $\eta$  on  $\mathbb{R}^d \times \Gamma$  that is concentrated on  $(x, \gamma)$  such that*

$$\begin{cases} \dot{\gamma}(t) = v(\gamma(t), t), & \text{a.e. } t > 0, \\ \gamma(0) = x \end{cases}$$

and

$$\sigma(t) = e(t) \# \eta, \quad t \geq 0.$$

**Proof.** We adapt the proof of Theorem 8.2.1 of [2]. Recall that  $v$  satisfies (1.7), and so

$$\int_0^T \int_{\mathbb{R}^d} |v(x, t)|^p d\sigma_t(x) dt < \infty$$

for each  $T > 0$ . From this inequality, Propositions 8.17 and 8.18 of [2] imply that  $\sigma^\epsilon(t) := \rho^\epsilon * \sigma(t)$  and  $v^\epsilon(t) := \rho^\epsilon * (v(t)\sigma(t))/\sigma^\epsilon(t)$  satisfy the continuity equation,

$$\int_{\mathbb{R}^d} |v(x, t)|^p d\sigma_t(x) \leq \int_{\mathbb{R}^d} |v^\epsilon(x, t)|^p d\sigma_t^\epsilon(x), \quad t > 0,$$

and that

$$\sigma^\epsilon(t) = \varphi^\epsilon(t) \# \sigma^\epsilon(0), \quad t \geq 0.$$

Here  $\{\rho^\epsilon\}_{\epsilon>0}$  is a family of positive mollifiers and the map  $\varphi^\epsilon : [0, \infty) \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  is the flow of the ODE

$$\begin{cases} \partial_t \varphi^\epsilon(t, x) = v^\epsilon(\varphi^\epsilon(t, x), t), & t > 0, \\ \varphi^\epsilon(0, x) = x. \end{cases}$$

Further denoting  $\varphi^\epsilon : x \mapsto \{\varphi^\epsilon(t, x) \in \mathbb{R}^d : t \geq 0\}$ , we define the measure on  $\mathbb{R}^d \times \Gamma$

$$\eta^\epsilon := (\text{id}_{\mathbb{R}^d} \times \varphi^\epsilon) \# \sigma^\epsilon(0)$$

for  $\epsilon > 0$ . We claim that  $\{\eta^\epsilon\}_{\epsilon>0}$  is tight.

Let  $r^1 : \mathbb{R}^d \times \Gamma \rightarrow \mathbb{R}^d; (x, \gamma) \mapsto x$  and  $r^2 : \mathbb{R}^d \times \Gamma \rightarrow \mathbb{R}^d; (x, \gamma) \mapsto \gamma - x$ . Further define  $r := r^1 \times r^2$  as mapping of  $\mathbb{R}^d \times \Gamma \rightarrow \mathbb{R}^d \times \Gamma$ . Note that  $r \in C(\mathbb{R}^d \times \Gamma)$  with inverse

$$r^{-1}(x, \gamma) = (x, \gamma + x).$$

In particular,  $r$  is a proper mapping. Further observe that  $r^1_\# \eta^\epsilon = \sigma^\epsilon(0) = \rho^\epsilon * \sigma(0)$  converges to  $\sigma(0)$  in  $\mathcal{M}_p$  and therefore is a tight sequence. Also defining

$$\mathcal{E}(\gamma) := \begin{cases} \int_0^\infty e^{-\delta t} |\dot{\gamma}(t)|^p dt, & \gamma \in AC_{p,\delta}(\mathbb{R}^d) \text{ and } \gamma(0) = 0, \\ +\infty, & \text{otherwise,} \end{cases}$$

we have by our above arguments and Tonelli’s theorem

$$\begin{aligned} \int_\Gamma \mathcal{E}(\gamma) d(r^2_\# \eta^\epsilon)(\gamma) &= \int_{\mathbb{R}^d} \mathcal{E}(\varphi^\epsilon(x)) d((\varphi^\epsilon - \text{id}_{\mathbb{R}^d}) \# \sigma^\epsilon(0))(x) \\ &= \int_{\mathbb{R}^d} \int_0^\infty e^{-\delta t} |\partial_t \varphi^\epsilon(t, x)|^p dt d\sigma_0^\epsilon(x) \\ &= \int_{\mathbb{R}^d} \int_0^\infty e^{-\delta t} |v^\epsilon(\varphi^\epsilon(t, x), t)|^p dt d\sigma_0^\epsilon(x) \\ &= \int_0^\infty \left\{ \int_{\mathbb{R}^d} e^{-\delta t} |v^\epsilon(\varphi^\epsilon(t, x), t)|^p d\sigma_0^\epsilon(x) \right\} dt \\ &= \int_0^\infty \left\{ \int_{\mathbb{R}^d} e^{-\delta t} |v^\epsilon(x, t)|^p d\sigma_t^\epsilon(x) \right\} dt \end{aligned}$$



$$\leq \int_0^\infty \left\{ \int_{\mathbb{R}^d} e^{-\delta t} |v(x, t)|^p d\sigma_t(x) \right\} dt < \infty.$$

Since the sublevel sets of  $\mathcal{E}$  are compact in  $\Gamma$  by Lemma A.1,  $\{r_\#^2 \eta^\epsilon\}_{\epsilon>0}$  is tight. By Lemma 5.2.2 of [2],  $\{\eta_\epsilon\}_{\epsilon>0}$  is consequently tight, as well.

It follows that there is a sequence of positive numbers  $\epsilon_k$  tending to 0 as  $k \rightarrow \infty$  and a probability measure  $\eta$  on  $\mathbb{R}^d \times \Gamma$  such that  $\eta^{\epsilon_k} \rightarrow \eta$  narrowly in  $\mathbb{R}^d \times \Gamma$  as  $k \rightarrow \infty$ . Moreover, for each  $f \in C_b(\mathbb{R}^d)$  and  $t \geq 0$

$$\begin{aligned} \int_{\mathbb{R}^d} f(x) d\sigma_t(x) &= \lim_{k \rightarrow \infty} \int_{\mathbb{R}^d} f(x) d\sigma_t^{\epsilon_k}(x) \\ &= \lim_{k \rightarrow \infty} \int_{\mathbb{R}^d \times \Gamma} f(\gamma(t)) d\eta^{\epsilon_k}(x, \gamma) \\ &= \int_{\mathbb{R}^d \times \Gamma} f(\gamma(t)) d\eta(x, \gamma). \end{aligned}$$

As a result  $\sigma(t) = e(t)_\# \eta$ , as claimed. By a slight modification of the argument given in Theorem 8.2.1 of [2], we conclude that  $\eta$  is concentrated on  $(x, \gamma)$  for which  $\dot{\gamma}(t) = v(\gamma(t), t)$  for almost every  $t > 0$  and  $\gamma(0) = x$ .  $\square$

**Remark A.4.** For the  $\sigma$  and  $\eta$  as in the above claim, Tonelli’s theorem implies

$$\begin{aligned} \int_{\mathbb{R}^d \times \Gamma} \left\{ \int_0^\infty e^{-\delta t} |\dot{\gamma}(t)|^p dt \right\} d\eta(x, \gamma) &= \int_0^\infty \int_{\mathbb{R}^d \times \Gamma} e^{-\delta t} |\dot{\gamma}(t)|^p d\eta(x, \gamma) dt \\ &= \int_0^\infty \int_{\mathbb{R}^d \times \Gamma} e^{-\delta t} |v(\gamma(t), t)|^p d\eta(x, \gamma) dt \\ &= \int_0^\infty \int_{\mathbb{R}^d} e^{-\delta t} |v(x, t)|^p d\sigma_t(x) dt \\ &= \int_0^\infty e^{-\delta t} \|\dot{\sigma}(t)\|^p dt < \infty. \end{aligned}$$

Consequently, for  $\eta$  almost every  $(x, \gamma) \in \mathbb{R}^d \times \Gamma$ ,  $\gamma \in AC_{p, \delta}(\mathbb{R}^d)$ .

## References

- [1] J.P. Aubin, H. Frankowska, *Set-Valued Analysis*, Birkhäuser, Boston, 1990.
- [2] Luigi Ambrosio, Nicola Gigli, Giuseppe Savaré, *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, second edition, *Lectures Math. ETH Zürich*, Birkhäuser Verlag, Basel, 2008.
- [3] L. Ambrosio, J. Feng, On a class of first order Hamilton–Jacobi equations in metric space, *J. Differential Equations* 256 (7) (2014) 2194–2245.
- [4] Viorel Barbu, Giuseppe da Prato, *Hamilton–Jacobi Equations in Hilbert Spaces*, *Res. Notes Math.*, vol. 86, Pitman (Advanced Publishing Program), Boston, MA, 1983.
- [5] Lorenzo Brasco, Curves of minimal action over metric spaces, *Ann. Mat. Pura Appl.* (4) 189 (1) (2010) 95–125.
- [6] Patrick Billingsley, *Convergence of Probability Measures*, second edition, *Wiley Ser. Probab. Stat.*, A Wiley Interscience Publication, John Wiley & Sons, Inc., New York, 1999.
- [7] Pierre Cardaliaguet, *Notes on mean field games*, unpublished notes.
- [8] Michael G. Crandall, Pierre-Louis Lions, Hamilton–Jacobi equations in infinite dimensions. III, *J. Funct. Anal.* 68 (2) (1986) 214–247.
- [9] Jin Feng, Truyen Nguyen, Hamilton–Jacobi equations in space of measures associated with a system of conservation laws, *J. Math. Pures Appl.* (9) 97 (4) (2012) 318–390.
- [10] Wendell H. Fleming, H. Mete Soner, *Controlled Markov Processes and Viscosity Solutions*, second edition, *Stoch. Model. Appl. Probab.*, vol. 25, Springer, New York, 2006.
- [11] Wilfrid Gangbo, Truyen Nguyen, Adrian Tudorascu, Hamilton–Jacobi equations in the Wasserstein space, *Methods Appl. Anal.* 15 (2) (2008) 155–183.
- [12] W. Gangbo, T. Nguyen, A. Tudorascu, Euler–Poisson systems as action-minimizing paths in the Wasserstein space, *Arch. Ration. Mech. Anal.* 192 (3) (2009) 419–452.
- [13] W. Gangbo, A. Świech, Optimal transport and large number of particles, *Discrete Contin. Dyn. Syst.* 34 (4) (2014) 1397–1441.
- [14] D. Gomes, L. Nurbekyan, On the minimizers of calculus of variations problems in Hilbert spaces, *Calc. Var. Partial Differential Equations* (2014), <http://dx.doi.org/10.1007/s00526-013-0705-6>, in press.
- [15] D. Gomes, L. Nurbekyan, Weak KAM theory on the  $d$ -dimensional torus, preprint, 2014.
- [16] N. Gozlan, C. Roberto, P.-M. Samson, Hamilton–Jacobi equations on metric spaces and transport-entropy inequalities, *Rev. Mat. Iberoam.* 30 (1) (2014) 133–163.
- [17] Peter J. Huber, Elvezio M. Ronchetti, *Robust Statistics*, second edition, *Wiley Ser. Probab. Stat.*, John Wiley & Sons, Inc., Hoboken, NJ, 2009.
- [18] R. Hynd, H. Kim, Value functions in the Wasserstein spaces: finite time horizons, preprint, 2013.
- [19] Adam M. Oberman, Ryo Takei, Alexander Vladimirovsky, Homogenization of metric Hamilton–Jacobi equations, *Multiscale Model. Simul.* 8 (1) (2009) 269–295.
- [20] Cedric Villani, *Optimal Transport. Old and New*, *Grundlehren Math. Wiss. (Fundamental Principles of Mathematical Sciences)*, vol. 338, Springer-Verlag, Berlin, 2009.