A VARIATIONAL METHOD FOR A CLASS OF PARABOLIC PDES

ALESSIO FIGALLI, WILFRID GANGBO, AND TÜRKAY YOLCU

Abstract. In this manuscript we extend De Giorgi's interpolation method to a class of parabolic equations which are not gradient flows but possess an entropy functional and an underlying Lagrangian. The new fact in the study is that not only the Lagrangian may depend on spatial variables, but it does not induce a metric. Assuming the initial condition to be a density function, not necessarily smooth, but solely of bounded first moments and finite "entropy", we use a variational scheme to discretize the equation in time and construct approximate solutions. Then De Giorgi's interpolation method is revealed to be a powerful tool for proving convergence of our algorithm. Finally we show uniqueness and stability in L^1 of our solutions.

1. Introduction

In the theory of existence of solutions of ordinary differential equations on a metric space, curves of maximal slope and minimizing movements play an important role. The minimizing movements in general are obtained via a discrete scheme. They have the advantage of providing an approximate solution of the differential equation by discretizing in time while not requiring the initial condition to be smooth. Then a clever interpolation method introduced by De Giorgi [7, 6] ensures compactness for the family of approximate solutions. Many recent works [3, 14] have used minimizing movement methods as a powerful tool for proving existence of solutions for some classes of partial differential equations (PDEs). So far, most of these studies concern PDEs which can be interpreted as gradient flow of an entropy functional with respect to a metric on the space of probability measures. This paper extends the minimizing movements and De Giorgi's interpolation method to include PDEs which are not gradient flows, but possess an entropy functional and an underlying Lagrangian which may be dependent of the spatial variables.

In the current manuscript $X \subset \mathbb{R}^d$ is an open set whose boundary is of zero measure. We denote by $\mathcal{P}_1^{ac}(X)$ the set of Borel probability densities on X of bounded first moments, endowed with the 1-Wasserstein distance W_1 (cfr. subsection 2.2). We consider distributional solutions of a class of PDEs of the form

$$\partial_t \varrho_t + \mathrm{div}(\varrho_t V_t) = 0, \quad \text{in} \quad \mathcal{D}'((0,T)\times\mathbb{R}^d) \tag{1.1}$$

(this implicitly means that we have imposed Neumann boundary condition), with

$$\varrho_t V_t := \varrho_t \nabla_p H\big(x, -\varrho_t^{-1}\nabla[P(\varrho_t)]\big) \quad \text{on} \quad (0,T)\times X$$

and

$$t \mapsto \varrho_t \in AC_1(0,T;\mathcal{P}_1^{ac}(X)) \subset C([0,T];\mathcal{P}_1^{ac}(X)).$$

By abuse of notation, ϱ_t will denote at the same time the solution at time t and the function $(t,x) \mapsto \varrho_t(x)$ defined over $(0,T)\times X$. (It will be clear from the context which one we are referring

Date: January 28, 2011.

Key words: mass transfer, Quasilinear Parabolic–Elliptic Equations, Wasserstein metric. AMS code: 35, 49J40, 82C40 and 47J25.

to.) We recall that the unknown ϱ_t is nonnegative, and can be interpreted as the density of a fluid, whose pressure is $P(\varrho_t)$. Here, the data H, U and P satisfy specific properties, which are stated in subsection 2.1.

We only consider solutions such that $\nabla[P(\varrho_t)] \in L^1((0,T)\times X)$, and is absolutely continuous with respect to ϱ_t. If ϱ_t satisfies additional conditions which will soon comment on, then $t \mapsto \mathcal{U}(\varrho_t) := \int_X U(\varrho_t)\,dx$ is absolutely continuous, monotone nonincreasing, and

$$\frac{d}{dt}\mathcal{U}(\varrho_t) = \int_X \langle \nabla[P(\varrho_t)], V_t\rangle\,dx. \tag{1.2}$$

The space to which the curve $t \mapsto \varrho_t$ belongs ensures that ϱ_t converges to ϱ_0 in $\mathcal{P}_1^{ac}(X)$ as $t \to 0$.

Solutions of our equation can be viewed as curves of maximal slope on a metric space contained in $\mathcal{P}_1(X)$. They include the so-called minimizing movements (cfr. [3] for a precise definition) obtained by many authors in case the Lagrangian does not depend on spatial variables (e.g. [13] when $H(p) = 1/2|p|^2$, [1, 3] when $H(x,p) \equiv H(p)$). These studies have been very recently extended to a special class of Lagrangian depending on spatial variables where the Hamiltonian assume the form $H(x,p) = \langle A^*(x)p, p\rangle$ [14]. In their pioneering work Alt and Luckhaus [2] consider differential equations similar to (1.1), imposing some assumptions not very comparable to ours. Their method of proof is very different from the ones used in the above cited references and is based on a Galerkin type approximation method.

Let us describe the strategy of the proof of our results. The first step is the existence part. Let $L(x,\cdot)$ be the Legendre transform of $H(x,\cdot)$, to which we refer as a Lagrangian. For a time step $h > 0$, let $c_h(x,y)$, the cost for moving a unit mass from a point x to a point y, be the minimal action $\min_\sigma \int_0^h L(\sigma,\dot\sigma)dt$. Here, the minimum is performed over the set of all paths (not necessarily contained in X) such that $\sigma(0) = x$ and $\sigma(h) = y$. The cost c_h provides a way of defining the minimal total work $\mathcal{C}_h(\varrho_0,\varrho)$ (cfr. (2.8)) for moving a mass of distribution ϱ_0 to another mass of distribution ϱ in time h. For measures which are absolutely continuous, the recent papers [4, 8, 9] give uniqueness of a minimizer in (2.8), which is concentrated on the graph of a function $T_h : \mathbb{R}^d \to \mathbb{R}^d$. Furthermore, $\mathcal{C}_h$ provides a natural way of interpolating between these measures: there exists a unique density $\bar\varrho_s$ such that $\mathcal{C}_h(\varrho_0,\varrho_h) = \mathcal{C}_s(\varrho_0,\bar\varrho_s) + \mathcal{C}_{h-s}(\bar\varrho_s,\varrho_h)$ for $s \in (0,h)$.

Assume for a moment that X is bounded. For a given initial condition $\varrho_0 \in \mathcal{P}_1^{ac}(X)$ such that $\mathcal{U}(\varrho_0) < +\infty$ we inductively construct $\{\varrho^h_{nh}\}_n$ in the following way: $\varrho^h_{(n+1)h}$ is the unique minimizer of $\mathcal{C}_h(\varrho^h_{nh},\varrho) + \mathcal{U}(\varrho)$ over $\mathcal{P}_1^{ac}(X)$. We refer to this minimization problem as a primal problem. Under the additional condition that $L(x,v) > L(x,0) \equiv 0$ for all $x, v \in \mathbb{R}^d$ such that $v \neq 0$, one has $c_h(x,x) < c_h(x,y)$ for $x \neq y$. As a consequence, under that condition the following maximum principle holds: if $\varrho_0 \le M$ then $\varrho^h_{nh} \le M$ for all $n \ge 0$.

We then study a problem, dual to the primal one, which provides us with a characterization and some important regularity properties of the minimizer $\varrho^h_{(n+1)h}$. These properties would have been harder to obtain studying only the primal problem. Having determined $\{\varrho^h_{nh}\}_{n\in\mathbb{N}}$, we consider two interpolating paths. The first one is the path $t \mapsto \bar\varrho^h_t$ such that

$$\mathcal{C}_h(\varrho^h_{nh},\varrho^h_{(n+1)h}) = \mathcal{C}_s(\varrho^h_{nh},\bar\varrho^h_{nh+s}) + \mathcal{C}_{h-s}(\bar\varrho^h_{nh+s},\varrho^h_{(n+1)h}), \qquad 0 < s < h.$$

The second path $t \mapsto \varrho_t^h$ is defined by

$$\varrho_{nh+s}^h := \arg\min\Big\{\mathcal{C}_s(\varrho_{nh}^h, \varrho) + \mathcal{U}(\varrho)\Big\}, \qquad 0 < s < h.$$

This interpolation was introduced by De Giorgi in the study of curves of maximal slopes when $\sqrt{\mathcal{C}_s}$ defines a metric. The path $\{\bar{\varrho}_t^h\}$ satisfies equation (3.42), which is a discrete analogue of the differential equation (1.1). Then we write a discrete energy inequality in terms of both paths $\{\bar{\varrho}_t^h\}$ and $\{\varrho_t^h\}$, and we prove that up to a subsequence both paths converge (in a sense to be made precise) to the same path ϱ_t. Furthermore, ϱ_t satisfies the energy inequality

$$\mathcal{U}(\varrho_0) - \mathcal{U}(\varrho_T) \geq \int_0^T dt \int_X \Big[L\big(x, V_t\big) + H\big(x, -\varrho_t^{-1}\nabla[P(\varrho_t)]\big)\Big]\varrho_t \, dx, \tag{1.3}$$

which thanks to the assumptions on H (cfr. subsection 2.1) implies for instance that $\nabla[P(\varrho_t)] \in L^1((0,T)\times X)$. The above inequality corresponds to what can be considered as one half of the chain rule:

$$\frac{d}{dt}\mathcal{U}(\varrho_t) \leq \int_X \langle V_t, \nabla[P(\varrho_t)]\rangle \, dx.$$

Here V_t is a velocity associated to the path $t \mapsto \varrho_t$, in the sense that equation (1.1) holds without yet the knowledge that $\varrho_t V_t = \varrho_t \nabla_p H\big(x, -\varrho_t^{-1}\nabla[P(\varrho_t)]\big)$. The current state of the art allows us to establish the reverse inequality yielding to the whole chain rule only if we know that

$$\int_0^T dt \int_X |V_t|^\alpha \varrho_t \, dx, \quad \int_0^T dt \int_X |\varrho_t^{-1}\nabla[P(\varrho_t)]|^{\alpha'}\varrho_t \, dx < +\infty \tag{1.4}$$

for some $\alpha \in (1, +\infty)$, $\alpha' = \alpha/(\alpha - 1)$. In that case, we can conclude that

$$\varrho_t V_t = \varrho_t \nabla_p H\big(x, -\varrho_t^{-1}\nabla[P(\varrho_t)]\big)$$

and

$$\frac{d}{dt}\mathcal{U}(\varrho_t) = \int_X \langle V_t, \nabla[P(\varrho_t)]\rangle \, dx.$$

In light of the energy inequality (3.43), a sufficient condition to have the inequality (1.4) is that $L(x,v) \sim |v|^\alpha$. This is what we later impose in this work.

Suppose now that X may be unbounded. As pointed out in remark 3.18, by a simple scaling argument we can solve equation (1.1) for general nonnegative densities, not necessarily of unit mass. Lemma 4.1 shows that if we impose the bound (4.1) on the negative part of U, then $\mathcal{U}(\varrho)$ is well-defined for $\varrho \in \mathcal{P}_1^{ac}(X)$. We assume that the initial condition $\varrho_0 \in \mathcal{P}_1^{ac}(X)$ and $\int_X |U(\varrho_0)|dx$ is finite, and we start our approximation argument by replacing X by $X_m := X \cap B_m(0)$ and ϱ_0 by $\varrho_0^m := \varrho_0 \chi_{B_m(0)}$. Here, $B_m(0)$ is the open ball of radius m, centered at the origin. The previous argument provides us with a solution of equation (1.1), starting at ϱ_0^m, for which we show that

$$\max_{t\in[0,T]}\left\{\int_{X_m} |x|\varrho_t^m \, dx + \int_{X_m} |U(\varrho_t^m)| \, dx\right\}$$

is bounded by a constant independent of m. Using the fact that for each m, ϱ^m satisfies the energy inequality (1.3), we obtain that a subsequence of $\{\varrho^m\}$ converges to a solution of equation (1.1) starting at ϱ_0. Moreover, as we will see, our approximation argument also allows to relax the regularity assumptions on the Hamiltonian H. This shows a remarkable feature of the existence scheme described before, as it allows to construct solutions of a highly nonlinear PDE as (1.1) by approximating at the same time the initial datum and the Hamiltonian (and the same strategy could also be applied to relax the assumptions on U, cfr. section 4). This completes the existence part.

In order to prove uniqueness of solution in equation (1.1) we make several additional assumptions on P and H. First of all, we assume that $L(x, v) > L(x, 0)$ for all $x, v \in \mathbb{R}^d$ such that $v \neq 0$ to ensure that the maximum principle holds. Next, let Q denote the inverse of P and set $u(t, \cdot) := P(\varrho_t)$. Then equation (1.1) is equivalent to

$$\partial_t Q(u) = \operatorname{div} \mathbf{a}(x, Q(u), \nabla u) \quad \text{in} \quad \mathcal{D}'((0,T) \times X), \tag{1.5}$$

which is a quasilinear elliptic-parabolic equation. Here $\mathbf{a}$ is given by equation (5.2). The study in [15] addresses contraction properties of solutions of equation (1.5) even when $\partial_t Q(u)$ is not a bounded measure but is merely a distribution, as in our case. Our vector field $\mathbf{a}$ does not necessarily satisfy the assumptions in [15]. (Indeed one can check that it violates drastically the strict monotonicity condition of [15], for large $Q(u)$.) For this reason, we only study uniqueness of solutions with bounded initial conditions even if, for this class of solution, $\mathbf{a}$ is still not strictly monotone in the sense of [2] or [15].

The strategy consists first in showing that there exists a Hamiltonian $\bar{H} \equiv \bar{H}(x, \varrho, m)$ (cfr. equation (5.3)) such that for each x, $-\mathbf{a}(x, \varrho, -m)$ is contained in the subdifferential of $\bar{H}(x, \cdot, \cdot)$ at (ϱ, m). Then, assuming $\bar{H}(x, \cdot, \cdot)$ convex and Q Lipschitz, we establish a contraction property for bounded solutions of (1.1). As a by product we conclude uniqueness of bounded solutions.

The paper is structured as follows: in section 2 we start with some preliminaries and set up the general framework for our study. The proof of the existence of solutions is then split into two cases. Section 3 is concerned with the case where X is bounded, and we prove existence of solutions of equations (1.1) by applying the discrete algorithm described before. In section 4 we relax the assumption that X is bounded: under the hypotheses that $\varrho_0 \in \mathcal{P}_1^{ac}(X)$ and $\int_X |U(\varrho_0)| dx$ is finite, we construct by approximation a solution of equation (1.1) as described above. Section 5 is concerned with uniqueness and stability in L^1 of bounded solutions of equation (1.1) when Q is Lipschitz. To achieve that goal, we impose the stronger condition (5.5) on the Hamiltonian H. We avoid repeating known facts as much as possible, while trying to provide all the necessary details for a complete proof.

2. Preliminaries, Notation and Definitions

2.1. **Main assumptions.** We fix a convex superlinear function $\theta : [0, +\infty) \to [0, +\infty)$ such that $\theta(0) = 0$. The main examples we have in mind are functions θ which are positive combinations of functions like $t \mapsto t^\alpha$ with $\alpha > 1$ (for functions like $t \mapsto t(\ln t)^+$ or e^t, cfr. remark 3.19). We consider a function $L : \mathbb{R}^d \times \mathbb{R}^d \mapsto \mathbb{R}$ which we call *Lagrangian*. We assume that:

- (L1) $L \in C^2(\mathbb{R}^d \times \mathbb{R}^d)$, and $L(x, 0) = 0$ for all $x \in \mathbb{R}^d$.
- (L2) The matrix $\nabla_{vv} L(x, v)$ is strictly positive definite for all $x, v \in \mathbb{R}^d$.
- (L3) There exist constants $A^*, A_*, C^* > 0$ such that

$$C^* \theta(|v|) + A^* \geq L(x, v) \geq \theta(|v|) - A_* \qquad \forall\, x, v \in \mathbb{R}^d.$$

Let us remark that the condition $L(x, 0) = 0$ is not restrictive, as we can always replace L by $L - L(x, 0)$, and this would not affect the study of the problem we are going to consider. We also note that (L1), (L2) and (L3) ensure that L is a so-called *Tonelli Lagrangian* (cfr. for instance

[8, Appendix B]). To prove a maximum principle for the solutions of (1.1), we will also need the assumption:

(L4) $L(x,v) \geq L(x,0)$ for all $x, v \in \mathbb{R}^d$.

The *global Legendre transform* $\mathcal{L} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ of L is defined by

$$\mathcal{L}(x,v) := (x, \nabla_v L(x,v)) .$$

We denote by $\mathbf{\Phi}^L : [0,+\infty) \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ the *Lagrangian flow* defined by

$$\begin{cases} \frac{d}{dt}\big[\nabla_v L\big(\mathbf{\Phi}^L(t,x,v)\big)\big] = \nabla_x L\big(\mathbf{\Phi}^L(t,x,v)\big), \\ \mathbf{\Phi}^L(0,x,v) = (x,v). \end{cases} \tag{2.1}$$

Furthermore, we denote by $\mathbf{\Phi}_1^L : [0,+\infty) \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ the first component of the flow: $\mathbf{\Phi}_1^L := \pi_1 \circ \mathbf{\Phi}^L$, $\pi_1(x,v) := x$.

The Legendre transform of L, called the *Hamiltonian* of L, is defined by

$$H(x,p) := \sup_{v \in \mathbb{R}^d} \big\{ \langle v, p \rangle - L(x,v) \big\}.$$

Moreover we define the Legendre transform of θ as

$$\theta^*(s) := \sup_{t \geq 0} \big\{ st - \theta(t) \big\}, \qquad s \in \mathbb{R}.$$

It is well-known that L satisfies (L1), (L2) and (L3) if and only if H satisfies the following conditions:

(H1) $H \in C^2(\mathbb{R}^d \times \mathbb{R}^d)$, and $H(x,p) \geq 0$ for all $x, p \in \mathbb{R}^d$.

(H2) The matrix $\nabla_{pp} H(x,p)$ is strictly positive definite for all $x, p \in \mathbb{R}^d$.

(H3) $\theta^* : \mathbb{R} \to [0,+\infty)$ is convex, superlinear at $+\infty$, and we have

$$-A^* + C^* \theta^* \left(\frac{|p|}{C^*} \right) \leq H(x,p) \leq \theta^*(|p|) + A_* \qquad \forall\, x, v \in \mathbb{R}^d.$$

Moreover (L4) is equivalent to:

(H4) $\nabla_p H(x,0) = 0$ for all $x \in \mathbb{R}^d$.

We also introduce some weaker conditions on L, which combined with (L3) make it a *weak Tonelli Lagrangian*:

($L1^w$) $L \in C^1(\mathbb{R}^d \times \mathbb{R}^d)$, and $L(x,0) = 0$ for all $x \in \mathbb{R}^d$.

($L2^w$) For each $x \in \mathbb{R}^d$, $L(x,\cdot)$ is strictly convex.

Under ($L1^w$), ($L2^w$) and (L3), the global Legendre transform is an homeomorphism, and the Hamiltonian associated to L satisfies (H3) and

($H1^w$) $H \in C^1(\mathbb{R}^d \times \mathbb{R}^d)$, and $H(x,p) \geq 0$ for all $x, p \in \mathbb{R}^d$.

($H2^w$) For each $x \in \mathbb{R}^d$, $H(x,\cdot)$ is strictly convex.

(Cfr. for instance [8, Appendix B].) In this paper we will mainly work assuming (L1), (L2) and (L3), except in section 4 where we relax the assumptions on L (and correspondingly that on H) to ($L1^w$), ($L2^w$) and (L3).

Let $U : [0,+\infty) \to \mathbb{R}$ be a given function such that

$$U \in C^2((0,+\infty)) \cup C([0,+\infty)), \qquad U'' > 0, \tag{2.2}$$

and

$$U(0)=0, \qquad \lim_{t\to+\infty}\frac{U(t)}{t}=+\infty. \tag{2.3}$$

We set $U(t)=+\infty$ for $t\in(-\infty,0)$, so that U remains convex and lower-semicontinuous on the whole $\mathbb{R}$. We denote by U^* the Legendre transform of U :

$$U^*(s):=\sup_{t\in\mathbb{R}}\big\{st-U(t)\big\}=\sup_{t\geq 0}\big\{st-U(t)\big\}. \tag{2.4}$$

When ϱ is a Borel probability density of $\mathbb{R}^d$ such $U^-(\varrho)\in L^1(\mathbb{R}^d)$ we define the *internal energy*

$$\mathcal{U}(\varrho):=\int_{\mathbb{R}^d}U(\varrho)\,dx.$$

If ϱ represents the *density* of a fluid, one interprets $P(\varrho)$ as a *pressure*, where

$$P(s):=U'(s)s-U(s). \tag{2.5}$$

Note that $P'(s)=sU''(s)$, so that P is increasing on $[0,+\infty)$.

2.2. Notation and definitions.

If ϱ is a probability density and $\alpha>0$, we write

$$M_\alpha(\varrho):=\int_{\mathbb{R}^d}|x|^\alpha\varrho(x)\,dx$$

for its moment of order α. If $X\subset\mathbb{R}^d$ is a Borel set, we denote by $\mathcal{P}^{ac}(X)$ the set of all Borel probability densities on X. If $\varrho\in\mathcal{P}^{ac}(X)$, we tacitly identify it with its extension defined to be 0 outside X. We denote by $\mathcal{P}(X)$ the set of Borel probability measures μ on $\mathbb{R}^d$ that are concentrated on X: $\mu(X)=1$. Finally, we denote by $\mathcal{P}^{ac}_\alpha(X)\subset\mathcal{P}^{ac}(X)$ the set of ϱ probability density on X such that $M_\alpha(\varrho)$ is finite. When $\alpha\geq 1$, this is a metric space when endowed with the Wasserstein distance W_α (cfr. equation (2.10) below). We denote by $\mathcal{L}^d$ the d–dimensional Lebesgue measure.

Let $u,v:X\subset\mathbb{R}^d\to\mathbb{R}\cup\{\pm\infty\}$. We denote by $u\oplus v$ the function $(x,y)\mapsto u(x)+v(y)$ where it is well-defined. The set of points x such that $u(x)\in\mathbb{R}$ is called the domain of u and denoted by dom$u$. We denote by $\partial_-u(x)$ the subdifferential of u at x. Similarly, we denote by $\partial^+u(x)$ the superdifferential of u at x. The set of point where u is differentiable is called the domain of ∇u and is denoted by dom∇u.

Let $u:\mathbb{R}^d\to\mathbb{R}\cup\{+\infty\}$. Its Legendre transform is $u^*:\mathbb{R}^d\to\mathbb{R}\cup\{+\infty\}$ defined by

$$u^*(y)=\sup_{x\in X}\big\{\langle x,y\rangle-u(x)\big\}.$$

In case $u:X\subset\mathbb{R}^d\to\mathbb{R}\cup\{+\infty\}$, its Legendre transform is defined by identifying u with its extension which takes the value $+\infty$ outside X.

Finally, for $f:(a,b)\to\mathbb{R}$, we set

$$\frac{d^+f}{dt}\Big|_{t=c}:=\limsup_{h\to0^+}\frac{f(c+h)-f(c)}{h}, \qquad \frac{d^-f}{dt}\Big|_{t=c}:=\liminf_{h\to0^-}\frac{f(c+h)-f(c)}{h}.$$

Definition 2.1 (c-transform)**.** Let $c : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$, let $X \subset \mathbb{R}^d$ and let $u, v : X \to \mathbb{R} \cup \{-\infty\}$. The *first c-transform* of u, $u^c : X \to \mathbb{R} \cup \{-\infty\}$, and the *second c-transform* of v, $v_c : X \to \mathbb{R} \cup \{-\infty\}$, are respectively defined by

$$u^c(y) := \inf_{x \in X} \{c(x,y) - u(x)\}, \qquad v_c(x) := \inf_{y \in X} \{c(x,y) - v(y)\}. \tag{2.6}$$

Definition 2.2 (c-concavity)**.** We say that $u : X \to \mathbb{R} \cup \{-\infty\}$ is *first c-concave* if there exists $v : X \to \mathbb{R} \cup \{-\infty\}$ such that $u = v_c$. Similarly, $v : X \to \mathbb{R} \cup \{-\infty\}$ is *second c-concave* if there exists $u : X \to \mathbb{R} \cup \{-\infty\}$ such that $v = u^c$.

For simplicity we will omit the words "first" and "second" when referring to c-transform and c-concavity.

For $h > 0$, we define the *action* $\mathcal{A}_h(\sigma)$ of an absolutely continuous curve $\sigma : [0,h] \to \mathbb{R}^d$ as

$$\mathcal{A}_h(\sigma) := \int_0^h L(\sigma(\tau), \dot{\sigma}(\tau))\, d\tau$$

and the *cost function*

$$c_h(x,y) := \inf_\sigma \Big\{ \mathcal{A}_h(\sigma) \,:\, \sigma \in W^{1,1}(0,h;\mathbb{R}^d), \sigma(0) = x, \sigma(h) = y \Big\}. \tag{2.7}$$

For $\mu_0, \mu_1 \in \mathcal{P}(\mathbb{R}^d)$, let $\Gamma(\mu_0, \mu_1)$ be the set of probability measures on $\mathbb{R}^d \times \mathbb{R}^d$ which have μ_0 and μ_1 as marginals. Set

$$\mathcal{C}_h(\mu_0, \mu_1) := \inf_\gamma \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} c_h(x,y) d\gamma(x,y) \,:\, \gamma \in \Gamma(\mu_0, \mu_1) \right\} \tag{2.8}$$

and

$$W_{\theta,h}(\mu_0, \mu_1) := h \inf_\gamma \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} \theta\left(\frac{|y-x|}{h}\right) d\gamma(x,y) \,:\, \gamma \in \Gamma(\mu_0, \mu_1) \right\}. \tag{2.9}$$

Remark 2.3. By remark 2.11 c_h is continuous. In particular, there always exists a minimizer for (2.8) (trivial if $\mathcal{C}_h$ is identically $+\infty$ on $\Gamma(\varrho_0, \varrho_1)$). We denote the set of minimizers by $\Gamma_h(\varrho_0, \varrho_1)$. Similarly, there is a minimizer for (2.9), and we denote the set of its minimizers by $\Gamma_h^\theta(\varrho_0, \varrho_1)$.

We also recall the definition of the α-Wasserstein distance, $\alpha \geq 1$:

$$W_\alpha(\mu_0, \mu_1) := \inf_\gamma \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} |y-x|^\alpha \, d\gamma(x,y) \,:\, \gamma \in \Gamma(\mu_0, \mu_1) \right\}^{1/\alpha}. \tag{2.10}$$

It is well-known (cfr. for instance [3]) that W_α metrizes the weak* topology of measures on bounded subsets of $\mathbb{R}^d$. Although we define W_α here for all $\alpha \geq 1$, only W_1 will be used except after section 3.5.

The following fact can be checked easily:

$$\mathcal{C}_h(\mu_0, \mu_2) \leq \mathcal{C}_{h-t}(\mu_0, \mu_1) + \mathcal{C}_t(\mu_1, \mu_2) \tag{2.11}$$

for all $t \in [0,h]$ and $\mu_0, \mu_1, \mu_2 \in \mathcal{P}(\mathbb{R}^d)$.

2.3. **Properties of enthalpy and pressure functionals.** In this subsection, we assume that (2.2) and (2.3) hold.

Lemma 2.4. *The following properties hold:*

(i) $U' : [0, +\infty) \to \mathbb{R}$ *is strictly increasing, and so invertible. Its inverse is of class* C^1 *and* $\lim_{t\to+\infty} U'(t) = +\infty$.

(ii) $U^* \in C^1(\mathbb{R})$ *is nonnegative, and* $(U^*)'(s) \geq 0$ *for all* $s \in \mathbb{R}$.

(iii) $\lim_{s\to+\infty}(U^*)'(s) = +\infty$.

(iv) $\lim_{s\to+\infty} \frac{U^*(s)}{s} = +\infty$.

(v) $P : [0, +\infty) \to [0, +\infty)$ *is strictly increasing, bijective,* $\lim_{t\to+\infty} P(t) = +\infty$*, and its inverse* $Q : [0, +\infty) \to [0, +\infty)$ *satisfies* $\lim_{s\to+\infty} Q(s) = +\infty$.

Proof: (i) Since U is convex and $U(0) = 0$, we have $U'(t) \geq U(t)/t$. This together with $U'' > 0$ and the superlinearity of U easily imply the result.
(ii) $U^* \geq 0$ follows from $U(0) = 0$. The remaining part is a consequence of $(U^*)'(U'(t)) = t$ for $t > 0$, together with $U^*(s) = 0$ (and so $(U^*)'(s) = 0$) for $s \leq U'(0^+)$.
(iii) Follows from (i) and the identity $(U^*)'(U'(t)) = t$ for $t > 0$.
(iv) Since U^* is convex and nonnegative we have $U^*(s) \geq \frac{s}{2}(U^*)'\left(\frac{s}{2}\right)$, so that the result follows from (iii).
(v) Observe that $P(t) = U^*(U'(t)) \geq 0$ by (ii). Since U' is monotone nondecreasing, for $t < 1$ we have $P(t) \leq tU'(1) - U(t)$. We conclude that $\lim_{t\to0^+} P(t) = 0$. The remaining statements follow. □

Remark 2.5. Let $X \subset \mathbb{R}^d$ be a bounded set, and let $\varrho \in \mathcal{P}^{ac}(X)$ be a probability density. Recall that we extend ϱ outside X by setting its value to be identically 0. If $R > 0$ is such that $X \subset B_R(0)$, we have $\int_{\mathbb{R}^d} \theta(|x|)\varrho(x)\,dx \leq \theta(R)$. Moreover, since by convexity $U(t) \geq U(1) + U'(1)(t-1) \equiv at + b$ for $t \geq 0$, $\int_{\mathbb{R}^d} U^-(\varrho)\,dx$ is bounded on $\mathcal{P}^{ac}(X)$ by $|a| + |b|\mathcal{L}^d(X)$. Hence $\mathcal{U}(\varrho)$ is always well-defined on $\mathcal{P}^{ac}(X)$, and is finite if and only if $U^+(\varrho) \in L^1(X)$.

The following lemma is a standard result of the calculus of variations, cfr. for instance [5] (for a more general result on unbounded domains, cfr. section 4):

Lemma 2.6. *Let* $X \subset \mathbb{R}^d$ *and suppose* $\{\varrho^n\}_{n\in\mathbb{N}} \subset \mathcal{P}^{ac}(X)$ *converges weakly to* ϱ *in* $L^1(X)$*. Assume that either* X *is bounded, or* X *is unbounded and* $U \geq 0$*. Then*

$$\liminf_{n\to\infty} \mathcal{U}(\varrho^n) \geq \mathcal{U}(\varrho).$$

2.4. **Properties of H and the cost functions.**

Lemma 2.7. *The following properties hold for* $0 < \bar{h} < h$ *and* $x, y \in \mathbb{R}^d$*:*

(i) $c_h(x, x) \leq 0$.

(ii) $c_h(x, y) \leq c_{\bar{h}}(x, y)$.

(iii)
$$C^* h\, \theta\left(\frac{|x-y|}{h}\right) + A^* h \geq c_h(x, y) \geq h\, \theta\left(\frac{|x-y|}{h}\right) - A_* h \geq -A_* h.$$

Proof: (i) Set $\sigma(t) \equiv x$ for $t \in [0, h]$ and recall that $L(x, 0) = 0$ to get $c_h(x, x) \leq \mathcal{A}_h(\sigma) = 0$.
(ii) Given $\sigma \in W^{1,1}(0, \bar{h}; \mathbb{R}^d)$ satisfying $\sigma(0) = x$ and $\sigma(\bar{h}) = y$, we can associate an extension to

$(\bar{h}, h]$, which we still denote σ, such that $\sigma(t) = y$ for $t \in (\bar{h}, h]$. We have $\sigma \in W^{1,1}(0, h; \mathbb{R}^d)$, $\sigma(0) = x$ and $\sigma(\bar{h}) = y$. Hence,

$$c_h(x, y) \leq \mathcal{A}_h(\sigma) = \mathcal{A}_{\bar{h}}(\sigma) + \int_{\bar{h}}^{h} L(y, 0)\, dt = \mathcal{A}_{\bar{h}}(\sigma).$$

Since $\sigma \in W^{1,1}(0, \bar{h}; \mathbb{R}^d)$ is arbitrary, this concludes the proof of (ii).
(iii) The first inequality is obtained using (L3) and $c_h(x, y) \leq \mathcal{A}_T(\sigma)$ with $\sigma(t) = (1-t/h)x+(t/h)y$, while the second one follows from Jensen's inequality. □

The next proposition can readily be derived from the standard theory of Hamiltonian systems (cfr. e.g. [8, Appendix B]):

Proposition 2.8. *Under the assumptions (L1), (L2) and (L3),* (2.7) *admits a minimizer* $\sigma_{x,y}$ *for any* $x, y \in \mathbb{R}^d$. *We have that* $\sigma_{x,y} \in C^2([0, h])$ *and satisfies the Euler-Lagrange equation*

$$(\sigma_{x,y}(\tau), \dot{\sigma}_{x,y}(\tau)) = \mathbf{\Phi}^L(\tau, x, \dot{\sigma}_{x,y}(0)) \qquad \forall \tau \in [0, h], \tag{2.12}$$

where $\mathbf{\Phi}^L$ *is the Lagrangian flow defined in equation* (2.1)*. Moreover, for any* $r > 0$ *and* $S \subset (0, +\infty)$ *a compact set, there exists a constant* $k_S(r)$*, depending on* S *and* r *only, such that* $||\sigma_{x,y}||_{C^2([0,h])} \leq k_S(r)$ *if* $|x|, |y| \leq r$ *and* $h \in S$.

Remark 2.9. Let σ be a minimizer of the problem (2.7), and set

$$p(\tau) := \nabla_v L\big(\sigma(\tau), \dot{\sigma}(\tau)\big).$$

(a) The Euler-Lagrange equation (2.12) implies that σ and p are of class C^1 and satisfy the system of ordinary differential equations

$$\dot{\sigma}(\tau) = \nabla_p H(\sigma(\tau), p(\tau)), \qquad \dot{p}(\tau) = -\nabla_x H(\sigma(\tau), p(\tau)) \tag{2.13}$$

(b) The Hamiltonian is constant along the integral curve $(\sigma(\tau), p(\tau))$, i.e. $H(\sigma(\tau), p(\tau)) = H(\sigma(0), p(0))$ for $\tau \in [0, h]$.

The following lemma is standard (cfr. for instance [8, Appendix B]):

Lemma 2.10. *Under the assumptions in proposition 2.8, let* σ *be a minimizer of (2.7), and define* $p_i := \nabla_v L(\sigma(i), \dot{\sigma}(i))$ *for* $i = 0, h$. *For* $r, m > 0$ *there exists a constant* $l_h(r, m)$*, depending on* h, r, m *only, such that if* $x, y \in B_r(0)$ *and* $w \in B_m(0)$*, then:*

(a) $c_h(x + w, y) \leq c_h(x, y) - \langle p_0, w\rangle + \frac{1}{2}\ell_h(r, m)|w|^2$;
(b) $c_h(x, y + w) \leq c_h(x, y) + \langle p_h, w\rangle + \frac{1}{2}\ell_h(r, m)|w|^2$.

Remark 2.11. This lemma says that $-p_0 \in \partial^+ c_h(\cdot, y)(x)$, and for $y \in B_r(0)$ the restriction of $c(\cdot, y)$ to $B_r(0)$ is $\ell_h(r, m)$-concave. Similarly, $p_h \in \partial^+ c_h(x, \cdot)(y)$, and for $x \in B_r(0)$ the restriction of $c(x, \cdot)$ to $B_r(0)$ is $\ell_h(r, m)$-concave.

Lemma 2.12. *Suppose (L1), (L2) and (L3) hold. Let* $a, b, r \in (0, +\infty)$ *be such that* $a < b$ *and set* $S = [a, b]$. *Then there exists a constant* $\tilde{k}_S(r)$*, depending on* S *and* r *only, such that*

$$|c_h(x, y) - c_{\bar{h}}(x, y)| \leq \tilde{k}_S(r)|h - \bar{h}|$$

for all $h, \bar{h} \in S$ *and all* $x, y \in \mathbb{R}^d$ *satisfying* $|x|, |y| \leq r$.

Proof: Let $k_S(r)$ be the constant appearing in proposition 2.8 and let

$$E_1 := \sup_{x,v}\{|L(x,v)| \ : |x|, |v| \leq k_S(r)\}, \qquad E_2 := \sup_{x,v}\Big\{|\nabla_v L(x,v)| \ : |x| \leq k_S(r), |v| \leq k_S(r)\frac{b}{a}\Big\}.$$

Fix $h, \bar h \in S$ such that $\bar h < h$. For $x, y \in \mathbb{R}^d$ such that $|x|, |y| \leq r$ we denote by σ a minimizer of (2.7). Define $\bar\sigma(t) = \sigma(t\bar h/h)$ for $t \in [0, \bar h]$. Then $\bar\sigma \in C^2([0,\bar h])$, $\bar\sigma(0) = x$ and $\bar\sigma(\bar h) = y$. Then

$$c_{\bar h}(x,y) \leq \int_0^{\bar h} L\big(\bar\sigma, \dot{\bar\sigma}\big)\, dt = \frac{\bar h}{h}\int_0^h L\Big(\sigma, \frac{h}{\bar h}\dot\sigma\Big)\, ds = \frac{\bar h}{h} c_h(x,y) + \frac{\bar h}{h}\int_0^h \Big(L\Big(\sigma, \frac{h}{\bar h}\dot\sigma\Big) - L(\sigma, \dot\sigma)\Big)\, ds.$$

This implies

$$c_{\bar h}(x,y) \leq \frac{\bar h}{h} c_h(x,y) + \frac{\bar h}{h} h E_2\Big(\frac{h}{\bar h} - 1\Big) k_S(r) = \frac{\bar h}{h} c_h(x,y) + (h - \bar h) E_2 k_S(r),$$

and so

$$c_{\bar h}(x,y) - c_h(x,y) \leq \frac{\bar h - h}{h} c_h(x,y) + (h - \bar h) E_2 k_S(r) \leq |h - \bar h|(E_1 + E_2 k_S(r)), \tag{2.14}$$

where we used the trivial bound $c_h(x,y) \leq E_1 h$. Since by lemma 2.7(ii) $c_h(x,y) \leq c_{\bar h}(x,y)$, (2.14) proves the lemma. □

2.5. **Total works and their properties.** In this subsection we assume that (2.2) and (2.3) hold.

Lemma 2.13. *The following properties hold:*

(i) *For any $\mu \in \mathcal{P}(\mathbb{R}^d)$ we have $\mathcal{C}_h(\mu,\mu) \leq 0$. In particular, for any $\mu, \bar\mu \in \mathcal{P}(\mathbb{R}^d)$, $\mathcal{C}_{\bar h}(\mu, \bar\mu) \leq \mathcal{C}_h(\mu, \bar\mu)$ if $h < \bar h$.*

(ii) *For any $h > 0$, $\mu, \bar\mu \in \mathcal{P}(\mathbb{R}^d)$,*

$$-A_* h \leq -A_* h + W_{\theta,h}(\mu,\bar\mu) \leq \mathcal{C}_h(\mu,\bar\mu) \leq C^* W_{\theta,h}(\mu,\bar\mu) + A^* h.$$

(iii) *For any $K > 0$ there exists a constant $C(K) > 0$ such that*

$$W_1(\mu,\bar\mu) \leq \frac{1}{K} W_{\theta,h}(\mu,\bar\mu) + \frac{C(K)}{K} h \qquad \forall\, h > 0,\ \mu, \bar\mu \in \mathcal{P}(\mathbb{R}^d). \tag{2.15}$$

Proof: (i) The first part follows from $c_h(x,x) \leq 0$, while the second statement is a consequence of the first one and $\mathcal{C}_{\bar h}(\mu,\bar\mu) \leq \mathcal{C}_h(\mu,\bar\mu) + \mathcal{C}_{\bar h - h}(\bar\mu,\bar\mu)$.
(ii) It follows directly from Lemma 2.7(iii).
(iii) Thanks to the superlinearity of h, for any $K > 0$ there exists a constant $C(K) > 0$ such that

$$\theta(s) \geq Ks - C(K) \qquad \forall\, s \geq 0. \tag{2.16}$$

Fix now $\gamma \in \Gamma_h^\theta(\mu_0, \mu_1)$. Then

$$\begin{aligned} W_1(\mu,\bar\mu) &\leq \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|\, d\gamma(x,y) \\ &\leq \frac{h}{K}\int_{\mathbb{R}^d\times\mathbb{R}^d} \Big[K\frac{|x-y|}{h} - C(K)\Big]\, d\gamma(x,y) + \frac{C(K)}{K} h \\ &\leq \frac{1}{K}\int_{\mathbb{R}^d\times\mathbb{R}^d} \theta\left(\frac{|x-y|}{h}\right)\, d\gamma(x,y) + \frac{C(K)}{K} h = \frac{1}{K} W_{\theta,h}(\mu,\bar\mu) + \frac{C(K)}{K} h. \end{aligned}$$

□

Lemma 2.14. *Let $h > 0$. Suppose that $\{\varrho^n\}_{n\in\mathbb{N}}$ converges weakly to ϱ in $L^1(\mathbb{R}^d)$ and that $\{M_1(\varrho^n)\}_{n\in\mathbb{N}}$ is bounded. Then $M_1(\varrho)$ is finite, and we have*

$$\liminf_{n\to\infty} \mathcal{C}_h(\bar\varrho, \varrho^n) \geq \mathcal{C}_h(\bar\varrho, \varrho) \qquad \forall\, \bar\varrho \in \mathcal{P}_1^{ac}(X).$$

Proof: The fact that $M_1(\varrho)$ is finite follows from the weak lower-semicontinuity in $L^1(\mathbb{R}^d)$ of M_1. Let now $\gamma^n \in \Gamma_h(\bar\varrho, \varrho^n)$. Since $\{M_1(\varrho^n)\}_{n\in\mathbb{N}}$ is bounded we have

$$\sup_{n\in\mathbb{N}} \int_{\mathbb{R}^d} \Big(|x| + |y|\Big)\gamma^n(dx, dy) < +\infty. \tag{2.17}$$

As $|x| + |y|$ is coercive, equation (2.17) implies that $\{\gamma^n\}_{n\in\mathbb{N}}$ admits a cluster point γ for the topology of the narrow convergence. Furthermore it is easy to see that $\gamma \in \Gamma(\bar\varrho, \varrho)$ and so, since c_h is continuous and bounded below, we get

$$\liminf_{n\to\infty} \mathcal{C}_h(\bar\varrho, \varrho^n) = \liminf_{n\to\infty} \int_{\mathbb{R}^d\times\mathbb{R}^d} c_h(x, y)\, d\gamma^n(x, y) \geq \int_{\mathbb{R}^d\times\mathbb{R}^d} c_h(x, y)\, d\gamma(x, y) \geq \mathcal{C}_h(\bar\varrho, \varrho).$$

□

3. Existence of solutions in a bounded domain

Throughout this section we assume that (2.2) and (2.3) hold. We recall that L satisfies (L1), (L2) and (L3). We also assume that $X \subset \mathbb{R}^d$ is an open bounded set whose boundary ∂X is of zero Lebesgue measure, and we denote by $\overline{X}$ its closure. The goal is to prove existence of distributional solutions to equation (1.1) by using an approximation by discretization in time. More precisely, in subsection 3.1 we construct approximate solutions at discrete times $\{h, 2h, 3h, \ldots\}$ by an implicit Euler scheme, which involves the minimization of a suitable functional. Then in subsection 3.2 we explicitly characterize the minimizer introducing a dual problem. We then study the properties of an augmented action functional which allows to prove a priori bounds on the De Giorgi's variational and geodesic interpolations (cfr. subsection 3.4). Finally, using these bounds we can take the limit as $h \to 0$ and prove existence of distributional solutions to equation (1.1) when θ behaves at infinity like t^α, $\alpha > 1$.

3.1. **The discrete variational problem.**

We fix a time step $h > 0$ and for simplicity of notation we set $c = c_h$. We fix $\varrho_0 \in \mathcal{P}^{ac}(X)$, and we consider the variational problem

$$\inf_{\varrho\in\mathcal{P}^{ac}(X)} \mathcal{C}_h(\varrho_0, \varrho) + \mathcal{U}(\varrho). \tag{3.1}$$

Lemma 3.1. *There exists a unique minimizer ϱ_* of problem* (3.1)*. Suppose in addition that (L4) holds. If $M \in (0, +\infty)$ and $\varrho_0 \leq M$, then $\varrho_* \leq M$. In other words, the maximum principle holds.*

Proof: Existence of a minimizer ϱ_* follows by classical methods in the calculus of variation, thanks to the lower-semicontinuity of the functional $\varrho \mapsto \mathcal{C}_h(\varrho_0, \varrho) + \mathcal{U}(\varrho)$ in the weak topology of measures and to the superlinearity of U (which implies that any limit point of a minimizing sequence still belongs to $\mathcal{P}^{ac}(X)$).

To prove uniqueness, let ϱ_1 and ϱ_2 be two minimizers, and take $\gamma_1 \in \Gamma_h(\varrho_0, \varrho_1)$, $\gamma_2 \in \Gamma_h(\varrho_0, \varrho_2)$ (cfr. remark 2.3). Then $\frac{\gamma_1+\gamma_2}{2} \in \Gamma\left(\varrho_0, \frac{\varrho_1+\varrho_2}{2}\right)$, so that

$$\mathcal{C}_h\left(\varrho_0, \frac{\varrho_1 + \varrho_2}{2}\right) \leq \int_{X\times X} c(x, y)\, d\left(\frac{\gamma_1 + \gamma_2}{2}\right) = \frac{\mathcal{C}_h(\varrho_0, \varrho_1) + \mathcal{C}_h(\varrho_0, \varrho_2)}{2}.$$

Moreover by strict convexity of U

$$\mathcal{U}\left(\frac{\varrho_1+\varrho_2}{2}\right) \leq \frac{\mathcal{U}(\varrho_1)+\mathcal{U}(\varrho_2)}{2}$$

with equality if and only if $\varrho_1 = \varrho_2$. This implies uniqueness.

Thanks to (L1) and (L4) one easily gets that $c_h(x,x) < c_h(x,y)$ for all $x, y \in X$, $x \neq y$. Thanks to this fact the proof of the maximum principle is a folklore which can be found in [18]. $\square$

3.2. **Characterization of minimizers via a dual problem.** The aim of this paragraph is to completely characterize the minimizer ϱ_* provided by lemma 3.1. We are going to identify a problem, dual to problem (3.1), and to use it to achieve that goal.

We define $\mathcal{E} \equiv \mathcal{E}_c$ to be the set of pairs $(u,v) \in C(\overline{X}) \times C(\overline{X})$ such that $u(x)+v(y) \leq c(x,y)$ for all $x, y \in \overline{X}$, and we write $u \oplus v \leq c$. We consider the functional

$$J(u,v) := \int_X u\varrho_0\,dx - \int_X U^*(-v)\,dx.$$

To alleviate the notation, we have omitted to display the ϱ_0 dependence in J.

We recall some well-known results:

Lemma 3.2. *Let $u \in C_b(X)$. Then $(u_c)^c \geq u$, $(u^c)_c \geq u$, $\big((u_c)^c\big)_c = u_c$, and $\big((u^c)_c\big)^c = u^c$. Moreover:*

- (i) *If $u = v^c$ for some $v \in C(\overline{X})$, then:*
  - (a) *There exists a constant $A = A(c,X)$, independent of u, such that u is A-Lipschitz and A-semiconcave.*
  - (b) *If $\bar{x} \in X$ is a point of differentiability of u, $\bar{y} \in \overline{X}$, and $u(\bar{x}) + v(\bar{y}) = c(\bar{x},\bar{y})$, then $\bar{x}$ is a point of differentiability of $c(\cdot,\bar{y})$ and $\nabla u(\bar{x}) = \nabla_x c(\bar{x},\bar{y})$. Furthermore $\bar{y} = \mathbf{\Phi}_1^L\big(h,\bar{x},\nabla_p H\big(\bar{x},-\nabla u(\bar{x})\big)\big)$, and in particular $\bar{y}$ is uniquely determined.*
- (ii) *If $v = u_c$ for some $u \in C(\overline{X})$, then:*
  - (a) *There exists a constant $A = A(c,X)$, independent of v, such that v is A-Lipschitz and A-semiconcave.*
  - (b) *If $\bar{x} \in \overline{X}$, $\bar{y} \in X$ is a point of differentiability of v, and $u(\bar{x}) + v(\bar{y}) = c(\bar{x},\bar{y})$, then $\bar{y}$ is a point of differentiability of $c(\bar{x},\cdot)$ and $\nabla v(\bar{y}) = \nabla_y c(\bar{x},\bar{y})$. Furthermore, $\bar{x} = \mathbf{\Phi}_1^L\big(-h,\bar{y},\nabla_p H\big(\bar{y},\nabla v(\bar{y})\big)\big)$, and in particular $\bar{y}$ is uniquely determined.*

In particular, if $K \subset \mathbb{R}$ is bounded, the set $\{v^c : v \in C(\overline{X}),\ v^c(X) \cap K \neq \emptyset\}$ is compact in $C(\overline{X})$, and weak* compact in $W^{1,\infty}(X)$.

Proof: Despite the fact that the assertions made in the lemma are now part of the folklore of the Monge-Kantorovich theory, we sketch the main steps of the proof.

The first part is classical, and can be found in [12, 16, 17].

Regarding (i)-(a), we observe that by remark 2.11 the functions $c(\cdot,y)$ are uniformly semiconcave for $y \in X$, so that u is semiconcave as the infimum of uniformly semiconcave functions (cfr. for instance [8, Appendix A]). In particular u is Lipschitz, with a Lipschitz constant bounded by $\|\nabla_x c\|_{L^\infty(\overline{X}\times\overline{X})}$.

To prove (i)-(b), we note that $\partial_- u(\bar{x}) \subset \partial_- c(\cdot,\bar{y})(\bar{x})$. Since by remark 2.11 $\partial^+ c(\cdot,\bar{y})(\bar{x})$ is nonempty, we conclude that $c(\cdot,\bar{y})$ is differentiable at $\bar{x}$ if u is. Hence

$$\nabla u(\bar{x}) = \nabla_x c(\bar{x},\bar{y}) = -\nabla_v L(\sigma(0),\dot{\sigma}(0))$$

where $\sigma : [0,h] \to X$ is (the unique curve) such that $c(\bar{x},\bar{y}) = \int_0^h L(\sigma,\dot{\sigma})\,dt$ (cfr. [8, Section 4 and Appendix B]). This together with equation (2.12) implies

$$\bar{y} = \mathbf{\Phi}_1^L\big(h,\bar{x},\nabla_p H\big(\bar{x},-\nabla u(\bar{x})\big)\big). \tag{3.2}$$

The proof of (ii) is analogous. □

Remark 3.3. By lemma 3.2, if $u = v^c$ for some $v \in C_b(\overline{X})$ we can uniquely define $\mathcal{L}^d$-a.e. a map $T : \mathrm{dom}\nabla u \to \overline{X}$ such that $u(x) + v(Tx) = c(x,Tx)$. This map is continuous on $\mathrm{dom}\nabla u$, and since ∇u can be extended to a Borel map on X we conclude that T can be extended to a Borel map on X, too. Moreover we have $\nabla u(x) = \nabla_x c(x,Tx)$ $\mathcal{L}^d$-a.e., and T is the unique optimal map pushing any density $\varrho \in \mathcal{P}^{ac}(X)$ forward to $\bar{\mu} := T_\#(\varrho\mathcal{L}^d) \in \mathcal{P}(X)$ (cfr. for instance [12, 16, 17]).

Lemma 3.4. *If $(u,v) \in \mathcal{E}$ and $\varrho \in \mathcal{P}^{ac}(X)$, then $J(u,v) \le \mathcal{C}_h(\varrho_0,\varrho) + \mathcal{U}(\varrho)$.*

Proof: Let $\gamma \in \Gamma(\varrho_0,\varrho)$. Since $U(\varrho(y)) + U^*(-v(y)) \ge -\varrho(y)v(y)$ and $(u,v) \in \mathcal{E}$, integrating the inequality we get

$$\int_X \big(U(\varrho(y)+U^*(-v(y))\big)\,dy \ge -\int_X \varrho(y)v(y)\,dy \ge -\int_{X\times X} c(x,y)\,d\gamma(x,y) + \int_X \varrho_0(x)u(x)\,dx. \tag{3.3}$$

Rearranging the expressions in equation (3.3) and optimizing over $\Gamma(\varrho_0,\varrho)$ we obtain the result. □

Lemma 3.5. *There exists $(u_*,v_*) \in \mathcal{E}$ maximizing $J(u,v)$ over $\mathcal{E}$ and satisfying $u_*^c = v_*$ and $(v_*)_c = u_*$. Furthermore:*

(i) *u_* and v_* are Lipschitz with a Lipschitz constant bounded by $||\nabla c||_{L^\infty(X\times X)}$.*
(ii) *$\varrho_{v_*} := (U^*)'(-v_*)$ is a probability density on X, and the optimal map T associated to u_* (cfr. remark 3.3) pushes $\varrho_0\mathcal{L}^d$ forward to $\varrho_{v_*}\mathcal{L}^d$.*

Proof: Note that if $u_*^c = v_*$ and $(v_*)_c = u_*$, then (i) is a direct consequence of lemma 3.2.

Before proving the first statement of the lemma, let us show that it implies (ii). Let $\varphi \in C(\overline{X})$ and set

$$v_\varepsilon := v_* + \varepsilon\varphi, \quad u_\varepsilon := (v_\varepsilon)_c.$$

Remark 3.3 says that for $\mathcal{L}^d$-a.e. $x \in X$ the equation $u_*(x) + v_*(y) = c(x,y)$ admits a unique solution Tx. As done in [10] (cfr. also [11]) we have that

$$\|u_\varepsilon - u_*\|_\infty \le \varepsilon\|\varphi\|_\infty, \qquad \lim_{\varepsilon\to 0}\frac{u_\varepsilon(x)-u_*(x)}{\varepsilon} = \varphi(Tx)$$

for $\mathcal{L}^d$-a.e. $x \in X$. Hence by the Lebesgue dominated convergence theorem

$$\lim_{\varepsilon\to 0}\int_X \frac{u_\varepsilon(x)-u_*(x)}{\varepsilon}\varrho_0(x)\,dx = -\int_X \varphi(Tx)\varrho_0(x)\,dx. \tag{3.4}$$

Since (u_*,v_*) maximizes J over $\mathcal{E}$, by equation (3.4) we obtain

$$0 = \lim_{\varepsilon\to 0}\frac{J(u_\varepsilon,v_\varepsilon)-J(u_*,v_*)}{\varepsilon} = -\int_X \varphi(Tx)\varrho_0(x)\,dx + \int_X (U^*)'(-v_*(x))\varphi(x)\,dx.$$

Therefore

$$\int_X \varphi(Tx)\varrho_0(x)\,dx = \int_X (U^*)'(-v_*(x))\varphi(x)\,dx. \tag{3.5}$$

Choosing $\varphi \equiv 1$ in equation (3.5) and recalling that $(U^*)' \geq 0$ (cfr. lemma 2.4(ii)) we discover that $\varrho_{v_*} := (U^*)'(-v_*)$ is a probability density on X. Moreover equation (3.5) means that T pushes $\varrho_0\mathcal{L}^d$ forward to $\varrho_{v_*}\mathcal{L}^d$. This proves (ii).

We eventually proceed with the proof of the first statement. Observe that the functional J is continuous on $\mathcal{E}$, which is a closed subset of $C(\overline{X}) \times C(\overline{X})$. Thus it suffices to show the existence of a compact set $\mathcal{E}' \subset \mathcal{E}$ such that $\mathcal{E}' \subset \{(u,v) : \ u^c = v, v_c = u\}$ and $\sup_{\mathcal{E}} J = \sup_{\mathcal{E}'} J$.

If $(u,v) \in \mathcal{E}$ then $u \leq v_c$, and so $J(u,v) \leq J(v_c, v)$. But as pointed out in lemma 3.2 $v \leq (v_c)^c$, and since by lemma 2.4(ii) $U^* \in C^1(\mathbb{R})$ is monotone nondecreasing we have $J(u,v) \leq J(v_c,v) \leq J(v_c,(v_c)^c)$. Set $\bar u = v_c$ and $\bar v = (v_c)^c$. Observe that by lemma 3.2 $\bar u = \bar v_c$ and $\bar v = \bar u^c$.

As $U^* \in C^1(\mathbb{R})$ and $(U^*)' \geq 0$, the functional $\lambda \mapsto e(\lambda) := \int_X U^*(-\bar v(x) + \lambda)\, dx$ is differentiable and

$$e'(\lambda) = \int_X (U^*)'(-\bar v(x) + \lambda)\, dx \geq 0.$$

Since by lemma 2.4(iv) U^* grows superlinearly at infinity, so does $e(\lambda)$. Hence

$$\lim_{\lambda \to +\infty} J(\bar u + \lambda, \bar v - \lambda) = \lim_{\lambda \to +\infty} \int_X \bar u \varrho_0\, dx + \lambda - e(\lambda) = -\infty. \tag{3.6}$$

Moreover, as $U^* \geq 0$ (cfr. lemma 2.4(ii)),

$$\lim_{\lambda \to -\infty} J(\bar u + \lambda, \bar v - \lambda) \leq \lim_{\lambda \to -\infty} \int_X \bar u \varrho_0 dx + \lambda = -\infty. \tag{3.7}$$

Since $\lambda \mapsto J(\bar u + \lambda, \bar v - \lambda)$ is differentiable, (3.6) and (3.7) imply that $J(\bar u + \lambda, \bar v - \lambda)$ achieves its maximum at a certain value $\bar\lambda$ which satisfies $1 = e'(\bar\lambda)$. Therefore we have

$$(\tilde u, \tilde v) := (\bar u + \bar\lambda, \bar v - \bar\lambda) \in \mathcal{E}, \qquad J(\bar u, \bar v) \leq J(\tilde u, \tilde v), \qquad \text{and} \qquad \int_X (U^*)'(-\tilde v)\, dx = 1.$$

This last inequality and the fact that $(U^*)'(-\tilde v)$ is continuous on the compact set $\overline{X}$ ensure the existence of a point $x \in \overline{X}$ such that $-\tilde v(x) = U'\big(1/\mathcal{L}^d(X)\big)$. In light of lemma 3.2 and the above reasoning we have established that the set

$$\mathcal{E}' := \{(u,v) \ : \ (u,v) \in \mathcal{E}, u^c = v, v_c = u, v(x) = -U'\big(1/\mathcal{L}^d(X)\big) \text{ for some } x \in \overline{X}\}$$

satisfies the required conditions. □

Set

$$\phi(\varrho) := \mathcal{C}_h(\varrho_0, \varrho) + \mathcal{U}(\varrho).$$

Lemma 3.6. *Let ϱ_* be the unique minimizer of ϕ provided by lemma 3.1, and let (u_*, v_*) be a maximizer of J obtained in lemma 3.5. Then $\varrho_* = (U^*)'(-v_*)$, and*

$$\max_{\mathcal{E}} J = J(u_*, v_*) = \phi(\varrho_*) = \min_{\mathcal{P}^{ac}(X)} \phi.$$

Proof: Let T be as in lemma 3.5(ii), and define $\varrho_{v_*} := (U^*)'(-v_*)$. Note that since T pushes $\varrho_0\mathcal{L}^d$ forward to $\varrho_{v_*}\mathcal{L}^d$, we have that $(\mathbf{id}\times T)_\#(\varrho_0\mathcal{L}^d) \in \Gamma(\varrho_0, \varrho_{v_*})$. Therefore, as $c(x,Tx) = u_*(x) + v_*(Tx)$ for $\varrho_0\mathcal{L}^d$-a.e. $x \in X$,

$$\begin{aligned} \mathcal{C}_h(\varrho_0, \varrho_{v_*}) &\leq \int_X c(x,Tx)\varrho_0(x)\, dx = \int_X \big(u_*(x) + v_*(Tx)\big)\varrho_0(x)\, dx \\ &= \int_X u_*(x)\varrho_0(x)\, dx + \int_X v_*(x)\varrho_{v_*}(x)\, dx. \end{aligned} \tag{3.8}$$

Since

$$\mathcal{C}_h(\varrho_0,\varrho_{v_*}) \geq \int_X u_*(x)\varrho_0(x)\,dx + \int_X v_*(x)\varrho_{v_*}(x)\,dx$$

trivially holds (as $u \oplus v \leq c$), inequality (3.8) is in fact an equality, and so

$$\mathcal{C}_h(\varrho_0,\varrho_{v_*}) + \mathcal{U}(\varrho_{v_*}) = \int_X u_*(x)\varrho_0(x)\,dx + \int_X \big(v_*(x)\varrho_{v_*}(x) + U(\varrho_{v_*})\big)\,dx. \tag{3.9}$$

Combining (3.9) with the equality $-v_*\varrho_{v_*} = U(\varrho_{v_*}) + U^*(-v_*)$ (which follows from $\varrho_{v_*} = (U^*)'(-v_*)$) we get

$$\mathcal{C}_h(\varrho_0,\varrho_*) + \mathcal{U}(\varrho_*) = J(u_*,v_*),$$

which together with lemma 3.4 gives that ϱ_{v_*} minimizes ϕ over $\mathcal{P}^{ac}(X)$ and $\sup_{\mathcal{E}} J = \phi(\varrho_{v_*})$. Since the minimizer of ϕ over $\mathcal{P}^{ac}(X)$ is unique (cfr. lemma 3.1), this concludes the proof. □

Remark 3.7. Thanks to lemma 3.2, on $\mathrm{dom}\nabla v_*$ we can uniquely define a map S by $u_*(Sy)+v_*(y) = c(Sy,y)$, and we have $\nabla v_*(y) = \nabla_y c(Sy,y)$. This map is the inverse of T up to a set of zero measure, it pushes $\varrho_*\mathcal{L}^d$ forward to $\varrho_0\mathcal{L}^d$, and

$$Sy = \mathbf{\Phi}_1^L\big(-h, y, \nabla_p H\big(y, \nabla v_*(y)\big)\big).$$

Moreover, thanks to lemma 3.6, $U'(\varrho_*) = -v_*$ is Lipschitz, and

$$-\nabla v_*(y) = \nabla[U'(\varrho_*)](y).$$

In particular $Sy = \mathbf{\Phi}_1^L\big(-h, y, \nabla_p H\big(y, -\nabla[U'(\varrho_*)](y)\big)\big)$.

We observe that the duality method allows to deduce in an easy way the Euler-Lagrange equation associated to the functional ϕ, by-passing many technical problems due to regularity issues. Moreover it also gives $\nabla_y c(Sy,y) = -\nabla[U'(\varrho_*)](y)$ $\mathcal{L}^d$-a.e. in X (and not only $\varrho_*\mathcal{L}^d$-a.e.).

3.3. **Augmented actions.** We now introduce the functional

$$\Phi(\tau,\varrho_0,\varrho) := \mathcal{C}_\tau(\varrho_0,\varrho) + \mathcal{U}(\varrho) \qquad \varrho_0,\varrho \in \mathcal{P}^{ac}(X),$$

and we define

$$\phi_\tau(\varrho_0) := \inf_{\varrho\in\mathcal{P}^{ac}(X)} \Phi(\tau,\varrho_0,\varrho).$$

The goal of this subsection is to study the properties of Φ and ϕ_τ, in the same spirit as in [3, Chapter 3].

In the sequel, we fix $\varrho_0 \in \mathcal{P}^{ac}(X)$. Lemma 3.6 provides existence of a unique minimizer of $\Phi(\tau,\varrho_0,\varrho)$ over $\mathcal{P}^{ac}(X)$, which we call ϱ_τ.

Remark 3.8. (i) Note that $\phi_\tau(\varrho_0) \leq \Phi(\tau,\varrho_0,\varrho_0) \leq \mathcal{U}(\varrho_0)$ (since $\mathcal{C}_\tau(\varrho_0,\varrho_0) \leq 0$, cfr. lemma 2.7).

(ii) Thanks lemma 3.9 below, $\tau \mapsto \phi_\tau$ is monotone nonincreasing on $(0,+\infty)$. Therefore setting $\phi_0(\varrho_0) = \mathcal{U}(\varrho_0)$ ensures that $\tau \mapsto \phi_\tau$ remains monotone nonincreasing on $[0,+\infty)$, and we have

$$\mathcal{U}(\varrho_0) - \mathcal{U}(\varrho_h) = \phi_0(\varrho_0) - \phi_h(\varrho_0) + \mathcal{C}_h(\varrho_0,\varrho_h).$$

Lemma 3.9. *The function $\tau \mapsto \phi_\tau(\varrho_0)$ is nonincreasing, and satisfies*

$$\frac{\mathcal{C}_{\tau_1}(\varrho_0,\varrho_{\tau_1}) - \mathcal{C}_{\tau_0}(\varrho_0,\varrho_{\tau_1})}{\tau_1-\tau_0} \leq \frac{\phi_{\tau_1}(\varrho_0) - \phi_{\tau_0}(\varrho_0)}{\tau_1-\tau_0} \leq \frac{\mathcal{C}_{\tau_1}(\varrho_0,\varrho_{\tau_0}) - \mathcal{C}_{\tau_0}(\varrho_0,\varrho_{\tau_0})}{\tau_1-\tau_0} \quad \forall\, 0<\tau_0\leq\tau_1. \tag{3.10}$$

The function $\tau \mapsto \phi_\tau(\varrho_0)$ *is Lipschitz on an interval of the form* $[\tau_0, \tau_1] \subset (0, +\infty)$, $\frac{d\phi_\tau(\varrho_0)}{d\tau} \in L^1_{loc}([0, +\infty))$, *and*

$$\phi_{\tau_1}(\varrho_0) - \phi_{\tau_0}(\varrho_0) = \int_{\tau_0}^{\tau_1} \frac{d\phi_\tau(\varrho_0)}{d\tau}(\tau)\, d\tau \qquad \forall\, 0 \le \tau_0 \le \tau_1. \tag{3.11}$$

Proof: It is an immediate consequence of the definition of ϕ_τ and ϱ_τ that, for all $\tau_0, \tau_1 > 0$,

$$\mathcal{C}_{\tau_1}(\varrho_0, \varrho_{\tau_0}) - \mathcal{C}_{\tau_0}(\varrho_0, \varrho_{\tau_0}) \ge \phi_{\tau_1}(\varrho_0) - \phi_{\tau_0}(\varrho_0).$$

This gives the second inequality in (3.10), which together with lemma 2.13(i) implies that $\tau \mapsto \phi_\tau(\varrho_0)$ is nonincreasing. The first inequality in (3.10) can be established in a similar manner.

Set $S := [\tau_0, \tau_1] \subset (0, +\infty)$, and let $r > 0$ be such that X is contained in the closed ball of radius r centered at the origin. Let $\tilde{k}_S(r)$ be the constant appearing in lemma 2.12, and fix $\gamma \in \Gamma_{\tau_0}(\varrho_0, \varrho_{\tau_0})$ (cfr. remark 2.3). Since ϱ_0 and ϱ_{τ_0} are supported inside $\bar{X}$ we have that γ is supported on $\bar{X} \times \bar{X}$, which implies that $|c_{\tau_0}(x, y) - c_{\tau_1}(x, y)| \le \tilde{k}_S(r)$ for γ-a.e. (x, y). We conclude that

$$\mathcal{C}_{\tau_0}(\varrho_0, \varrho_{\tau_0}) = \int_{\mathbb{R}^d \times \mathbb{R}^d} c_{\tau_0}\, d\gamma \le \int_{\mathbb{R}^d \times \mathbb{R}^d} \Big(c_{\tau_1} + \tilde{k}_S(r)|\tau_0 - \tau_1|\Big)\, d\gamma \le \mathcal{C}_{\tau_1}(\varrho_0, \varrho_{\tau_0}) + \tilde{k}_S(r)|\tau_0 - \tau_1|.$$

This together with lemma 2.13(i) implies

$$|\mathcal{C}_{\tau_0}(\varrho_0, \varrho_{\tau_0}) - \mathcal{C}_{\tau_1}(\varrho_0, \varrho_{\tau_0})| = \mathcal{C}_{\tau_0}(\varrho_0, \varrho_{\tau_0}) - \mathcal{C}_{\tau_1}(\varrho_0, \varrho_{\tau_0}) \le \tilde{k}_S(r)|\tau_0 - \tau_1|. \tag{3.12}$$

Similarly

$$|\mathcal{C}_{\tau_1}(\varrho_0, \varrho_{\tau_1}) - \mathcal{C}_{\tau_0}(\varrho_0, \varrho_{\tau_1})| \le \tilde{k}_S(r)|\tau_0 - \tau_1|. \tag{3.13}$$

We combine (3.10), (3.12) and (3.13) to conclude that $\tau \mapsto \phi_\tau(\varrho_0)$ is Lipschitz on $[\tau_0, \tau_1] \subset (0, +\infty)$. Now, since $\tau \mapsto \phi_\tau(\varrho_0)$ is nonincreasing, recalling remark 3.8 we have

$$\int_{\tau_0}^{\tau_1} \Big|\frac{d\phi_\tau(\varrho_0)}{d\tau}(\tau)\Big|\, d\tau = \phi_{\tau_0}(\varrho_0) - \phi_{\tau_1}(\varrho_0) \le \mathcal{U}(\varrho_0) - \phi_{\tau_1}(\varrho_0).$$

Since $\tau_1 > \tau_0 > 0$ are arbitrary, we conclude that $\frac{d\phi_\tau(\varrho_0)}{d\tau} \in L^1_{loc}([0, +\infty))$ and (3.11) holds. □

For $h > 0$, we denote by T_h the optimal map that pushes $\varrho_0 \mathcal{L}^d$ forward to $\varrho_h \mathcal{L}^d$ as provided by the previous subsection 3.2. We have

$$T_h x = \mathbf{\Phi}_1^L\big(h, x, \nabla_p H\big(x, -\nabla u_h(\bar{x})\big)\big),$$

with (u_h, v_h) a maximizer of $(u, v) \mapsto \int_X \varrho_0 u dx - \int_X U^*(-v)\, dx$ over the set of $(u, v) \in C(\bar{X}) \times C(\bar{X})$ such that $u \oplus v \le c_h$. We recall that u_h and v_h are Lipschitz (cfr. lemma 3.5) and $(U^*)'(-v_h) = \varrho_h$ (cfr. lemma 3.6). Moreover, if we define the interpolation map between ϱ_0 and ϱ_h by

$$T_h^s x := \mathbf{\Phi}_1^L\big(s, x, \nabla_p H\big(x, -\nabla u_h(\bar{x})\big)\big), \qquad s \in [0, h], \tag{3.14}$$

we have

$$c_h(x, T_h x) = \int_0^h L\big(\sigma_0^x(s), \dot{\sigma}_0^x(s)\big) ds, \qquad \text{with } \sigma_0^x(s) := T_h^s x. \tag{3.15}$$

Finally, since $v_h = -U'(\varrho_h)$, denoting by S_h the inverse of T_h we also have

$$\nabla_y c_h(S_h y, y) = -\nabla[U'(\varrho_h)](y) = \nabla_v L\big(\sigma_0^{S_h y}(h), \dot{\sigma}_0^{S_h y}(h)\big) \quad \text{for } \mathcal{L}^d\text{-a.e. } y \in X. \tag{3.16}$$

Lemma 3.10. *For $\mathcal{L}^1$-a.e. $h>0$ we have*

$$\frac{d\phi_t(\varrho_0)}{dt}|_{t=h} = -\int_X H\big(y, -\nabla[U'(\varrho_h)](y)\big)\varrho_h(y)\,dy. \tag{3.17}$$

Proof: For $|\varepsilon| \le h/2$, $s\in[0,h+\varepsilon]$ and $\mathcal{L}^d$-a.e. $x\in X$ we define

$$\sigma^x_\varepsilon(s) := \mathbf{\Phi}^L_1\left(\frac{sh}{h+\varepsilon}, x, \nabla_p H\big(x, -\nabla u_h(x)\big)\right).$$

Because u_h is a Lipschitz function we have

$$\sup_{x,\varepsilon}\{||\sigma^x_\epsilon||_{C^1[0,h+\epsilon]} \ : |\varepsilon|\le h/2,\ x\in X\} < +\infty \tag{3.18}$$

Since $\sigma^x_\varepsilon(0)=x$ and $\sigma_\varepsilon(h+\varepsilon)=T_h x$, by the definition of $\mathcal{C}_{h+\varepsilon}$ we get

$$\mathcal{C}_{h+\varepsilon}(\varrho_0,\varrho_h) \le \int_X \varrho_0(x)\int_0^{h+\varepsilon} L\big(\sigma^x_\varepsilon, \dot\sigma^x_\varepsilon\big)\,ds\,dx = \frac{h+\varepsilon}{h}\int_X \varrho_0(x)\int_0^h L\Big(\sigma^x_0, \frac{h}{h+\varepsilon}\dot\sigma^x_0\Big)\,ds\,dx \tag{3.19}$$

Moreover, since $L(x,\cdot)$ is convex,

$$L\big(\sigma^x_0,\dot\sigma^x_0\big) \ge L\Big(\sigma^x_0, \frac{h}{h+\varepsilon}\dot\sigma^x_0\Big) + \frac{\varepsilon}{h+\varepsilon}\Big\langle \nabla_v L\big(\sigma^x_0, \frac{h}{h+\varepsilon}\dot\sigma^x_0\big), \dot\sigma^x_0\Big\rangle. \tag{3.20}$$

Recall that

$$\Big\langle \nabla_v L\Big(\sigma^x_0, \frac{h}{h+\varepsilon}\dot\sigma^x_0\Big), \frac{h}{h+\varepsilon}\dot\sigma^x_0\Big\rangle = L\Big(\sigma^x_0, \frac{h}{h+\varepsilon}\dot\sigma^x_0\Big) + H\Big(\sigma^x_0, \nabla_v L\Big(\sigma^x_0, \frac{h}{h+\varepsilon}\dot\sigma^x_0\Big)\Big), \tag{3.21}$$

We combine equations (3.19- 3.21) to obtain

$$\mathcal{C}_{h+\varepsilon}(\varrho_0,\varrho_h) \le \mathcal{C}_h(\varrho_0,\varrho_h) - \frac{\varepsilon}{h+\varepsilon}\int_X \varrho_0(x)\int_0^h H\Big(\sigma^x_0, \nabla_v L\Big(\sigma^x_0, \frac{h}{h+\varepsilon}\dot\sigma^x_0\Big)\Big)\,ds\,dx. \tag{3.22}$$

Thanks to (3.18) we can apply the Lebesgue dominated convergence theorem and then use to the conservation of the Hamiltonian H (cfr. remark 2.9(ii)) to obtain

$$\begin{aligned}\lim_{\varepsilon\to 0}\int_X \varrho_0(x)\int_0^h H\Big(\sigma^x_0, \nabla_v L\Big(\sigma^x_0, \frac{h}{h+\varepsilon}\dot\sigma^x_0\Big)\Big)\,ds\,dx &= \int_X \varrho_0(x)\int_0^h H\Big(\sigma^x_0, \nabla_v L\big(\sigma^x_0, \dot\sigma^x_0\big)\Big)\,ds\,dx \\ &= h\int_X \varrho_0(x) H\Big(\sigma^x_0(h), \nabla_v L\big(\sigma^x_0(h), \dot\sigma^x_0(h)\big)\Big)dx.\end{aligned}$$

Recalling that $\sigma^x_0(h) = T_h x$ pushes $\varrho_0\mathcal{L}^d$ forward to $\varrho_h\mathcal{L}^d$ and exploiting (3.16) in the previous equality we conclude that

$$\lim_{\varepsilon\to 0}\int_X \varrho_0(x)\int_0^h H\Big(\sigma^x_0, \nabla_v L\Big(\sigma^x_0, \frac{h}{h+\varepsilon}\dot\sigma^x_0\Big)\Big)\,ds\,dx = \int_X H\big(y, -\nabla[U'(\varrho_h)](y)\big)\varrho_h(y)\,dy. \tag{3.23}$$

We now distinguish two cases, depending on the sign of ε.

1. If $0<\varepsilon<h/2$ we set $\tau_1 = h+\varepsilon$ and $\tau_0 = h$ in the second inequality in (3.10) to obtain

$$\frac{\phi_{h+\varepsilon}(\varrho_0)-\phi_h(\varrho_0)}{\varepsilon} \le \frac{\mathcal{C}_{h+\varepsilon}(\varrho_0,\varrho_h) - \mathcal{C}_h(\varrho_0,\varrho_h)}{\varepsilon}.$$

This together with (3.22) yields

$$\frac{\phi_{h+\varepsilon}(\varrho_0)-\phi_h(\varrho_0)}{\varepsilon} \le -\frac{1}{h+\varepsilon}\int_X \varrho_0(x)\int_0^h H\Big(\sigma^x_0, \nabla_v L\Big(\sigma^x_0, \frac{h}{h+\varepsilon}\dot\sigma^x_0\Big)\Big)\,ds\,dx.$$

Letting ε tend to 0 and exploiting (3.23) we conclude that

$$\frac{d^+\phi_t(\varrho_0)}{dt}|_{t=h} \le -\int_X H\big(y, -\nabla[U'(\varrho_h)](y)\big)\varrho_h(y)\,dy. \tag{3.24}$$

2. If $-h/2 < \varepsilon < 0$ we set $\tau_1 = h$ and $\tau_0 = h + \varepsilon$ in the first inequality in (3.10), and by rearranging the terms we obtain

$$\frac{\mathcal{C}_{h+\varepsilon}(\varrho_0, \varrho_h) - \mathcal{C}_h(\varrho_0, \varrho_h)}{\varepsilon} \le \frac{\phi_{h+\varepsilon}(\varrho_0) - \phi_h(\varrho_0)}{\varepsilon}.$$

This, together with (3.22) yields

$$-\frac{1}{h+\varepsilon}\int_0^h H\Big(\sigma_0^x, \nabla_v L\Big(\sigma_0^x, \frac{h}{h+\varepsilon}\dot\sigma_0^x\Big)\Big)\,ds\,dx \le \frac{\phi_{h+\varepsilon}(\varrho_0) - \phi_h(\varrho_0)}{\varepsilon}. \tag{3.25}$$

We combine (3.23) and (3.25) to get

$$-\int_X H\big(y, -\nabla[U'(\varrho_h)](y)\big)\varrho_h(y)\,dy \le \frac{d^-\phi_t(\varrho_0)}{dt}|_{t=h}. \tag{3.26}$$

Since by lemma 3.9 $\tau \mapsto \phi_\tau(\varrho_0)$ is locally Lipschitz on $(0, +\infty)$, it is differentiable $\mathcal{L}^1$-a.e. Hence (3.24) and (3.26) yield (3.17) at any differentiability point. □

3.4. **De Giorgi's variational and "geodesic" interpolations.** We fix $\varrho_0 \in \mathcal{P}^{ac}(X)$, a time step $h > 0$, and set $\varrho_0^h := \varrho_0$. We consider $\varrho_h \in P^{ac}(X)$ the (unique) minimizer of $\Phi(\tau, \varrho_0, \cdot)$ provided by lemma 3.1, and we interpolate between ϱ_0^h and ϱ_h^h along paths minimizing the action $\mathcal{A}_h$: thanks to [8, Theorem 5.1] there exists a the unique solution $\bar\varrho_s^h \in \mathcal{P}^{ac}(X)$ of

$$\mathcal{C}_s\big(\varrho_0^h, \bar\varrho_s^h\big) + \mathcal{C}_{h-s}\big(\bar\varrho_s^h, \varrho_h^h\big) = \mathcal{C}_h\big(\varrho_0^h, \varrho_h^h\big),$$

which is also given by (cfr. (3.14) for the definition of T_h^s)

$$\bar\varrho_s^h \mathcal{L}^d := (T_h^s)_\# \varrho_0 \mathcal{L}^d, \qquad 0 \le s \le h.$$

Moreover [8, Theorem 5.1] ensures that T_h^s is invertible $\bar\varrho_s$-a.e., so that in particular there exists a unique vector field V_s^h defined $\bar\varrho_s$-a.e. such that

$$V_s^h(T_h^s) = \partial_s T_h^s \qquad \bar\varrho_s\text{-a.e.}$$

Recall that by lemma 3.5(i) $||\nabla u_h||_{L^\infty(X)} \le ||\nabla c_h||_{L^\infty(X\times X)}$. Exploiting equation (3.14) and the fact that $\partial_s \mathbf{\Phi}^L$ maps bounded subsets of $\mathbb{R}^d \times \mathbb{R}^d$ onto bounded subsets of $\mathbb{R}^d \times \mathbb{R}^d$, we obtain that $\sup_{0\le s\le h} ||\partial_s T_h^s||_{L^\infty(\bar\varrho_s)} < +\infty$. Therefore $\sup_{0\le s\le h} ||V_s^h||_{L^\infty(\bar\varrho_s)} < +\infty$. Finally a direct computation gives that

$$\partial_s \bar\varrho_s^h + \mathrm{div}(\bar\varrho_s^h V_s^h) = 0 \tag{3.27}$$

in the sense of distribution on $(0, h) \times \mathbb{R}^d$. Observe that $\bar\varrho_0^h = \varrho_0$ and $\bar\varrho_h^h = \varrho_h^h$.

Remark 3.11. Note that although the range of T_h is contained in X, that of T_h^s may fail to be in that set. Indeed even if x and $T_h x$ are both in X, the Lagrangian flow provided by L and connecting x to $T_h x$ may not lie entirely in X.

We set

$$\varrho_s := \operatorname{argmin}\Big\{\mathcal{C}_s(\varrho_0,\varrho) + \mathcal{U}(\varrho) : \ \varrho \in \mathcal{P}^{ac}(X)\Big\}, \qquad 0 \le s \le h.$$

In a metric space $(\mathcal{S}, \mathrm{dist})$ with $s\mathcal{C}_s = \mathrm{dist}^2$, the interpolation $s \mapsto \varrho_s$ is due to De Giorgi [6] (cfr. also [3, 7]).

Theorem 3.12. *The following energy inequality holds:*

$$\mathcal{U}(\varrho_0) - \mathcal{U}(\varrho_h) \ge \int_0^h ds \int_{\mathbb{R}^d} L(y, V_s)\bar{\varrho}_s \, dy + \int_0^h ds \int_X H\big(x, -\nabla U'(\varrho_s)\big)\varrho_s \, dx.$$

Proof: By lemma 3.10 $s \mapsto \phi_s(\varrho_0)$ is locally Lipschitz on $[0,h]$, and (3.17) holds. Hence, by exploiting remark 3.8 we conclude that

$$\begin{aligned}
\int_0^h dt \int_X H\big(x, -\nabla[U'(\varrho_s)]\big)\varrho_s \, dx &\le \phi_0(\varrho_0) - \phi_h(\varrho_0) = \mathcal{U}(\varrho_0) - \mathcal{U}(\varrho_h) - \mathcal{C}_h(\varrho_0,\varrho_h) \\
&= \mathcal{U}(\varrho_0) - \mathcal{U}(\varrho_h) - \int_0^h dt \int_X L(T_h^s(x), \partial_s T_h^s(x))\varrho_0(x) \, dx \\
&= \mathcal{U}(\varrho_0) - \mathcal{U}(\varrho_h) - \int_0^h dt \int_{\mathbb{R}^d} L(y, V_s(y))\bar{\varrho}_s(y) \, dy.
\end{aligned}$$

(We remark that the last integral has to be taken on the whole $\mathbb{R}^d$, as we do not know in general that the measures $\bar{\varrho}_s$ are concentrated on X, cfr. remark 3.11.) □

We now iterate the argument above: lemma 3.6 ensures the existence of a sequence $\{\varrho_{kh}^h\}_{k=0}^{\infty} \subset \mathcal{P}^{ac}(X)$ such that

$$\varrho_{(k+1)h}^h := \operatorname{argmin}\Big\{\mathcal{C}_h(\varrho_{kh}^h, \varrho) + \mathcal{U}(\varrho) \ : \ \varrho \in \mathcal{P}^{ac}(X)\Big\}.$$

As above, we define

$$\varrho_{kh+s}^h := \operatorname{argmin}\Big\{\mathcal{C}_s(\varrho_{kh}^h, \varrho) + \mathcal{U}(\varrho) \ : \ \varrho \in \mathcal{P}_1^{ac}\Big\}, \qquad 0 \le s \le h. \tag{3.28}$$

The arguments used before can be applied to $(\varrho_{kh}, \varrho_{(k+1)h})$ in place of (ϱ_0, ϱ_h) to obtain a unique map $T_{kh} : X \to X$ such that $(\mathbf{id} \times T_{kh})_\#(\varrho_{kh}\mathcal{L}^d) \in \Gamma_h(\varrho_{kh}, \varrho_{(k+1)h})$. Moreover, for $s \in (0,h)$ we define $\bar{\varrho}_{kh+s}^h$ to be the interpolation along paths minimizing the action $\mathcal{A}_h$, that is $\bar{\varrho}_{kh+s}$ is the unique solution of

$$\mathcal{C}_s\big(\varrho_{kh}^h, \bar{\varrho}_{kh+s}^h\big) + \mathcal{C}_{h-s}\big(\bar{\varrho}_{kh+s}^h, \varrho_{(k+1)h}^h\big) = \mathcal{C}_h\big(\varrho_{kh}^h, \varrho_{(k+1)h}^h\big).$$

We denote by (u_{kh+s}^h, v_{kh+s}^h) the solution to the dual problem to (3.28) provided by lemma 3.5. Replacing u_h by u_{kh}^h in (3.14) we obtain the interpolation maps T_{kh}^s. As before, we consider the interpolation measures $\bar{\varrho}_{kh+s}^h\mathcal{L}^d := (T_{kh}^s)_\#\varrho_{kh}^h\mathcal{L}^d$, and we define $\bar{\varrho}_{kh+s}^h$-a.e. the velocities V_{kh+s}^h by $V_{kh+s}^h(T_{kh}^s) = \partial_s T_{kh}^s$. As in (3.27), one can easily see that the curve of densities $s \mapsto \bar{\varrho}_s^h$ satisfies the continuity equation

$$\partial_s \bar{\varrho}_s^h + \mathrm{div}(\bar{\varrho}_s^h V_s^h) = 0 \tag{3.29}$$

in the sense of distribution on $(0, +\infty) \times \mathbb{R}^d$.

Corollary 3.13. *For $h > 0$, for any $j \le k \in \mathbb{N}$, we have*

$$\mathcal{U}(\varrho^h_{jh}) - \mathcal{U}(\varrho^h_{kh}) \ge \int_{jh}^{kh} ds \int_{\mathbb{R}^d} L(y, V^h_s)\bar{\varrho}^h_s \, dy + \int_{jh}^{kh} ds \int_X H\big(x, -\nabla[U'(\varrho^h_s)]\big)\varrho^h_s \, dx.$$

Proof: The result is a direct consequence of theorem 3.12. □

3.5. **Stability property and existence of solutions.** As before, we fix $\varrho_0 \in \mathcal{P}^{ac}(X)$ and make the additional assumption that $\mathcal{U}(\varrho_0)$ is finite. We fix $T > 0$ and we want to prove existence of solutions to equation (1.1) on $[0, T]$. Recall that by lemma 2.13(i) $\mathcal{C}_s(\varrho, \varrho) \le 0$ for any $s \ge 0$, $\varrho \in \mathcal{P}^{ac}_1$. This together with the definition of ϱ^h_{kh+s} yields

$$\mathcal{C}_h\big(\varrho^h_{kh}, \varrho^h_{kh+s}\big) + \mathcal{U}\big(\varrho^h_{kh+s}\big) \le \mathcal{U}\big(\varrho^h_{kh}\big), \qquad 0 \le s \le h.$$

By adding over $k \in \mathbb{N}$ the above inequality, thanks to remark 2.5 we get

$$\sum_{k=0}^{\infty} \mathcal{C}_h\big(\varrho^h_{kh}, \varrho^h_{(k+1)h}\big) \le \mathcal{U}\big(\varrho^h_0\big) - \liminf_{n\to\infty} \mathcal{U}\big(\varrho^h_{nh}\big) \le \mathcal{U}\big(\varrho^h_0\big) + |a| + |b|\mathcal{L}^d(X). \tag{3.30}$$

Similarly, using corollary 3.13, the fact that $H \ge 0$ and (L3), for any $N > 0$ integer we have

$$\begin{aligned} \mathcal{U}(\varrho^h_0) &\ge \mathcal{U}(\varrho^h_{Nh}) + \int_0^{Nh} ds \int_{\mathbb{R}^d} L(y, V^h_s)\bar{\varrho}^h_s \, dy \varrho^h_s \, dx \\ &\ge \int_0^{Nh} ds \int_{\mathbb{R}^d} \theta(|V^h_s|)\bar{\varrho}^h_s \, dy \varrho^h_s \, dx - A_* Nh - |a| - |b|\mathcal{L}^d(X). \end{aligned} \tag{3.31}$$

We also recall that $v^h_t : X \to \mathbb{R}$ is a Lipschitz function (cfr. lemma 3.5(i)) which satisfies $v^h_t = -U'(\varrho^h_t)$, so that setting

$$\beta^h_t := U^*(-v^h_t) = P(\varrho^h_t)$$

we have

$$\varrho^h_t \nabla[U'(\varrho^h_t)] = -(U^*)'(-v^h_t)\nabla v^h_t = \nabla[U^*(-v^h_t)] = \nabla[P(\varrho^h_t)] = \nabla\beta^h_t \qquad \mathcal{L}^d\text{-a.e.} \tag{3.32}$$

We start with the following:

Lemma 3.14. *Let A_* be the constant provided in assumption (L3). We have*

$$\mathcal{U}(\varrho^h_t) \le \mathcal{U}(\varrho_0) + A_* t. \tag{3.33}$$

Moreover, for any $K > 0$ there exists a constant $C(K) > 0$ such that, for any $h \in (0, 1]$,

$$W_1(\bar{\varrho}^h_t, \varrho^h_t) \le \frac{C_0}{K} + 2\frac{A_* + C(K)}{K}h \qquad \forall\, t \in [0, T], \tag{3.34}$$

$$W_1(\bar{\varrho}^h_t, \varrho^h_{kh}) \le \frac{C_0}{K} + \frac{A_* + C(K)}{K}h \qquad \forall\, t \in [kh, (k+1)h],\ k \in \mathbb{N}, \tag{3.35}$$

$$W_1\big(\varrho^h_t, \varrho^h_s\big) \le \frac{C_0}{K} + 2\frac{A_* + C(K)}{K}[(t-s) + h] \qquad \forall\, 0 \le s \le t, \tag{3.36}$$

$$W_1\big(\bar{\varrho}^h_t, \bar{\varrho}^h_s\big) \le \frac{C_0}{K} + \frac{A_* + C(K)}{K}[(t-s) + h] \qquad \forall\, 0 \le s \le t. \tag{3.37}$$

Here C_0 is a positive constant independent of t, K, and $h \in (0, 1]$.

Proof: Let $t \in [kh,(k+1)h]$ for some $k \in \mathbb{N}$. Then by lemma 2.13(ii) we have

$$\mathcal{U}(\varrho_t^h) - A_*(t-kh) \le \mathcal{U}(\varrho_t^h) + \mathcal{C}_{t-kh}(\varrho_t^h, \varrho_{kh}^h) \le \mathcal{U}(\varrho_{kh}^h). \tag{3.38}$$

In particular $\mathcal{U}(\varrho_{(k+1)h}^h) \le \mathcal{U}(\varrho_{kh}^h) + A_* h$ for all $k \in \mathbb{N}$, so that adding over k we get

$$\begin{aligned} \mathcal{U}(\varrho_t^h) &\le \mathcal{U}(\varrho_{kh}^h) + A_*(t-kh) \le \mathcal{U}(\varrho_{(k-1)h}^h) + A_*[h + (t-kh)] \\ &\le \ldots \le \mathcal{U}(\varrho_0) + A_*[kh + (t-kh)] = \mathcal{U}(\varrho_0) + A_* t. \end{aligned}$$

This proves (3.33).

Now, since $\mathcal{C}_h \le \mathcal{C}_{t-kh}$ (cfr. lemma 2.13(i)), we have

$$\mathcal{C}_h(\varrho_{kh}^h, \varrho_t^h) \le \mathcal{U}(\varrho_{kh}^h) - \mathcal{U}(\varrho_t^h) \qquad \forall\, t \in [kh,(k+1)h],$$

which combined with equation (3.38) and remark 2.5 gives

$$\mathcal{C}_h(\varrho_{kh}^h, \varrho_t^h) \le \mathcal{U}(\varrho_0) + A_* h + |a| + |b|\mathcal{L}^d(X) \qquad \forall\, t \in [kh,(k+1)h]. \tag{3.39}$$

Moreover, as $\varrho_{kh}^h = \bar\varrho_{kh}^h$ for any $k \in \mathbb{N}$, using again lemma 2.13(ii) we get

$$\begin{aligned} \mathcal{C}_h(\varrho_{kh}^h, \varrho_{(k+1)h}^h) &= \mathcal{C}_{t-kh}(\varrho_{kh}^h, \bar\varrho_t^h) + \mathcal{C}_{(k+1)h-t}(\bar\varrho_t^h, \varrho_{(k+1)h}^h) \\ &\ge \mathcal{C}_{t-kh}(\varrho_{kh}^h, \bar\varrho_t^h) - A_* h \ge \mathcal{C}_h(\varrho_{kh}^h, \bar\varrho_t^h) - A_* h. \end{aligned}$$

Thanks to lemma 2.13(ii)-(iii), for any $K > 0$ there exists a constant $C(K) > 0$ such that

$$\begin{aligned} W_1(\varrho_{kh}^h, \varrho_t^h) &\le \frac{1}{K}\mathcal{C}_h(\varrho_{kh}^h, \varrho_t^h) + \frac{A_* + C(K)}{K} h, \\ W_1(\varrho_{kh}^h, \bar\varrho_t^h) &\le \frac{1}{K}\mathcal{C}_h(\varrho_{kh}^h, \bar\varrho_t^h) + \frac{A_* + C(K)}{K} h. \end{aligned}$$

Using the triangle inequality for W_1 and combining together the estimates above, (3.34) and (3.35) follow.

Finally, to prove equations (3.36) and (3.37), we observe that (3.30) combined with lemma 2.13(iii) gives

$$\begin{aligned} W_1\big(\varrho_{Nh}^h, \varrho_{Mh}^h\big) &\le \sum_{j=M}^{N-1} W_1\big(\varrho_{(j+1)h}^h, \varrho_{jh}^h\big) \\ &\le \frac{1}{K}\sum_{j=M}^{N-1} \mathcal{C}_h(\varrho_{(j+1)h}^h, \varrho_{jh}^h) + \frac{A_* + C(K)}{K} h(N-M) \\ &\le \frac{1}{K}\big[\mathcal{U}\big(\varrho_0^h\big) + |a| + |b|\mathcal{L}^d(X)\big] + \frac{A_* + C(K)}{K} h(N-M). \end{aligned}$$

Combining this estimate with (3.34) and (3.35) we obtain the desired result. □

Lemma 3.15. *(i) The curve $t \mapsto \bar\varrho_t^h \in \mathcal{P}_1^{ac}(\mathbb{R}^d)$ is continuous on $[0,T]$.*

(ii) The curve $t \mapsto \varrho_t^h \in \mathcal{P}^{ac}(X)$ is continuous on $[0,T]$ (with respect to W_1).

Proof:

(i) Thanks to (3.29) and (3.31), it is not difficult to show $t \mapsto \bar\varrho_t^h \in \mathcal{P}_1^{ac}(\mathbb{R}^d)$ is (uniformly) continuous on $[0,T]$ (cfr. [3, Chapter 8]).

(ii) Without loss of generality we can restrict the study of (ii) to the set $[0,h]$. Before starting our argument, let us recall that since X is bounded the weak* convergence coincides with the W_1 convergence on $\mathcal{P}^{ac}(X)$. Fix $s \in (0,h]$, and let $S \subset (0,h]$ be a closed interval. Lemma 2.12 yields that $t \mapsto \mathcal{C}_t(\varrho_o, \varrho)$ is Lipschitz on S, with a Lipschitz constant $\tilde{k}_S$ independent of ϱ. Since X is bounded and $\mathcal{P}^{ac}(X)$ is precompact for the weak* convergence, it is also precompact for W_1. Now, given $\varrho \in \mathcal{P}^{ac}(X)$, by definition of ϱ_t we have $\Phi(t, \varrho_0, \varrho_t) \le \Phi(t, \varrho_0, \varrho)$, and so

$$\Phi(s, \varrho_0, \varrho_t) \le \Phi(s, \varrho_0, \varrho) + 2|t-s|\tilde{k}_S \tag{3.40}$$

Let $\{\varrho_{t_n}\}_{n\in\mathbb{N}}$ be an arbitrary subsequence of $\{\varrho_t\}_{t\in S}$ converging to ϱ^* as $t_n \to s$. Then, thanks to 3.40 and the fact that $\Phi(s, \varrho_0, \cdot)$ is lower semicontinuous for the weak* topology we obtain

$$\Phi(s, \varrho_0, \varrho^*) \le \Phi(s, \varrho_0, \varrho),$$

that is ϱ^* minimizes $\Phi(s, \varrho_0, \cdot)$ over $\mathcal{P}^{ac}(X)$. Hence, thanks to lemma 3.1 $\varrho^* = \varrho_s$. Since the limit ϱ^* is independent of the subsequence $\{t_n\}_{n\in\mathbb{N}}$ and we are on a metric space, we conclude that $\{\varrho_t\}_{t\in S}$ converges to ϱ_s as $t \to s$. This prove that $t \mapsto \varrho_t^h$ is continuous at s.

It remains to show that $t \mapsto \varrho_t^h$ is continuous at 0. The fact that $c_t(x,x) \le 0$ (cfr. lemma 2.7) shows that $\mathcal{C}_t(\varrho_0, \varrho_0) \le 0$. Hence, using the definition of ϱ_t we obtain

$$\mathcal{C}_t(\varrho_0, \varrho_t) + \mathcal{U}(\varrho_t) \le \mathcal{C}_t(\varrho_0, \varrho_0) + \mathcal{U}(\varrho_0) \le \mathcal{U}(\varrho_0).$$

Combining the above estimate with 2.13(iii) we conclude that for each $K > 0$ there exists a constant $C(K)$ independent of t such that

$$-(A_* + C(K))t + KW_1(\varrho_0, \varrho_t) + \mathcal{U}(\varrho_t) \le \mathcal{U}(\varrho_0). \tag{3.41}$$

Let $\{\varrho_{t_n}\}_{n\in\mathbb{N}}$ is be subsequence of $\{\varrho_t\}_{t\in S}$ converging to ϱ^* as $t_n \to 0$. Then the lower semicontinuity of $\mathcal{U}$ with respect to the weak* topology together with (3.41) give

$$KW_1(\varrho_0, \varrho^*) + \mathcal{U}(\varrho^*) \le \mathcal{U}(\varrho_0).$$

Letting $K \to +\infty$ we obtain $W_1(\varrho_0, \varrho^*)$, that is $\varrho^* = \varrho_0$. By the arbitrariness of the sequence $\{t_n\}_{n\in\mathbb{N}}$ we conclude as before that $t \mapsto \varrho_t^h$ is continuous at 0. ☐

We can now prove the compactness of our discrete solutions.

Proposition 3.16. *There exists a sequence $h_n \to 0$, a density $\varrho \in \mathcal{P}^{ac}([0,T] \times X)$, and a Borel function $V : [0,T] \times X \to \mathbb{R}^d$ such that:*

(i) *The curves $t \mapsto \varrho_t^{h_n} \in \mathcal{P}^{ac}(X)$ and $t \mapsto \bar{\varrho}_t^{h_n} \in \mathcal{P}^{ac}(\mathbb{R}^d)$ converge to $t \mapsto \varrho_t := \varrho(t, \cdot)$ with respect to the narrow topology. Moreover the curve $t \mapsto \varrho_t := \varrho(t, \cdot)$ is uniformly continuous and $\lim_{t\to 0^+} \varrho_t = \varrho_0$ in $(\mathcal{P}^{ac}(X), W_1)$.*

(ii) *The vector-valued measures $\bar{\varrho}_t^n(x) V_t^{h_n}(x) dxdt$ converge narrowly to $\varrho_t(x) V_t(x) dxdt$, where $V_t := V(t, \cdot)$.*

(iii) *$\partial_t \varrho_t + \mathrm{div}(\varrho_t V_t) = 0$ holds on $(0,T) \times X$ in the sense of distribution.*

Proof: First of all, let us recall that the narrow topology is metrizable (cfr. [3, Chapter 5]).

Thanks to lemma 3.15 and the estimates (3.36) and (3.37), as $K > 0$ is arbitrary it is easy to see that the curves $t \mapsto \varrho_t^h$ and $t \mapsto \bar{\varrho}_t^h$ are equicontinuous with respect to the 1-Wasserstein distance. Since bounded sets with respect to W_1 are precompact with respect to the narrow topology on $\mathbb{R}^d$ (cfr. for instance [3, Chapter 7]), by Ascoli-Arzelà Theorem we can find a sequence $h_n \to 0$ such that $t \mapsto \varrho_t^{h_n} \in \mathcal{P}^{ac}(X)$ and $t \mapsto \bar{\varrho}_t^{h_n} \in \mathcal{P}^{ac}(\mathbb{R}^d)$ converge uniformly (locally in time) to a narrow-continuous curve $t \mapsto \mu_t \in \mathcal{P}(X)$ (which is the same for both $\varrho_t^{h_n}$ and $\bar{\varrho}_t^{h_n}$ thanks to (3.34)).

Moreover $t \mapsto \mu_t$ is supported in X as so is $\varrho_t^{h_n}$, and the initial condition $w^*\text{-}\lim_{t\to 0^+} \bar\varrho_t^{h_n} = \varrho_0$ holds in the limit.

Concerning the vector-valued measure $V_t^h \varrho_t^h$, recalling that $H \geq 0$, thanks to corollary 3.13 and remark 2.5 we have

$$\int_0^T dt \int_{\mathbb{R}^d} L(x, V_t^h)\bar\varrho_t^h \, dx \leq \mathcal{U}(\varrho_0) + |a| + |b|\mathcal{L}^d(X).$$

By (L3) this gives

$$\int_0^T dt \int_{\mathbb{R}^d} \theta(|V_t^h|)\bar\varrho_t^h \, dx \leq \mathcal{U}(\varrho_0) + |a| + |b|\mathcal{L}^d(X) + A_* T =: C_1.$$

The above inequality, together with the superlinearity of θ and the uniform convergence of $\bar\varrho^h$ to μ_t, implies easily that the vector-valued measure $V_t^h \bar\varrho^h$ have a limit point λ, which is concentrated on $[0,T] \times X$. Moreover, the superlinearity and the convexity of θ ensure that $\lambda \ll \mu$, and there exists a μ-measurable vector field $V : [0,T] \times \overline{X} \to \mathbb{R}^d$ such that $\lambda = V\mu$, and

$$\int_0^T dt \int_X \theta(|V_t|) \, d\mu_t \leq C_1.$$

To conclude the proof of (i) and (ii) we have to show that $\mu \ll \mathcal{L}^{d+1}$. We observe that thanks to equation (3.33)

$$\int_0^T dt \int_X U(\varrho_t^{h_n}) \, dx = \int_0^T \mathcal{U}(\varrho_t^{h_n}) \, dt \leq T\mathcal{U}(\varrho_0) + A_* \frac{T^2}{2},$$

so that by the superlinearity of U any limit point of ϱ^{h_n} is absolutely continuous. Hence $\mu = \varrho\mathcal{L}^d$, and (i) and (ii) are proved.

Finally from equation (3.29) we deduce that

$$\partial_t \bar\varrho_t^{h_n} + \operatorname{div}(\bar\varrho_t^{h_n} V_t^{h_n}) = 0 \qquad \text{on } (0,T) \times X \tag{3.42}$$

in the sense of distribution, so that (iii) follows taking the limit as $n \to \infty$. □

We are now ready to prove the following existence result. To simplify the notation, given two nonnegative functions f and g, we write $f \gtrsim g$ if there exists two nonnegative constants c_0, c_1 such that $c_0 f + c_1 \geq g$. If both $\gtrsim$ and $\lesssim$ hold, we write $f \sim g$.

Theorem 3.17. *Let $X \subset \mathbb{R}^d$ be an open bounded set whose boundary is of zero Lebesgue measure, and assume that H satisfies (H1), (H2) and (H3). Assume that U satisfies* (2.2) *and* (2.3)*, and let $\varrho_0 \in \mathcal{P}_1^{ac}(X)$ be such that $\mathcal{U}(\varrho_0)$ is finite. Let ϱ_t and V_t be as in proposition 3.16. Then we have $P(\varrho_t) \in L^1(0,T; W^{1,1}(X))$, $\nabla[P(\varrho_t)]$ is absolutely continuous with respect to ϱ_t, and*

$$\mathcal{U}(\varrho_0) - \mathcal{U}(\varrho_T) \geq \int_0^T dt \int_X \Big[L\big(x, V_t\big) + H\big(x, -\varrho_t^{-1}\nabla[P(\varrho_t)]\big)\Big]\varrho_t \, dx. \tag{3.43}$$

Furthermore, if $\theta(t) \sim t^\alpha$ for some $\alpha > 1$ and U satisfies the doubling condition

$$U(t+s) \leq C(U(t) + U(s) + 1) \qquad \forall\, t, s \geq 0, \tag{3.44}$$

then $\varrho_t \in AC_\alpha(0,T; \mathcal{P}_\alpha^{ac}(X))$,

$$\mathcal{U}(\varrho_{T_1}) - \mathcal{U}(\varrho_{T_2}) = -\int_{T_1}^{T_2} dt \int_X \big\langle \nabla[P(\varrho_t)], V_t \big\rangle \, dx, \tag{3.45}$$

for $T_1, T_2 \in [0,T]$. *In particular*

$$V_t(x) = \nabla_p H\big(x, -\varrho_t^{-1}\nabla[P(\varrho_t)]\big) \qquad \varrho_t\text{-a.e.},$$

and ϱ_t *is a distributional solution of equation* (1.1) *starting from* ϱ_0.
Suppose in addition that (H4) holds. If $\varrho_0 \le M$ *for some* $M \ge 0$ *then* $\varrho_t \le M$ *for every* $t \in (0,T)$ *(maximum principle).*

Proof: The maximum principle is a direct consequence of lemma 3.1. We first remark that the last part of the statement is a simple consequence of equations (3.43) and (3.45) combined with proposition 3.16(i)-(iii). So it suffices to prove equations (3.43) and (3.45).
We first prove (3.43). Corollary 3.13 implies that, if $T \in [kh_n, (k+1)h_n]$ for some $k \in \mathbb{N}$, since $L \ge -A_*$ and $H \ge 0$ we have

$$\mathcal{U}(\varrho_0^{h_n}) - \mathcal{U}(\varrho_{(k+1)h_n}^{h_n}) \ge \int_0^T dt \int_{\mathbb{R}^d} \Big[L(x, V_t^{h_n})\bar\varrho_t^{h_n} + H\big(x, -\nabla[U'(\varrho_t^{h_n})]\big)\varrho_t^{h_n}\Big]\,dx - A_* h_n.$$

We now consider two continuous functions $w, \bar w : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ with compact support. Then

$$\begin{aligned}\int_0^T dt \int_{\mathbb{R}^d} \Big[L(x, V_t^{h_n})\bar\varrho_t^{h_n} \;&+\; H\big(x, -\nabla[U'(\varrho_t^{h_n})]\big)\varrho_t^{h_n}\Big]\,dx \\ &\ge \int_0^T dt \int_{\mathbb{R}^d} \Big[\langle V_t^{h_n}, \bar w(t,x)\rangle \bar\varrho_t^{h_n} - H(x, \bar w(t,x))\bar\varrho_t^{h_n}\Big]\,dx \\ &+ \int_0^T dt \int_X \Big[\langle -\nabla[U'(\varrho_t^{h_n})], w(t,x)\rangle \varrho_t^{h_n} - L(x, w(t,x))\varrho_t^{h_n}\Big]\,dx.\end{aligned}$$

Thanks to proposition 3.16(i)-(ii) we immediately get

$$\begin{aligned}\lim_{n\to\infty} \int_0^T dt \int_{\mathbb{R}^d} &\Big[\langle V_t^{h_n}, \bar w(t,x)\rangle \bar\varrho_t^{h_n} - H(x, \bar w(t,x))\bar\varrho_t^{h_n}\Big]\,dx \\ &= \int_0^T dt \int_{\mathbb{R}^d} \Big[\langle V_t, \bar w(t,x)\rangle \varrho_t - H(x, \bar w(t,x))\varrho_t\Big]\,dx,\end{aligned}$$

so that taking the supremum among all continuous functions $\bar w : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ with compact support we obtain

$$\liminf_{n\to\infty} \int_0^T dt \int_{\mathbb{R}^d} L(x, V_t^{h_n})\bar\varrho_t^{h_n}\,dx \ge \int_0^T dt \int_{\mathbb{R}^d} L(x, V_t)\varrho_t\,dx.$$

Concerning the other term, we observe that, thanks to remark 2.5, as $L \ge -A_*$ we have that

$$\int_0^T dt \int_{\mathbb{R}^d} H\big(x, -\nabla[U'(\varrho_t^{h_n})]\big)\varrho_t^{h_n}\,dx$$

is uniformly bounded with respect to n. In particular, since by (H3) $H(x,p) \ge |p| - C_1$ for some constant C_1, thanks to equation (3.32) we get that

$$\int_0^T dt \int_{\mathbb{R}^d} |\nabla[P(\varrho_t^{h_n})]|\,dx = \int_0^T dt \int_{\mathbb{R}^d} |\nabla[U'(\varrho_t^{h_n})]|\varrho_t^{h_n}\,dx$$

is uniformly bounded. This implies that, up to a subsequence, the vector-valued measures $\nabla[P(\varrho_t^{h_n})]dxdt$ converges weakly to a measure ν of finite total mass. Therefore we obtain

$$\begin{aligned} +\infty &> \liminf_{n\to\infty}\int_0^T dt\int_{\mathbb{R}^d} H\big(x,-\nabla[U'(\varrho_t^{h_n})]\big)\varrho_t^{h_n}\,dx \\ &\geq \lim_{n\to\infty}\int_0^T dt\int_X \Big[\langle -\nabla[U'(\varrho_t^{h_n})], w(t,x)\rangle\varrho_t^{h_n} - L(x,w(t,x))\varrho_t^{h_n}\Big]\,dx \\ &= \int_0^T dt\int_X -\langle w(t,x),\nu(dt,dx)\rangle - \int_0^T dt\int_X L(x,w(t,x))\varrho_t\,dx. \end{aligned}$$

By the arbitrariness of w we easily get that the measure $\nu(dt,dx)$ is absolutely continuous with respect to $\varrho_t dxdt$, so that $\nu(dt,dx) = e_t(x)\varrho_t(x)dxdt$ for some Borel function $e:[0,T]\times X\to\mathbb{R}^d$. We now observe that by Fatou Lemma we also have

$$\int_0^T\left(\liminf_{n\to\infty}\int_{\mathbb{R}^d}|\nabla[P(\varrho_t^{h_n})]|\,dx\right)dt < +\infty,$$

which gives

$$\liminf_{n\to\infty}\int_{\mathbb{R}^d}|\nabla[P(\varrho_t^{h_n})]|\,dx < +\infty \qquad \text{for } t\in[0,T]\setminus\mathcal{N}, \tag{3.46}$$

with $\mathcal{L}^1(\mathcal{N})=0$. Hence, for any $t\in[0,T]\setminus\mathcal{N}$ there exists a subsequence $\varrho_t^{h_{n_k(t)}}$ such that

$$\liminf_{n\to\infty}\int_{\mathbb{R}^d}|\nabla[P(\varrho_t^{h_n})]|\,dx = \lim_{k\to\infty}\int_{\mathbb{R}^d}|\nabla[P(\varrho_t^{h_{n_k(t)}})]|\,dx,$$

and $P(\varrho_t^{h_{n_k(t)}})$ converges weakly in $BV(X)$ and $\mathcal{L}^d$-a.e. to a function β_t. As a consequence $D\beta_t = e_t\varrho_t\mathcal{L}^d$, so that $\beta_t\in W^{1,1}(X)$. Since Q is continuous, we deduce that $\varrho_t^{h_{n_k(t)}} = Q\big(P(\varrho_t^{h_{n_k(t)}})\big)$ converges $\mathcal{L}^d$-a.e. to $Q(\beta_t)$. Recalling that $\varrho_t^{h_{n_k(t)}}$ also converges weakly to ϱ_t, we obtain $Q(\beta_t)=\varrho_t$, that is $\beta_t = P(\varrho_t)$. Moreover, from the equality $\nabla\beta_t = e_t\varrho_t$, we get $\nabla[P(\varrho_t)] = e_t\varrho_t$. We have proved that $P(\varrho_t)\in L^1(0,T;W^{1,1}(X))$ and $\nabla[P(\varrho_t)]$ is absolutely continuous with respect to ϱ_t. Finally $\varrho_{(k+1)h_n}^{h_n}$ converges weakly* to ϱ_T, and the term $\mathcal{U}(\varrho_T^{h_n})$ is lower-semicontinuous under weak* convergence, and this concludes the proof of equation (3.43).

We now prove equation (3.45). Let us observe that the assumption $\theta\sim t^\alpha$ implies that $L(x,v)\gtrsim |v|^\alpha$ and $H(x,p)\gtrsim |p|^{\alpha'}$, where $\alpha'=\alpha/(\alpha-1)$. Hence, thanks to equation (3.43) we have

$$+\infty > \int_0^T dt\int_X L(x,V_t)\varrho_t\,dx \gtrsim \int_0^T ||V_t||^\alpha_{L^\alpha(\varrho_t)}\,dt \tag{3.47}$$

and

$$+\infty > \int_0^T dt\int_X H(x,-e_t)\varrho_t\,dx \gtrsim \int_0^T ||e_t||^{\alpha'}_{L^{\alpha'}(\varrho_t)}\,dt. \tag{3.48}$$

Since

$$\partial_t\varrho_t + \mathrm{div}\big(\varrho_t V_t\big) = 0,$$

(3.47) implies that the curve $t\mapsto\varrho_t$ is absolutely continuous with values in the α-Wasserstein space $P_\alpha(X)$, and we denote by $\bar V$ its velocity field of minimal norm (cfr. [3, Chapter 8]). Moreover, thanks to equation (3.48), $e_t\in L^{\alpha'}(\varrho_t)$ for a.e. $t\in(0,T)$.

Denoting by $|\varrho'|$ the metric derivative of the curve $t \mapsto \varrho_t$ (with respect to the α-Wasserstein distance, cfr. equation (2.10)), by (3.47) and [3, theorem 8.3.1] we have

$$|\varrho'|(t) \leq ||\bar{V}_t||_{L^\alpha(\varrho_t)} \leq ||V_t||_{L^\alpha(\varrho_t)} < +\infty. \tag{3.49}$$

Since $e_t\varrho_t = \nabla P(\varrho_t)$ with $P(\varrho_t) \in W^{1,1}(X)$ for a.e. t, we can apply [3, Theorem 10.4.6] to conclude that, for $\mathcal{L}^1$-a.e. t, $\mathcal{U}$ has a finite slope at $\varrho\mathcal{L}^d$, $|\partial\mathcal{U}|(\varrho_t) = ||e_t||_{L^{\alpha'}(\varrho_t)}$, and $e_t = \partial^o\mathcal{U}(\varrho_t)$. The last statement means that e_t is the element of minimal norm of the convex set $\partial\mathcal{U}(\varrho_t)$, and so it belongs to the closure of $\{\nabla\varphi : \varphi \in C_c^\infty(X)\}$ in $L^{\alpha'}(\varrho_t)$. Let $\Lambda \subset (0, T)$ be the set of t such that

(a) $\partial\mathcal{U}(\varrho_t) \neq \emptyset$;
(b) $\mathcal{U}$ is approximately differentiable at t;
(c) (8.4.6) of [3] holds.

We use equations (3.48), (3.49), and the fact that $|\partial\mathcal{U}|(\varrho_t) = ||e_t||_{L^{\alpha'}(\varrho_t)}$ for $\mathcal{L}^1$-a.e. $t \in (0, T)$, to conclude that

$$\int_0^T |\partial\mathcal{U}|(\varrho_t)|\varrho'|(t)dt \leq \frac{1}{\alpha'}\int_0^T ||e_t||^{\alpha'}_{L^{\alpha'}(\varrho_t)}dt + \frac{1}{\alpha}\int_0^T ||V_t||^\alpha_{L^\alpha(\varrho_t)}dt < +\infty. \tag{3.50}$$

By [3, Proposition 9.3.9] $\mathcal{U}$ is convex along α-Wasserstein geodesics, and so exploiting equation (3.50) and invoking [3, Proposition 10.3.18] we obtain that $\mathcal{L}^1\big((0, T) \setminus \Lambda\big) = 0$ and $t \mapsto \mathcal{U}(\varrho_t)$ is absolutely continuous. Thus its pointwise, distributional, and approximate derivatives coincide almost everywhere, and by [3, Proposition 10.3.18] and the fact that $e_t \in \partial\mathcal{U}(\varrho_t)$ we get

$$\frac{d}{dt}\mathcal{U}(\varrho_t) = \int_X \langle e_t, \bar{V}_t\rangle\varrho_t dx. \tag{3.51}$$

Because V and $\bar{V}$ are both velocity fields for $t \mapsto \varrho_t$ we have

$$\int_X \langle\nabla\phi, V_t - \bar{V}_t\rangle\varrho_t dx = 0$$

for all $\phi \in C_c^\infty(X)$ for $\mathcal{L}^1$-a.e. $t \in (0, T)$, and since e_t belongs to the closure of $\{\nabla\varphi : \varphi \in C_c^\infty(X)\}$ in $L^{\alpha'}(\varrho_t)$ we conclude by a density argument that

$$\int_X \langle e_t, V_t - \bar{V}_t\rangle\varrho_t dx = 0$$

for $\mathcal{L}^1$-a.e. $t \in (0, T)$. This together with equation (3.51) finally yields

$$\mathcal{U}(\varrho_{T_1}) - \mathcal{U}(\varrho_{T_2}) = -\int_{T_1}^{T_2} dt \int_X \langle e_t, \bar{V}_t\rangle\varrho_t dx = -\int_{T_1}^{T_2} dt \int_X \langle e_t, V_t\rangle\varrho_t dx, \tag{3.52}$$

as desired. □

Remark 3.18. If ϱ_0 is a general nonnegative integrable function on X which does not necessarily have unit mass, we can still prove existence of solutions to equation (1.1). Indeed, defining $c := \int_X \varrho_0\, dx$, we consider $\varrho_t^c \in \mathcal{P}^{ac}(X)$ a solution of equation (1.1) for the Hamiltonian $H^c(x, p) := cH(x, p/c)$ and the internal energy $U^c(t) := U(ct)$, starting from $\varrho_0^c := \varrho_0/c$. Then $\varrho_t := c\varrho_t^c$ solves equation (1.1). Moreover, using this scaling argument also at a discrete level, we can also construct discrete solutions starting from ϱ_0.

Remark 3.19. We believe that the above existence result could be extend to more general functions θ by introducing some Orlicz-type spaces as follows: for $\theta : [0, +\infty) \to [0, +\infty)$ convex, superlinear, and such that $\theta(0) = 0$, we define the Orlicz-Wasserstein distance

$$\mathcal{W}_\theta(\mu_0, \mu_1) := \inf \left\{ \lambda > 0 \ : \ \inf_{\gamma \in \Gamma(\mu_0,\mu_1)} \int_{X \times X} \theta\left(\frac{|x-y|}{\lambda}\right) d\gamma \leq 1 \right\}.$$

We also define the Orlicz-type norm

$$\|f\|_{\theta,\mu} := \inf \left\{ \lambda > 0 \ : \ \int_X \theta\left(\frac{|f|}{\lambda}\right) d\mu(x) \leq 1 \right\}.$$

It is not difficult to prove that the following dynamical formulation of the Orlicz-Wasserstein distance holds:

$$\mathcal{W}_\theta(\mu_0, \mu_1) = \inf \left\{ \int_0^1 \|V_t\|_{\theta,\mu_t} \, dt \ : \ \partial_t \mu_t + \operatorname{div}(\mu_t V_t) = 0 \right\}. \tag{3.53}$$

Now, in order to prove the identity (3.45) of the previous theorem in the case where θ does not necessarily behave as a power function, one should extend the results of [3] to this more general setting. We believe such an extension to be reachable although not straightforward. However this kind of effort goes beyond the scope of this paper.

4. Existence of solutions in unbounded domains for weak Tonelli Lagrangians

The aim of this section is to extend the existence result proved in the previous section to unbounded domains X, using an approximation argument where we construct our solutions in $X \cap B_m(0)$ for smoothed Lagrangians L_m, and then we let $m \to +\infty$. In order to be able to pass to the limit in the estimates and find a solution, we need to assume the existence of two constants $c > 0$ and $a \in \left(\frac{d}{d+1}, 1\right)$ such that

$$U^-(t) := \max\{-U(t), 0\} \leq ct^a \qquad \forall\, t \geq 0. \tag{4.1}$$

The above assumption, together with (2.2) and (2.3), are satisfied by positive multiples of the following functions: $t \ln t$, or t^α with $\alpha > 1$. Under this additional assumption we now prove some lemmas and proposition which easily allow to construct our solution as a limit of solutions in bounded domains (cfr. subsection 4.2).

Thanks to assumption (4.1) we can prove that if $M_1(\varrho)$ is finite then $\mathcal{U}^-(\varrho) := \int_{\mathbb{R}^d} U^-(\varrho)\, dx$ is finite, and so $\mathcal{U}(\varrho) = \int_{\mathbb{R}^d} U(\varrho)\, dx$ is well-defined.

Lemma 4.1. *There exists $C = C(d, a)$ such that $\mathcal{U}^-(\varrho) \leq C(M_1(\varrho)^a + 1)$. Consequently $\mathcal{U}(\varrho)$ is well defined whenever $M_1(\varrho)$ is finite. Furthermore C can be chosen so that*

$$\int_{B_R(0)^c} U^-(\varrho)\, dx \leq C M_1(\varrho)^a R^{d(1-a)-a} \qquad \forall\, R \geq 0.$$

Proof: We use assumption (4.1) to obtain

$$\begin{aligned}\int_{B_R(0)^c} U^-(\varrho)\,dx &\le c\int_{B_R(0)^c}\varrho^a dx = c\int_{B_R(0)^c}(|x|\varrho)^a\frac{1}{|x|^a}dx\\ &\le c\Big(\int_{B_R(0)^c}|x|\varrho dx\Big)^a\Big(\int_{B_R(0)^c}|x|^{-a/(1-a)}dx\Big)^{1-a}\\ &\le cM_1(\varrho)^a\Big(\int_R^{+\infty} r^{\left(d-1-\frac{a}{1-a}\right)}\mathrm{d}r\Big)^{1-a} =: c_1(d,a)M_1(\varrho)^aR^{d(1-a)-a}. \qquad (4.2)\end{aligned}$$

This proves the second statement of the lemma. Observing that

$$\int_{B_R(0)} U^-(\varrho)\,dx \le c\int_{B_R(0)}\varrho^a dx \le c\int_{B_R(0)}(1+\varrho)\,dx \le c\big(\mathcal{L}^d(B_R(0))+1\big) = \tilde{c}R^d + c, \qquad (4.3)$$

we combine (4.2) and (4.3) to conclude the proof. □

We now prove a lower-semicontinuity result.

Proposition 4.2. *Suppose that $\{\varrho_n\}_{n\in\mathbb{N}} \subset \mathcal{P}_1^{ac}(\mathbb{R}^d)$ converges weakly in $L^1(\mathbb{R}^d)$ to ϱ, and that $\sup_{n\in\mathbb{N}} M_1(\varrho_n) < +\infty$. Then $\varrho \in \mathcal{P}_1^{ac}(\mathbb{R}^d)$ and $\liminf_{n\to\infty}\mathcal{U}(\varrho_n) \ge \mathcal{U}(\varrho)$.*

Proof: The fact that $\varrho \in \mathcal{P}_1^{ac}(\mathbb{R}^d)$ follows from the lower-semicontinuity with respect to the weak L^1-topology of the first moment.

We now suppose without loss of generality that $\liminf_{n\to\infty}\mathcal{U}(\varrho_n)$ is finite. Fix $\varepsilon > 0$. We have to prove that $\liminf_{n\to\infty}\mathcal{U}(\varrho_n) \ge \mathcal{U}(\varrho) - \varepsilon$. By lemma 4.1 we can find $R > 0$ such that

$$\sup_{n\in\mathbb{N}}\int_{B_R(0)^c} U^-(\varrho_n)\,dx \le \varepsilon. \qquad (4.4)$$

By lemma 2.6 and the fact that U and $U^+ \ge 0$ are convex we get

$$\liminf_{n\to\infty}\int_{B_R(0)} U(\varrho_n)\,dx \ge \int_{B_R(0)} U(\varrho)\,dx, \qquad \liminf_{n\to\infty}\int_{B_R(0)^c} U^+(\varrho_n)\,dx \ge \int_{B_R(0)^c} U^+(\varrho)\,dx, \qquad (4.5)$$

Combining equations (4.4) and (4.5) we obtain $\liminf_{n\to\infty}\mathcal{U}(\varrho_n) \ge \mathcal{U}(\varrho) - \varepsilon$. This concludes the proof. □

We remark that, thanks to the above results, minimizing $\Phi(h,\varrho_0,\cdot)$ over $\mathcal{P}_1^{ac}(X)$ still makes sense even when X is unbounded, provided that $\varrho_0 \in \mathcal{P}_1^{ac}(X)$.

4.1. **Properties on moments in the case X is unbounded.** Fix $T > 0$, and for any $h > 0$ suppose that we are given a sequence $\{\varrho_k^h\}_{0\le k\le T/h} \subset \mathcal{P}_1^{ac}$, not necessarily produced by any minimization procedure, such that

$$\mathcal{C}_h(\varrho_k^h,\varrho_{k+1}^h) + \mathcal{U}(\varrho_{k+1}^h) \le \mathcal{U}(\varrho_k^h). \qquad (4.6)$$

Assume that

$$m^*(1) := \sup_h\Big\{M_1(\varrho_0^h) + \int_{\mathbb{R}^d}|U(\varrho_0^h)|\,dx\Big\} < +\infty. \qquad (4.7)$$

For instance, if $\varrho_0^h = \varrho_0$ for all $h > 0$, equation (4.7) holds if $M_1(\varrho_0)$ and $\int_{\mathbb{R}^d}|U(\varrho_0)|\,dx$ are both finite.

Set $\mathcal{U}^+(\varrho) := \int_X U^+(\varrho)\,dx$, where $U^+(t) := \max\{U(t), 0\}$. By equations (2.11) and (4.6)

$$\mathcal{C}_{lh}(\varrho_0^h, \varrho_l^h) \le \sum_{k=0}^{l-1} \mathcal{C}_h(\varrho_k^h, \varrho_{k+1}^h) \le \mathcal{U}(\varrho_0^h) + \mathcal{U}^-(\varrho_l^h) - \mathcal{U}^+(\varrho_l^h), \tag{4.8}$$

which together with lemma 2.13(ii), implies

$$-A_* hl + W_{\theta,lh}(\varrho_0^h, \varrho_l^h) + \mathcal{U}^+(\varrho_l^h) \le \mathcal{U}(\varrho_0^h) + \mathcal{U}^-(\varrho_l^h). \tag{4.9}$$

Lemma 4.3. *If $\varrho, \bar{\varrho} \in \mathcal{P}_1^{ac}$, then*

$$M_1(\bar{\varrho}) \le [A_* + C(1)]h + \mathcal{C}_h(\varrho, \bar{\varrho}) + M_1(\varrho) \qquad \forall\, h > 0,$$

where $C(1)$ is the constant provided by lemma 2.13(iii).

Proof: We have

$$|y| \le |y - x| + |x|$$

so that integrating the above inequality with respect to $\gamma \in \Gamma(\varrho, \bar{\varrho})$ we obtain

$$M_1(\bar{\varrho}) \le \int_{\mathbb{R}^d \times \mathbb{R}^d} |y - x|\, d\gamma(x, y) + M_1(\varrho), \tag{4.10}$$

and since $\gamma \in \Gamma(\varrho, \bar{\varrho})$ is arbitrary we conclude that

$$M_1(\bar{\varrho}) \le W_1(\varrho, \bar{\varrho}) + M_1(\varrho).$$

This together with lemma 2.13(ii)-(iii) gives the desired estimate. □

The following proposition shows that $M_1(\varrho_k^h)$ is uniformly bounded for $kh \le T$, provided that it is bounded for $k = 0$.

Proposition 4.4. *Suppose (4.6) and (4.7) hold. Then there exists a constant $\bar{C}$, depending on $m^*(1)$ and T only, such that the following holds:*

$$M_1(\varrho_k^h) + \int_{\mathbb{R}^d} |U(\varrho_k^h)|\, dx \le \bar{C} \qquad \forall\, k, h, \text{ with } kh \le T.$$

Proof: We recall that by assumption $\varrho_k^h \in \mathcal{P}_1^{ac}$ for all k, h, so that $M_1(\varrho_k^h) < +\infty$. Suppose $kh \le T$. By lemma 4.3 and by equation (4.8)

$$M_1(\varrho_k^h) \le \mathcal{C}_{kh}(\varrho_0^h, \varrho_k^h) + [A_* + C(1)]hk + M_1(\varrho_0^h) \le \mathcal{U}(\varrho_0^h) - \mathcal{U}(\varrho_k^h) + [A_* + C(1)]hk + M_1(\varrho_0^h). \tag{4.11}$$

Let C be the constant provided by lemma 4.1. We use that lemma and equation (4.11) to obtain

$$M_1(\varrho_k^h) + \mathcal{U}^+(\varrho_k^h) \le \mathcal{U}^+(\varrho_0^h) + C\Big(1 + M_1^a(\varrho_k^h)\Big) + [A_* + C(1)]hk + M_1(\varrho_0^h). \tag{4.12}$$

Define for $t \ge 0$

$$f(t) := \sup_{m \ge 0}\{m :\ m - C(m^a + 1) \le t\}.$$

Observe that $f(t) \ge t$, and f is nondecreasing. Thus, recalling that $M_1(\varrho_k^h) < +\infty$, by equation (4.12) we get

$$M_1(\varrho_k^h) \le f\Big(\mathcal{U}^+(\varrho_0^h) + [A_* + C(1)]T + M_1(\varrho_0^h)\Big) := f_0 \tag{4.13}$$

and

$$\mathcal{U}^+(\varrho_k^h) \le \mathcal{U}^+(\varrho_0^h) + C\Big(1 + M_1^a(\varrho_k^h)\Big) + [A_* + C(1)]T + M_1(\varrho_0^h).$$

By lemma 4.1 and (4.13)

$$\mathcal{U}^-(\varrho_k^h) \le \tilde{C}\big(f_0^a + 1\big) \qquad \text{for } kh \le T,$$

where $\tilde{C}$ depend on C, T, $m^*(1)$, A_* and $C(1)$ only. This concludes the proof. □

Remark 4.5. It is easy to check that the estimates proved in this subsection depend on L only through the function θ and the constants A^*, A_*, C^* appearing in (L3). Hence such estimates are uniform if $\{L_m\}_{m\in\mathbb{N}}$ is a sequence of Lagrangians satisfying (L1), (L2) and (L3) with the same function θ and the same constants A^*, A_*, C^*.

4.2. **Existence of solutions.** In this paragraph we briefly sketch how to prove existence of solutions in the case when X is not necessarily bounded and L satisfies (L1^w), (L2^w) and (L3), leaving the details to the interested reader. We remark that our approximation argument could also be used to relax some of the assumptions on U.

Let $X \subset \mathbb{R}^d$ be an open set whose boundary has zero Lebesgue measure. We fix $\varrho_0 \in \mathcal{P}^{ac}(X)$, and we assume that $M_1(\varrho_0)$ and $\int_X |U(\varrho_0)|\,dx$ are both finite. Let us remark that under assumption (4.1) we have that $\int_X U^-(\varrho_0)\,dx$ is controlled by $M_1(\varrho)$ (cfr. lemma 4.1). Hence the finiteness of $M_1(\varrho_0)$ and $\int_X |U(\varrho_0)|\,dx$ is equivalent to assume that both $M_1(\varrho_0)$ and $\mathcal{U}(\varrho_0)$ are finite.

Assuming that L satisfies (L1^w), (L2^w) and (L3), we consider a sequence of Lagrangians $\{L_m\}_{m\in\mathbb{N}}$ converging to L in $C^1(\mathbb{R}^d \times \mathbb{R}^d)$ and which satisfy (L1), (L2) and (L3) with the same function θ as for L and constants A^*+1, A_*+1, C^*+1 (we slightly increase the constants of L to ensure that one can construct such a sequence). We denote by H_m the Hamiltonians associated to L_m. Consider now the increasing sequence of bounded sets X_m defined as

$$X_m := X \cap B_m(0),$$

and observe that, for each $m \in \mathbb{N}$, the set X_m is open and its boundary has zero Lebesgue measure (since $\partial X_m \subset \partial X \cup \partial B_m(0)$). We now apply the variational scheme in X_m starting from $\varrho_0^m := \varrho_0 \chi_{B_m(0)}$ (cfr. remark 3.18) with Lagrangian L_m. In this way we construct approximate discrete solutions $\{\varrho_{kh}^{h,m}\}$ on X_m which satisfy the discrete energy inequality

$$\mathcal{U}(\rho_0^m) - \mathcal{U}(\rho_{(k+1)h}^{h,m}) \ge \int_0^T \int_{\mathbb{R}^d} \Big[L_m(x, V_t^{h,m})\bar{\varrho}_t^{h,m} + H_m\big(x, -\nabla[U'(\varrho_t^{h,m})]\big)\varrho_t^{h,m}\Big]\,dx\,dt - A_* h.$$

Moreover, thanks to proposition 4.4 (cfr. remark 4.5) the measures $\{\varrho_{kh}^{h,m}\}$ have uniformly bounded first moments for all k, h, m, with $kh \le T$. This fact together with lemma 4.1 implies that also $\mathcal{U}^-(\varrho_{kh}^{h,m})$ is uniformly bounded. Therefore, taking the limit as $h \to 0$ (cfr. subsection 3.5) we obtain a family of curves $t \mapsto \varrho_t^m$ satisfying the energy bound (3.43) and such that

$$\sup_{m\in\mathbb{N},\, t\in[0,T]} \Big\{M_1(\varrho_t^m) + \int_{\mathbb{R}^d} |U(\varrho_t^m)|\,dx\Big\} < +\infty.$$

(Indeed $\mathcal{U}^-(\varrho_t^m)$ are uniformly bounded, and $t \mapsto \mathcal{U}(\varrho_t^m)$ is bounded too, cfr. equation (3.33).) Moreover

$$\partial_t \varrho_t^m + \operatorname{div}(\rho_t^m V_t^m) = 0 \qquad \text{on } (0,T) \times \mathbb{R}^d,$$

with

$$\sup_{m\in\mathbb{N}} \int_0^T \int_{\mathbb{R}^d} \theta(|V_t^m|)\varrho_t^m\,dx\,dt < +\infty$$

(by equation (3.43)), which implies a uniform continuity in time of the curves $[0,T] \ni t \mapsto \varrho_t^m$. Thanks to these bounds, it is not difficult to take the limit as $m \to +\infty$ (cfr. the arguments in subsection 3.5) and find a uniformly continuous curve $t \mapsto \varrho_t$ which satisfies

$$\partial_t \varrho_t + \operatorname{div}(\varrho_t V_t) = 0 \qquad \text{on } (0,T) \times \mathbb{R}^d$$

in the sense of distributions and

$$\mathcal{U}(\rho_0) - \mathcal{U}(\rho_T) \geq \int_0^T \int_{\mathbb{R}^d} \Big[L(x, V_t)\varrho_t + H\big(x, -\nabla[U'(\varrho_t)]\big)\varrho_t \Big]\, dx\, dt.$$

(Here we used that $\mathcal{U}(\rho_0^m) \to \mathcal{U}(\rho_0)$ and proposition 4.2.) Once this estimate is established, the proof of (3.45) is the same as in the bounded case. Hence we obtain:

Theorem 4.6. *Let $X \subset \mathbb{R}^d$ be an open set whose boundary is of zero Lebesgue measure, and assume that H satisfies (H1^w), (H2^w) and (H3). Assume the U satisfies* (2.2), (2.3) *and* (4.1), *and let $\varrho_0 \in \mathcal{P}_1^{ac}(X)$ be such that $M_1(\varrho_0)$ and $\mathcal{U}(\varrho_0)$ are both finite. Then there exists a narrowly continuous curve $t \mapsto \varrho_t \in \mathcal{P}_1^{ac}(X)$ on $[0,T]$ starting from ϱ_0 and a Borel time-dependent vector field V_t on $\mathbb{R}^d$ such that $M_1(\varrho_t)$ is bounded on $[0,T]$ (so that in particular $\mathcal{U}^-(\rho_t)$ is bounded), the continuity equation*

$$\partial_t \varrho_t + \operatorname{div}(\varrho_t V_t) = 0 \qquad \text{on } (0,T) \times \mathbb{R}^d$$

holds in the sense of distributions, and (3.43) *holds true. Moreover $\nabla[P(\varrho_t)] \in L^1(0,T; L^1(X))$ and $\nabla[P(\varrho_t)]$ is absolutely continuous with respect to ϱ_t.*

Furthermore, if $\theta(t) \sim t^\alpha$ for some $\alpha > 1$ and U satisfies the doubling condition (3.44) then ϱ_t is a solution of (1.1) *starting from ϱ_0, that is*

$$V_t(x) = \nabla_p H\big(x, -\varrho_t^{-1}(x)\nabla[P(\varrho_t(x))]\big) \qquad \varrho_t\text{-}a.e.$$

Moreover (1.2) (or equivalently (3.45)) holds. Finally, if (H4) holds and $\varrho_0 \leq M$ for some $M \geq 0$, then $\varrho_t \leq M$ for all $t \in [0,T]$ (maximum principle).

Remark 4.7. When $\theta(t) \sim t^\alpha$ with $\alpha > 1$, it is not difficult to see that if $\int_X |x|^\alpha \varrho_0\, dx$ is finite so is $\int_X |x|^\alpha \varrho_t\, dx$ (here ϱ_t is any limit curve constructed using the minimizing movement scheme). Hence one can generalize lemma 4.1 proving that the α-moment of ϱ controls $\mathcal{U}^-(\rho)$ assuming only that condition (4.1) holds for some $a \in (\frac{d}{d+\alpha}, 1)$, and the above theorem still holds under this weaker assumption on U. In particular if $\varrho_0 \in \mathcal{P}_\alpha(X)$ then $\varrho \in AC_\alpha(0,T;\mathcal{P}_\alpha(X))$.

Remark 4.8 (Extension to manifolds). The above existence theorem can be easily extended to Riemannian manifolds. Indeed in the compact case the proof is more or less exactly the same, while in the noncompact case one has to replace the first moment by $\int_X d(x, x_0)\varrho_t\, d\mathrm{vol}(x)$, where x_0 is any (fixed) point in M, d denotes the Riemannian distance, and vol is the volume measure.

5. Uniqueness of solutions

Throughout this section we assume that H satisfies (H1), (H2^w), (H3) and (H4). We further assume that U satisfies (2.2) and (2.3), $X \subset \mathbb{R}^d$ is an open set whose boundary is of zero Lebesgue measure, and we denote by $\overline{X}$ its closure. We suppose that either X is bounded or X is unbounded but condition (4.1) holds. We suppose that $\theta(t) \sim t^\alpha$ for some $\alpha > 1$ and U satisfies the doubling condition (3.44). Our goal is to prove uniqueness of distributional solutions of equation (1.1) when the initial condition ϱ_0 is bounded. The ellipticity conditions we impose seem to be different from what is usually imposed in the literature. Our proof of uniqueness of solution follows the same line as that of [15], except that most of our assumptions are not always comparable with the ones there. In the sequel

$$\Omega := (0,T) \times X, \qquad \tilde{\Omega} := (0,T) \times \Omega.$$

5.1. **A new Hamiltonian.** We consider the density function ϱ_t of equation (1.1) provided by theorem 4.6, which satisfies the property that $\nabla[P(\varrho_t)] \in L^1(\Omega)$ and is absolutely continuous with respect to ϱ_t. If we set $u(t,\cdot) := P(\varrho_t)$ we have

$$\partial_t Q(u) = \operatorname{div} \mathbf{a}(x, Q(u), \nabla u) \quad \text{in} \quad \mathcal{D}'(\Omega), \tag{5.1}$$

where

$$\mathbf{a}(x,s,m) := \begin{cases} -\nabla_m \bar{H}(x,s,-m) & \text{if } s > 0 \\ 0 & \text{if } s = 0, m = \vec{0}, \end{cases} \tag{5.2}$$

and $\bar{H} : \mathbb{R}^d \times [0,+\infty) \times \mathbb{R}^d \to [0,+\infty)$ is defined by

$$\bar{H}(x,s,m) := \begin{cases} s^2 H(x, \frac{m}{s}) & \text{if } s > 0 \\ 0 & \text{if } s = 0, m = \vec{0} \\ +\infty & \text{if } s = 0, m \neq \vec{0} \end{cases} \tag{5.3}$$

Here, $\vec{0} := (0,\dots,0)$.

For each $x \in \mathbb{R}^d$, $\bar{H}(x,\cdot,\cdot)$ is of class $C^2((0,+\infty)\times\mathbb{R}^d)$, and the gradient of $\bar{H}(x,\cdot,\cdot)$ at (s,m) is given by

$$\nabla \bar{H}(x,s,m) = \begin{pmatrix} 2sH(x,\frac{m}{s}) - s\big\langle \nabla_p H(x,\frac{m}{s}), \frac{m}{s}\big\rangle \\ s\nabla_p H(x,\frac{m}{s}) \end{pmatrix}$$

for $s > 0$ and $m \in \mathbb{R}^d$. Observe that

$$\nabla^2 \bar{H}(x,\cdot,\cdot) = \begin{pmatrix} 2H - 2\big\langle \nabla_p H, \frac{m}{s}\big\rangle + \langle \nabla_{pp} H \cdot \frac{m}{s}, \frac{m}{s}\rangle & \nabla_p H - \nabla_{pp} H \cdot \frac{m}{s} \\ \nabla_p H - \nabla_{pp} H \cdot \frac{m}{s} & \nabla_{pp} H \end{pmatrix}. \tag{5.4}$$

Here $H, \nabla_p H, \nabla_{pp} H$ are all evaluated at $\big(x, \frac{m}{s}\big)$. Since $H(x,\cdot)$ is convex we have that

$$\langle \nabla^2 \bar{H}(x,\cdot,\cdot)(0,\lambda),(0,\lambda)\rangle = \langle \nabla_{pp} H \cdot \lambda, \lambda\rangle \geq 0.$$

for $\lambda \in \mathbb{R}^d$. Hence, the matrix in equation (5.4) is nonnegative definite if and only if for every $\lambda \in \mathbb{R}^d$

$$\begin{aligned} 0 &\leq \langle \nabla^2 \bar{H}(x,\cdot,\cdot)(1,\lambda),(1,\lambda)\rangle \\ &= 2H - \langle \nabla_p H, \frac{m}{s}\rangle + \langle \nabla_{pp} H \cdot \frac{m}{s}, \frac{m}{s}\rangle + 2\langle \nabla_p H, \lambda\rangle - 2\langle \nabla_{pp} H \cdot \frac{m}{s}, \lambda\rangle + \langle \nabla_{pp} H \cdot \lambda, \lambda\rangle \\ &= 2H - 2\big\langle \nabla_p H, \lambda - \frac{m}{s}\big\rangle + \langle \nabla_{pp} H \cdot \big(\lambda - \frac{m}{s}\big), \lambda - \frac{m}{s}\rangle. \end{aligned}$$

Equivalently, $\bar{H}(x,\cdot,\cdot)$ is convex on $(0,+\infty)\times\mathbb{R}^d$ if and only if

$$2H - 2\langle \nabla_p H, w\rangle + \langle \nabla_{pp} H \cdot w, w\rangle \geq 0 \qquad \forall\, w \in \mathbb{R}^d. \tag{5.5}$$

This is what we assume in the sequel.

Remark 5.1. $H(x,p) = |p|^r$ satisfies condition (5.5) if and only if $r \geq 2$. If $A(x)$ is a symmetric non-negative definite matrix then $H(x,p) = \langle A(x)p,p\rangle$ satisfies condition (5.5). Moreover, by linearity, if H_1 and H_2 satisfy condition (5.5) so does $H_1 + H_2$.

Remark 5.2. Suppose assumption (5.5) holds.

(a) Since $\bar H \geq 0$ we have that $(0, \vec 0)$ belongs to the subdifferential of $\bar H(x, \cdot, \cdot)$ at $(0, \vec 0)$. In other words $-\mathbf{a}(x, 0, \vec 0)$ belongs to the subdifferential of $\bar H(x, \cdot, \cdot)$ at $(0, \vec 0)$.

(b) The convexity of $\bar H(x, \cdot, \cdot)$ is equivalent to

$$\langle \mathbf{a}(x, s_1, m_1) - \mathbf{a}(x, s_2, m_2), m_1 - m_2 \rangle \geq -(s_1 - s_2)\Big\{ 2\Big(s_1 H\big(x, -\frac{m_1}{s_1}\big) - s_2 H\big(x, -\frac{m_2}{s_2}\big)\Big) + \big\langle \nabla_p H\big(x, -\frac{m_1}{s_1}\big), m_1 \big\rangle - \big\langle \nabla_p H\big(x, -\frac{m_2}{s_2}\big), m_2 \big\rangle \Big\}$$

5.2. **Additional properties satisfied by bounded solutions.** We assume that (5.5) holds. Let $\varrho_t \in AC_1(0, T; \mathcal{P}_1^{ac}(X))$ be a solution of equation (1.1) satisfying (1.2) such that $t \mapsto \mathcal{U}(\varrho_t)$ is absolutely continuous, monotone nonincreasing, and $\nabla[P(\varrho_t)] \in L^1(\Omega)$ and is absolutely continuous with respect to ϱ_t. Observe that ϱ_t satisfies in fact equation (3.43) and the inequality there becomes an equality. Suppose there exists a constant $M > 0$ such that $\varrho_t \leq M$. Because $\theta(t) \sim t^\alpha$, (H3) implies that for $\bar c > 0$ sufficiently small

$$\bar c \Big(\big| \varrho_t^{-1} \nabla[P(\varrho)] \big|^{\alpha'} - 1 \Big) \leq H(x, -\varrho_t^{-1} \nabla[P(\varrho_t)]),$$

so that multiplying both sides of the above inequality by ϱ_t we deduce

$$\bar c \Big(M^{1-\alpha'} \big| \nabla[P(\varrho_t)] \big|^{\alpha'} - \varrho_t \Big) \leq \varrho_t H(x, -\varrho_t^{-1} \nabla[P(\varrho_t)]) \tag{5.6}$$

Taking $\bar c > 0$ small enough, (L3) ensures

$$\bar c (\varrho_t |V_t|^\alpha - \varrho_t) \leq \varrho_t L(x, V_t), \qquad \bar c |\varrho_t V_t|^\alpha \leq M^{\alpha - 1} \varrho_t \Big(\bar c + L(x, V_t) \Big). \tag{5.7}$$

Using the fact that equality holds in equation (3.43) and exploiting equations (5.6) and (5.7), it is easy to show that existence of a constant C_M, which depends only on M and θ, such that

$$\int_0^T dt \int_X \Big| \nabla[P(\varrho_t)] \Big|^{\alpha'} dx, \quad \int_0^T dt \int_X \varrho |V_t|^\alpha \varrho_t \, dx, \quad \int_0^T \int_X |\varrho_t V_t|^\alpha \varrho_t dx \leq C_M \tag{5.8}$$

where $V_t := \nabla_p H\big(x, -\varrho_t^{-1} \nabla[P(\varrho_t)]\big)$. Also, choosing C_M large enough and using (L3), (H3) and equation (5.8), we have

$$\int_0^T dt \int_X \varrho_t \big| H(x, -\varrho_t^{-1} \nabla[P(\varrho_t)]) \big| \, dx, \quad \int_0^T dt \int_X \varrho_t \big| L(x, V_t) \big| \, dx \leq C_M. \tag{5.9}$$

Remark 5.3. Since $\varrho_t \in AC_1(0, T; \mathcal{P}_1^{ac}(X))$, U is strictly convex, and $t \mapsto \mathcal{U}(\varrho_t)$ is absolutely continuous, we have $\varrho_t \in C([0, T]; L^1(X))$.

Observe that by equation (5.8) we have that $u(t, \cdot) = P(\varrho_t)$ satisfies $\nabla u \in L^{\alpha'}(\Omega)$, while the last inequality in (5.8) gives $\mathbf{a}(\cdot, Q(u), \nabla u) \in L^\alpha(\Omega)$. Since ϱ_t satisfies equation (1.1), by an approximation argument and thanks to remark 5.3 we have

$$\int_\Omega Q(u) \partial_t \mathcal{E} = \int_\Omega \langle \mathbf{a}(x, Q(u), \nabla u), \nabla \mathcal{E} \rangle \tag{5.10}$$

for any $\mathcal{E} \in W^{1,\alpha'}(\Omega)$ such that $\mathcal{E}(t, \cdot) \equiv 0$ for t near 0 and T.

As in [15], for $\eta \in C^2(\mathbb{R})$ convex monotone nondecreasing such that η' and η'' are bounded we define $q_\eta, \eta_* : \mathbb{R}^2 \to \mathbb{R}$ by

$$q_\eta(z, z^o) := \int_{z^o}^{z} \eta'(s - z^o) Q'(s)\, ds, \quad z, z^o \in \mathbb{R},$$

$$\eta_*(w, z^o) := \sup_{z \in \mathbb{R}} \{\eta'(z - z^o)(w - Q(z)) + q_\eta(z, z^o)\}, \quad w, z^o \in \mathbb{R}.$$

Lemma 5.4. *Suppose $v^o \in W^{1,\alpha'}(X) \cap L^\infty(X)$ and $\gamma \in C_c^\infty((0,T) \times \mathbb{R}^d)$ is nonnegative. Then*

$$\int_\Omega -q_\eta(u, v^o)\partial_t \gamma + \langle \mathbf{a}(x, Q(u), \nabla u), \nabla[\eta'(u - v^o)\gamma] \rangle \le 0. \tag{5.11}$$

Proof: The proof is identical to that of [15, Lemma 1]. □

5.3. **Uniqueness of bounded solutions.** In this subsection, for $i = 1, 2$, we consider $\varrho_t^i \in AC_1(0, T; \mathcal{P}_1^{ac}(X))$ solutions of equation (1.1) satisfying (1.2) and such that $t \mapsto \mathcal{U}(\varrho_t^i)$ is absolutely continuous and monotone nonincreasing. We impose that $\nabla[P(\varrho_t^i)] \in L^1(\Omega)$ is absolutely continuous with respect to ϱ_t^i. We further assume existence of a constant $M > 0$ such that $\varrho_t^i \le M$. The goal of the subsection is to show that

$$t \mapsto \int_X |\varrho_t^1 - \varrho_t^2| dx \quad \text{is monotone nondecreasing.}$$

Once such an estimate is proved, it extends immediately to solutions whose initial datum belongs to L^1 and has bounded first moment, and which are constructed by approximation (cfr. section 4) as a limit of solutions with bounded initial data. We are neither claiming any uniqueness result in a more general setting nor we are claiming to be able to provide optimal conditions under which uniqueness fail.

We define u_1, u_2 on $\tilde{\Omega}$ by

$$u_1(t_1, t_2, x) := P(\varrho^1(t_1, x)), \qquad u_2(t_1, t_2, x) := P(\varrho^2(t_2, x)).$$

If $r \in \mathbb{R}$ we set $r^+ = \max\{0, r\}$ and $r^- = \max\{0, r\}$.

To achieve the main goal of this subsection, we first prove a lemma whose proof is more or less a repetition of the arguments presented on [15, pages 31-33]. Since $\mathbf{a}$ does not satisfy the assumptions imposed in that paper, we felt the need to show that the arguments there go through.

Lemma 5.5. *If $\min_{[0,M]} P' > 0$ and $\tilde{\gamma} \in C_c^\infty((0,T)^2)$ is nonnegative, then*

$$-\int_{\tilde{\Omega}} \big(Q(u_1) - Q(u_2)\big)^+ (\partial_{t_1}\tilde{\gamma} + \partial_{t_2}\tilde{\gamma}) \le 0. \tag{5.12}$$

Proof: Let $f_n \in C_c^\infty(\mathbb{R}^d)$ be such that $0 \le f_n \le 1$, $f_n(x) = 1$ for $|x| \le n$, $f_n(x) = 0$ for $|x| \ge n+2$, and $|\nabla f_n| \le 1$. Let $\eta \in C^2(\mathbb{R})$ be a convex nonnegative function such that $\eta(z) = 0$ for $z \le 0$, $\eta(z) = z - 1/2$ for $z \ge 1$. Set

$$\eta_\delta^+(z) := \delta\eta\left(\frac{z}{\delta}\right), \quad \eta_\delta^-(z) := \delta\eta\left(-\frac{z}{\delta}\right), \quad q_\delta^\pm := q_{\eta_\delta^\pm},$$

so that

$$(\eta_\delta^-)'(z) = -(\eta_\delta^+)'(-z). \tag{5.13}$$

We fix t_2 and apply lemma 5.4 to

$$v^o = u_2(\cdot, t_2, \cdot) \equiv u_2(t_2, \cdot), \quad \eta = \eta_\delta^+, \quad \gamma = \tilde{\gamma}(\cdot, t_2) f_n.$$

Then, we integrate the subsequent inequality with respect to t_2 over $(0, T)$ to obtain

$$\int_{\tilde{\Omega}} -q_\delta^+(u_1, u_2)\partial_{t_1}(\tilde{\gamma} f_n) + \Big\langle \mathbf{a}(x, Q(u_1), \nabla u_1), \nabla[(\eta_\delta^+)'(u_1 - u_2)\tilde{\gamma} f_n] \Big\rangle \leq 0. \tag{5.14}$$

Similarly,

$$\int_{\tilde{\Omega}} -q_\delta^-(u_2, u_1)\partial_{t_2}(\tilde{\gamma} f_n) + \Big\langle \mathbf{a}(x, Q(u_2), \nabla u_2), \nabla[(\eta_\delta^-)'(u_2 - u_1)\tilde{\gamma} f_n] \Big\rangle \leq 0. \tag{5.15}$$

We exploit equations (5.13), (5.14) and (5.15) to obtain

$$\int_{\tilde{\Omega}} \tilde{\gamma} \Big\langle \mathbf{a}(x, Q(u_1), \nabla u_1) - \mathbf{a}(x, Q(u_2), \nabla u_2), (\nabla u_1 - \nabla u_2)(\eta_\delta^+)''(u_1 - u_2) f_n + (\eta_\delta^+)'(u_1 - u_2)\nabla f_n \Big\rangle$$
$$\leq \int_{\tilde{\Omega}} \Big(q_\delta^+(u_1, u_2)\partial_{t_1}\tilde{\gamma} + q_\delta^-(u_2, u_1)\partial_{t_2}\tilde{\gamma} \Big) f_n$$

This together with remark 5.2(b) yields

$$\int_{\tilde{\Omega}} f_n (\eta_\delta^+)''(u_1 - u_2)(Q(u_2) - Q(u_1))(E_1 - E_2)\tilde{\gamma} + \int_{\tilde{\Omega}} R_n^1 \leq \int_{\tilde{\Omega}} (q_\delta^+(u_1, u_2)\partial_{t_1}\tilde{\gamma} + q_\delta^-(u_2, u_1)\partial_{t_2}\tilde{\gamma}) f_n, \tag{5.16}$$

where

$$R_n^1 := \tilde{\gamma} \Big\langle \mathbf{a}(x, Q(u_1), \nabla u_1) - \mathbf{a}(x, Q(u_2), \nabla u_2), (\eta_\delta^+)'(u_1 - u_2)\nabla f_n \Big\rangle.$$

$$E_i(t_1, t_2, x) := \varrho^i(t_i, x)\Big(2H(x, -e^i(t_i, x)) + \Big\langle \nabla_p H\big(x, -e^i(t_i, x)\big), e^i(t_i, x) \Big\rangle \Big), \qquad i = 1, 2,$$

with $\varrho^i(t_i, x)e^i(t_i, x) := \nabla[P(\varrho^i)](t_i, x)$. We observe that, thanks to 5.9, it is not difficult to show that $E_1, E_2 \in L^1(\tilde{\Omega})$.

Now, the second inequality in (5.8) gives

$$V_t^1 := \nabla_p H\big(x, -(\varrho_t^1)^{-1}\nabla[P(\varrho_t^1)]\big) \in L^\alpha(\varrho_t^1) \subset L^1(\varrho_t^1),$$

and so $\mathbf{a}(x, Q(u_1), \nabla u_1) \in L^1(\tilde{\Omega})$. Similarly $\mathbf{a}(x, Q(u_2), \nabla u_2) \in L^1(\tilde{\Omega})$. Hence

$$|R_n^1| \leq A^1 |\nabla f_n| \leq A^1$$

where $A^1 \in L^1(\tilde{\Omega})$. Since $|\nabla f_n| \to 0$ as $n \to \infty$, we use the dominated convergence theorem to conclude that $\int_{\tilde{\Omega}} T_n^1 \to 0$ as $n \to \infty$. Since u_1 and u_2 are bounded, we may apply the Lebesgue dominated convergence theorem to the first term in the left hand side of (5.16) and to the right hand side, to conclude that

$$-\int_{\tilde{\Omega}} (\eta_\delta^+)''(u_1 - u_2)|Q(u_1) - Q(u_2)||E_1 - E_2|\tilde{\gamma} \leq \int_{\tilde{\Omega}} \Big(q_\delta^+(u_1, u_2)\partial_{t_1}\tilde{\gamma} + q_\delta^-(u_2, u_1)\partial_{t_2}\tilde{\gamma} \Big). \tag{5.17}$$

Recall that u_1 and u_2 have their ranges contained in the compact set $[0, P(M)]$. Moreover, since $\min_{[0,M]} P' > 0$, Q is Lipschitz on $[0, P(M)]$ for some Lipschitz constant $\bar{C}_M$. Then (5.17) gives

$$-\bar{C}_M \int_{\tilde{\Omega}} (\eta_\delta^+)''(u_1 - u_2)|u_1 - u_2||E_1 - E_2|\tilde{\gamma} \leq \int_{\tilde{\Omega}} \Big(q_\delta^+(u_1, u_2)\partial_{t_1}\tilde{\gamma} + q_\delta^-(u_2, u_1)\partial_{t_2}\tilde{\gamma} \Big). \tag{5.18}$$

Recalling that

$$|q_\delta^\pm(z,z^o)| \le (Q(z)-Q(z^o))^\pm, \quad |(\eta_\delta^+)'(z)| \le z^+, \quad |z(\eta_\delta^+)''(z)| \le \sup_{a\in\mathbb{R}} |a\eta''(a)|, \tag{5.19}$$

and that, as $\delta \to 0^+$,

$$q_\delta^\pm(z,z^o) \to (Q(z)-Q(z^o))^\pm, \quad (\eta_\delta^+)'(z) \to z^+, \quad z(\eta_\delta^+)''(z) \to 0, \tag{5.20}$$

we conclude the proof of the lemma by combining (5.18), (5.19) and (5.20). □

Theorem 5.6. *Suppose H satisfies (H1), (H2^w), (H3) and (H4). Suppose U satisfies* (2.2), (2.3) *and the doubling condition* (3.44). *Assume* $\min_{[0,M]} P' > 0$ *for any $M > 0$, $X \subset \mathbb{R}^d$ is an open set whose boundary is of zero Lebesgue measure, and $\theta(t) \sim t^\alpha$ with $\alpha > 1$. Suppose for $i = 1, 2$ that $\varrho_t^i \in AC_\alpha(0,T;\mathcal{P}_\alpha^{ac}(X))$ are solutions of equation (1.1) satisfying (1.2). Assume further that $t \mapsto \mathcal{U}(\varrho_t^i)$ is absolutely continuous, monotone nonincreasing, and $\nabla[P(\varrho_t^i)] \in L^1(\Omega)$ and is absolutely continuous with respect to ϱ_t^i. If ϱ_0^1, ϱ_0^2 are bounded then $t \mapsto \int_X |\varrho_t^1 - \varrho_t^2| dx$ is monotone nondecreasing.*

Proof: As shown in [15] this theorem is a direct consequence of equation (5.12). □

Acknowledgments The collaboration on this manuscript started during Spring 2008 while the three authors were visiting IPAM–UCLA, whose financial support and hospitality are gratefully acknowledged. WG gratefully acknowledges the support provided by NSF grants DMS-03-54729 and DMS-06-00791. TY gratefully acknowledges RAs support provided by NSF grants DMS-03-54729 and DMS-06-00791. It is a pleasure to express our gratitude to G. Savaré for fruitful discussions.

References

[1] M. Agueh. Existence of solutions to degenerate parabolic equations via the Monge-Kantorovich theory. *Adv. Differential Equations,* **10** no.3 309-360, 2005.
[2] H-W. Alt and S. Luckhaus. Quasilinear elliptic–parabolic differential equations. *Math. Z.,* **183** 311-341, 1983.
[3] L. Ambrosio, N. Gigli and G. Savaré. Gradient flows in metric spaces and the Wasserstein spaces of probability measures. Lectures in Mathematics, ETH Zurich, Birkhäuser, 2005.
[4] P. Bernard and B. Buffoni. Optimal mass transportation and Mather theory. *J. Eur. Math. Soc.,* **9** no.1 85-121, 2007,.
[5] B. Dacorogna. Direct methods in the calculus of variations. Springer-Verlag, Berlin, 1989.
[6] E. De Giorgi.New problems on minimizing movements. In Boundary Value Problems for PDE and Applications, C. Baiocchi and J.L. Lions, eds., Masson 81-98, 1993.
[7] E. De Giorgi, A. Mariono and M. Tosques. Problems of evolutions in metric spaces and maximal decreasing curve. *Atti. Accad. Naz. Lincei Rend. Cl. Sci. Mat. Natur. (8)*, **68** 180-187, 1980.
[8] A. Fathi and A. Figalli. Optimal transportation on non-compact manifolds. *Israel J. Math.*, **175** no.1 1-59, 2010.
[9] A. Figalli. Existence, uniqueness, and regularity of optimal transport maps. *SIAM J. Math. Anal.,* **39** no.1 126-137, 2007.
[10] W. Gangbo. Quelques problèmes d'analyse convexe. *Rapport d'habilitation à diriger des recherches*, Jan. 1995. Available at `http://www.math.gatech.edu/~gangbo/publications/`).
[11] W. Gangbo and R. McCann. Optimal maps in Monge's mass transport problem. *C. R. Acad. Sci. Paris Série I Math.*, **321** 1653-1658, 1995.
[12] W. Gangbo and R. McCann. The geometry of optimal transportation. *Acta Math.*, **177** 113-161, 1996.
[13] R. Jordan, D. Kinderlehrer and F. Otto The variational formulation of the Fokker–Planck equation. *SIAM J. Math Anal.*, **29** 1-17, 1998.
[14] S. Lisini. Nonlinear diffusion equations with variable coefficients as gradient flows in Wasserstein spaces. *ESAIM Control Optim. Calc. Var.*, **15**, no. 3 712-740, 2009.

[15] F. Otto. L^1–Contraction and Uniqueness for Quasilinear Elliptic–Parabolic Equations. *J. Differential Equations*, **131** 20-38, 1996.
[16] C. Villani. Topics in optimal transportation. Graduate Studies in Mathematics **58**, American Mathematical Society, 2003.
[17] C. Villani. Optimal transport, old and new. Grundlehren der mathematischen Wissenschaften.Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 338. Springer-Verlag, Berlin, 2009.
[18] T. Yolcu. Parabolic Systems with an underlying Lagrangian. *PhD Dissertation, Georgia Tech.*, 2009.

Departement of Mathematics, The University of Texas at Austin, Austin TX, USA
E-mail address: figalli@math.utexas.edu

Georgia Institute of Technology, Atlanta GA, USA
E-mail address: gangbo@math.gatech.edu

Georgia Institute of Tecnology, Atlanta GA, USA
E-mail address: tyolcu@math.gatech.edu