On a class of modified Wasserstein distances induced by concave mobility functions defined on bounded intervals

Stefano Lisini,* Antonio Marigonda†

Abstract

We study a new class of distances between Radon measures similar to those studied in [13]. These distances (more correctly pseudo-distances because can assume the value $+\infty$) are defined generalizing the dynamical formulation of the Wasserstein distance by means of a concave mobility function. We are mainly interested in the physical interesting case (not considered in [13]) of a concave mobility function defined in a bounded interval. We state the basic properties of the space of measures endowed with this pseudo-distance. Finally, we study in detail two cases: the set of measures defined in $\mathbb{R}^d$ with finite moments and the set of measures defined in a bounded convex set. In the two cases we give sufficient conditions for the convergence of sequences with respect to the distance and we prove a property of boundedness.

KEYWORDS: generalized Wasserstein distance, mobility function.

2000 Mathematics Subject Classification: 49J27, 49J40.

1 Introduction

In [13], Dolbeault, Nazaret and Savaré introduce and study the basic properties of a new class of distances between non-negative Radon measures on $\mathbb{R}^d$. These distances are defined generalizing the dynamical characterization of the Wasserstein distance. We briefly recall that the Wasserstein distance between two non-negative measures with the same mass can be defined as a relaxed optimal transportation problem (see [26], [2], [27] for a reference on this interesting topic)

$$W_p(\mu_0,\mu_1) := \inf\left\{ \left(\int_{\mathbb{R}^d\times\mathbb{R}^d} |y-x|^p \,\mathrm{d}\Sigma \right)^{\frac{1}{p}} : \ \Sigma \in \Gamma(\mu_0,\mu_1) \right\} \tag{1}$$

where $\Gamma(\mu_0,\mu_1)$ is the set of all *transport plans* between μ_0 and μ_1: they are non-negative measures Σ on $\mathbb{R}^d\times\mathbb{R}^d$ with the same mass of μ_0 and μ_1 whose first and second marginals are respectively μ_0 and μ_1, i.e. $\Sigma(B\times\mathbb{R}^d) = \mu_0(B)$ and $\Sigma(\mathbb{R}^d\times B) = \mu_1(B)$ for all Borel set B of $\mathbb{R}^d$.

In [4], Benamou and Brenier prove that the Wasserstein distance defined in (1) can be characterized, for absolutely continuous measures with respect to the Lebesgue measure $\mathscr{L}^d$, with compactly supported

*Address: Dipartimento di Matematica "F. Casorati", Università degli Studi di Pavia, Via Ferrata 1, 27100 Pavia, Italy, e-mail: stefano.lisini@unipv.it

†Address: Dipartimento di Informatica, Università degli Studi di Verona, Strada Le Grazie 15, 37134 Verona, Italy, e-mail: antonio.marigonda@univr.it

smooth densities, as follows

$$
\begin{aligned}
W_p^p(\mu_0,\mu_1) = \inf\Big\{ &\int_0^1 \int_{\mathbb{R}^d} |\boldsymbol{v}_t(x)|^p \rho_t(x)\,\mathrm{d}x\,\mathrm{d}t \ : \\
&\partial_t \rho_t + \nabla\cdot(\rho_t \boldsymbol{v}_t) = 0 \text{ in } \mathbb{R}^d\times(0,1), \quad \mu_0 = \rho_{|t=0}\mathscr{L}^d, \quad \mu_1 = \rho_{|t=1}\mathscr{L}^d \Big\}.
\end{aligned} \tag{2}
$$

The proof of the dynamical characterization for general non-negative Borel measures was given in [2] where the continuity equation in (2) was considered in distributional sense.

The generalization of (2) studied in [13], roughly speaking, replace the mobility coefficient ρ in (2) with a non-linear one $h(\rho)$, where $h : [0,+\infty) \to [0,+\infty)$ is a concave increasing function such that $h(0) = 0$ (particularly important examples are the functions $h(\rho) = \rho^\alpha$, $\alpha \geq 0$) and the new "distance" is defined modifying (2) as follows

$$
\begin{aligned}
W_{p,h}^p(\mu_0,\mu_1) = \inf\Big\{ &\int_0^1 \int_{\mathbb{R}^d} |\boldsymbol{v}_t(x)|^p h\big(\rho_t(x)\big)\,\mathrm{d}x\,\mathrm{d}t : \\
&\partial_t \rho_t + \nabla\cdot(h(\rho_t)\,\boldsymbol{v}_t) = 0 \text{ in } \mathbb{R}^d\times(0,1), \quad \mu_0 = \rho_{|t=0}\mathscr{L}^d, \quad \mu_1 = \rho_{|t=1}\mathscr{L}^d \Big\}.
\end{aligned} \tag{3}
$$

This "definition" is not rightly stated because it is necessary to specify the spaces where ρ and $\boldsymbol{v}$ has to belong, and the notion of solution of the modified continuity equation in (3). The right framework is that of Radon measures and distributional solutions.

The motivation for studying distances defined like in (3) arises from physical problems. Indeed many interesting models are described by partial differential equations whose solutions can be seen as trajectory of the gradient flow of a suitable energy functional with respect to this distance (see for instance the introduction of [13] and [10]).

On the other hand, the concave mobility $h(\rho) \geq 0$ considered in [13] is defined on the unbounded interval $[0,+\infty)$ and has to be necessarily non-decreasing. If we want to consider non-monotone concave mobilities $h(\rho) \geq 0$, then the domain of h has to be a bounded interval. This case, not considered in [13], is physically interesting. Indeed, examples of equations that can be modeled as gradient flows with respect to this kind of distances are a version of Cahn-Hilliard equation [14], some equation modelling chemotaxis with prevention of overcrowding [7, 8, 12], equations describing the relaxation of gas of fermions [21, 20, 15, 16, 9, 11], studies of phase segregation [18, 25], and studies of thin liquid films [5].

The principal example of mobility function in the papers cited above is

$$
h(\rho) = \rho(1-\rho), \qquad \text{defined in } [0,1],
$$

or $h(\rho) = 1-\rho^2$ defined in $[-1,1]$, mainly for the Cahn-Hilliard equation, the relaxation of fermion gas and the chemotaxis with overcrowding prevention. A more general example is of the form $h(\rho) = (\rho-a)^\alpha(b-\rho)^\beta$ defined in $[a,b]$ for some $\alpha,\beta \in [0,1]$. In the previous examples, if $a < 0$ then the density could be negative at some points and we have to consider signed measures instead of non-negative measures.

In this paper we will show that almost all the properties of the distance studied in [13] can be extended to this case.

As previously observed, in order to give a precise meaning of the dynamical characterization (2) and to define in a rigorous way the modified distance (3), the right framework is that of time dependent families of Radon measures and distributional solutions of the continuity equation. Following the explanation given in the introduction of [13], we replace ρ_t by a continuous curve $t \in [0,1] \mapsto \mu_t$ ($\mu_t = \rho_t \mathscr{L}^d$ in the absolutely continuous case) in the space $\mathcal{M}^+(\mathbb{R}^d)$ of nonnegative Radon measures in $\mathbb{R}^d$ endowed with the usual weak* topology. We replace the vector field $\boldsymbol{v}_t$ in (2) with a time dependent family of vector measures $\boldsymbol{\nu}_t := \boldsymbol{v}_t\mu_t \ll \mu_t$. The continuity equation in (2) can be written in terms of the couple $(\mu,\boldsymbol{\nu})$

$$
\partial_t \mu_t + \nabla\cdot\boldsymbol{\nu}_t = 0 \quad \text{in the sense of distributions in } \mathscr{D}'(\mathbb{R}^d\times(0,1)), \tag{4}
$$

and it is a linear equation. Since $\boldsymbol{v}_t = \mathrm{d}\boldsymbol{\nu}_t/\mathrm{d}\mu_t$ is the density of $\boldsymbol{\nu}_t$ with respect to μ_t, the action functional which has to be minimized in (2) is

$$\int_0^1 \Phi(\mu_t, \boldsymbol{\nu}_t)\,\mathrm{d}t, \qquad \Phi(\mu, \boldsymbol{\nu}) := \int_{\mathbb{R}^d} \left|\frac{\mathrm{d}\boldsymbol{\nu}}{\mathrm{d}\mu}\right|^p \mathrm{d}\mu. \tag{5}$$

In the case of absolutely continuous measures with respect to $\mathscr{L}^d$, i.e. $\mu = \rho\mathscr{L}^d$ and $\boldsymbol{\nu} = \boldsymbol{w}\mathscr{L}^d$, the functional Φ can be expressed as

$$\Phi(\mu, \boldsymbol{\nu}) := \int_{\mathbb{R}^d} \phi(\rho, \boldsymbol{w})\,\mathrm{d}\mathscr{L}^d(x), \quad \phi(\rho, \boldsymbol{w}) := \rho\left|\frac{\boldsymbol{w}}{\rho}\right|^p. \tag{6}$$

Denoting by $\mathcal{CE}(0,1)$ the class of measure-valued distributional solutions $(\mu, \boldsymbol{\nu})$ of the continuity equation (4), we can state the dynamical characterization of the Wasserstein distance as follows

$$W_p^p(\mu_0, \mu_1) = \inf\left\{\int_0^1 \Phi(\mu_t, \boldsymbol{\nu}_t)\,\mathrm{d}t : (\mu, \boldsymbol{\nu}) \in \mathcal{CE}(0,1),\ \mu_{|t=0} = \mu_0,\, \mu_{|t=1} = \mu_1\right\} \tag{7}$$

(as already observed, the Benamou-Brenier characterization (7) for Borel non-negative measures was proven in [2]). We observe that the function ϕ defined in (6) is p-homogeneous w.r.t. $\boldsymbol{w}$, is convex with respect to $(\rho, \boldsymbol{w})$, and positively 1-homogeneous with respect to $(\rho, \boldsymbol{w})$. By the 1-homogeneity it is immediate to check that the functional Φ in (6) is independent on the Lebesgue measure, in the sense that if $\gamma \in \mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$ is another reference measure such that $\mathrm{supp}(\gamma) = \mathbb{R}^d$ and $\mu = \tilde{\rho}\gamma$ and $\boldsymbol{\nu} = \tilde{\boldsymbol{w}}\gamma$, then

$$\Phi(\mu, \boldsymbol{\nu}) = \int_{\mathbb{R}^d} \phi(\tilde{\rho}, \tilde{\boldsymbol{w}})\,\mathrm{d}\gamma. \tag{8}$$

We explain the main idea of [13] for state rigorously the intuitive "definition" (3). Given a concave mobility function $h : (a,b) \to (0, +\infty)$, we consider still the linear continuity equation (4) and modify the action density ϕ in the following way: $\phi : (a,b) \times \mathbb{R}^d \to [0, +\infty)$

$$\phi(\rho, \boldsymbol{w}) := h(\rho)\left|\frac{\boldsymbol{w}}{h(\rho)}\right|^p. \tag{9}$$

The concavity of h is a necessary and sufficient condition for the convexity of ϕ in (9) (see [24] and Theorem 2.1). We observe that ϕ still satisfies the p-homogeneity with respect to $\boldsymbol{w}$ and is globally convex, but it is no longer positively 1-homogeneous with respect to $(\rho, \boldsymbol{w})$. Hence, in order to consider the integral functional Φ like (8) it is necessary to precise the reference measure $\gamma \in \mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$ for that ρ and $\boldsymbol{w}$ are the densities of μ and $\boldsymbol{\nu}$ respectively. Defining

$$\Phi(\mu, \boldsymbol{\nu}) = \int_{\mathbb{R}^d} \phi(\rho, \boldsymbol{w})\mathrm{d}\gamma$$

when $\mu = \rho\gamma$, $\boldsymbol{\nu} = \boldsymbol{w}\gamma$, and defining Φ suitably on the singular part of μ and $\boldsymbol{\nu}$ with respect to γ, (see Definition 2.5) the definition of the generalized Wasserstein distance associated to (ϕ, γ) is therefore

$$\mathcal{W}^p_{\phi,\gamma}(\mu_0, \mu_1) := \inf\left\{\int_0^1 \Phi(\mu_t, \boldsymbol{\nu}_t)\,\mathrm{d}t : (\mu, \boldsymbol{\nu}) \in \mathcal{CE}(0,1),\ \mu_{|t=0} = \mu_0,\, \mu_{|t=1} = \mu_1\right\}. \tag{10}$$

Particularly important for the applications are the following choices of γ:

- $\gamma := \mathscr{L}^d_{|\Omega} = \chi_\Omega\mathscr{L}^d$, with Ω an open subset of $\mathbb{R}^d$;
- $\gamma := e^{-V}\mathscr{L}^d$ for some C^1 potential $V : \mathbb{R}^d \to \mathbb{R}$;

- $\gamma := \mathscr{H}^k|_{\mathbb{M}}$, where $\mathbb{M}$ is a smooth k-dimensional manifold embedded in $\mathbb{R}^d$ with the Riemannian metric induced by the Euclidean distance and $\mathscr{H}^k$ denotes the k-dimensional Hausdorff measure.

In the paper [10], the authors used this kind of distance in the case $\gamma = \mathscr{L}^d_{|\Omega}$ in order to study the problem of the convexity of integral functionals along geodesics induced by the distance. The forthcoming paper [22] will be devoted to the study of forth orders equations (Cahn-Hilliard type with nonlinear mobility and thin-film like equations), with the proof of the existence of solutions by means of the minimizing movements approximation scheme (see [2]) for the distance like (10) and a first order integral functional.

We conclude this introduction stating the principal properties obtained in this paper for the distance like (10) with $h : (a, b) \to (0, +\infty)$, referring to Section 3 for the precise definitions and the complete statements. We recall that the choice of consider the mobility with bounded domain (a, b) allow to consider also the distance between signed measures.

- The space $\mathcal{M}_{\mathrm{loc}}(\mathbb{R}^d)$ endowed with the distance $\mathcal{W}_{\phi,\gamma}$ is a complete pseudo-metric space (the distance can assume the value $+\infty$), inducing as strong as, or stronger topology than the weak* one. Bounded sets with respect to $\mathcal{W}_{\phi,\gamma}$ are weakly* relatively compact. The distance $\mathcal{W}_{\phi,\gamma}$ is lower semi continuous with respect to the weak* convergence.
- In order to avoid that the distance could be $+\infty$ we consider the space $\mathcal{M}[\sigma] := \{\mu \in \mathcal{M}_{\mathrm{loc}}(\mathbb{R}^d) : \mathcal{W}_{\phi,\gamma}(\mu, \sigma) < +\infty\}$ for a given measure $\sigma \in \mathcal{M}_{\mathrm{loc}}(\mathbb{R}^d)$. The space $\mathcal{M}[\sigma]$ turns out to be a complete metric space.
- $\mathcal{M}[\sigma]$ is a geodesic space and the geodesic are unique if h is strictly concave.
- If $\tilde{\mathrm{m}}_{-q}(\gamma) < +\infty$, where q is the conjugate exponent of p and the generalized momentum is defined in Definition 2.9, then $\mu(\mathbb{R}^d) = \sigma(\mathbb{R}^d)$ for every $\mu \in \mathcal{M}[\sigma]$.

Finally, in Section 4 we give sufficient conditions on the measures μ_0, μ_1 in order to have finiteness of the distance $\mathcal{W}_{\phi,\gamma}(\mu_0, \mu_1)$, and we prove two results: one for the all space $\mathbb{R}^d$ with the Lebesgue measure as a reference, the other one for convex bounded domains in $\mathbb{R}^d$. In the two cases we study also the relation between the weak-$*$ convergence of measures and the convergence with respect to the distance $\mathcal{W}_{\phi,\gamma}$.

2 Preliminaries

In this Section we introduce the necessary tools in order to define in the next Section the modified Wasserstein distance and prove its basic properties. The contents are an adaptation of Sections 2-4 of [13].

2.1 Notation

Let X be a topological space, $A \subset X$, $f : X \to \mathbb{R} \cup \{\pm\infty\}$ be a function. We denote by:

$\mathrm{int}(A)$, $\bar{A}$, ∂A	the *interior*, the *closure* and the *boundary* of A, respectively;
$\chi_A : X \to \{0,1\}$	the *characteristic function* of A, namely $\chi_A(x) = 1$ if $x \in A$, $\chi_A(x) = 0$ if $x \notin A$;
$\mathrm{dom}(f) := \{x \in X : f(x) \in \mathbb{R}\}$	the *(effective) domain* of f;
$\mathrm{epi}(f) := \{(x,\alpha) \in X \times \mathbb{R} : \alpha \geq f(x)\}$	the *epigraph* of f;
$\mathrm{hypo}(f) := \{(x,\beta) \in X \times \mathbb{R} : \beta \leq f(x)\}$	the *hypograph* of f.
$\mathscr{L}^d$	the *Lebesgue measure* on $\mathbb{R}^d$;
$\mathcal{M}_{\mathrm{loc}}(\mathbb{R}^d)$	the set of *signed Radon measures* on $\mathbb{R}^d$;
$\mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$	the set of *non-negative Radon measures* on $\mathbb{R}^d$;
$\mathcal{M}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^h)$	the set of $\mathbb{R}^h$-valued Radon measures on $\mathbb{R}^d$.

We say that f is *lower semicontinuous or l.s.c.* (resp. *upper semicontinuous or u.s.c.*) iff $\mathrm{epi}(f)$ (resp. $\mathrm{hypo}(f)$) is closed in $X \times \mathbb{R}$. If (X,d) is a metric space, this is equivalent to say that f is l.s.c. (resp. u.s.c.) iff $f(x) \leq \liminf\limits_{y\to x} f(y)$ (resp. $f(x) \geq \limsup\limits_{y\to x} f(y)$).

2.1.1 Push-forward of measures

Given a Borel measure μ on a topological space X, and a Borel map $T : X \to Y$, with values in a topological space Y, we define the image measure of μ through the map T, denoted by $\nu = T_\#\mu$, by $\nu(B) := \mu(T^{-1}(B))$, for any Borel measurable set $B \subset Y$, or equivalently

$$\int_Y \zeta(y)\,d\nu(y) = \int_X \zeta(T(x))\,d\mu(x)\,, \quad \forall \zeta \in \mathcal{C}^0_b(Y). \tag{11}$$

If X and Y are domains of $\mathbb{R}^d$, the map T is sufficiently smooth and the measures μ and ν are absolutely continuous with respect to Lebesgue measure with densities $\widetilde{\rho}$ and ρ respectively, then $\nu = T_\#\mu$ is equivalent, by the change of variables theorem, to

$$\rho(T(x))\,\det(DT(x)) = \widetilde{\rho}(x). \tag{12}$$

The formula (12) for the densities holds in a very greater generality (see [2, Lemma 5.5.3]).

2.2 Convex Analysis

In this subsection we recall some concepts from convex analysis, our main reference is [24].

Definition 2.1 (Recession functional)**.** Let $f : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ be a proper convex function. The *recession functional* f^∞ of f is the positively homogeneous proper convex function defined by (cfr. [24, Theorem 8.5, p.66]):

$$f^\infty(y) := \sup\{f(x+y) - f(x) : \, x \in \mathrm{dom} f\}.$$

If f is l.s.c, then f^∞ is l.s.c., and for any $x \in \mathrm{dom}(f)$ it holds:

$$f^\infty(y) := \lim_{\lambda \to +\infty} \frac{f(x + \lambda y) - f(x)}{\lambda}.$$

We have that:

1. if $0 \in \mathrm{dom}(f)$, it holds $f^\infty(y) := \lim_{\lambda \to +\infty} \dfrac{f(\lambda y)}{\lambda}$ for all $y \in \mathbb{R}^N$.
2. if $0 \notin \mathrm{dom}(f)$, it holds $f^\infty(y) := \lim_{\lambda \to +\infty} \dfrac{f(\lambda y)}{\lambda}$ for all $y \in \mathrm{dom}(f)$.

Definition 2.2 (Concave-convex functions)**.** Let $C \subset \mathbb{R}^k$, $D \subset \mathbb{R}^d$ be convex sets, and $\tilde{f} : C \times D \to \mathbb{R} \cup \{\pm\infty\}$ be a function. We will call $\tilde{f}$ a *concave-convex* function if:

1. for each $z \in D$ the map $r \mapsto \tilde{f}(r, z)$ is concave,
2. for each $r \in C$ the map $z \mapsto \tilde{f}(r, z)$ is convex.

Given a concave-convex function $\tilde{f} : C \times D \to \mathbb{R}$, we define its *lower extension* $\tilde{f}_1 : \mathbb{R}^k \times \mathbb{R}^d \to \mathbb{R} \cup \{\pm\infty\}$ by setting:

$$\tilde{f}_1(r, z) = \begin{cases} \tilde{f}(r, z) & \text{if} \quad r \in C, \, z \in D \\ +\infty & \text{if} \quad r \in C, \, z \notin D \\ -\infty & \text{if} \quad r \notin C \end{cases}$$

$\tilde{f}_1$ is still a concave-convex function.

Theorem 2.1 (Partial Legendre)**.** *Let $f : \mathbb{R}^k \times \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be convex and l.s.c. Then the function defined by:*

$$\tilde{f}(r, z) := \sup_{w \in \mathbb{R}^d} [\langle z, w \rangle - f(r, w)]$$

is a concave-convex function from $\mathbb{R}^k \times \mathbb{R}^d$ to $\mathbb{R} \cup \{\pm\infty\}$. For every fixed r, the function $z \mapsto \tilde{f}(r, z)$ is l.s.c. Conversely, given any concave-convex function $\tilde{f} : \mathbb{R}^k \times \mathbb{R}^d \to \mathbb{R} \cup \{\pm\infty\}$, the function defined by:

$$f(r, w) := \sup_{z \in \mathbb{R}^d} [\langle z, w \rangle - \tilde{f}(r, z)]$$

is a convex map and for every fixed r, the function $w \mapsto f(r, w)$ is l.s.c. Moreover, if $\mathrm{dom}(\tilde{f}) = C \times D$ and $\tilde{f}$ agrees with its lower extension, then f is l.s.c.

Proof. See [24, Theorem 33.1]. □

2.3 Action function

Definition 2.3 (Admissible action density functions)**.** Let $\phi : \mathbb{R} \times \mathbb{R}^d \to [0, +\infty]$ be a l.s.c. nonnegative proper convex function, $1 < p < \infty$. We say that ϕ is an *admissible action density of order* p if it satisfies the following two properties:

(F1) $\boldsymbol{w} \mapsto \phi(\cdot, \boldsymbol{w})$ is p-homogeneus, i.e. for every given $\rho \in \mathbb{R}$ such that $\{\rho\} \times \mathbb{R}^d \cap \mathrm{dom}(\phi) \neq \emptyset$ we have $\phi(\rho, 0) = 0$ and for every $\lambda \neq 0$, $\boldsymbol{w} \in \mathbb{R}^d$ we have $\phi(\rho, \lambda \boldsymbol{w}) = |\lambda|^p \phi(\rho, \boldsymbol{w})$ (both sides may be $+\infty$).

(F2) there exists $\rho_0 \in \mathbb{R}$ such that $\{\rho_0\} \times \mathbb{R}^d \subseteq \mathrm{dom}(\phi)$ and $\phi(\rho_0, \boldsymbol{w}) > 0$ for all $\boldsymbol{w} \neq 0$.

The set of all admissible action densities of order p will be denoted by $\mathcal{A}_p$. Given $a, b \in \mathbb{R}$, $a < b$ we will denote by $\mathcal{A}_p(a, b)$ the set of action densities in $\mathcal{A}_p$ such that $\mathrm{int}(\mathrm{dom}(\phi)) = (a, b) \times \mathbb{R}^d$.
Let q be the conjugate exponent of p. We construct the *partial dual* $\mathcal{A}_q^*$ of $\mathcal{A}_p$ as follows. For all $\phi \in \mathcal{A}_p$, we define the concave-convex function $\tilde{\phi} : \mathrm{dom}(\phi) \to \mathbb{R} \cup \{+\infty\}$ by setting:

$$\frac{1}{q}\tilde{\phi}(\rho, \boldsymbol{z}) := \sup_{\boldsymbol{w} \in \mathbb{R}^d} \Big\{ \langle \boldsymbol{z}, \boldsymbol{w} \rangle - \frac{1}{p}\phi(\rho, \boldsymbol{w}) \Big\}. \tag{13}$$

We will call the lower extension of $\tilde{\phi}$ the *marginal conjugate* of ϕ and we will still denote it by $\tilde{\phi}$. We observe that $\tilde{\phi}$ is q-homogeneous with respect to the second variable and $\tilde{\phi}(\rho, \boldsymbol{z}) \geq 0$. We define:

$$\mathcal{A}_q^* := \{\tilde{\phi} : \ \tilde{\phi} \text{ is the marginal conjugate of } \phi, \ \phi \in \mathcal{A}_p\}$$

and it is easy to check that $\mathrm{int}(\mathrm{dom}(\tilde{\phi})) = (a, b) \times \mathbb{R}^d$ if $\phi \in \mathcal{A}_p(a, b)$.

The following proposition can be proved exactly as Theorem 3.1 of [13].

Proposition 2.1 (ϕ-norm)**.** *Let $1 < p < +\infty$, q be the conjugate exponent of p and $\phi \in \mathcal{A}_p$. Then:*

1. *For every $\rho \in \mathbb{R}$ such that $\{\rho\} \times \mathbb{R}^d \subset \mathrm{dom}(\phi)$, the functions $\boldsymbol{w} \mapsto \phi(\rho, \boldsymbol{w})^{1/p}$ and $\boldsymbol{z} \mapsto \tilde{\phi}(\rho, \boldsymbol{z})^{1/q}$ are norms on $\mathbb{R}^d$ each one dual of the other. We have:*

$$\|\boldsymbol{z}\|_{(\phi,\rho)*} := \tilde{\phi}(\rho, \boldsymbol{z})^{1/q} = \sup_{\boldsymbol{w} \neq 0} \frac{\langle \boldsymbol{w}, \boldsymbol{z} \rangle}{\phi(\rho, \boldsymbol{w})^{1/p}}, \qquad \|\boldsymbol{w}\|_{(\phi,\rho)} := \phi(\rho, \boldsymbol{w})^{1/p} = \sup_{\boldsymbol{z} \neq 0} \frac{\langle \boldsymbol{w}, \boldsymbol{z} \rangle}{\tilde{\phi}(\rho, \boldsymbol{z})^{1/q}}. \tag{14}$$

2. *The restriction to $\mathrm{dom}(\tilde{\phi})$ of the marginal conjugate $\tilde{\phi}$ of ϕ takes its values in $[0, +\infty)$ and it is a concave-convex function.*

3. *Given $\rho_0, \rho_1 \in \mathbb{R}$ with $[\rho_0, \rho_1] \times \mathbb{R}^d \subseteq \mathrm{dom}(\phi)$, there exists a constant $C = C(\rho_0, \rho_1)$ such that for every $\rho \in [\rho_0, \rho_1]$ it holds:*

$$C^{-1}|\boldsymbol{w}|^p \leq \phi(\rho, \boldsymbol{w}) \leq C|\boldsymbol{w}|^p, \qquad C^{-1}|\boldsymbol{z}|^q \leq \tilde{\phi}(\rho, \boldsymbol{z}) \leq C|\boldsymbol{z}|^q, \qquad \forall \boldsymbol{w}, \boldsymbol{z} \in \mathbb{R}^d.$$

Equivalently, a function ϕ belongs to $\mathcal{A}_p$ if and only if it admits the dual representation formula

$$\frac{1}{p}\phi(\rho, \boldsymbol{w}) := \sup_{\boldsymbol{z} \in \mathbb{R}^d} \Big\{ \langle \boldsymbol{z}, \boldsymbol{w} \rangle - \frac{1}{q}\tilde{\phi}(\rho, \boldsymbol{z}) \Big\}, \tag{15}$$

where $\tilde{\phi} : \mathbb{R} \times \mathbb{R}^d \to [0, \infty)$ is (the lower extension of) a nonnegative concave-convex function which is q-homogeneous with respect to $\boldsymbol{z}$.

Lemma 2.1. *Let $\phi \in \mathcal{A}_p$. Then the recession functional is p-homogeneous with respect to the second variable, i.e.* $\phi^\infty(\rho, \lambda \boldsymbol{w}) = |\lambda|^p \phi^\infty(\rho, \boldsymbol{w})$ *for* $\lambda \in \mathbb{R}$. *Moreover, for* $\rho \neq 0$ *it is possible to express* $\phi^\infty(\rho, \boldsymbol{w}) = \rho \varphi^\infty(\boldsymbol{w}/\rho)$, *where* $\varphi^\infty : \mathbb{R}^d \to [0, +\infty]$ *is convex p-homogeneous function such that* $\varphi^\infty(\boldsymbol{w}) > 0$ *if* $\boldsymbol{w} \neq 0$.

Proof. We notice that $(0,0)$ may not belong in general to $\mathrm{dom}(\phi)$, however we have:

$$\phi^\infty(\rho, \boldsymbol{w}) \quad := \lim_{\lambda \to +\infty} \frac{\phi(\bar\rho + \lambda\rho, \lambda\boldsymbol{w}) - \phi(\bar\rho, 0)}{\lambda} = \lim_{\lambda \to +\infty} \frac{\phi(\bar\rho + \lambda\rho, \lambda\boldsymbol{w})}{\lambda} = \lim_{\lambda \to +\infty} \lambda^{p-1}\phi(\bar\rho + \lambda\rho, \boldsymbol{w}),$$

for every $\bar\rho \in \mathbb{R}$ such that $(\bar\rho, 0) \in \mathrm{dom}(\phi)$, and such $\bar\rho$ exists by definition of the class $\mathcal{A}_p$. Hence ϕ^∞ is still p-homogeneous with respect to $\boldsymbol{w}$. The other statement follows from the definition of the class $\mathcal{A}_p$. □

We notice that in the case of $\phi \in \mathcal{A}_p(a,b)$ we have $\phi^\infty(0,0) = 0$ and $\phi^\infty(\rho, \boldsymbol{w}) = +\infty$ for $(\rho, \boldsymbol{w}) \neq (0,0)$.

One of the most interesting example of admissible density function in $\mathcal{A}_p(a,b)$ is the following:

Definition 2.4. Let $p > 1$ and q its conjugate exponent. Let $h : \mathbb{R} \to [0, +\infty) \cup \{-\infty\}$ be an u.s.c. concave function with $\mathrm{int}(\mathrm{dom}(h)) = (a,b)$, $a, b \in \mathbb{R}$, $a < b$, $h(\rho) > 0$ for every $\rho \in (a,b)$. Define $\tilde\phi_h(\rho, \boldsymbol{z}) = h(\rho)|\boldsymbol{z}|^q$ on $\mathbb{R} \times \mathbb{R}^d$. We have that this is a concave-convex map which is q-homogeneous with respect to $\boldsymbol{z}$. Hence, it is the marginal conjugate of the l.s.c. convex map $\phi_h \in \mathcal{A}_p(a,b)$ defined by

$$\phi_h(\rho, \boldsymbol{w}) = \begin{cases} \dfrac{|\boldsymbol{w}|^p}{h(\rho)^{p-1}} & \text{if } \rho \in \mathrm{dom}(h),\ h(\rho) \neq 0 \\ 0 & \text{if } h(\rho) = 0,\ \boldsymbol{w} = 0 \\ +\infty & \text{if } h(\rho) = 0,\ \boldsymbol{w} \neq 0 \text{ or } h(\rho) = -\infty. \end{cases} \tag{16}$$

Such function h is called *mobility function.*

The following proposition shows that every admissible function ϕ is bounded from above by an admissible function of the type (16).

Proposition 2.2. *If* $\phi \in \mathcal{A}_p(a,b)$, *then there exists a concave function h such that* $\mathrm{int}(\mathrm{dom}(h)) = (a,b)$, $h(r) > 0$ *for every* $r \in (a,b)$ *and*

$$\phi(r, \boldsymbol{w}) \leq \phi_h(r, \boldsymbol{w}). \tag{17}$$

Proof. Let us define

$$h(r) := \inf_{|\boldsymbol{z}|=1} \tilde\phi(r, \boldsymbol{z}),$$

where $\tilde\phi$ is defined in (13). By the q-homogeneity of $\tilde\phi$ we have

$$\tilde\phi(r, \boldsymbol{z}) \geq h(r)|\boldsymbol{z}|^q.$$

Then, by the representation (15) for ϕ and ϕ_h, we obtain

$$\frac{1}{p}\phi(r, \boldsymbol{w}) \leq \sup_{\boldsymbol{z} \in \mathbb{R}^d} \left\{ \langle \boldsymbol{w}, \boldsymbol{z} \rangle - \frac{1}{q} h(r) |\boldsymbol{z}|^q \right\} = \frac{1}{p}\phi_h(r, \boldsymbol{w}).$$

□

2.4 Action functional

Given an admissible action density function ϕ and a reference measure γ on $\mathbb{R}^d$, we can define the corresponding action functional.

Definition 2.5 (ϕ-Action functional)**.** Let $\gamma \in \mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$ be a reference measure and $\phi \in \mathcal{A}_p$.
For every $\mu \in \mathcal{M}_{\rm loc}(\mathbb{R}^d)$ and $\boldsymbol{\nu} \in \mathcal{M}_{\rm loc}(\mathbb{R}^d;\mathbb{R}^d)$ such that $\mathrm{supp}(\mu)$ and $\mathrm{supp}(\boldsymbol{\nu})$ are contained in $\mathrm{supp}(\gamma)$ we can write their Lebesgue decomposition $\mu = \rho\gamma + \mu^\perp$, $\boldsymbol{\nu} = \boldsymbol{w}\gamma + \boldsymbol{\nu}^\perp$. Introducing a nonnegative Radon measure $\sigma \in \mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$ such that $\mu^\perp \ll \sigma$ and $\boldsymbol{\nu}^\perp \ll \sigma$ (e.g. take $\sigma = |\mu^\perp| + |\boldsymbol{\nu}^\perp|$) and using the notation $\mu^\perp = \rho^\perp\sigma$ and $\boldsymbol{\nu}^\perp = \boldsymbol{w}^\perp\sigma$, we define the action functional Φ associated to ϕ by

$$\Phi(\mu,\boldsymbol{\nu}|\gamma) = \Phi^a(\mu,\boldsymbol{\nu}|\gamma) + \Phi^\infty(\mu,\boldsymbol{\nu}|\gamma) := \int_{\mathbb{R}^d} \phi(\rho,\boldsymbol{w})\,d\gamma + \int_{\mathbb{R}^d} \phi^\infty(\rho^\perp,\boldsymbol{w}^\perp)\,d\sigma.$$

Since ϕ^∞ is 1-homogeneous, the definition does not depend on σ.

We collect in the following theorem some properties of convex functionals on measures. The proof can be found in [13] (see also [3] for functionals defined on measures).

Theorem 2.2 (Properties of Φ)**.** *Let $\phi \in \mathcal{A}_p$ and Φ as in Definition 2.5.*

1. Lower semicontinuity. *If three sequences $(\gamma_n)_{n\in\mathbb{N}} \subset \mathcal{M}^+_{\rm loc}(\mathbb{R}^k)$, $(\mu_n)_{n\in\mathbb{N}} \subset \mathcal{M}_{\rm loc}(\mathbb{R}^d)$, $(\boldsymbol{\nu}_n)_{n\in\mathbb{N}} \subset \mathcal{M}_{\rm loc}(\mathbb{R}^d,\mathbb{R}^d)$ weakly* converge to γ, μ, $\boldsymbol{\nu}$ respectively, then $\Phi(\mu,\boldsymbol{\nu}|\gamma) \le \liminf_{n\to+\infty} \Phi(\mu_n,\boldsymbol{\nu}_n|\gamma_n)$.*

2. Monotonicity w.r.to γ. *Assume that $(0,0) \in \mathrm{dom}(\phi)$ (in this case by homogeneity we have $\phi(0,0) = 0$) and let $\gamma_1,\gamma_2 \in \mathcal{M}^+_{\rm loc}(\mathbb{R}^k)$ be such that $\gamma_1 \le \gamma_2$. Then $\Phi(\mu,\boldsymbol{\nu}|\gamma_2) \le \Phi(\mu,\boldsymbol{\nu}|\gamma_1)$ for every $(\mu,\boldsymbol{\nu})$ such that $\mathrm{supp}(\mu) \cup \mathrm{supp}(\boldsymbol{\nu}) \subseteq \mathrm{supp}(\gamma_i)$, $i = 1,2$.*

3. Monotonicity w.r.to convolution. *Let $k \in C^\infty_c(\mathbb{R}^d)$ be a convolution kernel satisfying $k(x) \ge 0$ for all $x \in \mathbb{R}^d$ and $\int_{\mathbb{R}^d} k(x)\,dx = 1$. Then $\Phi(\mu * k, \boldsymbol{\nu} * k|\gamma * k) \le \Phi(\mu,\boldsymbol{\nu}|\gamma)$.*

The following example shows that the statement on monotonicity with respect to the reference measure may fail if $(0,0) \notin \mathrm{dom}(\phi)$.

Example 1 (Non-monotonicity). Let $d = 1$. We define $\phi : \mathbb{R}\times\mathbb{R} \to \mathbb{R}\cup\{+\infty\}$ to be $\phi(r,\boldsymbol{v}) = |\boldsymbol{v}|^2$ if $r \in [3/2,2]$ and $+\infty$ elsewhere. Define $\gamma_2 = 3/2\gamma_1 = \chi_{[1,2]}(x)\mathscr{L}^1$ and set $\mu = \boldsymbol{\nu} = \gamma_2 = 3/2\gamma_1$. Then

$$\Phi(\mu,\boldsymbol{\nu}|\gamma_2) = \int_{\mathbb{R}} \phi(1,1)\,d\gamma_2 = +\infty.$$

$$\Phi(\mu,\boldsymbol{\nu}|\gamma_1) = \int_{\mathbb{R}} \phi(3/2,3/2)\,d\gamma_1 = \frac{3}{2}.$$

Hence $\gamma_1 < \gamma_2$ but $\Phi(\mu,\boldsymbol{\nu}|\gamma_1) < \Phi(\mu,\boldsymbol{\nu}|\gamma_2)$.

When $\phi \in \mathcal{A}_p(a,b)$, the finiteness of the corresponding action functional $\Phi(\mu,\boldsymbol{\nu}|\gamma)$, force the absolute continuity of μ with respect to γ and a boundedness of the density of μ with respect to γ . We state this important property in the following proposition.

Proposition 2.3. *Let $\phi \in \mathcal{A}_p(a,b)$ and $\gamma \in \mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$ a fixed reference measure. Let $\mu \in \mathcal{M}_{\rm loc}(\mathbb{R}^d)$, $\boldsymbol{\nu} \in \mathcal{M}_{\rm loc}(\mathbb{R}^d;\mathbb{R}^d)$ be such that $\Phi(\mu,\boldsymbol{\nu}|\gamma) < +\infty$. Then $\mu \ll \gamma$, $\boldsymbol{\nu} \ll \gamma$ and*

$$\Phi(\mu,\boldsymbol{\nu}|\gamma) = \Phi^a(\mu,\boldsymbol{\nu}|\gamma) = \int_{\mathbb{R}^d} \phi(\rho,\boldsymbol{w})\,d\gamma, \tag{18}$$

where $\mu = \rho\gamma$ and $\boldsymbol{\nu} = \boldsymbol{w}\gamma$. Moreover we have

$$a \le \rho(x) \le b \qquad \text{for } \gamma\text{-a.e. } x \in \mathbb{R}^d. \tag{19}$$

Proof. Since $\phi \in \mathcal{A}_p(a,b)$ and $p > 1$, by the definition of ϕ^∞ and Lemma 2.1, it is easy to check that $\phi^\infty(\rho, \boldsymbol{w}) = +\infty$ if $(\rho, \boldsymbol{w}) \neq (0,0)$, and $\phi^\infty(0,0) = 0$. If $\mu = \rho\gamma + \mu^\perp$, $\boldsymbol{\nu} = \boldsymbol{w}\gamma + \boldsymbol{\nu}^\perp$ and $\sigma \in \mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$ such that $\mu^\perp = \rho^\perp\sigma$ and $\boldsymbol{\nu}^\perp = \boldsymbol{w}^\perp\sigma$, we can represent

$$\Phi^\infty(\mu, \boldsymbol{\nu}|\gamma) = \int_{\mathbb{R}^d} \phi^\infty(\rho^\perp, \boldsymbol{w}^\perp)\, d\sigma.$$

In order to have $\Phi^\infty(\mu, \boldsymbol{\nu}|\gamma) < \infty$, we must have $\rho^\perp(x) = 0$ and $\boldsymbol{w}^\perp(x) = 0$ for σ-a.e. $x \in \mathbb{R}^d$. This implies that $\rho \ll \gamma$ and $\boldsymbol{\nu} \ll \gamma$ and (18) holds. The last statement follows from $\int_{\mathbb{R}^d} \phi(\rho, \boldsymbol{w})\, d\gamma < +\infty$. □

2.5 Continuity equation

In this Subsection we collect the basic facts on the measure solutions of the continuity equation. It is an adaptation of [13] and [2], with the novelty that here we consider signed measures instead of non-negative measures.

Definition 2.6. Given $T > 0$, we consider the *continuity equation*:

$$\partial_t \mu_t + \mathrm{div}(\boldsymbol{\nu}_t) = 0, \qquad \text{in } \mathbb{R}^d \times (0,T), \tag{20}$$

where $\mu_t, \boldsymbol{\nu}_t$ are Borel families of measures in $\mathcal{M}_{\mathrm{loc}}(\mathbb{R}^d)$ and $\mathcal{M}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^d)$ respectively, defined for $t \in (0,T)$ satisfying

$$\int_0^T |\mu_t|(B(0,R))\, dt < +\infty, \qquad V_R := \int_0^T |\boldsymbol{\nu}_t|(B(0,R)\, dt < +\infty \qquad \forall\, R > 0, \tag{21}$$

and the equation (20) holds in the sense of distributions, i.e.

$$\int_0^T \int_{\mathbb{R}^d} \partial_t \zeta(x,t)\, d\mu_t(x)\, dt + \int_0^T \int_{\mathbb{R}^d} \nabla_x(\zeta(x,t))\, d\boldsymbol{\nu}_t(x)\, dt = 0, \qquad \text{for every } \zeta \in C^1_c(\mathbb{R}^d \times (0,T)). \tag{22}$$

We recall that, thanks to the disintegration theorem, we can identify $(\boldsymbol{\nu}_t)_{t\in[0,T]}$ with the measure $\boldsymbol{\nu} = \int_0^T \boldsymbol{\nu}_t\, dt \in \mathcal{M}_{\mathrm{loc}}(\mathbb{R}^d \times (0,T);\mathbb{R}^d)$ defined by:

$$\langle \boldsymbol{\nu}, \zeta \rangle = \int_0^T \left(\int_{\mathbb{R}^d} \zeta(x,t) d\boldsymbol{\nu}_t \right) dt, \qquad \forall \zeta \in C^0_c(\mathbb{R}^d \times (0,T);\mathbb{R}^d).$$

Similarly, we can identify $(\mu_t)_{t\in[0,T]}$ with a measure $\mu = \int_0^T \mu_t\, dt \in \mathcal{M}_{\mathrm{loc}}(\mathbb{R}^d \times (0,T))$.

Lemma 2.2. *Let $T > 0$ and $(\mu_t, \boldsymbol{\nu}_t)_{t\in(0,T)}$ be a Borel family of measures satisfying (21) and (22). Then there exists a unique weakly* continuous curve $[0,T] \ni t \mapsto \tilde{\mu}_t \in \mathcal{M}_{\mathrm{loc}}(\mathbb{R}^d)$ such that $\mu_t = \tilde{\mu}_t$ for $\mathcal{L}^1$-a.e. $t \in (0,T)$; if $\zeta \in C^1_c(\mathbb{R}^d \times (0,T))$ and $t_1, t_2 \in [0,T]$ with $t_1 \le t_2$ we have:*

$$\int_{\mathbb{R}^d} \zeta(t_2, x)\, d\mu_{t_2} - \int_{\mathbb{R}^d} \zeta(t_1, x)\, d\mu_{t_1} = \int_{t_1}^{t_2} \left(\int_{\mathbb{R}^d} \partial_t \zeta(t,x)\, d\mu_t(x) + \int_{\mathbb{R}^d} \nabla_x(\zeta(t,x))\, d\boldsymbol{\nu}_t(x) \right) dt.$$

Moreover if $\tilde{\mu}_s(\mathbb{R}^d) \in \mathbb{R}$ for some $s \in [0,T]$ and $\lim_{R\to+\infty} R^{-1}V_R = 0$, then the total mass $\tilde{\mu}_t(\mathbb{R}^d) \in \mathbb{R}$ and is constant.

Definition 2.7 (Solution of continuity equation)**.** Let $T > 0$, we denote by $\mathcal{CE}(0,T)$ the set of time-dependent measures $(\mu_t, \boldsymbol{\nu}_t)_{t\in[0,T]}$ such that

1. $t \mapsto \mu_t$ is weakly* continuous in $\mathcal{M}_{\mathrm{loc}}(\mathbb{R}^d)$ satisfying (21);

2. $(\boldsymbol{\nu}_t)_{t\in[0,T]}$ is a Borel family satisfying (21);

3. $(\mu,\boldsymbol{\nu})$ satisfies (22).

Given $\mu^1,\mu^2 \in \mathcal{M}_{\rm loc}(\mathbb{R}^d)$, we denote the set of solutions connecting μ^1 to μ^2 (possibly empty) by $\mathcal{CE}(0,T,\mu^1 \to \mu^2) = \{(\mu,\boldsymbol{\nu}) \in \mathcal{CE}(0,T) : \mu_0 = \mu^1, \mu_T = \mu^2\}$. Given $\gamma \in \mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$ reference measure and $\phi \in \mathcal{A}_p$, we denote by $\mathcal{CE}_{\phi,\gamma}(0,T;\mu^1 \to \mu^2) = \{(\mu,\boldsymbol{\nu}) \in \mathcal{CE}(0,T;\mu^1 \to \mu^2) : \int_0^T \Phi(\mu_t,\boldsymbol{\nu}_t|\gamma)\,dt < +\infty\}$, which is the set of solutions of the continuity equation connecting μ^1 to μ^2 with finite energy. We also use the notation $\mathcal{CE}_{\phi,\gamma}(0,T) := \{(\mu,\boldsymbol{\nu}) \in \mathcal{CE}(0,T) : \int_0^T \Phi(\mu_t,\boldsymbol{\nu}_t|\gamma)\,dt < +\infty\}$.

Lemma 2.3. *The following properties hold:*

1. *(Time rescaling) Let $\tau : [0,T'] \to [0,T]$ be a strictly increasing absolutely continuous map with absolutely continuous inverse $s = \tau^{-1}$. Then $(\mu,\boldsymbol{\nu})$ is a distributional solution of (22) iff $(\hat\mu,\hat{\boldsymbol{\nu}})$, where $\hat\mu = \mu\circ\tau$ and $\hat{\boldsymbol{\nu}} = \tau'(\boldsymbol{\nu}\circ\tau)$ is a distributional solution of (22) on $(0,T')$.*

2. *(Gluing solution) Let $(\mu^1,\boldsymbol{\nu}^1) \in \mathcal{CE}(0,T_1)$, $(\mu^2,\boldsymbol{\nu}^2) \in \mathcal{CE}(0,T_2)$ with $\mu^1_{T_1} = \mu^2_0$. Then the new family $(\mu_t,\boldsymbol{\nu}_t)_{t\in(0,T_1+T_2)}$, defined by $(\mu_t,\boldsymbol{\nu}_t) = (\mu^1_t,\boldsymbol{\nu}^1_t)$ for $0 \le t \le T_1$ and $(\mu_t,\boldsymbol{\nu}_t) = (\mu^2_{t-T_1},\boldsymbol{\nu}^2_{t-T_1})$ for $T_1 \le t \le T_2$, belongs to $\mathcal{CE}(0,T_1+T_2)$.*

2.5.1 Conservation of the mass for solutions with finite energy

In this paragraph we prove that, under a condition on the generalized moments of the reference measure γ and for $\phi \in \mathcal{A}_p(a,b)$, the total (signed) mass conserves for solutions of the continuity equation with finite energy.

Definition 2.8 (Upper uniform concave bound)**.** Let $\phi \in \mathcal{A}_p(a,b)$. Fixing $\bar\rho := (a+b)/2$ we use the notation

$$\|\boldsymbol{w}\| := \|\boldsymbol{w}\|_{(\phi,\bar\rho)}, \qquad \|\boldsymbol{z}\|_* := \|\boldsymbol{z}\|_{(\phi,\bar\rho)*}, \tag{23}$$

where the norms above (equivalents to the euclidean one) are defined in (14). We consider the set:

$$\mathcal{H} := \{g : \mathbb{R} \to \mathbb{R}\cup\{-\infty\} : g \text{ is u.s.c. and concave, } g(\rho) \ge \tilde\phi(\rho,\boldsymbol{z}/\|\boldsymbol{z}\|_*)\ \forall \boldsymbol{z} \ne 0\}.$$

This set is nonempty, and we can define:

$$\mathrm{h}(\rho) = \inf\{g(\rho) : g \in \mathcal{H}\},$$

which turns out to be the smallest u.s.c. concave function greater than or equal to $\sup\{\tilde\phi(\rho,\boldsymbol{z}) : \|\boldsymbol{z}\|_* = 1\}$. Since $\mathrm{int}(\mathrm{dom}(\mathrm{h})) = (a,b)$ we obtain that

$$\mathrm{h}_{max} := \sup_{\rho\in\mathbb{R}} \mathrm{h}(\rho) < +\infty. \tag{24}$$

By homogeneity property it is immediate to prove that

$$\tilde\phi(\rho,\boldsymbol{z}) \le \mathrm{h}(\rho)\|\boldsymbol{z}\|_*^q \qquad \text{and} \qquad \|\boldsymbol{w}\| \le \mathrm{h}(\rho)^{1/q}\phi(\rho,\boldsymbol{w})^{1/p}. \tag{25}$$

When ϕ is given as in Definition 2.4, we have $\mathrm{h}(\rho) = C\cdot h(\rho)$, where $C := \max\{|\boldsymbol{z}|/\|\boldsymbol{z}\|_* : \boldsymbol{z} \ne 0\}$, and $|\cdot|$ is the euclidean norm.

Definition 2.9. Let $\gamma \in \mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$, $r \in \mathbb{R}$. We define the *generalized r-th momentum* $\tilde{\mathrm{m}}_r(\gamma)$ of γ by setting:

$$\tilde{\mathrm{m}}_r(\gamma) := \gamma(B(0,1)) + \int_{\mathbb{R}^d\setminus B(0,1)} |x|^r\,d\gamma(x).$$

We observe that if $\tilde{\mathrm{m}}_r(\gamma) < +\infty$ then $\tilde{\mathrm{m}}_s(\gamma) < +\infty$ for every $s \le r$.

Proposition 2.4 (Mass conservation)**.** *Let $p > 1$, q its conjugate exponent and $\phi \in \mathcal{A}_p(a,b)$. Let $r \in \mathbb{R}$ such that $r \ge -q$ and $\gamma \in \mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$ be a reference measure satisfying $\tilde{\mathrm{m}}_r(\gamma) < +\infty$.*
If $(\mu_t, \boldsymbol{\nu}_t)_{t\in[0,T]} \in \mathcal{CE}_{\phi,\gamma}(0,T)$ and $\mu_0(\mathbb{R}^d) \in \mathbb{R}$, then $\mu_t(\mathbb{R}^d) = \mu_0(\mathbb{R}^d)$ for every $t \in [0,T]$.

Proof. We consider a cutoff function $\zeta \in C^\infty_c(\mathbb{R}^d)$ such that $\zeta(x) = 1$ if $|x| \le 1$, $\zeta(x) = 0$ if $|x| \ge 2$ and $|\nabla\zeta(x)| \le 1$ for all $x \in \mathbb{R}^d$. We consider the family $\zeta_R(x) = \zeta(x/R)$, for $R > 0$, that obviously satisfies $|\nabla\zeta_R(x)| \le 1/R$ for all $x \in \mathbb{R}^d$.

Using the notations of Definition 2.8, for every $t_1, t_2 \in [0,T]$, $t_1 < t_2$, by Proposition 2.1 and (25) we have

$$\begin{aligned}\Big|\int_{\mathbb{R}^d} \zeta_R\,d\mu_{t_1} - \int_{\mathbb{R}^d} \zeta_R\,d\mu_{t_2}\Big| &\le \int_{t_1}^{t_2}\int_{\mathbb{R}^d} |\nabla\zeta_R \cdot \boldsymbol{w}_t|\,d\gamma\,dt \\ &\le \int_{t_1}^{t_2}\int_{\mathbb{R}^d} \tilde{\phi}(\rho_t, \nabla\zeta_R)^{1/q}\phi(\rho_t, \boldsymbol{w}_t)^{1/p}\,d\gamma\,dt \\ &\le \Big(\int_{t_1}^{t_2}\int_{B_{2R}\setminus B_R} \mathrm{h}(\rho_t)\|\nabla\zeta_R\|_*^q\,d\gamma\,dt\Big)^{1/q}\Big(\int_{t_1}^{t_2}\int_{\mathbb{R}^d} \phi(\rho_t, \boldsymbol{w}_t)\,d\gamma\,dt\Big)^{1/p}.\end{aligned}$$

Since $\int_0^T \Phi(\mu_t, \boldsymbol{\nu}_t|\gamma)\,dt < +\infty$, by (24) and the equivalence of $\|\cdot\|_*$ with the euclidean norm, the last inequality shows that there exists $C > 0$ such that

$$\Big|\int_{\mathbb{R}^d} \zeta_R\,d\mu_{t_1} - \int_{\mathbb{R}^d} \zeta_R\,d\mu_{t_2}\Big| \le C\Big(\frac{1}{R^q}\gamma(B_{2R}\setminus B_R)\Big)^{1/q}. \tag{26}$$

Since $\tilde{\mathrm{m}}_r(\gamma) < +\infty$ shows that $\lim_{R\to+\infty} R^r\gamma(B_{2R}\setminus B_R) = 0$ we have that $\lim_{R\to+\infty}\frac{1}{R^q}\gamma(B_{2R}\setminus B_R) = 0$ if $r \ge -q$. Then the conservation of the mean follows from (26). ☐

Example 2. When $\phi \in \mathcal{A}_p(a,b)$ with $a < 0$ and $b > 0$, if $\gamma = \mathscr{L}^d$ and $d > q$ in general solutions of the continuity equations with finite energy could not conserve the mass.

Let $\varepsilon > 0$ such that $a + \varepsilon < 0$ and $b - \varepsilon > 0$ and consider an initial measure with compact support and mass different from 0, $\mu_0 = \rho_0\mathscr{L}^d$, such that $a + \varepsilon \le \rho_0 \le b - \varepsilon$. We define the curve, for $t \ge 0$,

$$\mu_t := \rho_t\mathscr{L}^d, \quad \rho_t(x) := e^{-dt}\rho_0(e^{-t}x), \quad \boldsymbol{\nu}_t := \boldsymbol{w}_t\mathscr{L}^d = x\rho_t(x)\mathscr{L}^d. \tag{27}$$

It is easy to check that $(\mu, \boldsymbol{\nu}) \in \mathcal{CE}(0,+\infty)$, $\mu_t(\mathbb{R}^d) = \mu_0(\mathbb{R}^d)$ and $a + \varepsilon \le \rho_t \le b - \varepsilon$. By 3 of Proposition 2.1 we have that $\phi(\rho_t, \boldsymbol{w}_t) \le C|\boldsymbol{w}_t|^p$. By a simple computation we obtain that

$$\int_{\mathbb{R}^d} |\boldsymbol{w}_t(x)|^p\,\mathrm{d}x = \int_{\mathbb{R}^d} |x|^p e^{-tdp}|\rho_0(e^{-t}x)|^p\,\mathrm{d}x = \int_{\mathbb{R}^d} |y|^p e^{t((1-d)p+d)}|\rho_0(y)|^p\,\mathrm{d}y$$

and then

$$\int_0^{+\infty} \phi(\rho_t, \boldsymbol{w}_t)\mathrm{d}x\,\mathrm{d}t < +\infty$$

when $d > q$.

The curve $(\mu_t, \boldsymbol{\nu}_t)$ can be reparametrized between $[0,1]$ setting $s = \frac{2}{\pi}\arctan t$, $t \in (0,+\infty)$ and $\eta_s = \rho_{\tan(\frac{\pi}{2}s)} = \rho_t$. It is not difficult to check that the energy is still finite and η_s connect μ_0 with the null measure.

2.5.2 Compactness for solutions with finite energy

In this section we prove a compactness result for signed solutions of the continuity equation. This result is a useful tool in order to obtain existence of geodesics of the distance defined in the next Section and its lower semi-continuity with respect to weak* convergence.

Proposition 2.5 (Compactness)**.** *Let $\phi \in \mathcal{A}_p(a,b)$ and $\gamma^n, \gamma \in \mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$ be a sequence such that $\gamma^n \rightharpoonup^* \gamma$. If $(\mu^n, \boldsymbol{\nu}^n)$ is a sequence in $\mathcal{CE}_{\phi,\gamma^n}(0,T)$ satisfying*

$$\sup_{n\in\mathbb{N}} \int_0^T \Phi(\mu^n_t, \boldsymbol{\nu}^n_t|\gamma^n)\, dt < +\infty, \tag{28}$$

then there exists a subsequence (still indexed by n) and a couple $(\mu, \boldsymbol{\nu}) \in \mathcal{CE}_{\phi,\gamma}(0,T)$ such that $\mu^n_t \rightharpoonup^* \mu_t$ in $\mathcal{M}_{\rm loc}(\mathbb{R}^d)$ for all $t \in [0,T]$, $\boldsymbol{\nu}^n \rightharpoonup^* \boldsymbol{\nu}$ in $\mathcal{M}_{\rm loc}(\mathbb{R}^d \times (0,T); \mathbb{R}^d)$, and

$$\int_0^T \Phi(\mu_t, \boldsymbol{\nu}_t|\gamma)\, dt \le \liminf_{n\to+\infty} \int_0^T \Phi(\mu^n_t, \boldsymbol{\nu}^n_t|\gamma^n)\, dt. \tag{29}$$

If along the subsequence $\tilde{\rm m}_{-q}(\gamma^n) < +\infty$ for all n and $\tilde{\rm m}_{-q}(\gamma) < +\infty$, and $\mu^n_0(\mathbb{R}^d) \to \mu_0(\mathbb{R}^d) \in \mathbb{R}$, then $\mu^n_t(\mathbb{R}^d) \to \mu_t(\mathbb{R}^d)$ for every $t \in [0,T]$.

Proof. By Proposition 2.3 we have $\mu^n = \rho^n\gamma^n$, $\boldsymbol{\nu}^n = \boldsymbol{w}^n\gamma^n$ and $|\rho^n| \le c := \max\{|a|,|b|\}$. Then there exists a subsequence (still indexed by n) and ρ such that $\rho^n \rightharpoonup \rho$ weakly in $L^1_{\rm loc}(\mathbb{R}^d \times [0,T])$. On the other hand, by (25), for every bounded Borel set $B \subset \mathbb{R}^d$ and for every $t_1, t_2 \in [0,T]$, $t_1 < t_2$ we have

$$\begin{aligned}\int_{t_1}^{t_2}\int_B \|\boldsymbol{w}^n\|\, d\gamma^n\, dt &\le \int_{t_1}^{t_2}\int_B {\rm h}(\rho^n)^{1/q}\phi(\rho^n, \boldsymbol{w}^n)^{1/p}\, d\gamma^n\, dt \\ &\le \Big(\int_{t_1}^{t_2}\int_B {\rm h}(\rho^n)\, d\gamma^n\, dt\Big)^{1/q}\Big(\int_{t_1}^{t_2}\int_B \phi(\rho^n, \boldsymbol{w}^n)\, d\gamma^n\, dt\Big)^{1/p}.\end{aligned}$$

By (24), (28) and the equivalence of $\|\cdot\|$ with the euclidean norm, the last inequality shows that there exist $C > 0$ such that

$$\int_{t_1}^{t_2}\int_B \|\boldsymbol{w}^n\|\, d\gamma^n\, dt \le C((t_2 - t_1)\gamma^n(B))^{1/q},$$

By this estimate there exist $\boldsymbol{\nu} \in \mathcal{M}_{\rm loc}(\mathbb{R}^d \times [0,T], \mathbb{R}^d)$ and a subsequence such that $\boldsymbol{\nu}^n \rightharpoonup^* \boldsymbol{\nu}$. By the lower semicontinuity property of Theorem 2.2 we obtain (29). Reasoning as in the proof of Lemma 4.5 of [13] we obtain that $(\mu, \boldsymbol{\nu})$ satisfies the continuity equation.
Finally, by Proposition 2.4 $\mu^n_t(\mathbb{R}^d)$ and $\mu_t(\mathbb{R}^d)$ do not depend on $t \in [0,T]$. □

3 The modified Wasserstein distance

In this Section we give the rigorous definition of the modified Wasserstein distance illustrated in the introduction. We deal only with the case of the distance induced by an action density function $\phi \in \mathcal{A}_p(a,b)$ for $a, b \in \mathbb{R}$ and we refer to [13] for the case $\phi \in \mathcal{A}_p(0,+\infty)$.

The proofs are almost all omitted because follows exactly as in [13, Section 5] from the results of the previous Sections.

Definition 3.1. Given a reference measure $\gamma \in \mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$, an admissible action density function $\phi \in \mathcal{A}_p(a,b)$ and the corresponding action functional Φ of Definition 2.5, for $\mu^0, \mu^1 \in \mathcal{M}_{\rm loc}(\mathbb{R}^d)$ we define

$$\mathcal{W}_{\phi,\gamma}(\mu^0,\mu^1) := \inf\left\{\Big(\int_0^1 \Phi(\mu_s, \boldsymbol{\nu}_s|\gamma)\, {\rm d}s\Big)^{1/p} : (\mu, \boldsymbol{\nu}) \in \mathcal{CE}_{\phi,\gamma}(0,1;\mu^0 \to \mu^1)\right\}. \tag{30}$$

$\mathcal{W}_{\phi,\gamma}(\mu^0,\mu^1) = +\infty$ if the set of connecting curves $\mathcal{CE}_{\phi,\gamma}(0,1;\mu^0 \to \mu^1)$ is empty.

By the compactness Proposition 2.5 we obtain the existence of constant speed minimizing geodesics. Precisely, following the proof of [13, Thm. 5.4] and Theorem 5.11 of [13] we can prove the following result.

Proposition 3.1 (Existence of geodesics, convexity and uniqueness of geodesics)**.** *Given $\gamma \in \mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$ and $\phi \in \mathcal{A}_p(a,b)$, for every $\mu_0,\ \mu_1 \in \mathcal{M}_{\rm loc}(\mathbb{R}^d)$ such that $\mathcal{W}_{\phi,\gamma}(\mu_0,\mu_1) < +\infty$ there exists a minimizing couple $(\mu,\boldsymbol{\nu})$ in* (30) *and the curve $(\mu_s)_{s\in[0,1]}$ is a constant speed geodesic for $\mathcal{W}_{\phi,\gamma}$, thus satisfying*

$$\mathcal{W}_{\phi,\gamma}(\mu_t,\mu_s) = |t-s|\mathcal{W}_{\phi,\gamma}(\mu_0,\mu_1) \qquad \forall s,t \in [0,1].$$

We have the characterization

$$\mathcal{W}_{\phi,\gamma}(\mu^0,\mu^1) = \inf\left\{\int_0^1 \Big(\Phi(\mu_s,\boldsymbol{\nu}_s|\gamma)\Big)^{1/p} \mathrm{d}s : (\mu,\boldsymbol{\nu}) \in \mathcal{CE}(0,1;\mu^0 \to \mu^1)\right\}. \tag{31}$$

Moreover $\mathcal{W}^p_{\phi,\gamma} : \mathcal{M}_{\rm loc}(\mathbb{R}^d) \times \mathcal{M}_{\rm loc}(\mathbb{R}^d) \to [0,+\infty]$ is convex, i.e. for every $\mu_i^j \in \mathcal{M}_{\rm loc}(\mathbb{R}^d)$, $i,j = 0,1$, and $\tau \in [0,1]$, if $\mu_i^\tau = (1-\tau)\mu_i^0 + \tau\mu_i^1$,

$$\mathcal{W}^p_{\phi,\gamma}(\mu_0^\tau,\mu_1^\tau) \le (1-\tau)\mathcal{W}^p_{\phi,\gamma}(\mu_0^0,\mu_1^0) + \tau\mathcal{W}^p_{\phi,\gamma}(\mu_0^1,\mu_1^1). \tag{32}$$

If ϕ is strictly convex then for every $\mu_0,\mu_1 \in \mathcal{M}_{\rm loc}(\mathbb{R}^d)$ with $\mathcal{W}_{\phi,\gamma}(\mu_0,\mu_1) < +\infty$ there exists a unique *minimizer $(\mu,\boldsymbol{\nu}) \in \mathcal{CE}_{\phi,\gamma}(0,1;\mu_0 \to \mu_1)$ of* (30).

Proposition 3.2. *Given $\gamma \in \mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$ and $\phi \in \mathcal{A}_p(a,b)$, we have that $\mathcal{W}_{\phi,\gamma}$ is a pseudo-distance on $\mathcal{M}_{\rm loc}(\mathbb{R}^d)$; i.e. $\mathcal{W}_{\phi,\gamma}$ satisfies the axiom of the distance but can assume the value $+\infty$.*
The topology induced by $\mathcal{W}_{\phi,\gamma}$ on $\mathcal{M}_{\rm loc}(\mathbb{R}^d)$ is stronger than or equivalent to the weak one.*
Bounded sets with respect to $\mathcal{W}_{\phi,\gamma}$ are weakly relatively compact.*

Proof. The verification of the axioms of the distance is straightforward except for the triangular inequality where we use the gluing of solutions of Lemma 2.3 and the characterization (31).

In order to prove the topological property, reasoning as in the proof of Proposition 2.4 we obtain that

$$\Big|\int_{\mathbb{R}^d} \zeta\, d\mu_1 - \int_{\mathbb{R}^d} \zeta\, d\mu_0\Big| \le \sup|\nabla\zeta|(\mathrm{h}_{max}\gamma(\mathrm{supp}(\zeta)))^{1/q}\mathcal{W}_{\phi,\gamma}(\mu_0,\mu_1) \tag{33}$$

for every $\zeta \in C^1_c(\mathbb{R}^d)$. Since $C^1_c(\mathbb{R}^d)$ is dense in $C_c(\mathbb{R}^d)$ we obtain the assertion on the topology induced by the distance and on the relative compactness. □

The following lower semi-continuity result can be proved exactly as Theorem 5.6 of [13] by using the compactness Proposition 2.5.

Proposition 3.3 (Lower semi-continuity)**.** *If $\gamma^n \rightharpoonup^* \gamma$ in $\mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$, $\mu_0^n \rightharpoonup^* \mu_0$, $\mu_1^n \rightharpoonup^* \mu_1$ in $\mathcal{M}_{\rm loc}(\mathbb{R}^d)$ and $\phi^n,\phi \in \mathcal{A}_p(a,b)$, such that $\phi^n \le \phi^{n+1}$ and ϕ^n converges pointwise to ϕ, then*

$$\liminf_{n\to+\infty} \mathcal{W}_{\phi^n,\gamma^n}(\mu_0^n,\mu_1^n) \ge \mathcal{W}_{\phi,\gamma}(\mu_0,\mu_1). \tag{34}$$

The following completeness result can be proved as in Theorem 5.7 of [13] ad using Proposition 3.3. The final assertion about the equality of the signed mass follows from Proposition 2.4.

Proposition 3.4 (Completeness and equality of the mass)**.** *Given $\gamma \in \mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$ and $\phi \in \mathcal{A}_p(a,b)$, we have that the space $\mathcal{M}_{\rm loc}(\mathbb{R}^d)$ endowed with the pseudo-distance $\mathcal{W}_{\phi,\gamma}$ is complete.*
Given a measure $\sigma \in \mathcal{M}_{\rm loc}(\mathbb{R}^d)$, the space $\mathcal{M}[\sigma] := \big\{\mu \in \mathcal{M}_{\rm loc}(\mathbb{R}^d) : \mathcal{W}_{\phi,\gamma}(\mu,\sigma) < +\infty\big\}$ is a complete metric space.
If $\tilde{\mathrm{m}}_{-q}(\gamma) < +\infty$ then $\mu(\mathbb{R}^d) = \sigma(\mathbb{R}^d)$ for every $\mu \in \mathcal{M}[\sigma]$.

The following results follows from 3 and 4 of Theorem 2.2.

Proposition 3.5 (Monotonicity)**.** *If* $\phi_1 \le \phi_2$ *then*

$$\mathcal{W}_{\phi_1,\gamma}(\mu_0,\mu_1) \le \mathcal{W}_{\phi_2,\gamma}(\mu_0,\mu_1),$$

for every μ_0, $\mu_1 \in \mathcal{M}_{\rm loc}(\mathbb{R}^d)$.
Moreover, if $(0,0) \in \mathrm{dom}(\phi_i)$, $i = 1,2$ *and* $\gamma_1 \le \gamma_2$ *then*

$$\mathcal{W}_{\phi_1,\gamma_2}(\mu_0,\mu_1) \le \mathcal{W}_{\phi_2,\gamma_1}(\mu_0,\mu_1),$$

for every μ_0, $\mu_1 \in \mathcal{M}_{\rm loc}(\mathbb{R}^d)$.

Proposition 3.6 (Approximation by convolution)**.** *Let* $k \in C_c^\infty(\mathbb{R}^d)$ *be a nonnegative convolution kernel, with* $\int_{\mathbb{R}^d} k(x)\,\mathrm{d}x = 1$ *and* $\mathrm{supp}(k) = \overline{B}_1(0)$*, and let* $k_\varepsilon(x) := \varepsilon^{-d}k(x/\varepsilon)$*. For every* $\mu_0,\mu_1 \in \mathcal{M}(\mathbb{R}^d)$

$$\mathcal{W}_{\phi,\gamma*k_\varepsilon}(\mu_0 * k_\varepsilon, \mu_1 * k_\varepsilon) \le \mathcal{W}_{\phi,\gamma}(\mu_0,\mu_1) \quad \forall\,\varepsilon > 0; \tag{35}$$

$$\lim_{\varepsilon\to 0} \mathcal{W}_{\phi,\gamma*k_\varepsilon}(\mu_0 * k_\varepsilon, \mu_1 * k_\varepsilon) = \mathcal{W}_{\phi,\gamma}(\mu_0,\mu_1). \tag{36}$$

The following proposition deals with a control of the moments and a comparison between the convergence with respect to $\mathcal{W}_{\phi,\gamma}$ and the standard Wasserstein distance defined in (1).

Proposition 3.7. *Let* $\gamma \in \mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$ *be satisfying* $\tilde{\mathrm{m}}_r(\gamma) < +\infty$ *for some* $r \in \mathbb{R}$ *and* $\phi \in \mathcal{A}_p(a,b)$*.*
If μ_0, $\mu_1 \in \mathcal{M}_{\rm loc}(\mathbb{R}^d)$ *satisfy* $\mathcal{W}_{\phi,\gamma}(\mu_0,\mu_1) < +\infty$*, then, setting* $C := \max\{|a|,|b|\}$*, we have*

$$\tilde{\mathrm{m}}_\delta(|\mu_i|) \le C\tilde{\mathrm{m}}_r(\gamma), \qquad \text{for } i = 0,1, \quad \forall\,\delta \le r. \tag{37}$$

If $r \ge 1$ *and* $a \ge 0$ *then the convergence with respect to* $\mathcal{W}_{\phi,\gamma}$ *in* $\mathcal{M}[\sigma]$*, for some non-negative measure* σ *satisfying* $\sigma(\mathbb{R}^d) < +\infty$*, implies the convergence with respect to the* r*-Wasserstein distance* W_r*.*

Proof. Denoting by $1 \vee |x| = \max\{1,|x|\}$, given $C = \max\{|a|,|b|\}$, $\delta \le r$ and a Borel set $A \subset \mathbb{R}^d$, by Proposition 2.3 we obtain

$$\int_A (1\vee|x|)^\delta\,\mathrm{d}|\mu_i|(x) = \int_A (1\vee|x|)^\delta|\rho_i(x)|\,\mathrm{d}\gamma(x) \le C\int_A (1\vee|x|)^\delta\,\mathrm{d}\gamma(x) \le C\int_A (1\vee|x|)^r\,\mathrm{d}\gamma(x). \tag{38}$$

Choosing $A = \mathbb{R}^d$ in (38) we obtain (37).
If μ_n is a sequence in $\mathcal{M}[\sigma]$ converging to μ with respect to $\mathcal{W}_{\phi,\gamma}$, then, by Proposition 3.2, μ_n weakly* converges to μ and, by Proposition 3.4, $\mu_n(\mathbb{R}^d) = \mu(\mathbb{R}^d) = \mu(\sigma)$ because of the assumption on the moment of γ and $r \ge 1$. By (38) with $\delta = 0$ we have that the sequence μ_n is tight and then μ_n narrowly converges to μ. Since (38) implies that the r-moments of μ_n are uniformly equiintegrable we obtain that (see Lemma 5.1.7 of [2]) $\tilde{\mathrm{m}}_r(\mu_n)$ converges to $\tilde{\mathrm{m}}_r(\mu)$ and we conclude. □

In particular the previous Proposition applies to the case $\gamma(\mathbb{R}^d) < +\infty$.
In the next proposition we state a simple comparison with the standard Wasserstein distance (1).

Proposition 3.8 (Comparison with Wasserstein distance)**.** *Let* $p > 1$, $\phi \in \mathcal{A}_p(0,M)$, $\Omega \subset \mathbb{R}^d$ *an open convex set and* $\gamma_\Omega = \chi_\Omega\mathscr{L}^d$*. If* μ_i, $i = 0,1$*, are two absolutely continuous measures with respect to* γ_Ω*,* $\mu_i = \rho_i\gamma_\Omega$*, such that* $0 \le \rho_i(x) \le M' < M$*,* $\tilde{\mathrm{m}}_p(\mu_i) < +\infty$ *for* $i = 0,1$ *and* $\mu_0(\mathbb{R}^d) = \mu_1(\mathbb{R}^d)$*, then there exists a constant* C*, depending only on* M'*,* ϕ *and* p*, such that*

$$\mathcal{W}_{\phi,\gamma_\Omega}(\mu_0,\mu_1) \le CW_p(\mu_0,\mu_1) < +\infty, \tag{39}$$

where W_p *denotes the standard* p*-Wasserstein distance.*

Proof. Let h be given by Proposition 2.2. Since h is concave and positive on $(0, M)$, we have that

$$h(\rho) \geq \frac{h(M')}{M'}\rho, \qquad \forall\, \rho \in (0, M'),$$

and, consequently,

$$\phi(\rho, \boldsymbol{w}) \leq \frac{|\boldsymbol{w}|^p}{h(\rho)^{p-1}} \leq \left(\frac{M'}{h(M')}\right)^{p-1} \frac{|\boldsymbol{w}|^p}{\rho^{p-1}} \qquad \forall \rho \in (0, M'). \tag{40}$$

Since the p-moments of μ_0 and μ_1 are finite, taking the geodesic interpolant μ_t between μ_0 and μ_1 for the standard p-Wasserstein distance $W_p(\mu_0, \mu_1)$, and denoting by ρ_t the density of μ_t, we have that $\rho_t \leq M'$ (see the proof of [13, Theorem 5.24]). Since Ω is convex, the support of μ_t belongs to $\overline{\Omega}$ and, denoting by $\psi(\rho, \boldsymbol{w}) := \frac{|\boldsymbol{w}|^p}{\rho^{p-1}}$, we have that $\mathcal{W}_{\psi,\gamma_\Omega} = W_p$ for all the measures with support in $\overline{\Omega}$. Then, by (40), and recalling Proposition 3.5 we obtain (39). ☐

4 Measures at finite distance and convergence

In this section we give sufficient conditions for the finiteness of the distance between two measures. We study also the relation between the convergence with respect to the distance and the weak-$*$ one. The first result concerns measures defined on the whole space $\mathbb{R}^d$ with the reference measure $\gamma = \mathscr{L}^d$, whereas the second one deals with measures defined on a bounded convex domain Ω with the reference measure $\gamma_\Omega = \mathscr{L}^d_{|\Omega}$.

4.1 The case of reference measure $\mathscr{L}^d$

Theorem 4.1 (Connectivity in $\mathbb{R}^d$)**.** *Let $p > 1$ and $\phi \in \mathcal{A}_p(0, M)$. If μ_i, $i = 0, 1$, are two absolutely continuous measures $\mu_i = \rho_i \mathscr{L}^d$, such that $0 \leq \rho_i(x) \leq M$, $\tilde{\mathrm{m}}_p(\mu_i) < +\infty$ for $i = 0, 1$ and $\mu_0(\mathbb{R}^d) = \mu_1(\mathbb{R}^d)$, then there exists a constant $C > 0$ depending only on ϕ, d and p such that*

$$\mathcal{W}_{\phi,\mathscr{L}^d}(\mu_0, \mu_1) \leq C(\tilde{\mathrm{m}}_p(\mu_0) + \tilde{\mathrm{m}}_p(\mu_1)) < +\infty. \tag{41}$$

We observe that the inequality (41) holds in the case of the standard Wasserstein distance (it is a very easy consequence of the definition (1)).

Proof. Let h be given by Proposition 2.2. Since h is concave and non-negative, there exists $\tilde{h} : [0, M] \to [0, +\infty)$ of the form $\tilde{h}(\rho) = A\rho(M/B - B\rho)$ for $A, B > 0$ such that $\tilde{h}(\rho) \leq h(\rho)$ in $[0, M]$. Hence $\mathcal{W}_{\phi,\mathscr{L}^d}(\mu_0, \mu_1) \leq \mathcal{W}_{\phi_h,\mathscr{L}^d}(\mu_0, \mu_1) \leq \mathcal{W}_{\phi_{\tilde{h}},\mathscr{L}^d}(\mu_0, \mu_1)$. Thanks to this observation, it is sufficient to prove the result under the assumption that $M = 1$, $h(\rho) = \rho(1 - \rho)$ and $0 \leq \rho_i(x) \leq 1$, for $i = 0, 1$.

Defining

$$\tilde{\mu}_i = \tilde{\rho}_i \mathscr{L}^d = 2\mathrm{Id}_\# \mu_i, \tag{42}$$

where Id denotes the identity map in $\mathbb{R}^d$, we prove that there exists a constant $C_{p,d}$, depending only on p and d, such that

$$W_{\phi_h,\mathscr{L}^d}(\mu_i, \tilde{\mu}_i) < C_{p,d} \tilde{\mathrm{m}}_p(\mu_i) \quad \text{ for } i = 0, 1. \tag{43}$$

Indeed, for $t \in [0, 1]$, taking $T_t(x) := (1 + t^p)x$ and $\mu_t := (T_t)_\# \mu_i = \rho_t \mathscr{L}^d$, by (12) we have that $\rho_t(y) = \frac{1}{(1+t^p)^d} \rho_i \left(\frac{y}{1+t^p}\right)$. Defining $\boldsymbol{v}_t(x) := \dot{T}_t \circ T_t^{-1}(x) = \frac{(pt^{p-1})}{1+t^p} x$, and $\boldsymbol{w}_t = \boldsymbol{v}_t \rho_t$, $\boldsymbol{\nu}_t = \boldsymbol{w}_t \mathscr{L}^d$ it is easy to check that $(\mu_t, \boldsymbol{\nu}_t)_{t\in(0,1)} \in \mathcal{CE}(0, 1; \mu_i \to \tilde{\mu}_i)$. By elementary computations, using the definition of μ_t and $\boldsymbol{v}_t$, we have

$$\int_0^1 \int_{\mathbb{R}^d} \frac{|\boldsymbol{w}_t(x)|^p}{(\rho_t(x)(1-\rho_t(x)))^{p-1}}\,dx\,dt = \int_0^1 \int_{\mathbb{R}^d} \frac{|\boldsymbol{v}_t(x)|^p \rho_t(x)}{(1-\rho_t(x))^{p-1}}\,dx\,dt$$
$$= \int_0^1 \int_{\mathbb{R}^d} \frac{|\boldsymbol{v}_t(x)|^p}{(1-\rho_t(x))^{p-1}}\,d\mu_t(x)\,dt$$
$$= \int_0^1 \int_{\mathbb{R}^d} \frac{|\boldsymbol{v}_t(T_t(x))|^p}{(1-\rho_t(T_t(x)))^{p-1}}\,d\mu_i(x)\,dt$$
$$= \int_0^1 \int_{\mathbb{R}^d} \frac{(pt^{p-1})^p|x|^p}{(1-(1+t^p)^{-d}\rho_i(x))^{p-1}}\,d\mu_i(x)\,dt.$$

Since $\rho_i(x) \le 1$ and $(1+t^p)^d \ge 1 + dt^p$ we have

$$\frac{1}{1-(1+t^p)^{-d}\rho_i(x)} \le \frac{1}{1-(1+t^p)^{-d}} = \frac{(1+t^p)^d}{(1+t^p)^d - 1} \le \frac{(1+t^p)^d}{dt^p}.$$

Then

$$\int_0^1 \int_{\mathbb{R}^d} \frac{(pt^{p-1})^p|x|^p}{(1-(1+t^p)^{-d}\rho_i(x))^{p-1}}\,d\mu_i(x)\,dt \le \int_0^1 \int_{\mathbb{R}^d} \frac{(pt^{p-1})^p(1+t^p)^{d(p-1)}}{(dt^p)^{p-1}}|x|^p\,d\mu_i(x)\,dt$$
$$\le \tilde{\mathrm{m}}_p(\mu_i) \int_0^1 p^p d^{1-p}(1+t^p)^{d(p-1)}\,dt,$$

and (43) follows with $C_{p,d} = \displaystyle\int_0^1 p^p d^{1-p}(1+t^p)^{d(p-1)}\,dt$.

Finally, by the triangular inequality, we have

$$\mathcal{W}_{\phi_h,\mathscr{L}^d}(\mu_0,\mu_1) \le \mathcal{W}_{\phi_h,\mathscr{L}^d}(\mu_0,\tilde{\mu}_0) + \mathcal{W}_{\phi_h,\mathscr{L}^d}(\tilde{\mu}_0,\tilde{\mu}_1) + \mathcal{W}_{\phi_h,\mathscr{L}^d}(\tilde{\mu}_1,\mu_1). \tag{44}$$

Since by (12) we have $\tilde{\rho}_i(x) = 2^{-d}\rho_i(x/2) \le 2^{-d}$ and $\tilde{\mathrm{m}}_p(\tilde{\mu}_i) = 2^p\tilde{\mathrm{m}}_p(\mu_i) < +\infty$, by Proposition 3.8 applied to $\tilde{\mu}_0, \tilde{\mu}_1$, and observing that $W_p(\tilde{\mu}_0,\tilde{\mu}_1) \le \tilde{\mathrm{m}}_p(\tilde{\mu}_0) + \tilde{\mathrm{m}}_p(\tilde{\mu}_1)$ (it is a simple consequence of the definition (1)), by (43) and (44) we obtain (41) . □

Given $M > 0$ and $c > 0$ we define the set of measures

$$\mathcal{M}^+_{M,c}(\mathbb{R}^d) := \{\mu \in \mathcal{M}^+(\mathbb{R}^d) : \mu = \rho\mathscr{L}^d,\ 0 \le \rho \le M,\ \mu(\mathbb{R}^d) = c,\ \tilde{\mathrm{m}}_p(\mu) < +\infty\}.$$

Theorem 4.2. *Let $p > 1$ and $\phi \in \mathcal{A}_p(0,M)$. If $(\mu_n)_{n\in\mathbb{N}}$ is a sequence in $\mathcal{M}^+_{M,c}(\mathbb{R}^d)$ weakly-$*$ convergent to $\mu \in \mathcal{M}^+_{M,c}(\mathbb{R}^d)$, such that*

$$\tilde{\mathrm{m}}_p(\mu_n) \to \tilde{\mathrm{m}}_p(\mu), \tag{45}$$

then

$$\lim_{n\to+\infty} \mathcal{W}_{\phi,\mathscr{L}^d}(\mu_n,\mu) = 0.$$

Proof. Let $\bar{\mu} = \bar{\rho}\mathscr{L}^d \in \mathcal{M}^+_{M,c}$ be a fixed auxiliary measure such that $M' := \sup\bar{\rho} < M$.
For every $\lambda \in (0,1)$, we define the convex combinations $\mu^\lambda_n := (1-\lambda)\mu_n + \lambda\bar{\mu}$ and $\mu^\lambda := (1-\lambda)\mu + \lambda\bar{\mu}$. Denoting by ρ^λ_n the density of μ^λ_n with respect to $\mathscr{L}^d$ we have $\rho^\lambda_n \le 1 - \lambda(M - M')$. By Proposition 3.8 and the convexity of the p-power of the standard p-Wasserstein distance (Proposition 3.1 applied to $\phi(\rho,\boldsymbol{w}) = |\boldsymbol{w}|^p/\rho^{p-1}$ or [26]) we have

$$\mathcal{W}^p_{\phi,\mathscr{L}^d}(\mu^\lambda_n,\mu^\lambda) \le CW^p_p(\mu^\lambda_n,\mu^\lambda) \le C(1-\lambda)W^p_p(\mu_n,\mu). \tag{46}$$

By the convergence of the p-moments (45) and the weak-$*$ convergence we have (see [2] or [26])

$$\lim_{n\to+\infty} W_p(\mu_n,\mu)=0. \tag{47}$$

Moreover for the convexity of $\mathcal{W}^p_{\phi,\mathscr{L}^d}$ (Proposition 3.1) we have

$$\mathcal{W}^p_{\phi,\mathscr{L}^d}(\mu_n,\mu_n^\lambda)\le\lambda\mathcal{W}^p_{\phi,\mathscr{L}^d}(\mu_n,\bar\mu),\qquad \mathcal{W}^p_{\phi,\mathscr{L}^d}(\mu,\mu^\lambda)\le\lambda\mathcal{W}^p_{\phi,\mathscr{L}^d}(\mu,\bar\mu). \tag{48}$$

Since

$$\mathcal{W}_{\phi,\mathscr{L}^d}(\mu_n,\mu)\le\mathcal{W}_{\phi,\mathscr{L}^d}(\mu_n,\mu_n^\lambda)+\mathcal{W}_{\phi,\mathscr{L}^d}(\mu_n^\lambda,\mu^\lambda)+\mathcal{W}_{\phi,\mathscr{L}^d}(\mu^\lambda,\mu), \tag{49}$$

by (46), (47) and (48) we have

$$\limsup_{n\to+\infty}\mathcal{W}_{\phi,\mathscr{L}^d}(\mu_n,\mu)\le\lambda^{1/p}\Big(\sup_n\mathcal{W}_{\phi,\mathscr{L}^d}(\mu_n,\bar\mu)+\mathcal{W}_{\phi,\mathscr{L}^d}(\mu,\bar\mu)\Big). \tag{50}$$

By (45) and Theorem 4.1 we obtain

$$\sup_n\mathcal{W}_{\phi,\mathscr{L}^d}(\mu_n,\bar\mu)<+\infty. \tag{51}$$

Since $\lambda>0$ is arbitrary, (50) and (51) imply

$$\limsup_{n\to+\infty}\mathcal{W}_{\phi,\mathscr{L}^d}(\mu_n,\mu)=0$$

and we conclude. □

We recall that the convergence with respect to the standard Wasserstein distance W_p is equivalent to the weak-$*$ convergence and the convergence of the p-moments $\tilde{\mathrm{m}}_p$ (see [26] or [2]). Consequently, Theorem 4.2 states that the convergence with respect to W_p in $\mathcal{M}^+_{M,c}(\mathbb{R}^d)$ implies the convergence with respect to $\mathcal{W}_{\phi,\mathscr{L}^d}$ for every $\phi\in\mathcal{A}_p(0,M)$. We observe that this property is not true in the case of $\phi\in\mathcal{A}_p(0,+\infty)$, where only a result like Proposition 3.8 hold (see Theorem 5.24 of [13]).

4.2 The case of the reference measure $\chi_\Omega\mathscr{L}^d$ with Ω bounded convex

When the reference measure is $\gamma_\Omega:=\chi_\Omega\mathscr{L}^d$, where Ω is a bounded convex smooth domain, we have the following result of finiteness of the distance and of boundedness of the space of admissible measures.

Theorem 4.3. *Let $\phi\in\mathcal{A}_2(a,b)$ and $\gamma_\Omega:=\chi_\Omega\mathscr{L}^d$ with $\Omega\subset\mathbb{R}^d$ a bounded convex smooth domain. For every $c\in(a\mathscr{L}^d(\Omega),b\mathscr{L}^d(\Omega))$ we define the set of measures*

$$\mathcal{M}_{(a,b),c}(\Omega):=\{\mu\in\mathcal{M}(\overline{\Omega}):\mu=\rho\gamma_\Omega,\ a\le\rho\le b,\ \mu(\overline{\Omega})=c\}.$$

The space $\mathcal{M}_{(a,b),c}(\Omega)$ endowed with the distance $\mathcal{W}_{\phi,\gamma_\Omega}$ is bounded. In particular $\mathcal{W}_{\phi,\gamma_\Omega}(\mu_0,\mu_1)<+\infty$ for every $\mu_0,\mu_1\in\mathcal{M}_{(a,b),c}(\Omega)$.

Proof. Defining $\mu_\infty:=\frac{c}{\mathscr{L}^d(\Omega)}\gamma_\Omega$, we prove that

$$\sup_{\mu_0\in\mathcal{M}_{(a,b),c}(\Omega)}\mathcal{W}_{\phi,\gamma_\Omega}(\mu_0,\mu_\infty)<+\infty. \tag{52}$$

Let h be given by Proposition 2.2.

For $\mu_0=\rho_0\gamma_\Omega\in\mathcal{M}_{(a,b),c}(\Omega)$, let $\rho:(0,+\infty)\times\Omega\to\mathbb{R}$ be the solution of Cauchy-Neumann problem for the heat equation

$$\begin{cases}\partial_t\rho-\Delta\rho=0 & \text{in } (0,+\infty)\times\Omega\\ \rho(0,\cdot)=\rho_0 & \text{in } \Omega\\ \nabla\rho\cdot\boldsymbol{n}=0 & \text{on } (0,\infty)\times\partial\Omega.\end{cases} \tag{53}$$

We use the notation $\rho_t := \rho(t, \cdot)$ and $S_t(\mu_0) := \rho_t \gamma_\Omega$.

Defining the convex function $U : (a, b) \to \mathbb{R}$ by

$$U''(r) = \frac{1}{h(r)}, \qquad U'((a + b)/2) = 0, \quad U((a + b)/2) = 0 \tag{54}$$

and the entropy functional

$$\mathscr{U}(\rho) = \int_\Omega U(\rho(x))\, dx,$$

we have the following entropy dissipation inequality

$$\mathscr{U}(\rho_T) - \mathscr{U}(\rho_0) \leq -\int_0^T \int_\Omega \frac{|\nabla \rho_s|^2}{h(\rho_s)}\, dx\, ds. \tag{55}$$

The inequality (55) can be obtained, in the case of smooth initial datum, with a simple computation and, in the general case, by a convolution approximation argument.

By Lemma 4.1, observing that in our case $\rho_\infty = \frac{c}{\mathscr{L}^d(\Omega)}$, we can prove that there exists $T > 0$, independent on μ_0, such that

$$\rho_t \leq \rho_\infty + \frac{b - \rho_\infty}{2}, \qquad \forall t \geq T. \tag{56}$$

By the triangular inequality we have that

$$\mathcal{W}_{\phi,\gamma_\Omega}(\mu_0, \mu_\infty) \leq \mathcal{W}_{\phi,\gamma_\Omega}(\mu_0, S_T(\mu_0)) + \mathcal{W}_{\phi,\gamma_\Omega}(S_T(\mu_0), \mu_\infty). \tag{57}$$

Since h is concave and Ω is bounded, it is not difficult to see that $\mathscr{U}$ is bounded in $\mathcal{M}_{(a,b),c}(\Omega)$, and recalling (17) we have $\mathcal{W}_{\phi,\gamma_\Omega}(\mu_0, S_T(\mu_0)) \leq \int_0^T \int_\Omega \frac{|\nabla \rho_s|^2}{h(\rho_s)}\, dx\, ds$, consequently (55) implies that

$$\sup_{\mu_0 \in \mathcal{M}_{(a,b),c}(\Omega)} \mathcal{W}_{\phi,\gamma_\Omega}(\mu_0, S_T(\mu_0)) < +\infty. \tag{58}$$

Since

$$\mathcal{W}_{\phi,\gamma_\Omega}(\mu, \nu) = \mathcal{W}_{\tilde{\phi},\gamma_\Omega}(\mu - a\gamma_\Omega, \nu - a\gamma_\Omega), \qquad \text{where } \tilde{\phi}(r, \boldsymbol{w}) := \phi(r + a, \boldsymbol{w}), \tag{59}$$

considering the new densities $\tilde{\rho} := \rho - a$, and using (56), by Proposition 3.8 we obtain

$$\sup_{\mu_0 \in \mathcal{M}_{(a,b),c}(\Omega)} \mathcal{W}_{\phi,\gamma_\Omega}(S_T(\mu_0), \mu_\infty) \leq C \sup_{\mu_0 \in \mathcal{M}_{(a,b),c}(\Omega)} W_2(S_T(\mu_0) - a\gamma_\Omega, \mu_\infty - a\gamma_\Omega) < +\infty, \tag{60}$$

because of the boundedness of the Wasserstein distance on the set of measures defined on the bounded convex set Ω. Finally, we conclude by (57), (58) and (60). □

Also in this case, following the proof of Theorem 4.2, and using the equality (59), Proposition 3.8 and Theorem 4.3, we can prove the following Theorem.

Theorem 4.4. *Let $\phi \in \mathcal{A}_2(a, b)$ and $\gamma_\Omega := \chi_\Omega \mathscr{L}^d$ with $\Omega \subset \mathbb{R}^d$ a bounded convex smooth domain. If $(\mu_n)_{n \in \mathbb{N}}$ is a sequence in $\mathcal{M}_{(a,b),c}(\Omega)$ weakly-$*$ convergent to $\mu \in \mathcal{M}_{(a,b),c}(\Omega)$, then*

$$\lim_{n \to +\infty} \mathcal{W}_{\phi,\gamma_\Omega}(\mu_n, \mu) = 0.$$

We recall that the space of non-negative measures with fixed mass $c > 0$, supported on a bounded convex open set, is bounded with respect to the standard Wasserstein distance (easy consequence of the definition), and the convergence with respect to the standard Wasserstein distance is equivalent to the weak* one on this set. Theorems 4.3 and 4.4 state that the analogous properties hold for the space $\mathcal{M}_{(a,b),c}(\Omega)$ endowed with the distance $\mathcal{W}_{\phi,\gamma_\Omega}$.

4.2.1 Appendix: decay for heat equation

In this appendix we recall a standard result on the asymptotic behavior of the heat equation. Since it seems not simple to find it in this form, we also give a proof.

Lemma 4.1. *Let Ω be a convex smooth domain of $\mathbb{R}^d$. If $\rho_0 : \Omega \to [a, b]$, and $\rho : (0, +\infty) \times \Omega \to \mathbb{R}$ denotes the solution of the problem*

$$\begin{cases} \partial_t \rho - \Delta \rho = 0 & \text{in } (0, +\infty) \times \Omega \\ \rho(0, \cdot) = \rho_0 & \text{in } \Omega \\ \nabla \rho \cdot \boldsymbol{n} = 0 & \text{on } (0, \infty) \times \partial\Omega, \end{cases} \tag{61}$$

then there exist two constants $C >$ and $\lambda > 0$, depending only on a, b and Ω such that

$$||\rho_s - \rho_\infty||_{L^\infty(\Omega)} \le C e^{-\lambda s}, \qquad \forall\, s \ge 0, \tag{62}$$

where $\rho_s := \rho(s, \cdot)$ and $\rho_\infty := \frac{1}{\mathscr{L}^d(\Omega)} \int_\Omega \rho_0(x)\, dx$.

Proof. Since $\partial_t(\rho_t - \rho_\infty) - \Delta(\rho_t - \rho_\infty) = 0$ with homogeneous Neumann boundary conditions, multiplying this equation by $\rho_t - \rho_\infty$ and integrating by parts we obtain the identity

$$\frac{d}{dt} ||\rho_t - \rho_\infty||^2_{L^2(\Omega)} + 2||\nabla \rho_t||^2_{L^2(\Omega)} = 0. \tag{63}$$

By Poincaré's inequality, there exists a constant C_P depending only on Ω such that

$$||\nabla \rho_t||^2_{L^2(\Omega)} \ge C_P ||\rho_t - \rho_\infty||^2_{L^2(\Omega)}, \tag{64}$$

and from (63) we immediately obtain the $L^2(\Omega)$ exponential decay

$$||\rho_t - \rho_\infty||_{L^2(\Omega)} \le e^{-C_P t} ||\rho_0 - \rho_\infty||_{L^2(\Omega)}, \qquad \forall\, t \ge 0. \tag{65}$$

The $L^2(\Omega) - W^{1,\infty}(\Omega)$ interpolation inequality (see for instance [6, Complements of Chapter IX] or [23]), states that there exist a constant C depending only on Ω such that

$$||\rho_t - \rho_\infty||_{L^\infty(\Omega)} \le C ||\rho_t - \rho_\infty||^{2/(d+2)}_{L^2(\Omega)} ||\rho_t - \rho_\infty||^{d/(d+2)}_{W^{1,\infty}(\Omega)} \qquad \forall\, t \ge 0. \tag{66}$$

In order to get a uniform bound of the L^∞ norm of the gradient, we define $v(t, x) := \rho_t^2(x) + t|\nabla \rho_t(x)|^2$, which solves the problem

$$\begin{cases} \partial_t v - \Delta v \le 0 & \text{in } (0, +\infty) \times \Omega \\ v(0, \cdot) = \rho_0^2 & \text{in } \Omega \\ \nabla v \cdot \boldsymbol{n} \le 0 & \text{on } (0, \infty) \times \partial\Omega. \end{cases} \tag{67}$$

Indeed, by a simple computation we have that v satisfies the partial differential inequality in (67). In order to obtain the boundary condition satisfied by v we have $\nabla v \cdot \boldsymbol{n} = \nabla \rho^2 \cdot \boldsymbol{n} + t \nabla |\nabla \rho|^2 \cdot \boldsymbol{n} = t \nabla |\nabla \rho|^2 \cdot \boldsymbol{n}$ because of the boundary condition in (61). Moreover, by the smoothness and the convexity of Ω, we have that $\nabla |\nabla \rho|^2 \cdot \boldsymbol{n} \le 0$ (see for instance [19, Lemma 5.1]).

The maximum principle for problem (67) (see for instance [17]) states that $v(t, x) \le ||\rho_0^2||_{L^\infty(\Omega)}$. In particular we have

$$\sqrt{t} ||\nabla \rho_t||_{L^\infty(\Omega)} \le ||\rho_0||_{L^\infty(\Omega)} \le \max(|a|, |b|). \tag{68}$$

The inequality (62) follows from (66) and (68) (for $t \ge 1$ for instance) and (65), recalling that $||\rho_t - \rho_\infty||_{L^\infty(\Omega)} \le 2 \max(|a|, |b|)$. □

Acknowledgements: The authors would like to thank Giuseppe Savaré for useful suggestions on this work.

References

[1] L. Ambrosio, G. Buttazzo, *Weak lower semicontinuous envelope of functionals defined on a space of measures*, Ann. Mat. Pura Appl., 150 (1988), pp.311–339.

[2] L. Ambrosio, N. Gigli, G. Savaré, Gradient flows in Metric Spaces and in the Space of Probability Measures, Birkäuser Verlag, Basel 2005

[3] L. Ambrosio, N. Fusco, D. Pallara, Functions of bounded variation and free discontinuity problems, Oxford Mathematical Monographs, Claredon Press, Oxford, 2000.

[4] J.-D. Beneamou, Y. Brenier,*A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numer. Math., 84 (2000), pp.375–393.

[5] A. Bertozzi, *The mathematics of moving contact lines in thin liquid films*, Notices Amer. Math. Soc., 45 (1998), pp. 689-697.

[6] H. Brezis. Analyse fonctionnelle. Masson, Paris, 1983.

[7] M. Burger, M. di Francesco and Y. Dolak, *The Keller-Segel model for chemotaxis with prevention of overcrowding: linear vs. nonlinear diffusion*, SIAM J. Math. Anal., 38 (2006), pp. 1288–1315.

[8] M. Burger and M. Di Francesco, *Large time behavior of nonlocal aggregation models with nonlinear diffusion*, Netw. Heterog. Media, 3 (2008), pp. 749–785.

[9] J. A. Carrillo, P. Laurençot and J. Rosado, *Fermi-Dirac-Fokker-Planck Equation: Well-posedness & Long-time Asymptotics*, J. Differential Equations, 247 (2009), pp. 2209–2234.

[10] J.A. Carrillo, S. Lisini, G. Savaré, D. Slepčev, *Nonlinear mobility continuity equations and generalized displacement convexity* arXiv:0901.3978v1 [math.AP] to appear on J. Funct. Anal.

[11] J. A. Carrillo, J. Rosado and F. Salvarani, *1D nonlinear Fokker-Planck equations for fermions and bosons*, Appl. Math. Lett., 21 (2008), pp. 148-154.

[12] M. Di Francesco and J. Rosado, *Fully parabolic Keller-Segel model for chemotaxis with prevention of overcrowding*, Nonlinearity, 21 (2008), pp. 2715–2730.

[13] J. Dolbeault, B. Nazaret, G. Savaré, *A new class of transport distances between measures*, Calc. Var. Partial Differential Equations, 34 (2009), pp. 193–231.

[14] C. M. Elliott, H. Garcke, *On the Cahn-Hilliard equation with degenerate mobility*, SIAM J. Math. Anal., 27 (1996), pp. 404–423.

[15] T. D. Frank, *Classical Langevin equations for the free electron gas and blackbody radiation*, J. Phys. A, 37 (2004), pp. 3561–3567.

[16] T. D. Frank, Nonlinear Fokker-Planck Equations, Springer Series in Synergetics, Springer, 2005.

[17] A. Friedman, Partial differential equations of parabolic type. Prentice-Hall, 1964.

[18] G. Giacomin and J. Lebowitz, *Phase segregation dynamics in particle systems with long range interactions. I. Macroscopic limits*, J. Statist. Phys., 87 (1997), pp. 37–61.

[19] U. Gianazza, G. Savaré and G. Toscani, *The Wasserstein gradient flow of the Fisher information and the quantum drift-diffusion equation*, Arch. Ration. Mech. Anal., 194 (2009), pp. 133–220.

[20] G. Kaniadakis, *Generalized Boltzmann equation describing the dynamics of bosons and fermions*, Phys. Lett. A, 203 (1995), pp. 229–234.

[21] G. Kaniadakis, P. Quarati, *Kinetic equation for classical particles obeying an exclusion principle*, Phys. Rev. E, 48 (1993), pp. 4263–4270.

[22] S. Lisini, D. Matthes, G. Savaré, (paper in preparation) (2009).

[23] L. Nirenberg, *On elliptic partial differential equations*, Ann. Scuola Norm. Sup. Pisa, 3 (1959), pp. 115–162.

[24] R. T. Rockafellar, Convex Analysis, Princeton University, Princeton, 1970.

[25] D. Slepčev, *Coarsening in nonlocal interfacial systems*, SIAM J. Math. Anal., 40 (2008), pp. 1029–1048.

[26] C. Villani, Topics in optimal transportation. Graduate Studies in Mathematics, 58. American Mathematical Society, Providence, RI, 2003.

[27] C. Villani, Optimal transport. Old and new. Grundlehren der Mathematischen Wissenschaften, 338. Springer-Verlag, Berlin, 2009.