GRADIENT FLOWS FOR NON-SMOOTH INTERACTION POTENTIALS

J. A. CARRILLO, S. LISINI, E. MAININI

Abstract. We deal with a nonlocal interaction equation describing the evolution of a particle density under the effect of a general symmetric pairwise interaction potential, not necessarily in convolution form. We describe the case of a convex (or λ-convex) potential, possibly not smooth at several points, generalizing the results of [CDFLS]. We also identify the cases in which the dynamic is still governed by the continuity equation with well-characterized nonlocal velocity field.

1. Introduction

Let us consider a distribution of particles, represented by a Borel probability measure μ on $\mathbb{R}^d$. We introduce the interaction potential $\mathbf{W} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$. The value $\mathbf{W}(x, y)$ describes the interaction of two particles of unit mass at the positions x and y. The total energy of a distribution μ under the effect of the potential is given by the interaction energy functional, defined by

$$\mathcal{W}(\mu) := \frac{1}{2} \int_{\mathbb{R}^d\times\mathbb{R}^d} \mathbf{W}(x, y)\, d(\mu \times \mu)(x, y). \tag{1.1}$$

We assume that $\mathbf{W}$ satisfies the following assumptions:

i) $\mathbf{W}$ is symmetric, i.e.

$$\mathbf{W}(x, y) = \mathbf{W}(y, x) \quad \text{for every } x, y \in \mathbb{R}^d; \tag{1.2}$$

ii) $\mathbf{W}$ is a λ-convex function for some $\lambda \in \mathbb{R}$, i.e.

$$\text{there exists } \lambda \in \mathbb{R} \text{ such that} \quad (x, y) \mapsto \mathbf{W}(x, y) - \frac{\lambda}{2}(|x|^2 + |y|^2) \quad \text{is convex;} \tag{1.3}$$

iii) $\mathbf{W}$ satisfies the quadratic growth condition at infinity, i.e.

$$\text{there exists } C > 0 \text{ such that} \quad \mathbf{W}(x, y) \le C(1 + |x|^2 + |y|^2) \quad \text{for every } x, y \in \mathbb{R}^d. \tag{1.4}$$

We are interested in the evolution problem given by the continuity equation

$$\partial_t\mu_t + \operatorname{div}(\boldsymbol{v}_t\mu_t) = 0, \qquad \text{in } (0,\infty) \times \mathbb{R}^d, \tag{1.5}$$

describing the dynamics of the particle density μ_t, under the mutual attractive-repulsive interaction described by functional (1.1). For any t, μ_t is a Borel probability measure and the velocity vector field $\boldsymbol{v}_t$ enjoys a nonlocal dependence on μ_t. For instance, in the basic model represented by a C^1 potential $\mathbf{W}$ which depends only on the difference of its variables, so that we may write $\mathbf{W}(x, y) = W(x - y)$, the velocity is given by convolution:

$$\boldsymbol{v}_t = -\nabla W * \mu_t. \tag{1.6}$$

2010 *Mathematics Subject Classification.* 49K20, 35F99.

Under the assumptions (1.2), (1.3), (1.4) and $\mathbf{W}(x,y) = W(x-y)$, in general W is not differentiable but only subdifferentiable, therefore it is reasonable to consider a velocity field of the form

$$\boldsymbol{v}_t = -\boldsymbol{\eta}_t * \mu_t, \tag{1.7}$$

where, for any t, $\boldsymbol{\eta}_t$ represents a Borel measurable selection in the subdifferential of W, and we will write $\boldsymbol{\eta}_t \in \partial W$. Unlike the case (1.6), in general such selection may depend on t. We stress that, for fixed t, the map $x \mapsto \boldsymbol{\eta}_t(x)$ needs to be pointwise defined, since the solutions we consider are probability measures, and since this model typically presents concentration phenomena when starting with absolutely continuous initial data.

In this paper, we are going to analyse equations of the form (1.5)-(1.7) as the gradient flow of the interaction energy (1.1) in the space of Borel probability measures with finite second moment, endowed with the metric-differential structure induced by the so-called Wasserstein distance. This interpretation coming from the optimal transport theory was introduced in [O1, O2] for nonlinear diffusion equations and generalized for a large class of functionals including potential, interaction, and internal energy by different authors [CMV, AGS, CMV2], see [V] for related information.

The gradient flow interpretation allows to construct solutions by means of variational schemes based on the euclidean optimal transport distance as originally introduced in [JKO] for the linear Fokker-Planck equation. The convergence of these variational schemes for general functionals was detailed in [AGS]. The results in this monograph apply to the interaction equation (1.5)-(1.6), with a C^1 smooth nonnegative potential verifying the convexity assumption (1.3) and a growth condition at infinity weaker than (1.4).

On the other hand, these equations have appeared in the literature as simple models of inelastic interactions [MY, BCP, BCCP, LT, T] in which the asymptotic behavior of the equations is given by a total concentration towards a unique Dirac Delta point measure. The typical potential in these models was a power law, $\mathbf{W}(x,y) = |x-y|^\alpha$, $\alpha \geq 0$. Moreover, it was noticed in [LT] that the convergence towards this unique steady state was in finite time for certain range of exponents in the one dimensional case.

Also these equations appear in very simplified swarming or population dynamics models for collective motion of individuals, see [MEBS, BL, BCL, KSUB, BCLR] and the references therein. The interaction potential models the long-range attraction and the short-range repulsion typical in animal groups. In case the potential is fully attractive, equation (1.5) is usually referred as the aggregation equation. For the aggregation equation, finite time blow-up results for weak-L^p solutions, unique up to the blow-up time, have been obtained in the literature [BCL, BLR, CR]. In fact, those results conjectured that solutions tend to concentrate and form Dirac Deltas in finite time under suitable conditions on the interaction potential. On the other hand, the confinement of particles is shown to happen for short-range repulsive long-range attractive potentials under certain conditions [CDFLS2]. Some singular stationary states such as uniform densities on spheres have been identified as stable/unstable for radial perturbations in [BCLR] with sharp conditions on the potential. Finally, in the one dimensional case, stationary states formed by finite number of particles and smooth stationary profiles are found whose stability have been studied in [FR1, FR2] in a suitable sense.

A global-in-time well-posedness theory of measure weak solutions has been developed in [CDFLS] for interaction potentials of the form $\mathbf{W}(x,y) = W(x-y)$ satisfying the assumptions (1.2),(1.3), (1.4), and additionally being C^1-smooth except possibly at the origin. The convexity condition (1.3) restricts the possible singularities of the potential at the origin since it implies that W is Lipschitz, and therefore the possible singularity cannot be worse than $|x|$ locally at the origin. Nevertheless, for a class of potentials in which the local behavior at the origin is like $|x|^\alpha$, $1 \le \alpha < 2$, the solutions converge towards a Dirac Delta with the full mass at the center of mass of the solution. The condition for blow-up is more general and related to the Osgood criterium for uniqueness of ODEs [BCL, CDFLS, BLR]. Note that the center of mass of the solution is preserved, at least formally, due to the symmetry assumption (1.2).

In this work, we push the ideas started in [CDFLS] further in the direction of giving conditions on the interaction potential to have a global-in-time well-posedness theory of measure solutions. The solutions constructed in Section 2 will be *gradient flow solutions*, as in [AGS], built via the variational schemes based on the optimal transport Wasserstein distance. The crucial point for the analysis in this framework is the identification of the velocity field in the continuity equation satisfied by the limiting curve of measures from the approximating variational scheme. In order to identify it, we need to characterize the sub-differential of the functional defined in (1.1) with respect to the differential structure induced by the Wasserstein metric. The Wasserstein sub-differential of the functional $\mathcal{W}$, which is rigorously introduced in Section 2, is defined through variations along transport maps. It turns out that that the element of minimal norm in this sub-differential, which will be denoted by $\partial^o\mathcal{W}(\cdot)$, is the element that governs the dynamics. Actually, it gives the velocity field via the relation $\boldsymbol{v}_t = -\partial^o\mathcal{W}(\mu_t)$ for a.e. $t \in (0,\infty)$, which corresponds to the notion of *gradient flow solution*. This notion will be discussed in Section 2, where we will give the precise definition and recall from [AGS, Chapter 11] the main properties, such as the semigroup generation.

In Section 3, we give a characterization of the subdifferential in the general case of the interaction potential $\mathbf{W}(x,y)$ satisfying only the basic assumptions (1.2),(1.3), and (1.4). However, the element of minimal norm in the subdifferential is not fully identified and cannot be universally characterized. Nevertheless, the global well-posedness of the evolution semigroup in measures is obtained.

A distinguished role will be played by the case of a kernel function $\mathbf{W}(x,y)$ which depends only on the difference $x-y$ of its arguments. Hence we will often consider one of the following additional assumptions.

iv) $\mathbf{W}$ depends on the difference of its arguments, i.e.

$$\text{there exists } W : \mathbb{R}^d \to \mathbb{R} \text{ such that } \mathbf{W}(x,y) = W(x-y) \quad \text{for every } x, y \in \mathbb{R}^d; \tag{1.8}$$

v) $\mathbf{W}$ satisfies *iv)* above and W is radial, i.e.

$$\begin{gathered} \text{there holds (1.8) and there exists } w : [0,+\infty) \to \mathbb{R} \text{ such that} \\ \mathbf{W}(x,y) = W(x-y) = w(|x-y|) \quad \text{for every } x, y \in \mathbb{R}^d. \end{gathered} \tag{1.9}$$

The radial hypothesis is frequently made in models, and corresponds to an interaction between particles which depends only on their mutual distance vector. In case $\mathbf{W}(x,y)$ is also radial and convex, we can fully generalize the identification of the element of minimal norm in the subdifferential of the interaction energy done in [CDFLS], regardless of the number of nondifferentiability points of W. We complement our results with explicit examples showing the sharpness of these characterizations.

Before stating the results and in order to fix notations we recall the characterization of subdifferential for λ-convex functions. Given a λ-convex function $V : \mathbb{R}^k \to \mathbb{R}$ (i.e. the map $z \mapsto V(z) - \lambda|z|^2/2$ is convex) a vector $\boldsymbol{\eta} \in \mathbb{R}^k$ belongs to the subdifferential of V at the point

$z_1 \in \mathbb{R}^k$ if

$$V(z_2) - V(z_1) \geq \langle \boldsymbol{\eta}, z_2 - z_1 \rangle + \frac{\lambda}{2}|z_2 - z_1|^2 \quad \text{for every } z_2 \in \mathbb{R}^k, \tag{1.10}$$

and we write $\boldsymbol{\eta} \in \partial V(z_1)$. In this case, for every $z_1 \in \mathbb{R}^k$, we have that $\partial V(z_1)$ is a non empty closed convex subset of $\mathbb{R}^k$. We denote by $\partial^o V(z_1)$ the unique element of minimal euclidean norm in $\partial V(z_1)$. λ-convexity of V is also equivalent to

$$V((1-t)z_1 + tz_2) \leq tV(z_1) + (1-t)V(z_2) - \frac{\lambda}{2} t(1-t)|z_2 - z_1|^2 \tag{1.11}$$

for every $t \in [0, 1]$ and $z_1, z_2 \in \mathbb{R}^k$. A map $\mathbb{R}^d \ni z \mapsto \boldsymbol{\eta}(z) \in \mathbb{R}^d$ is a selection in the subdifferential of V if $\boldsymbol{\eta}(z) \in \partial V(z)$ for any $z \in \mathbb{R}^d$ (and we write $\boldsymbol{\eta} \in \partial V$). Any such selection is λ-monotone, i.e.,

$$\langle \boldsymbol{\eta}(z_1) - \boldsymbol{\eta}(z_2), z_1 - z_2 \rangle \geq \lambda |z_1 - z_2|^2 \quad \text{for every } z_1, z_2 \in \mathbb{R}^k. \tag{1.12}$$

The main results. Let us give a brief summary of the results contained in this paper. The main theorem deals with radial-convex potentials and reads as follows. *Let* $\mathbf{W}$ *satisfy the three basic assumptions above:* (1.2), (1.3), *and* (1.4). *If in addition* $\boldsymbol{W}$ *satisfies* (1.8), (1.9) *and is convex (that is,* $\lambda \geq 0$ *in* (1.3)*), then there exists a unique* gradient flow solution *to the equation*

$$\partial_t \mu_t - \operatorname{div}((\partial^o W * \mu_t)\mu_t) = 0. \tag{1.13}$$

This solution is the gradient flow of the energy $\mathcal{W}$, in the sense that the velocity field in (1.13) satisfies

$$\partial^o W * \mu_t = \partial^o \mathcal{W}(\mu_t).$$

On the other hand, when omitting the radial hypothesis (1.9), or when letting the potential be λ-convex but not convex, we show that the evolution of the system under the effect of the potential, that is the gradient flow of $\mathcal{W}$, is characterized by (1.5)-(1.7), where $\boldsymbol{\eta}_t$ is a Borel anti-symmetric selection in ∂W. The corresponding rigorous statement is found in Section 2.

About this last result, let us remark that the velocity vector field is still written in terms of a suitable selection $\boldsymbol{\eta}_t$ in the subdifferential of W, but such selection $\boldsymbol{\eta}_t$ is not in general the minimal one in ∂W, and it is not a priori independent of t. By this characterization we also recover the result of [CDFLS], where the only non smoothness point for W is the origin: in such case, for any t we are left with $\boldsymbol{\eta}_t(x) = \nabla W(x)$ for $x \neq 0$ and $\boldsymbol{\eta}_t(0) = 0$, by anti-symmetry. We stress that, due to the nonlocal structure of the problem, the task of identifying the velocity vector field becomes much more involved when W has several non smoothness points, even if it is λ-convex. Later in Section 4, we will analyse some particular examples, showing that in general it is not possible to write the velocity field in terms of a single selection in ∂W.

Finally, when omitting also the assumption (1.8), we break the convolution structure: in this more general case we show that the velocity is given in terms of elements of $\partial_1 \mathbf{W}$, or equivalently of $\partial_2 \mathbf{W}$ by symmetry, where $\partial_1 \mathbf{W}$ and $\partial_2 \mathbf{W}$ denotes the partial subdifferentials of $\mathbf{W}$ with respect to the first d variables or the last d variables respectively. Indeed, we have

$$\boldsymbol{v}_t(x) = -\int_{\mathbb{R}^d} \boldsymbol{\eta}_t(x, y)\, d\mu_t(y),$$

where $\boldsymbol{\eta}_t \in \partial_1 \mathbf{W}$. An additional joint subdifferential condition is also present in this case, for the rigorous statement we still refer to the next section.

Pointwise particle model and asymptotic behavior. In the model case of a system of N point particles the dynamics are governed by a system of ordinary differential equations. In this case equation (1.5)-(1.7) corresponds to

$$\frac{dx_i(t)}{dt} = \sum_{j=1}^{N} m_j\, \boldsymbol{\eta}_t(x_j(t) - x_i(t)), \qquad i = 1, \dots, N, \tag{1.14}$$

where $x_i(t)$ is the position of the i-th particle and m_i is its mass. It is shown in [CDFLS] that if the attractive strength of the potential at the origin is sufficiently high, all the particles collapse to the center of mass in finite time. We will remark how this result is still working under our hypotheses and under the same non-Osgood criterium as in [BCL, CDFLS] for fully attractive potentials. For non-convex non-smooth repulsive-attractive potentials, albeit λ-convex, the analysis leads to non-trivial sets of stationary states with singularities that cannot be treated by the theory in [CDFLS]. Our analysis shows that a very wide range of asymptotic states is indeed possible, we give different explicit examples.

Plan of the paper. In the following Section 2 we introduce the optimal transport framework and the basic properties of our energy functional, in particular the subdifferentiability and the λ-convexity along geodesics. We briefly explain what is a gradient flow in the metric space $\mathscr{P}_2(\mathbb{R}^d)$ and we introduce the notion of gradient flow solution. We present the general well-posedness result of [AGS] and show how it will apply to our interaction models, stating our main results. In Section 3, we make a fine analysis on the Wasserstein subdifferential of $\mathcal{W}$ and find a first characterization of its element of minimal norm. In Section 4 we particularize the characterization to the case of assumption (1.8), which is the convolution case. In particular, we have the strongest result in the case of assumption (1.9). Section 5 gives examples of non-smooth non-convex repulsive-attractive potentials, albeit λ-convex, leading to non-trivial sets of stationary states. Finally, the Appendix is devoted to recall technical concepts from the differential calculus in Wasserstein spaces which are needed in Section 3.

2. Gradient flow of the interaction energy and main results

2.1. **Probability measures and Wasserstein distance.** We denote by $\mathscr{P}(\mathbb{R}^d)$ the space of Borel probability measures over $\mathbb{R}^d$ and by $\mathscr{P}_2(\mathbb{R}^d)$ the corresponding subset of measures with finite second moment, i.e.,

$$\mathscr{P}_2(\mathbb{R}^d) := \left\{ \mu \in \mathscr{P}(\mathbb{R}^d) : \int_{\mathbb{R}^d} |x|^2 \, d\mu(x) < +\infty \right\}.$$

The convergence of probability measures is considered in the narrow sense defined as the weak convergence in the duality with continuous and bounded functions over $\mathbb{R}^d$. A sequence $\{\mu_n\}_{n\in\mathbb{N}} \subset \mathscr{P}(\mathbb{R}^d)$ is said to be *tight* if for any $\varepsilon > 0$ there exists a compact set $K_\varepsilon \in \mathbb{R}^d$ such that $\sup_n \mu_n(\mathbb{R}^d \setminus K_\varepsilon) < \varepsilon$. We recall that Prokhorov theorem (see for instance [B]) entails that tight sequences admit narrow limit points.

Given $\mu, \nu \in \mathscr{P}_2(\mathbb{R}^d)$ and $\gamma \in \Gamma(\mu, \nu)$, where

$$\Gamma(\mu,\nu) := \{\gamma \in \mathscr{P}_2(\mathbb{R}^d \times \mathbb{R}^d) : \gamma(\Omega \times \mathbb{R}^d) = \mu(\Omega), \gamma(\mathbb{R}^d \times \Omega) = \nu(\Omega), \quad \text{for every Borel set } \Omega \subset \mathbb{R}^d\},$$

the euclidean quadratic transport cost between μ and ν with respect to the transport plan γ is defined by

$$\mathcal{C}(\mu,\nu\,;\gamma) = \left(\int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2\,d\gamma(x,y)\right)^{1/2}.$$

The "Wasserstein distance" between μ and ν is defined by

$$d_W(\mu,\nu) = \inf_{\gamma\in\Gamma(\mu,\nu)} \mathcal{C}(\mu,\nu\,;\gamma). \tag{2.1}$$

It is well known that the inf in (2.1) is attained by a minimizer. The minimizers in (2.1) are called optimal plans. We denote by $\Gamma_o(\mu,\nu) \subset \Gamma(\mu,\nu)$ the set of optimal plans between μ and ν. It is also well known that $\mu_n \to \mu$ in $\mathscr{P}_2(\mathbb{R}^d)$ (i.e. $d_W(\mu_n,\mu)\to 0$) if and only if μ_n narrowly converge to μ and $\int_{\mathbb{R}^d}|x|^2\,d\mu_n(x) \to \int_{\mathbb{R}^d}|x|^2\,d\mu(x)$. The space $\mathscr{P}_2(\mathbb{R}^d)$ endowed with the distance d_W is a complete and separable metric space. For all the details on Wasserstein distance and optimal transportation, we refer to [AGS, V].

We recall the push forward notation for a map $\mathbf{s} : (\mathbb{R}^d)^m \to (\mathbb{R}^d)^k$, $m,k\geq 1$, and a measure $\mu \in \mathscr{P}((\mathbb{R}^d)^m)$: the measure $\mathbf{s}_\#\mu \in \mathscr{P}((\mathbb{R}^d)^k)$ is defined by $\mathbf{s}_\#\mu(A) = \mu(\mathbf{s}^{-1}(A))$, where A is a Borel set. A transport plan $\gamma\in\Gamma(\mu,\nu)$ may be induced by a map $\mathbf{s}:\mathbb{R}^d\to\mathbb{R}^d$ such that $\mathbf{s}_\#\mu = \nu$. This means that $\gamma = (\mathbf{i},\mathbf{s})_\#\mu$, where $\mathbf{i}:\mathbb{R}^d\to\mathbb{R}^d$ denotes the identity map over $\mathbb{R}^d$ and $(\mathbf{i},\mathbf{s}):\mathbb{R}^d\to\mathbb{R}^d\times\mathbb{R}^d$ is the product map, whose image is the graph of $\mathbf{s}$. Finally, π^j will stand for the projection map on the j-th component of a product space. Hence, if γ is a probability measure over a product space (for instance a transport plan), $\pi^j_\#\gamma$ is its j-th marginal.

2.2. **Wasserstein subdifferential of the interaction energy.** We introduce the notion of subdifferential of $\mathcal{W}$ in the Wasserstein framework. Let us first discuss some elementary properties of $\mathbf{W}$ and $\mathcal{W}$.

Proposition 2.1. *If $\mathbf{W}$ satisfies assumptions* (1.3) *and* (1.4), *then there exists $K>0$ such that*

$$|\mathbf{W}(x,y)| \leq K(1+|x|^2+|y|^2) \quad \textit{for every } x,y\in\mathbb{R}^d. \tag{2.2}$$

Moreover, if $\boldsymbol{\eta}\in\partial\mathbf{W}$ is a selection in the subdifferential of $\mathbf{W}$, then there exists $M>0$ such that

$$|\boldsymbol{\eta}(x,y)| \leq M(1+|x|+|y|) \qquad \textit{for every } x,y\in\mathbb{R}^d. \tag{2.3}$$

Proof. The estimate (2.2) is a direct consequence of (1.3) and (1.4). Letting $z_1=(x_1,y_1)$ and $z_2=(x_2,y_2)$ in $\mathbb{R}^d\times\mathbb{R}^d$, by (2.2) and the inequality

$$\mathbf{W}(z_1+z_2) - \mathbf{W}(z_1) \geq \langle \boldsymbol{\eta}(z_1), z_2\rangle + \frac{\lambda}{2}\,|z_2|^2,$$

we have

$$\langle \boldsymbol{\eta}(z_1), z_2\rangle \leq \tilde{K}(1+|z_1|^2+|z_2|^2),$$

where $\tilde{K}$ depends only on K and λ. Dividing by $|z_2|$ and taking the supremum among all z_2 such that $|z_2| = \max\{|z_1|,1\}$ we obtain $|\boldsymbol{\eta}(z_1)| \leq 2\tilde{K}(1+|z_1|)$, which implies (2.3). □

Under assumptions (1.2), (1.3) and (1.4), as a consequence of (2.2), functional $\mathcal{W}$ is well defined and finite on $\mathscr{P}_2(\mathbb{R}^d)$. Moreover, it satisfies the following λ-convexity property.

Proposition 2.2. *Under assumptions* (1.2)*,* (1.3) *and* (1.4)*, the functional* $\mathcal{W} : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ *is lower semicontinuous with respect to the* d_W *metric and enjoys the following property: for every* $\mu, \nu \in \mathscr{P}_2(\mathbb{R}^d)$ *and every* $\gamma \in \Gamma(\mu, \nu)$ *it holds*

$$\mathcal{W}(\theta^\gamma(t)) \le (1-t)\mathcal{W}(\mu) + t\mathcal{W}(\nu) - \frac{\lambda}{2} t(1-t)\mathcal{C}^2(\mu, \nu; \gamma), \tag{2.4}$$

where θ^γ *denotes the interpolating curve* $t \in [0,1] \mapsto \theta^\gamma(t) = ((1-t)\pi^1 + t\pi^2)_\# \gamma \in \mathscr{P}_2(\mathbb{R}^d)$*.*

The d_W lower semicontinuity follows from standard arguments. For the convexity along interpolating curves we refer to [AGS, §9.3]. In particular, since every constant speed Wasserstein geodesic is of the form θ^γ where γ is an optimal plan, then $\mathcal{W}$ is λ-convex along every Wasserstein geodesic. We adapt from [AGS] the definition of the Wasserstein subdifferential. Still in the framework of assumptions (1.2), (1.3) and (1.4), we have

Definition 2.3 (**The Wasserstein subdifferential of** $\mathcal{W}$)**.** *Let* $\mu \in \mathscr{P}_2(\mathbb{R}^d)$*. We say that the vector field* $\boldsymbol{\xi} \in L^2(\mathbb{R}^d, \mu; \mathbb{R}^d)$ *belongs to* $\partial\mathcal{W}(\mu)$*, the Wasserstein subdifferential of the* λ*-convex functional* $\mathcal{W} : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ *at the point* μ*, if for every* $\nu \in \mathscr{P}_2(\mathbb{R}^d)$ *there exists* $\gamma \in \Gamma_o(\mu, \nu)$ *such that*

$$\mathcal{W}(\nu) - \mathcal{W}(\mu) \ge \int_{\mathbb{R}^d} \langle \boldsymbol{\xi}(x), y - x\rangle \, d\gamma(x,y) + \frac{\lambda}{2} \mathcal{C}^2(\mu, \nu\,; \gamma). \tag{2.5}$$

We say that that $\boldsymbol{\xi} \in \partial_S \mathcal{W}(\mu)$*, the strong subdifferential of* $\mathcal{W}$ *at the point* μ*, if for every* $\nu \in \mathscr{P}_2(\mathbb{R}^d)$ *and for every admissible plan* $\gamma \in \Gamma(\mu, \nu)$*,* (2.5) *holds.*

Notice that $\partial\mathcal{W}(\mu)$ and $\partial_S\mathcal{W}(\mu)$ are convex subsets of $L^2(\mathbb{R}^d, \mu; \mathbb{R}^d)$. We also define the metric slope of the functional $\mathcal{W}$ at the point μ as follows:

$$|\partial\mathcal{W}|(\mu) := \limsup_{\nu \to \mu \text{ in } \mathscr{P}_2(\mathbb{R}^d)} \frac{(\mathcal{W}(\nu) - \mathcal{W}(\mu))^+}{d_W(\nu, \mu)},$$

where $(a)^+$ denotes the positive part of the real number a. Since $\mathcal{W}$ is λ-convex we know that, whenever $\partial\mathcal{W}(\mu) \ne \emptyset$,

$$|\partial\mathcal{W}|(\mu) = \min\left\{ \|\boldsymbol{\xi}\|_{L^2(\mathbb{R}^d, \mu; \mathbb{R}^d)} : \boldsymbol{\xi} \in \partial\mathcal{W}(\mu) \right\}. \tag{2.6}$$

Moreover, the minimizer of the norm in (2.6) is unique and we denote it by $\partial^o\mathcal{W}(\mu)$ (see [AGS, Chapter 10]). Later on we will also show that, in our case, $\partial\mathcal{W}(\mu) \ne \emptyset$ for any $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ (see Theorem 3.1). The element of minimal norm in the subdifferential plays a *crucial role*, since it is known to be the velocity vector field of the evolution equation associated to the gradient flow of the functional (1.1) under certain conditions as reviewed next.

2.3. **Gradient flow solution.** As already shown in [CDFLS], we are forced to consider measure solutions. In fact, in case of attractive radial potentials verifying assumptions (1.2), (1.3), (1.4) and (1.9), with w increasing, it was shown in [BCL, BLR] that weak-L^p solutions blow-up in finite time. Moreover, these weak-L^p solutions can be uniquely continued as measure solutions, as proved in [CDFLS], leading to a total collapse in a single Dirac's Delta at the center of mass in finite time. Furthermore, particle solutions, i.e., solutions corresponding to an initial data composed by a finite number of atoms, remain particle solutions for all times for the evolution of (1.5). Summarizing, we can only expect that a regular solution enjoys local in time existence.

Before stating a theorem about gradient flow solutions, we recall that a curve $t \in [0,\infty) \mapsto \mu_t \in \mathscr{P}_2(\mathbb{R}^d)$ is locally absolutely continuous with locally finite energy, and we denote it by $\mu \in AC^2_{\rm loc}([0,\infty);\mathscr{P}_2(\mathbb{R}^d))$, if the restriction of μ to the interval $[0,T]$ is absolutely continuous for every $T>0$ and its metric derivative, which exists for a.e. $t>0$ defined by

$$|\mu'|(t) := \lim_{s\to t} \frac{d_W(\mu_s,\mu_t)}{|t-s|},$$

belongs to $L^2(0,T)$ for every $T>0$. Thanks to Proposition 2.2, functional $\mathcal{W}$ satisfies the assumptions of [AGS, §11.2]. Therefore, we can apply [AGS, Theorem 11.2.1] and directly deduce the following well-posedness result, see [AGS, §11.2] for details.

Theorem 2.4. *Let* $\mathbf{W}$ *satisfy the hypotheses* (1.2), (1.3), *and* (1.4). *For any initial datum* $\bar\mu \in \mathscr{P}_2(\mathbb{R}^d)$, *there exists a unique curve* $\mu \in AC^2_{\rm loc}([0,\infty);\mathscr{P}_2(\mathbb{R}^d))$ *satisfying*

$$\partial_t \mu_t + \operatorname{div}(\boldsymbol{v}_t \mu_t) = 0 \ \textit{in}\ \mathcal{D}'((0,\infty)\times\mathbb{R}^d),$$

where $\mathcal{D}'$ *denotes the space of distributions,*

$$\begin{aligned} \boldsymbol{v}_t &= -\partial^o\mathcal{W}(\mu_t), && \textit{for a.e. } t>0,\\ \|\boldsymbol{v}_t\|_{L^2(\mu_t)} &= |\mu'|(t) && \textit{for a.e. } t>0, \end{aligned}$$

with $\mu_0 = \bar\mu$. *The energy identity*

$$\int_a^b \int_{\mathbb{R}^d} |\boldsymbol{v}_t(x)|^2 \, d\mu_t(x)\, dt + \mathcal{W}(\mu_b) = \mathcal{W}(\mu_a)$$

holds for all $0 \le a \le b < \infty$. *Moreover, the solution is given by a* λ*-contractive semigroup* S *acting on* $\mathscr{P}_2(\mathbb{R}^d)$, *that is* $\mu_t = S[\mu_0](t)$ *with*

$$d_W(S[\mu_0](t), S[\nu_0](t)) \le e^{-\lambda t} d_W(\mu_0,\nu_0), \quad \forall \mu_0,\nu_0 \in \mathscr{P}_2(\mathbb{R}^d)\,.$$

The unique curve μ given by Theorem 2.4 is called *gradient flow solution* for equation

$$\partial_t \mu_t = \operatorname{div}(\partial^o\mathcal{W}(\mu_t)\mu_t), \tag{2.7}$$

starting from $\mu_0 = \bar\mu$. Let us remark that weak measure solutions as defined in [CDFLS] are equivalent to gradient flow solutions as shown therein.

2.4. **Summary of the main results.** Characterizing the element of minimal norm $\partial^o\mathcal{W}(\mu)$ is now essential to link the constructed solutions to the sought equation (1.5)-(1.6) or (1.5)-(1.7). This characterization was done in [CDFLS] for potentials satisfying (1.2), (1.3), (1.4) and (1.8), being the potential W C^1-smooth except possibly at the origin. Under those assumptions, the authors identified $\partial^o\mathcal{W}(\mu)$ as $\partial^o W * \mu$, i.e.

$$\partial^o W * \mu(x) = \int_{x\neq y} \nabla W(x-y)\, d\mu(y).$$

The main results in the present work will show that this characterization can be generalized to *convex* potentials satisfying assumptions (1.4) and the radial hypothesis (1.9), regardless of the number of points of non-differentiability of the potential W. In Theorem 4.5, we show that under those conditions, the formula $\partial^o\mathcal{W}(\mu) = \partial^o W * \mu$ also holds, and the equation takes the form (1.13), which generalizes the standard form of the interaction potential evolution (1.5)-(1.6). In the most general case, i.e. for potentials satisfying only (1.2), (1.3) and (1.4), we will obtain a characterization in terms of generic Borel measurable selections in $\partial\mathbf{W}$. Precisely, in the next two sections we are going to prove the following results:

- Let $\mathbf{W}$ satisfy (1.2), (1.3) and (1.4). Let $\mu \in \mathscr{P}_2(\mathbb{R}^d)$. There holds

$$\partial^o \mathcal{W}(\mu) = \int_{\mathbb{R}^d} \boldsymbol{\eta}(\cdot, y)\, d\mu(y) \tag{2.8}$$

  for some selection $\boldsymbol{\eta} \in \partial_1 \mathbf{W}$ having the form

$$\boldsymbol{\eta}(x,y) = \frac{1}{2}\left(\boldsymbol{\eta}^1(x,y) + \boldsymbol{\eta}^2(y,x)\right), \tag{2.9}$$

  with the couple $(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2)$ belonging to the joint subdifferential $\partial \mathbf{W}$. This is shown in Theorem 3.3.
- Under the additional assumption (1.8), let $\mu \in \mathscr{P}_2(\mathbb{R}^d)$. Then we have

$$\partial^o \mathcal{W}(\mu) = \boldsymbol{\eta} * \mu \tag{2.10}$$

  for some $\boldsymbol{\eta} \in \partial W$. This is the characterization following from Corollary 4.2.
- Finally, when the further condition (1.9) holds, and the potential is convex (not only λ-convex for a negative λ) there holds

$$\partial^o \mathcal{W}(\mu) = \partial^o W * \mu \qquad \text{for all } \mu \in \mathscr{P}_2(\mathbb{R}^d). \tag{2.11}$$

  Here, $\partial^o W$ is the element of minimal norm of the subdifferential of W. This is proven in the subsequent Theorem 4.5.

Remark 2.5. In (2.8) and (2.10) the selection $\boldsymbol{\eta}$ depends in general on μ, as we will show in Section 4. Assuming in addition that (1.9) holds and W is convex, the selection $\boldsymbol{\eta}$ in (2.10) is always given by $\partial^o W$ and thus it does not depend on μ. Substituting (2.8), (2.10) or (2.11) in (2.7) and applying Theorem 2.4, one obtains the corresponding well-posedness result. Therefore, in the case of (2.10), the dynamics will be governed by a velocity field of the form (1.7), where the selection depends in general on t via μ_t. Similarly for the case of (2.8). On the other hand, in the case of (2.11), we stress that the selection corresponding to the velocity $\boldsymbol{v}(t)$ in (1.7) does not depend on t.

Remark 2.6. The joint subdifferential constraint (2.9) has a natural interpretation: there is a symmetry in the interaction between particles (action-reaction law).

3. Characterization of the element of minimal norm in the subdifferential

In this section and in the next we analyze the Wasserstein subdifferential of $\mathcal{W}$, we characterize its element of minimal norm and we prove the main core results of this work.

Let us give a quick sketch of the strategy of the present section. We preliminarily show in Theorem 3.1 that if $(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2) \in \partial \mathbf{W}$ is a Borel measurable selection in the subdifferential of $\mathbf{W}$ and $\boldsymbol{\eta}$ is given by (2.9), the vector field $x \mapsto \int_{\mathbb{R}^d} \boldsymbol{\eta}(x,y)\, d\mu(y)$ belongs to the (strong) subdifferential of $\mathcal{W}$ at μ. Then, the main goal is to show that indeed the element of minimal norm in $\partial \mathcal{W}(\mu)$ is necessarily of the same form, as stated in the subsequent Theorem 3.3. In order to prove Theorem 3.3, we use the approach of [AGS, §10.3], trying to characterize $\partial^o \mathcal{W}(\mu)$ as limit of approximating strong subdifferentials. This is achieved with two parallel regularization steps. We first introduce the approximating measures μ_τ, given as solutions to

$$\min_{\nu \in \mathscr{P}_2(\mathbb{R}^d)} \mathcal{W}(\nu) + \frac{1}{2\tau} d_W^2(\nu, \mu).$$

By the theory developed in [AGS, §10.3], there are elements in the strong subdifferentials $\partial_S \mathcal{W}(\mu_\tau)$ that converge to $\partial^o \mathcal{W}(\mu)$ as $\tau \to 0$ (see Proposition 3.13 below). However, the strong subdifferential is fully characterized and it is a singleton when $\mathbf{W}$ is C^1 (thanks to the simple result we give with Proposition 3.2). Therefore, we need to further introduce, the approximating functionals $\mathcal{W}_n$, see (3.11), by taking the *Moreau-Yosida* approximation $\mathbf{W}_n$ of $\mathbf{W}$, defined by (3.6) below. Then we apply the above minimization procedure to $\mathcal{W}_n$, obtaining approximating measures μ_τ^n. This way, having at disposal the explicit form of $\partial_S \mathcal{W}_n(\mu_\tau^n)$, we will pass to the limit (both in τ and n), in the sense specified in Definition 3.8, and obtain the desired characterization of $\partial^o \mathcal{W}(\mu)$.

Theorem 3.1. *Let* $\mathbf{W}$ *satisfy assumptions* (1.2), (1.3), *and* (1.4). *Consider a Borel measurable selection* $(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2) \in \partial \mathbf{W}$, *i.e.,* $(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ *is a Borel measurable function such that* $(\boldsymbol{\eta}^1(x,y), \boldsymbol{\eta}^2(x,y)) \in \partial \mathbf{W}(x,y)$ *for every* $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$. *For any* $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, *the map*

$$\boldsymbol{\xi}(x) := \frac{1}{2} \int_{\mathbb{R}^d} (\boldsymbol{\eta}^1(x,y) + \boldsymbol{\eta}^2(y,x)) \, d\mu(y) \tag{3.1}$$

belongs to $\partial_S \mathcal{W}(\mu)$. *In particular* $\partial_S \mathcal{W}(\mu)$ *is not empty and the metric slope* $|\partial \mathcal{W}|(\mu)$ *is finite.*

Proof. Let $\mu \in \mathscr{P}_2(\mathbb{R}^d)$. First of all, by Jensen inequality we have

$$\int_{\mathbb{R}^d} |\boldsymbol{\xi}(x)|^2 \, d\mu(x) \leq \frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left(|\boldsymbol{\eta}^1(x,y)|^2 + |\boldsymbol{\eta}^2(y,x)|^2 \right) \, d\mu(y) \, d\mu(x),$$

thus $\boldsymbol{\xi} \in L^2(\mathbb{R}^d, \mu; \mathbb{R}^d)$ thanks to the estimate (2.3).

Since $\mathbf{W}$ is λ-convex, we use the corresponding of (1.10) for $\mathbf{W}$, so that

$$\begin{aligned} \mathbf{W}(y_1, y_2) - \mathbf{W}(x_1, x_2) \geq{}& \langle (\boldsymbol{\eta}^1(x_1, x_2), \, \boldsymbol{\eta}^2(x_1, x_2)), (y_1 - x_1, y_2 - x_2) \rangle + \\ &+ \frac{\lambda}{2} |(y_1 - x_1, y_2 - x_2)|^2, \quad \text{for every } (x_1, x_2), (y_1, y_2) \in \mathbb{R}^d \times \mathbb{R}^d. \end{aligned} \tag{3.2}$$

Let $\mu, \nu \in \mathscr{P}_2(\mathbb{R}^d)$ and $\gamma \in \Gamma(\mu, \nu)$. We show that inequality (2.5) holds. Considering the measure $\gamma \times \gamma$, we can write

$$\mathcal{W}(\nu) - \mathcal{W}(\mu) = \frac{1}{2} \int_{(\mathbb{R}^d)^4} (\mathbf{W}(y_1, y_2) - \mathbf{W}(x_1, x_2)) \, d(\gamma \times \gamma)(x_1, y_1, x_2, y_2).$$

Hence, by (3.2),

$$\begin{aligned} \mathcal{W}(\nu) - \mathcal{W}(\mu) \geq{}& \frac{1}{2} \int_{(\mathbb{R}^d)^4} \big[\langle \boldsymbol{\eta}^1(x_1, x_2), (y_1 - x_1) \rangle + \langle \boldsymbol{\eta}^2(x_1, x_2), (y_2 - x_2) \rangle \big] d(\gamma \times \gamma)(x_1, y_1, x_2, y_2) \\ &+ \frac{\lambda}{4} \int_{(\mathbb{R}^d)^4} (|y_1 - x_1|^2 + |y_2 - x_2|^2) \, d(\gamma \times \gamma)(x_1, y_1, x_2, y_2). \end{aligned}$$

The last term is $\frac{\lambda}{2}\,\mathcal{C}^2(\mu,\nu;\gamma)$, so that a change of variables gives

$$\begin{aligned}
\mathcal{W}(\nu)-\mathcal{W}(\mu) \geq & \frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d\times\mathbb{R}^d}\langle \boldsymbol{\eta}^1(x_1,x_2),(y_1-x_1)\rangle\,d\gamma(x_1,y_1)\,d\mu(x_2)\\
&+\frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d\times\mathbb{R}^d}\langle \boldsymbol{\eta}^2(x_1,x_2),(y_2-x_2)\rangle\,d\gamma(x_2,y_2)\,d\mu(x_1)+\frac{\lambda}{2}\,\mathcal{C}^2(\mu,\nu;\gamma)\\
=&\frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d\times\mathbb{R}^d}\langle \boldsymbol{\eta}^1(x,z),(y-x)\rangle\,d\gamma(x,y)\,d\mu(z)\\
&+\frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d\times\mathbb{R}^d}\langle \boldsymbol{\eta}^2(z,x),(y-x)\rangle\,d\gamma(x,y)\,d\mu(z)+\frac{\lambda}{2}\,\mathcal{C}^2(\mu,\nu;\gamma)\\
=&\int_{\mathbb{R}^d\times\mathbb{R}^d}\left\langle\frac{1}{2}\int_{\mathbb{R}^d}(\boldsymbol{\eta}^1(x,z)+\boldsymbol{\eta}^2(z,x))\,d\mu(z),(y-x)\right\rangle d\gamma(x,y)+\frac{\lambda}{2}\,\mathcal{C}^2(\mu,\nu;\gamma)
\end{aligned}$$

as desired. □

In the case of a smooth interaction function $\mathbf{W}$, there is a complete characterization of the strong subdifferential $\partial_S\mathcal{W}(\mu)$ which is single valued.

Proposition 3.2 (The smooth case). *Let $\mu\in\mathscr{P}_2(\mathbb{R}^d)$. If $\mathbf{W}\in C^1(\mathbb{R}^d\times\mathbb{R}^d)$ satisfies the assumptions (1.2), (1.3), and (1.4), then the strong Wasserstein subdifferential is a singleton and it is of the form*

$$\begin{aligned}
\partial_S\mathcal{W}(\mu) &= \left\{\int_{\mathbb{R}^d}\nabla_1\mathbf{W}(\cdot,y)\,d\mu(y)\right\} = \left\{\int_{\mathbb{R}^d}\nabla_2\mathbf{W}(y,\cdot)\,d\mu(y)\right\}\\
&= \left\{\frac{1}{2}\int_{\mathbb{R}^d}\big(\nabla_1\mathbf{W}(\cdot,y)+\nabla_2\mathbf{W}(y,\cdot)\big)\,d\mu(y)\right\},
\end{aligned}\tag{3.3}$$

where ∇_1 (resp. ∇_2) are the gradients with respect to the first-d (second-d) variables of $\mathbb{R}^d\times\mathbb{R}^d$.

Proof. Since $\partial\mathbf{W}(x,y)=\{\nabla\mathbf{W}(x,y)\}$, by Theorem 3.1 and the symmetry of $\mathbf{W}$ we have that the right hand sides of (3.3) are contained in $\partial_S\mathcal{W}(\mu)$.

In order to prove the opposite inclusion, assume that $\boldsymbol{\xi}\in L^2(\mathbb{R}^d,\mu;\mathbb{R}^d)$ belongs to $\partial_S\mathcal{W}(\mu)$. Let $\mathbf{s}\in L^2(\mathbb{R}^d,\mu;\mathbb{R}^d)$ be an arbitrary vector field, $\nu=\mathbf{s}_{\#}\mu$ and $\mu_t=(\mathbf{i}+t\mathbf{s})_{\#}\mu$. Writing (2.5) in correspondence of the plan

$$\gamma_t=(\mathbf{i},\mathbf{i}+t\mathbf{s})_{\#}\mu$$

between μ and μ_t, we have

$$\mathcal{W}(\mu_t)-\mathcal{W}(\mu)\geq\int_{\mathbb{R}^d\times\mathbb{R}^d}\langle\boldsymbol{\xi}(x),y-x\rangle\,d\gamma_t(x,y)+\frac{\lambda}{2}\,\mathcal{C}^2(\mu,\mu_t;\gamma_t).$$

Hence, for every $t>0$

$$\begin{aligned}
\frac{1}{2t}\int_{\mathbb{R}^d\times\mathbb{R}^d}&(\mathbf{W}(x+t\mathbf{s}(x),y+t\mathbf{s}(y))-\mathbf{W}(x,y))\,d(\mu\times\mu)(x,y)\\
&\geq\int_{\mathbb{R}^d}\langle\boldsymbol{\xi}(x),\mathbf{s}(x)\rangle\,d\mu(x)+\frac{\lambda}{2}\,t\|\mathbf{s}\|^2_{L^2(\mu)},
\end{aligned}$$

and, by a direct computation

$$
(3.4)\quad \frac{1}{2t}\int_{\mathbb{R}^d\times\mathbb{R}^d}\big(\mathbf{W}(x+t\mathbf{s}(x),y+t\mathbf{s}(y))-\mathbf{W}(x,y)\big)d(\mu\times\mu)(x,y)
$$

$$
\geq \frac{\lambda}{4t}\int_{\mathbb{R}^d\times\mathbb{R}^d}\big(|(x+t\mathbf{s}(x),y+t\mathbf{s}(y))|^2-|(x,y)|^2\big)d(\mu\times\mu)(x,y)
$$

$$
+\int_{\mathbb{R}^d}\langle\boldsymbol{\xi}(x),\mathbf{s}(x)\rangle\,d\mu(x)-\lambda\int_{\mathbb{R}^d}\langle x,\mathbf{s}(x)\rangle\,d\mu(x)\,.
$$

Since $\mathbf{W}$ is λ-convex, the map

$$
t\mapsto\frac{1}{t}\left(\mathbf{W}(x+t\mathbf{s}(x),y+t\mathbf{s}(y))-\mathbf{W}(x,y)\right)-\frac{\lambda}{2t}\left(|(x+t\mathbf{s}(x),y+t\mathbf{s}(y))|^2-|(x,y)|^2\right)
$$

is nondecreasing in t, for $t>0$. Taking advantage of the C^1 regularity and the quadratic growth of $\mathbf{W}$, by the monotone convergence theorem, we can pass to the limit in (3.4) as t goes to 0, obtaining

$$
(3.5)\qquad \frac{1}{2}\int_{\mathbb{R}^d\times\mathbb{R}^d}\langle\nabla\mathbf{W}(x,y),(\mathbf{s}(x),\mathbf{s}(y))\rangle\,d(\mu\times\mu)(x,y)\geq\int_{\mathbb{R}^d}\langle\boldsymbol{\xi},\mathbf{s}\rangle\,d\mu.
$$

Since by the symmetry of $\mathbf{W}$ we have $\nabla_1\mathbf{W}(x,y)=\nabla_2\mathbf{W}(y,x)$ for any $x,y\in\mathbb{R}^d$, we can write

$$
\begin{aligned}
&\frac{1}{2}\int_{\mathbb{R}^d\times\mathbb{R}^d}\langle\nabla\mathbf{W}(x,y),(\mathbf{s}(x),\mathbf{s}(y))\rangle\,d(\mu\times\mu)(x,y)\\
&\quad=\frac{1}{2}\int_{\mathbb{R}^d\times\mathbb{R}^d}\langle\nabla_1\mathbf{W}(x,y),\mathbf{s}(x)\rangle\,d(\mu\times\mu)(x,y)+\frac{1}{2}\int_{\mathbb{R}^d\times\mathbb{R}^d}\langle\nabla_1\mathbf{W}(y,x),\mathbf{s}(y)\rangle\,d(\mu\times\mu)(x,y)\\
&\quad=\int_{\mathbb{R}^d\times\mathbb{R}^d}\langle\nabla_1\mathbf{W}(x,y),\mathbf{s}(x)\rangle\,d(\mu\times\mu)(x,y).
\end{aligned}
$$

This way, (3.5) becomes

$$
\int_{\mathbb{R}^d}\left\langle\int_{\mathbb{R}^d}\nabla_1\mathbf{W}(x,y)\,d\mu(y),\mathbf{s}(x)\right\rangle\,d\mu(x)\geq\int_{\mathbb{R}^d}\langle\boldsymbol{\xi},\mathbf{s}\rangle\,d\mu,
$$

that is

$$
\int_{\mathbb{R}^d}\left\langle\int_{\mathbb{R}^d}\nabla_1\mathbf{W}(x,y)\,d\mu(y)-\boldsymbol{\xi}(x),\mathbf{s}(x)\right\rangle\,d\mu(x)\geq 0.
$$

Since $\mathbf{s}\in L^2(\mathbb{R}^d,\mu;\mathbb{R}^d)$ is arbitrary we conclude that $\boldsymbol{\xi}(x)=\int_{\mathbb{R}^d}\nabla_1\mathbf{W}(x,y)\,d\mu(y)$ as elements of $L^2(\mathbb{R}^d,\mu;\mathbb{R}^d)$. □

The next theorem contains the main result of this section. It provides the desired characterization of the element of minimal norm in the subdifferential of $\mathcal{W}$ in the nonsmooth case.

Theorem 3.3. *Let $\mathbf{W}$ satisfy assumptions (1.2), (1.3), and (1.4). Let $\mu\in\mathscr{P}_2(\mathbb{R}^d)$ and $\boldsymbol{\xi}=\partial^o\mathcal{W}(\mu)$. Then there exists a Borel measurable selection $(\boldsymbol{\eta}^1,\boldsymbol{\eta}^2)\in\partial\mathbf{W}$ such that*

$$
\boldsymbol{\xi}(x)=\frac{1}{2}\int_{\mathbb{R}^d}(\boldsymbol{\eta}^1(x,y)+\boldsymbol{\eta}^2(y,x))\,d\mu(y)\qquad\text{for }\mu\text{-a.e. }x\in\mathbb{R}^d.
$$

Remark 3.4. Theorem 3.3 characterizes only the element of minimal norm in $\partial\mathcal{W}(\mu)$. However the property stating that every element of $\partial\mathcal{W}(\mu)$ is of the form (3.1) for a suitable Borel selection of the subdifferential of $\mathbf{W}$ could be in general very difficult and we do not know if it is true. On the other hand, a consequence of Theorem 3.3 is that the element of minimal norm is a strong subdifferential, thanks to Theorem 3.1, and in the C^1 case it is the element given by Proposition 3.2.

The proof of Theorem 3.3 needs several preliminary results. The first step is the introduction of a sequence of regularized interaction functionals, obtained by taking the *Moreau-Yosida* approximation of the potential $\mathbf{W}$. We recall that the *Moreau-Yosida* approximation of the λ-convex function $\mathbf{W}$ is defined as

$$\mathbf{W}_n(x,y) := \inf_{(v,w)\in\mathbb{R}^d\times\mathbb{R}^d}\Big\{\mathbf{W}(v,w)+\frac{n}{2}\,|(x-v,y-w)|^2\Big\},\quad n\in\mathbb{N},\quad n>\lambda^-, \tag{3.6}$$

where $\lambda^- := \max\{0,-\lambda\}$. Notice that λ-convexity of $\mathbf{W}$ implies that for $n>\lambda^-$ the function being minimized in (3.6) is strictly convex and coercive for every $(x,y)\in\mathbb{R}^d\times\mathbb{R}^d$, so that the minimizer is uniquely attained. We have $\mathbf{W}_n(x,y)\le\mathbf{W}(x,y)$ for every $(x,y)\in\mathbb{R}^d\times\mathbb{R}^d$. It is well known that $\mathbf{W}_n\in C^1(\mathbb{R}^d\times\mathbb{R}^d)$, $\nabla\mathbf{W}_n$ is globally Lipschitz, and the sequence $\{\mathbf{W}_n\}_{n\in\mathbb{N}}$ converges pointwise and monotonically to $\mathbf{W}$ as $n\to\infty$. Moreover, we have the following

Proposition 3.5. *Let* $\mathbf{W}$ *satisfy assumptions* (1.2)*,* (1.3)*, and* (1.4)*. Then there exist* $\Lambda\le\lambda$*,* $\bar K>0$ *and* $\bar M>0$ *such that, for any* $n>\lambda^-$*,*

$$\mathbf{W}_n \text{ is } \Lambda\text{-convex}, \tag{3.7}$$

$$|\mathbf{W}_n(x,y)|\le\bar K(1+|x|^2+|y|^2)\quad \text{for every } x,y\in\mathbb{R}^d, \tag{3.8}$$

$$|\nabla_i\mathbf{W}_n(x,y)|\le\bar M(1+|x|+|y|),\quad i=1,2,\quad \text{for every } x,y\in\mathbb{R}^d. \tag{3.9}$$

Proof. Let $n>\lambda^-$. For any $z=(x,y)\in\mathbb{R}^d\times\mathbb{R}^d$, let $p_n(z)$ denote the unique solution to the minimization problem in (3.6). Then, since $0\in\partial(\mathbf{W}(\cdot)+n|z-(\cdot)|^2/2)$ at the point $p_n(z)$, there holds the subdifferential relation

$$n(z-p_n(z))\in\partial\mathbf{W}(p_n(z)).$$

Therefore, taking $z_1\in\mathbb{R}^d\times\mathbb{R}^d$, $z_2\in\mathbb{R}^d\times\mathbb{R}^d$ and applying the λ-monotonicity property (1.12) to $\partial\mathbf{W}$, it follows that

$$n\,\langle z_2-p_n(z_2)-(z_1-p_n(z_1)),p_n(z_2)-p_n(z_1)\rangle\ge\lambda|p_n(z_2)-p_n(z_1)|^2,$$

and we deduce

$$|p_n(z_1)-p_n(z_2)|\le\frac{1}{1+\lambda/n}\,|z_1-z_2|. \tag{3.10}$$

On the other hand, using the definition of $\mathbf{W}_n$ and the λ-convexity of $\mathbf{W}$ it is not difficult to see that for any $t\in[0,1]$

$$\mathbf{W}_n(tz_1+(1-t)z_2)\le t\mathbf{W}_n(z_1)+(1-t)\mathbf{W}_n(z_2)-\frac{\lambda}{2}t(1-t)|p_n(z_1)-p_n(z_2)|^2.$$

Combined with (3.10), the latter yields

$$\mathbf{W}_n(tz_1+(1-t)z_2)\le t\mathbf{W}_n(z_1)+(1-t)\mathbf{W}_n(z_2)+\frac{\lambda^-}{2}\frac{1}{(1+\lambda/n)^2}t(1-t)|z_1-z_2|^2.$$

Notice that $\lambda < 0$ implies that $1/(1+\lambda/n)^2$ is decreasing with respect to n, for $n > \lambda^-$. In particular, choosing $\Lambda = -\lambda^-/(1+\lambda/\bar{n})^2$, where $\bar{n}$ denotes the smallest integer strictly larger than λ^-, we see that $\mathbf{W}_n$ is Λ-convex for any $n > \lambda^-$. The estimate (3.8) easily follows from (3.7) and (1.4), since $\mathbf{W}_n \le \mathbf{W}$. Eventually, exploiting the subdifferential inequality $\mathbf{W}_n(z_1+z_2) - \mathbf{W}_n(z_1) \ge \langle \nabla \mathbf{W}_n(z_1), z_2 \rangle + \Lambda |z_2|^2/2$ and the estimate (3.8) we have

$$\langle \nabla \mathbf{W}_n(z_1), z_2 \rangle \le \tilde{K}(1 + |z_1|^2 + |z_2|^2),$$

where $\tilde{K}$ depends only on $\bar{K}$ and Λ. As in the proof of Proposition 2.1, we conclude that (3.9) holds with $\bar{M}$ depending only on $\bar{K}$ and Λ. □

We define the approximating interaction functionals $\mathcal{W}_n : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ by

$$\mathcal{W}_n(\mu) := \frac{1}{2} \int_{\mathbb{R}^d \times \mathbb{R}^d} \mathbf{W}_n(x,y)\, d(\mu \times \mu)(x,y), \qquad n \in \mathbb{N}, \quad n > \lambda^-. \tag{3.11}$$

Since $\mathbf{W}$ might enjoy a negative quadratic behavior at infinity, it is not true that $\mathcal{W}$ is lower semicontinuous also with respect to the narrow convergence. By the way, it is shown in [CDFLS, §2] that one can choose τ_0 small enough (depending only on $\mathbf{W}$) such that for any $\tau < \tau_0$ and for any $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, the functional

$$\nu \in \mathscr{P}_2(\mathbb{R}^d) \mapsto \mathcal{W}(\nu) + \frac{1}{2\tau}\, d_W^2(\nu, \mu), \tag{3.12}$$

is lower semicontinuous with respect to the narrow convegence. Moreover, for $\tau < \tau_0$, minimizers do exist for (3.12). The same facts hold true in our case, and moreover when considering functional $\mathcal{W}_n$ the constant τ_0 can be chosen independently of n, as stated in the following

Proposition 3.6. *Let $\mathbf{W}$ satisfy assumptions* (1.2), (1.3), *and* (1.4). *Let $n > \lambda^-$ and $\mu \in \mathscr{P}_2(\mathbb{R}^d)$. There exists $\tau_0 > 0$, depending only on $\mathbf{W}$, such that for any $\tau < \tau_0$ the functionals*

$$\nu \in \mathscr{P}_2(\mathbb{R}^d) \mapsto \mathcal{W}_n(\nu) + \frac{1}{2\tau}\, d_W^2(\nu, \mu)$$

as well as the functional in (3.12) *are lower semicontinuous with respect to the narrow convergence and admit minimizers in $\mathscr{P}_2(\mathbb{R}^d)$.*

Proof. These results are proven in [CDFLS, Lemma 2.3, Lemma 2.5]. By investigating the proof of [CDFLS, Lemma 2.3], it is evident that the very same arguments can be applied to our functions $\mathbf{W}_n$ and $\mathbf{W}$, since they are continuous and satisfy the quadratic bounds (3.8) and (2.2) respectively, and the value τ_0 depends only on $\bar{K}$ in (3.8) (resp. K in (2.2)). In particular, τ_0 depends only on $\mathbf{W}$ and it is independent of n. □

We prove the following more general lower semicontinuity property.

Proposition 3.7. *Let $\mathbf{W}$ satisfy assumptions* (1.2), (1.3), *and* (1.4). *Let τ_0 be as in* Proposition 3.6. *Let $\tau < \tau_0$ and let $\nu \in \mathscr{P}_2(\mathbb{R}^d)$. For any $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ and for any sequence $\{\mu_n\}_{n\in\mathbb{N}}$ such that μ_n narrowly converges to μ and $\sup_n \int_{\mathbb{R}^d} |x|^2\, d\mu_n(x) < +\infty$, there holds*

$$\mathcal{W}(\mu) + \frac{1}{2\tau}\, d_W^2(\mu, \nu) \le \liminf_n \left(\mathcal{W}_n(\mu_n) + \frac{1}{2\tau}\, d_W^2(\mu_n, \nu) \right).$$

Moreover, for $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ there holds $\mathcal{W}_n(\mu) \to \mathcal{W}(\mu)$.

Proof. Since $\mathbf{W}_n \geq \mathbf{W}_k$ if $n \geq k > \lambda^-$, by Proposition 3.6 we have

$$\liminf_{n\to\infty} \left(\mathcal{W}_n(\mu_n) + \frac{1}{2\tau}\, d_W^2(\mu_n,\nu) \right) \geq \liminf_{n\to\infty} \left(\mathcal{W}_k(\mu_n) + \frac{1}{2\tau}\, d_W^2(\mu_n,\nu) \right)$$
$$\geq \mathcal{W}_k(\mu) + \frac{1}{2\tau}\, d_W^2(\mu,\nu)$$

for any fixed $k > \lambda^-$. Now we shall pass to the limit as $k \to +\infty$. Notice that $\mathbf{W}_k \nearrow \mathbf{W}$ pointwise and monotonically, and thus by the monotone convergence theorem, $\mathcal{W}_k(\mu)$ converges to $\mathcal{W}(\mu)$. Both statements are proven. □

We recall a suitable notion of convergence of a sequence of vector fields $\boldsymbol{\xi}_n \in L^2(\mathbb{R}^d, \mu_n; \mathbb{R}^d)$.

Definition 3.8. *Let $\{\mu_n\}_{n\in\mathbb{N}} \subset \mathscr{P}(\mathbb{R}^d)$ narrowly converge to $\mu \in \mathscr{P}(\mathbb{R}^d)$ and let $\boldsymbol{\xi}_n \in L^2(\mathbb{R}^d, \mu_n; \mathbb{R}^d)$. We say that $\boldsymbol{\xi}_n$ weakly converge to $\boldsymbol{\xi} \in L^2(\mathbb{R}^d, \mu; \mathbb{R}^d)$ if*

$$\int_{\mathbb{R}^d} \langle \boldsymbol{\xi}_n, \zeta\rangle \, d\mu_n \to \int_{\mathbb{R}^d} \langle \boldsymbol{\xi}, \zeta\rangle \, d\mu, \qquad \forall \zeta \in C_0(\mathbb{R}^d; \mathbb{R}^d), \tag{3.13}$$

where $C_0(\mathbb{R}^d; \mathbb{R}^d)$ is the space of continuous functions vanishing at infinity. We say that the convergence is strong if (3.13) *holds and*

$$\int_{\mathbb{R}^d} |\boldsymbol{\xi}_n|^2 \, d\mu_n \to \int_{\mathbb{R}^d} |\boldsymbol{\xi}|^2 \, d\mu.$$

Remark 3.9. Let $\mathcal{M}^d$ denote the space of d-dimensional vector Radon measures over $\mathbb{R}^d$ with finite total variation. By Riesz representation theorem, $\mathcal{M}^d$ is the dual space of $C_0(\mathbb{R}^d; \mathbb{R}^d)$, the total variation being the dual norm, see for instance [AFP]. Since $C_0(\mathbb{R}^d; \mathbb{R}^d)$ is a separable Banach space, the weak* topology of $\mathcal{M}^d$ is metrizable on bounded sets.

If $\mu \in \mathscr{P}(\mathbb{R}^d)$ and $\boldsymbol{\xi} \in L^2(\mathbb{R}^d, \mu; \mathbb{R}^d)$, then $\boldsymbol{\xi}\mu \in \mathcal{M}^d$, since its total variation is $|\boldsymbol{\xi}\mu|(\mathbb{R}^d) = \int_{\mathbb{R}^d} |\boldsymbol{\xi}| d\mu < +\infty$. Letting μ_n narrowly converge to μ and $\boldsymbol{\xi}_n \in L^2(\mathbb{R}^d, \mu_n; \mathbb{R}^d)$, it is clear that the weak convergence $\boldsymbol{\xi}_n \to \boldsymbol{\xi}$ in the sense of Definition 3.8 is exactly the weak* convergence in $\mathcal{M}^d$ of the vector measures $\boldsymbol{\xi}_n\mu_n$ to the vector measure $\boldsymbol{\xi}\mu$. Therefore $\boldsymbol{\xi}_n$ weakly converge to $\boldsymbol{\xi}$ in the sense of Definition 3.8 if and only if $\mathbf{d}(\boldsymbol{\xi}_n\mu_n, \boldsymbol{\xi}\mu) \to 0$, where $\mathbf{d}$ is a distance which metrizes the weak* topology on $B_R := \{\boldsymbol{\nu} \in \mathcal{M}^d : |\boldsymbol{\nu}|(\mathbb{R}^d) \leq R\}$ and $R = \sup_n \int_{\mathbb{R}^d} |\boldsymbol{\xi}_n| \, d\mu_n$.

We also need to define the barycentric projection.

Definition 3.10 (Disintegration and barycenter). *Let $\mu, \nu \in \mathscr{P}(\mathbb{R}^d)$. Given $\beta \in \Gamma(\mu,\nu)$, we denote by $\{\beta_x\}_{x\in\mathbb{R}^d}$ the Borel family of measures in $\mathscr{P}(\mathbb{R}^d)$ such that $\beta = \int_{\mathbb{R}^d} \beta_x \, d\mu(x)$, which disintegrates β with respect to μ. The notation above means that the integral of a Borel function $\varphi : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ such that $\varphi \in L^1(\beta)$, can be sliced as*

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} \varphi(x,y)\, d\beta(x,y) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \varphi(x,y)\, d\beta_x(y)\, d\mu(x).$$

The barycentric projection of $\beta \in \Gamma(\mu,\nu)$ is defined by

$$\bar{\beta}(x) := \int_{\mathbb{R}^d} y \, d\beta_x(y).$$

The family $\{\beta_x\}_{x\in\mathbb{R}^d}$ is uniquely determined for μ-a.e. $x \in \mathbb{R}^d$, for more details about disintegration see [AFP, Theorem 2.28]. Notice that if $\nu \in \mathscr{P}_2(\mathbb{R}^d)$, then $\bar\beta \in L^2(\mu)$. Indeed, by Jensen inequality we have

$$
\begin{aligned}
\int_{\mathbb{R}^d} |\bar\beta(x)|^2\,d\mu(x) &= \int_{\mathbb{R}^d} \left|\int_{\mathbb{R}^d} y\,d\beta_x(y)\right|^2 d\mu(x) \le \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |y|^2\,d\beta_x(y)\,d\mu(x) \\
&= \int_{\mathbb{R}^d\times\mathbb{R}^d} |y|^2\,d\beta(x,y) = \int_{\mathbb{R}^d} |y|^2\,d\nu(y).
\end{aligned} \tag{3.14}
$$

We can prove the following simple

Proposition 3.11. *Let $\{\mu_n\} \subset \mathscr{P}_2(\mathbb{R}^d)$, $\{\nu_n\} \subset \mathscr{P}_2(\mathbb{R}^d)$ be sequences with uniformly bounded second moments and narrowly converging to μ and ν respectively. For every sequence $\{\gamma_n\}$ such that $\gamma_n \in \Gamma_o(\mu_n,\nu_n)$, there exists a subsequence $\{\gamma_{n_k}\}$ and $\gamma \in \Gamma(\mu,\nu)$ such that γ_{n_k} narrowly converge to γ. Moreover, $\gamma \in \Gamma_o(\mu,\nu)$ and*

$$\bar\gamma_{n_k} \to \bar\gamma \quad \textit{weakly in the sense of } \text{Definition 3.8} \textit{ as } k \to +\infty.$$

Proof. By the assumptions on the moments of μ_n and ν_n we have

$$\sup_n \int_{\mathbb{R}^d\times\mathbb{R}^d} (|x|^2 + |y|^2)\,d\gamma_n(x,y) < +\infty, \tag{3.15}$$

hence $\{\gamma_n\}$ is tight. By Prokhorov theorem it admits a narrow limit point $\gamma \in \Gamma(\mu,\nu)$. For the proof of optimality of γ, see [AGS, Proposition 7.1.3]. Without relabeling the subsequence, let γ_n narrowly converge to γ. Let $(\gamma_n)_x$ be the disintegration of γ_n with respect to μ_n and γ_x be the disintegration of γ with respect to μ. Let $\zeta \in C_0(\mathbb{R}^d;\mathbb{R}^d)$ and $f : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ the function defined by $f(x,y) = \langle \zeta(x), y\rangle$. Since f is continuous and satisfies $|f(x,y)| \le C|y|$ for every $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$ and since (3.15) holds, by [AGS, Lemma 5.1.7] we have that $\int_{\mathbb{R}^d\times\mathbb{R}^d} f(x,y)\,d\gamma_n(x,y) \to \int_{\mathbb{R}^d\times\mathbb{R}^d} f(x,y)\,d\gamma(x,y)$ as $n \to +\infty$. Using this property and the definition of barycenter, we have

$$
\begin{aligned}
\int_{\mathbb{R}^d} \langle \zeta, \bar\gamma_n\rangle\,d\mu_n &= \int_{\mathbb{R}^d} \left\langle \zeta(x), \int_{\mathbb{R}^d} y\,d(\gamma_n)_x(y)\right\rangle d\mu_n(x) \\
&= \int_{\mathbb{R}^d\times\mathbb{R}^d} \langle \zeta(x), y\rangle\,d\gamma_n(x,y) \to \int_{\mathbb{R}^d\times\mathbb{R}^d} \langle \zeta(x), y\rangle\,d\gamma(x,y) \\
&= \int_{\mathbb{R}^d} \left\langle \zeta(x), \int_{\mathbb{R}^d} y\,d\gamma_x(y)\right\rangle d\mu(x) = \int_{\mathbb{R}^d} \langle \zeta, \bar\gamma\rangle\,d\mu
\end{aligned}
$$

as $n \to +\infty$. □

We recall a definition from [AGS, Chapter 10].

Definition 3.12 (Rescaled plan). *Let* $\mathbf{W}$ *satisfy assumptions* (1.2), (1.3), *and* (1.4). *Let $\tau < \tau_0$, as in* Proposition 3.6. *Given $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, let*

$$\mu_\tau \in \operatorname{argmin}\left\{ \mathcal{W}(\nu) + \frac{1}{2\tau} d_W(\nu,\mu) : \nu \in \mathscr{P}_2(\mathbb{R}^d)\right\}.$$

Given $\hat\gamma_\tau \in \Gamma_o(\mu_\tau,\mu)$, we define the rescaled plan as

$$\gamma_\tau := \left(\pi^1, \frac{\pi^2 - \pi^1}{\tau}\right)_{\#} \hat\gamma_\tau.$$

Next we introduce an abstract result about approximation of the minimal selection in the subdifferential of $\mathcal{W}$. The argument is indeed a direct consequence of the analysis in [AGS, §10.3], but requires the concept of plan subdifferential. Since this is a technical definition, we prefer to postpone a discussion at the end of the paper. Therefore, the proof of the following proposition is given in the appendix.

Proposition 3.13. *Let* $\mathbf{W}$ *satisfy assumptions* (1.2), (1.3), *and* (1.4). *Let* μ, μ_τ *and* γ_τ *be as in* Definition 3.12. *Then,*

$$\lim_{\tau\to 0}\frac{d_W^2(\mu_\tau,\mu)}{\tau^2}=\int_{\mathbb{R}^d}|\partial^o\mathcal{W}(\mu)|^2\,d\mu<+\infty. \tag{3.16}$$

Moreover, denoting by $\bar\gamma_\tau$ *the barycenter of* γ_τ, *we have that* $\bar\gamma_\tau\in\partial_S\mathcal{W}(\mu_\tau)$ *and*

$$\bar\gamma_\tau\to\partial^o\mathcal{W}(\mu)\quad \textit{strongly in the sense of } \text{Definition 3.8 } \textit{as } \tau\to 0.$$

Making use of Proposition 3.13 we can prove the following

Lemma 3.14. *Let* $\mu\in\mathscr{P}_2(\mathbb{R}^d)$. *Let* $\mathbf{W}$ *satisfy assumptions* (1.2), (1.3), *and* (1.4). *There exist a sequence* $\{\mu_n\}_{n\in\mathbb{N}}\subset\mathscr{P}_2(\mathbb{R}^d)$ *and a strictly increasing sequence* $\{k(n)\}_{n\in\mathbb{N}}\subset\mathbb{N}$ *such that* μ_n *narrowly converge to* μ, $\sup_n\int_{\mathbb{R}^d}|x|^2\,d\mu_n(x)<+\infty$ *and*

$$\int_{\mathbb{R}^d}\nabla_1\mathbf{W}_{k(n)}(\cdot,y)\,d\mu_n(y)\to\partial^o\mathcal{W}(\mu)\quad \textit{weakly in the sense of } \text{Definition 3.8 } \textit{as } n\to+\infty.$$

Proof. Let τ_0 be as in Proposition 3.6, and consider a measure $\mu\in\mathscr{P}_2(\mathbb{R}^d)$. Let, for $\tau<\tau_0$,

$$\mu_\tau^h\in\operatorname{argmin}\left\{\mathcal{W}_h(\nu)+\frac{1}{2\tau}\,d_W^2(\nu,\mu):\nu\in\mathscr{P}_2(\mathbb{R}^d)\right\},\quad h\in\mathbb{N},\quad h>\lambda^-.$$

We claim that there exists $\tilde\tau\in(0,\tau_0)$ such that

$$\sup_{h>\lambda^-,\,\tau\le\tilde\tau}\int_{\mathbb{R}^d}|x|^2\,d\mu_\tau^h(x)<+\infty. \tag{3.17}$$

Indeed, since μ_τ^h satisfies the above minimality property, letting $\gamma_\tau^h\in\Gamma_o(\mu_\tau^h,\mu)$, a direct estimate using (3.8) and $\mathbf{W}_h\le\mathbf{W}$ shows that

$$\begin{aligned}\int_{\mathbb{R}^d}|x|^2\,d\mu_\tau^h(x)&=\int_{\mathbb{R}^d\times\mathbb{R}^d}|x|^2\,d\gamma_\tau^h(x,y)\le 2\int_{\mathbb{R}^d\times\mathbb{R}^d}|y|^2\,d\gamma_\tau^h(x,y)+2d_W^2(\mu_\tau^h,\mu)\\&\le 2\int_{\mathbb{R}^d}|y|^2\,d\mu(y)+4\tau\left(\mathcal{W}_h(\mu)-\mathcal{W}_h(\mu_\tau^h)\right)\\&\le 2\int_{\mathbb{R}^d}|y|^2\,d\mu(y)+4\tau\mathcal{W}(\mu)+4\tau\bar K\left(1+2\int_{\mathbb{R}^d}|x|^2\,d\mu_\tau^h(x)\right).\end{aligned}$$

The claim follows choosing $\tilde\tau$ smaller than $\frac{1}{8\bar K}$.

We are going to construct the desired sequence $\{\mu_n\}$ by extracting it from the family $\{\mu_\tau^h\}$, and showing the suitable convergence properties of the associated optimal transport plans, rescaled plans and barycenters. For, we let $\{\tau(n)\}_{n\in\mathbb{N}}\subset(0,\tilde\tau)$ be a vanishing sequence and we proceed in some steps.

Step 1. For any fixed $n\in\mathbb{N}$, by (3.17) the sequence $\{\mu_{\tau(n)}^h\}_{h\in\mathbb{N},\,h>\lambda^-}$ has uniformly bounded second moments, hence it is tight and by Prokhorov theorem it has narrow limit points as $h\to+\infty$. With a standard diagonal argument, we can find a strictly increasing sequence $\{h(m)\}_{m\in\mathbb{N}} \subset \mathbb{N}$ such that, for any $n \in \mathbb{N}$, the sequence $\{\mu^{h(m)}_{\tau(n)}\}_{m\in\mathbb{N}}$ narrowly converges as $m \to +\infty$. Let $\mu_{\tau(n)}$ denote the corresponding limit point. Proposition 3.7 yields, for any $\nu \in \mathscr{P}_2(\mathbb{R}^d)$,

$$\begin{aligned}\mathcal{W}(\mu_{\tau(n)}) + \frac{1}{2\tau(n)}\, d^2_W(\mu_{\tau(n)}, \mu) &\le \liminf_{m\to\infty} \left(\mathcal{W}_{h(m)}(\mu^{h(m)}_{\tau(n)}) + \frac{1}{2\tau(n)}\, d^2_W(\mu^{h(m)}_{\tau(n)}, \mu)\right) \\ &\le \liminf_{m\to\infty} \left(\mathcal{W}_{h(m)}(\nu) + \frac{1}{2\tau(n)}\, d^2_W(\nu, \mu)\right) \\ &= \mathcal{W}(\nu) + \frac{1}{2\tau(n)}\, d^2_W(\nu, \mu).\end{aligned}$$

This shows that, for any $n \in \mathbb{N}$,

$$\mu_{\tau(n)} \in \operatorname{argmin}\left\{\mathcal{W}(\nu) + \frac{1}{2\tau(n)}\, d^2_W(\nu, \mu) : \nu \in \mathscr{P}_2(\mathbb{R}^d)\right\}. \tag{3.18}$$

Step 2. Let $\hat{\gamma}^{h(m)}_{\tau(n)} \in \Gamma_o(\mu^{h(m)}_{\tau(n)}, \mu)$. Applying Proposition 3.11 to the sequence $\{\hat{\gamma}^{h(m)}_{\tau(n)}\}_{m\in\mathbb{N}}$ we find a corresponding narrow limit point $\hat{\gamma}_{\tau(n)} \in \Gamma_o(\mu_{\tau(n)}, \mu)$. By possibly extracting from $\{h(m)\}$ a subsequence, that we do not relabel, we have the narrow convergence $\hat{\gamma}^{h(m)}_{\tau(n)} \to \hat{\gamma}_{\tau(n)}$ as $m \to +\infty$, for any $n \in \mathbb{N}$. As in the proof of Proposition 3.11, we also have

$$\lim_{m\to+\infty} \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|\, d\hat{\gamma}^{h(m)}_{\tau(n)}(x,y) = \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|\, d\hat{\gamma}_{\tau(n)}(x,y),$$

since $(x, y) \mapsto |x-y|$ is continuous and has a linear growth. Therefore, it is possible to extract from $\{h(m)\}$ another subsequence, still not relabeled, such that

$$\left|\int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|\, d\hat{\gamma}^{h(m)}_{\tau(n)}(x,y) - \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|\, d\hat{\gamma}_{\tau(n)}(x,y)\right| \le \tau(n) \quad \text{if } m > n. \tag{3.19}$$

Step 3. If $\gamma^{h(m)}_{\tau(n)}$ denotes the rescaled of plan $\hat{\gamma}^{h(m)}_{\tau(n)}$ (see Definition 3.12) it easily follows that $\gamma^{h(m)}_{\tau(n)} \to \gamma_{\tau(n)}$ as $m \to +\infty$, where $\gamma_{\tau(n)}$ is the rescaled of $\hat{\gamma}_{\tau(n)}$. Moreover, for any $n \in \mathbb{N}$ the sequence of plans $\{\gamma^{h(m)}_{\tau(n)}\}_{m\in\mathbb{N}}$ has uniformly bounded second moments, as seen from the same property for $\{\hat{\gamma}^{h(m)}_{\tau(n)}\}_{m\in\mathbb{N}}$ and by definition of rescaled plan, so that reasoning as in the proof of Proposition 3.11 we get the convergence of the respective barycenters, that is,

$$\bar{\gamma}^{h(m)}_{\tau(n)} \to \bar{\gamma}_{\tau(n)} \quad \text{weakly in the sense of Definition 3.8 as } m \to +\infty. \tag{3.20}$$

Step 4. We conclude by combining the estimates of the previous steps and the application of Proposition 3.13. Indeed, since (3.18) holds, $\mu, \mu_{\tau(n)}$ and $\gamma_{\tau(n)}$ match the assumptions of Proposition 3.13, which can be applied by passing to the limit as $\tau \to 0$ along the sequence $\tau(n)$. Using the definition of barycenter and rescaled plan (denoting by $(\gamma^{h(m)}_{\tau(n)})_x$ the disintegration of $\gamma^{h(m)}_{\tau(n)}$ with respect to $\mu^{h(m)}_{\tau(n)}$), thanks to (3.19) we find the estimate

$$\begin{aligned}\int_{\mathbb{R}^d} |\bar{\gamma}^{h(m)}_{\tau(n)}(x)|\, d\mu^{h(m)}_{\tau(n)}(x) &\le \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |y|\, d(\gamma^{h(m)}_{\tau(n)})_x(y)\, d\mu^{h(m)}_{\tau(n)}(x) = \int_{\mathbb{R}^d\times\mathbb{R}^d} |y|\, d\gamma^{h(m)}_{\tau(n)}(x,y)\\ &= \int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{|y-x|}{\tau(n)} d\hat{\gamma}^{h(m)}_{\tau(n)}(x,y) \le 1 + \int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{|y-x|}{\tau(n)} d\hat{\gamma}_{\tau(n)}(x,y)\\ &\le 1 + \frac{d_W(\mu_{\tau(n)},\mu)}{\tau(n)},\end{aligned}$$

for any $m,n \in \mathbb{N}$ such that $m > n$. But the right hand side is converging as $n \to \infty$, due to Proposition 3.13, then it is bounded by some constant $R > 0$. This shows that the family of vector measures $\{\bar{\gamma}^{h(m)}_{\tau(n)}\mu^{h(m)}_{\tau(n)} : n, m \in \mathbb{N}, m > n\}$ is contained in the ball $B_R := \{\boldsymbol{\nu} \in \mathcal{M}^d : |\boldsymbol{\nu}|(\mathbb{R}^d) \le R\}$. By the arguments in Remark 3.9, the weak* convergence of vector measures is metrizable on B_R, and letting $\mathbf{d}$ denote a corresponding distance, property (3.20) translates into

$$\mathbf{d}(\bar{\gamma}^{h(m)}_{\tau(n)}\mu^{h(m)}_{\tau(n)}, \bar{\gamma}_{\tau(n)}\mu_{\tau(n)}) \to 0 \quad \text{as } m \to +\infty, \text{ for any } n \in \mathbb{N}.$$

Therefore, there exists a strictly increasing sequence $\{k(n)\}_{n\in\mathbb{N}} \subset \{h(m)\}_{m\in\mathbb{N}}$ such that

$$\mathbf{d}(\bar{\gamma}^{k(n)}_{\tau(n)}\mu^{k(n)}_{\tau(n)}, \bar{\gamma}_{\tau(n)}\mu_{\tau(n)}) < \frac{1}{n}.$$

Then we have

$$\begin{aligned}\mathbf{d}(\bar{\gamma}^{k(n)}_{\tau(n)}\mu^{k(n)}_{\tau(n)}, \partial^o\mathcal{W}(\mu)\mu) &\le \mathbf{d}(\bar{\gamma}^{k(n)}_{\tau(n)}\mu^{k(n)}_{\tau(n)}, \bar{\gamma}_{\tau(n)}\mu_{\tau(n)}) + \mathbf{d}(\bar{\gamma}_{\tau(n)}\mu_{\tau(n)}, \partial^o\mathcal{W}(\mu)\mu)\\ &\le \frac{1}{n} + \mathbf{d}(\bar{\gamma}_{\tau(n)}\mu_{\tau(n)}, \partial^o\mathcal{W}(\mu)\mu).\end{aligned}$$

Invoking Proposition 3.13, we know that $\bar{\gamma}_{\tau(n)}$ converge to $\partial^o\mathcal{W}(\mu)$ strongly (and thus weakly) in the sense of Definition 3.8 as $n \to +\infty$. Hence, letting $\gamma_n := \gamma^{k(n)}_{\tau(n)}$, $\mu_n := \pi^1_\#\gamma_n$ and $\bar{\gamma}_n$ the barycenter of γ_n, passing to the limit as $n \to \infty$, we see that $\bar{\gamma}_n$ weakly converge to $\partial^o\mathcal{W}(\mu)$ in the sense of Definition 3.8.

By Proposition 3.13, applied to $\mathbf{W}_{k(n)}$, for any n there holds $\bar{\gamma}_n \in \partial_S\mathcal{W}_{k(n)}(\mu_n)$. Since $\mathbf{W}_{k(n)} \in C^1(\mathbb{R}^d \times \mathbb{R}^d)$, by the characterization of strong subdifferential of Proposition 3.2 we have

$$\bar{\gamma}_n(x) = \int_{\mathbb{R}^d} \nabla_1 \mathbf{W}_{k(n)}(x,y)\, d\mu_n(y)$$

and the proof is concluded. □

Proof of **Theorem 3.3**. Let $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ and $\boldsymbol{\xi} := \partial^o\mathcal{W}(\mu)$. Let $\{\mu_n\} \subset \mathscr{P}_2(\mathbb{R}^d)$ and $\{k(n)\} \subset \mathbb{N}$ be the sequences given by Lemma 3.14. In particular, μ_n narrowly converge to μ and

$$\sup_n \int_{\mathbb{R}^d} |x|^2\, d\mu_n(x) < +\infty. \tag{3.21}$$

Let us consider the maps $(\mathbf{i}, \nabla\mathbf{W}_{k(n)}) : \mathbb{R}^d \times \mathbb{R}^d \longrightarrow (\mathbb{R}^d)^4$ given by

$$(\mathbf{i}, \nabla\mathbf{W}_{k(n)})(x,y) = (x, y, \nabla\mathbf{W}_{k(n)}(x,y)),$$

and let us introduce the measures

$$\nu_n := (\mathbf{i}, \nabla \mathbf{W}_{k(n)})_{\#}(\mu_n \times \mu_n).$$

Since

$$\int_{(\mathbb{R}^d)^4} \left(|x|^2 + |y|^2 + |v|^2 + |w|^2\right) d\nu_n(x,y,v,w) =$$
$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \left(|x|^2 + |y|^2 + |\nabla_1 \mathbf{W}_{k(n)}(x,y)|^2 + |\nabla_2 \mathbf{W}_{k(n)}(x,y)|^2\right) d\mu_n(x)\, d\mu_n(y),$$

from (3.9) and (3.21) we infer that the second moments of ν_n are uniformly bounded. Then the sequence $\{\nu_n\}$ is tight and by Prokhorov theorem there exists a subsequence (that we do not relabel) which narrowly converges to some $\nu \in \mathscr{P}((\mathbb{R}^d)^4)$. Moreover, for $\zeta \in C_0(\mathbb{R}^d;\mathbb{R}^d)$, we have

$$\lim_{n\to\infty} \int_{(\mathbb{R}^d)^4} \langle v, \zeta(x)\rangle\, d\nu_n(x,y,v,w) = \int_{(\mathbb{R}^d)^4} \langle v, \zeta(x)\rangle\, d\nu(x,y,v,w), \tag{3.22}$$

due to the linear growth of the integrand (as in the proof of Proposition 3.11). As a consequence,

$$\int_{(\mathbb{R}^d)^4} \langle v, \zeta(x)\rangle\, d\nu(x,y,v,w) = \int_{(\mathbb{R}^d)^4} \langle w, \zeta(y)\rangle\, d\nu(x,y,v,w), \tag{3.23}$$

since this identity holds for ν_n (being $\mathbf{W}_n$ symmetric).

The narrow convergence of measures implies that $\mathrm{supp}(\nu)$ is contained in the Kuratowski minimum limit of the supports of ν_n (see for instance [AGS, Proposition 5.1.8]), i.e. for every $(x,y,\boldsymbol{\eta}) \in \mathrm{supp}(\nu)$ there exists a sequence $(x_n, y_n, \boldsymbol{\eta}_n) \in \mathrm{supp}(\nu_n)$ such that $(x_n, y_n, \boldsymbol{\eta}_n)$ converges to $(x,y,\boldsymbol{\eta})$. Since, by definition of ν_n, $\mathrm{supp}(\nu_n) \subset \mathrm{graph}(\partial \mathbf{W}_{k(n)})$, then $\mathrm{supp}(\nu) \subset \mathrm{graph}(\partial \mathbf{W})$. Indeed $\boldsymbol{\eta}_n \in \partial \mathbf{W}_{k(n)}(x_n, y_n)$ and passing to the limit in the subdifferential inequality we obtain that $\boldsymbol{\eta} \in \partial \mathbf{W}(x,y)$.

Disintegrating ν with respect to $\mu \times \mu$, we obtain the measurable family of measures $(x,y) \mapsto \nu_{x,y}$ such that

$$\nu = \int_{\mathbb{R}^d \times \mathbb{R}^d} \nu_{x,y}\, d(\mu \times \mu)(x,y).$$

Since $\mathrm{supp}(\nu) \subset \mathrm{graph}(\partial \mathbf{W})$, it follows that $\mathrm{supp}(\nu_{x,y}) \subset \partial \mathbf{W}(x,y)$.

Making use of (3.22) and (3.23) we have,

$$\begin{aligned}
\lim_{n\to\infty} \int_{(\mathbb{R}^d)^4} \langle v,\zeta(x)\rangle\, d\nu_n(x,y,v,w) &= \int_{(\mathbb{R}^d)^4} \langle v, \zeta(x)\rangle\, d\nu(x,y,v,w) \\
&= \frac{1}{2}\int_{(\mathbb{R}^d)^4} \langle v, \zeta(x)\rangle\, d\nu(x,y,v,w) + \frac{1}{2}\int_{(\mathbb{R}^d)^4} \langle w, \zeta(y)\rangle\, d\nu(x,y,v,w) \\
&= \frac{1}{2}\int_{\mathbb{R}^d} \left\langle \int_{\mathbb{R}^d}\int_{\mathbb{R}^d\times\mathbb{R}^d} v\, d\nu_{x,y}(v,w)\, d\mu(y), \zeta(x) \right\rangle d\mu(x) \\
&\quad + \frac{1}{2}\int_{\mathbb{R}^d} \left\langle \int_{\mathbb{R}^d}\int_{\mathbb{R}^d\times\mathbb{R}^d} w\, d\nu_{x,y}(v,w)\, d\mu(x), \zeta(y) \right\rangle d\mu(y) \\
&= \int_{\mathbb{R}^d} \left\langle \int_{\mathbb{R}^d}\int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{1}{2}\left(v\, d\nu_{x,y} + w\, d\nu_{y,x}\right)(v,w)\, d\mu(y), \zeta(x) \right\rangle d\mu(x).
\end{aligned} \tag{3.24}$$

We define

$$(\boldsymbol{\eta}^1(x,y),\boldsymbol{\eta}^2(x,y)) := \bar{\nu}(x,y) = \left(\int_{\mathbb{R}^d\times\mathbb{R}^d} v\,d\nu_{x,y}(v,w), \int_{\mathbb{R}^d\times\mathbb{R}^d} w\,d\nu_{x,y}(v,w)\right),$$

that is, $(\boldsymbol{\eta}^1,\boldsymbol{\eta}^2)$ is the barycenter of ν. This way, (3.24) becomes

$$\lim_{n\to\infty}\int_{(\mathbb{R}^d)^4}\langle v,\zeta(x)\rangle\,d\nu_n(x,y,v,w) = \int_{\mathbb{R}^d}\left\langle \int_{\mathbb{R}^d}\frac{1}{2}\,\left(\boldsymbol{\eta}^1(x,y)+\boldsymbol{\eta}^2(y,x)\right)\,d\mu(y),\zeta(x)\right\rangle\,d\mu(x).$$

Now, invoking Lemma 3.14 we have the weak convergence in the sense of Definition 3.8 of $\int_{\mathbb{R}^d}\nabla_1\mathbf{W}_{k(n)}(\cdot,y)\,d\mu_n(y)$ to $\boldsymbol{\xi}=\partial^o\mathcal{W}(\mu)$ as $n\to+\infty$. Using the definition of ν_n, it follows that for any $\zeta\in C_0(\mathbb{R}^d;\mathbb{R}^d)$

$$\begin{aligned}\lim_{n\to\infty}\int_{(\mathbb{R}^d)^4}\langle v,\zeta(x)\rangle\,d\nu_n(x,y,v,w) &= \lim_{n\to\infty}\int_{\mathbb{R}^d}\left\langle\int_{\mathbb{R}^d}\nabla_1\mathbf{W}_{k(n)}(x,y)\,d\mu(y),\zeta(x)\right\rangle\,d\mu_n(x)\\ &= \int_{\mathbb{R}^d}\langle\boldsymbol{\xi}(x),\zeta(x)\rangle\,d\mu(x).\end{aligned}$$

We conclude

$$\boldsymbol{\xi}(x) = \int_{\mathbb{R}^d}\frac{1}{2}\,\left(\boldsymbol{\eta}^1(x,y)+\boldsymbol{\eta}^2(y,x)\right)\,d\mu(y) \qquad \text{for } \mu\text{-a.e. } x\in\mathbb{R}^d.$$

On the other hand, we proved that $\operatorname{supp}(\nu_{x,y})\subset\partial\mathbf{W}(x,y)$. Since $(\boldsymbol{\eta}^1,\boldsymbol{\eta}^2)$ is the barycenter of ν and $\partial\mathbf{W}(x,y)$ is convex, we have that $\left(\boldsymbol{\eta}^1(x,y),\boldsymbol{\eta}^2(x,y)\right)\in\partial\mathbf{W}(x,y)$. □

4. The convolution case

4.1. **The case of W depending on the difference.** In the case of assumption (1.8) we can particularize the results above.

Lemma 4.1. *Let* $\mathbf{W}$ *satisfy* (1.2), (1.3)*, and* (1.4)*. Let moreover* (1.8) *hold. Then,* $(\boldsymbol{\eta}^1,\boldsymbol{\eta}^2)\in\partial\mathbf{W}$ *if and only if there exists* $\boldsymbol{\eta}\in\partial W$ *such that* $(\boldsymbol{\eta}^1,\boldsymbol{\eta}^2)=(\boldsymbol{\eta},-\boldsymbol{\eta})$.

Proof. Assume that W is convex, the general case follows considering $x\mapsto W(x)-\frac{\lambda}{2}|x|^2$. If $\boldsymbol{\eta}\in\partial W$ we have for every $(\tilde{x},\tilde{y})\in\mathbb{R}^d\times\mathbb{R}^d$

$$\begin{aligned}\mathbf{W}(\tilde{x},\tilde{y})-\mathbf{W}(x,y) = W(\tilde{x}-\tilde{y})-W(x-y) &\geq \langle\boldsymbol{\eta}(x-y),\tilde{x}-\tilde{y}-(x-y)\rangle\\ &= \langle(\boldsymbol{\eta}(x-y),-\boldsymbol{\eta}(x-y)),(\tilde{x},\tilde{y})-(x,y)\rangle,\end{aligned}$$

which means that $(\boldsymbol{\eta}(x-y),-\boldsymbol{\eta}(x-y))\in\partial\mathbf{W}(x,y)$ by making the abuse of notation $\boldsymbol{\eta}(x,y)\equiv\boldsymbol{\eta}(x-y)$.

On the other hand, if $(\boldsymbol{\eta}^1,\,\boldsymbol{\eta}^2)\in\partial\mathbf{W}$, then for every $(\tilde{x},\tilde{y})\in\mathbb{R}^d\times\mathbb{R}^d$ we have

$$\begin{aligned}W(\tilde{x}-\tilde{y})-W(x-y) = \mathbf{W}(\tilde{x},\tilde{y})-\mathbf{W}(x,y) &\geq \left\langle\left(\boldsymbol{\eta}^1(x,y),\boldsymbol{\eta}^2(x,y)\right),((\tilde{x},\tilde{y})-(x,y))\right\rangle\\ &= \langle\boldsymbol{\eta}^1(x,y),(\tilde{x}-x)\rangle+\langle\boldsymbol{\eta}^2(x,y),(\tilde{y}-y)\rangle.\end{aligned} \tag{4.1}$$

Assuming in particular that $x-y=\tilde{x}-\tilde{y}$ the inequality above reduces to

$$0\geq\langle\boldsymbol{\eta}^1(x,y)+\boldsymbol{\eta}^2(x,y),(\tilde{y}-y)\rangle,$$

and the arbitrariness of $\tilde{y}$ implies that $\boldsymbol{\eta}^1=-\boldsymbol{\eta}^2$. Using this relation in (4.1) we obtain

$$W(\tilde{x}-\tilde{y})-W(x-y)\geq\langle\boldsymbol{\eta}^1(x,y),(\tilde{x}-\tilde{y})-(x-y)\rangle,$$

which means that $\boldsymbol{\eta}^1\in\partial W$. □

Applying the main results of the general case we obtain the following

Corollary 4.2. *Let the assumptions of* Lemma 4.1 *hold. If* $\boldsymbol{\eta}$ *is a Borel measurable anti-symmetric selection of* ∂W*, then*

$$\boldsymbol{\eta} * \mu \in \partial_S \mathcal{W}(\mu) \qquad \text{for any } \mu \in \mathscr{P}_2(\mathbb{R}^d), \tag{4.2}$$

where $(\boldsymbol{\eta} * \mu)(x) = \int_{\mathbb{R}^d} \boldsymbol{\eta}(x-y)\, d\mu(y)$.

Conversely, if $\mu \in \mathscr{P}_2(\mathbb{R}^d)$*, then there exists a Borel measurable anti-symmetric selection* $\boldsymbol{\eta} \in \partial W$ *such that* $\partial^o \mathcal{W}(\mu) = \boldsymbol{\eta} * \mu$.

Proof. Let $\boldsymbol{\eta}$ be a Borel measurable antisymmetric selection in ∂W. In particular $\boldsymbol{\eta}(x-y) = -\boldsymbol{\eta}(y-x)$ for every $x, y \in \mathbb{R}^d$ and by Lemma 4.1 $(\boldsymbol{\eta}^1(x,y), \boldsymbol{\eta}^2(x,y)) := (\boldsymbol{\eta}(x-y), \boldsymbol{\eta}(y-x)) \in \partial \mathbf{W}(x,y)$. By applying Theorem 3.1 to the Borel measurable selection $(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2)$ just defined, we get (4.2).

Conversely, let $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, $\boldsymbol{\xi} = \partial^o \mathcal{W}(\mu)$ and $(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2) \in \partial \mathbf{W}$ be the selection given by Theorem 3.3. By Lemma 4.1 we obtain that $\boldsymbol{\eta}^2 = -\boldsymbol{\eta}^1$, then

$$\boldsymbol{\xi}(x) = \int_{\mathbb{R}^d} \frac{1}{2}(\boldsymbol{\eta}^1(x-y) - \boldsymbol{\eta}^1(y-x))\, d\mu(y) \qquad \text{for } \mu\text{-a.e. } x \in \mathbb{R}^d.$$

By choosing $\boldsymbol{\eta}(z) = \frac{1}{2}(\boldsymbol{\eta}^1(z) - \boldsymbol{\eta}^1(-z))$, we conclude. □

Remark 4.3. Under the assumptions of Lemma 4.1, we observe that if $\boldsymbol{\eta}$ is a Borel measurable selection of ∂W such that (4.2) holds, then $\boldsymbol{\eta}$ is antisymmetric.

Indeed, supposing for simplicity that W is convex, in such a case $\boldsymbol{\eta}$ satisfies

$$\mathcal{W}(\nu) - \mathcal{W}(\mu) \geq \int_{\mathbb{R}^d} \left\langle \int_{\mathbb{R}^d} \boldsymbol{\eta}(x-y)\, d\mu(y), z - x \right\rangle d\gamma(x,z)$$

for any $\gamma \in \Gamma(\mu, \nu)$. Choosing $\mu = \delta_{x_1}$ and $\nu = \delta_{x_3}$, $\gamma = \mu \times \nu$, the inequality becomes

$$0 \geq \langle \boldsymbol{\eta}(0), x_2 - x_1 \rangle,$$

and since this must hold for any $x_1, x_2 \in \mathbb{R}^d$, we deduce $\boldsymbol{\eta}(0) = 0$. Moreover, taking into account that $\boldsymbol{\eta}(0) = 0$, if $\mu = \frac{1}{2}\delta_{x_1} + \frac{1}{2}\delta_{x_2}$, $\nu = \frac{1}{2}\delta_{x_3} + \frac{1}{2}\delta_{x_4}$, with $|x_1 - x_3| \leq |x_1 - x_4|$ and $|x_2 - x_4| \leq |x_2 - x_3|$, and $\gamma = \frac{1}{2}(\delta_{x_1} \times \delta_{x_3}) + \frac{1}{2}(\delta_{x_2} \times \delta_{x_4})$, the subdifferential inequality reduces to

$$W(x_3 - x_4) - W(x_1 - x_2) \geq \langle \boldsymbol{\eta}(x_1 - x_2), x_3 - x_1 \rangle + \langle \boldsymbol{\eta}(x_2 - x_1), x_4 - x_2 \rangle.$$

In particular, for $x_3 - x_4 = x_1 - x_2$ we get

$$0 \geq \langle \boldsymbol{\eta}(x_1 - x_2) + \boldsymbol{\eta}(x_2 - x_1), x_3 - x_1 \rangle,$$

which yields $\boldsymbol{\eta}(x_1 - x_2) = -\boldsymbol{\eta}(x_2 - x_1)$ for any $x_1, x_2 \in \mathbb{R}^d$.

Remark 4.4. Still under the assumptions of Lemma 4.1, if $\mu \ll \mathcal{L}^d$ we can conclude that $\boldsymbol{\eta} * \mu \in \partial \mathcal{W}(\mu)$ for any Borel selection $\boldsymbol{\eta}$ in ∂W.

Indeed, in this case the set where ∂W is not a singleton is μ-negligible. That is, in the convolution integral we can restrict to the points where W has a gradient (there is no need to select), and in that case $\nabla W * \mu$ belongs to the Wasserstein subdifferential of $\mathcal{W}$ at μ (it is actually its minimal selection), as shown in [CDFLS].

4.2. **The radial case.** In the radial case, we are able to give an explicit characterization of the minimal selection of the Wasserstein subdifferential. Before stating our theorem, we recall that $\partial^o W(x) = \operatorname{argmin}\{|\mathbf{y}| : \mathbf{y} \in \partial W(x)\}$. We have the following

Theorem 4.5. *Let* $\mathbf{W}$ *be convex and satisfying assumptions* (1.4) *and* (1.9)*. Then*

$$\partial^o \mathcal{W}(\mu) = \partial^o W * \mu \qquad \forall \mu \in \mathscr{P}_2(\mathbb{R}^d). \tag{4.3}$$

Proof. Let $\mu \in \mathscr{P}_2(\mathbb{R}^d)$. By Corollary 4.2, we know that $\partial^o \mathcal{W}(\mu)$ has the form of a convolution with an anti-symmetric selection in the subdifferential of W. Hence, in order to find the explicit form of $\partial^o \mathcal{W}(\mu)$, we have to minimize the quantity

$$\|\boldsymbol{\eta} * \mu\|^2_{L^2(\mathbb{R}^d,\mu;\mathbb{R}^d)} = \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} \boldsymbol{\eta}(x-y)\, d\mu(y) \right|^2 d\mu(x)$$

among all measurable anti-symmetric selections $\boldsymbol{\eta} \in \partial W$. Clearly the above quantity can also be written as

$$\begin{aligned} \|\boldsymbol{\eta} * \mu\|^2_{L^2(\mathbb{R}^d,\mu;\mathbb{R}^d)} &= \int_{\mathbb{R}^d} \left\langle \int_{\mathbb{R}^d} \boldsymbol{\eta}(x-y)\, d\mu(y), \int_{\mathbb{R}^d} \boldsymbol{\eta}(x-z)\, d\mu(z) \right\rangle d\mu(x) \\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \langle \boldsymbol{\eta}(x-y), \boldsymbol{\eta}(x-z)\rangle \, d\mu(x)\, d\mu(y)\, d\mu(z). \end{aligned} \tag{4.4}$$

Let us first discuss the consequences of our assumptions. Since W is radial, given a generic anti-symmetric measurable selection $\boldsymbol{\eta} \in \partial W$, we have the representation

$$\boldsymbol{\eta}(v) = \frac{\eta(v)}{|v|}\, v, \qquad \eta(v) \in \partial w(|v|), \qquad \eta(v) = \eta(-v), \quad \forall v \in \mathbb{R}^d, \tag{4.5}$$

where w is the profile function in (1.9). Notice that in general η is not radial, but it satisfies $\eta(v) = \eta(-v)$ since $\boldsymbol{\eta}$ is anti-symmetric. If $\boldsymbol{\eta} = \partial^o W$, then η is radial and precisely $\eta(v) = \partial^o w(|v|)$, where $\partial^o w$ denotes the minimal selection in the subdifferential ∂w. Moreover, the convexity hypothesis implies that $\eta(v) \geq 0$, and in particular

$$\eta(v) \geq \partial^o w(|v|) \geq 0, \quad \forall v \in \mathbb{R}^d. \tag{4.6}$$

Convexity also yields monotonicity of the subdifferential, that is, for any $x, y, z \in \mathbb{R}^d$,

$$\langle x - y, \boldsymbol{\eta}(x-z) - \boldsymbol{\eta}(y-z)\rangle \geq 0. \tag{4.7}$$

Next, define

$$\mathcal{I}_\mu(\overline{\boldsymbol{\eta}}, \widetilde{\boldsymbol{\eta}}) := \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \langle \overline{\boldsymbol{\eta}}(x-y), \widetilde{\boldsymbol{\eta}}(x-z)\rangle \, d\mu(x)\, d\mu(y)\, d\mu(z),$$

where $\overline{\boldsymbol{\eta}}$ and $\widetilde{\boldsymbol{\eta}}$ are any two anti-symmetric selections in the subdifferential of W. It is readily seen that

$$\mathcal{I}_\mu(\overline{\boldsymbol{\eta}}, \widetilde{\boldsymbol{\eta}}) = \mathcal{I}_\mu(\widetilde{\boldsymbol{\eta}}, \overline{\boldsymbol{\eta}}), \tag{4.8}$$

simply exchanging y with z in the integral above. We claim that

$$\mathcal{I}_\mu(\overline{\boldsymbol{\eta}}, \widetilde{\boldsymbol{\eta}}) \geq \mathcal{I}_\mu(\partial^o W, \widetilde{\boldsymbol{\eta}}) \geq 0. \tag{4.9}$$

For, let us consider the following representation of $\mathcal{I}_\mu$ obtained by exchanging x with y and using anti-symmetry of $\overline{\boldsymbol{\eta}}$:

$$\begin{aligned}\mathcal{I}_\mu(\overline{\boldsymbol{\eta}},\widetilde{\boldsymbol{\eta}}) &= \frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \langle \overline{\boldsymbol{\eta}}(x-y), \widetilde{\boldsymbol{\eta}}(x-z)\rangle \, d\mu(x)\, d\mu(y)\, d\mu(z) \\ &\quad + \frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \langle \overline{\boldsymbol{\eta}}(y-x), \widetilde{\boldsymbol{\eta}}(y-z)\rangle \, d\mu(y)\, d\mu(x)\, d\mu(z) \\ &= \frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \langle \overline{\boldsymbol{\eta}}(x-y), \widetilde{\boldsymbol{\eta}}(x-z) - \widetilde{\boldsymbol{\eta}}(y-z)\rangle \, d\mu(x)\, d\mu(y)\, d\mu(z).\end{aligned}$$

Thanks to the monotonicity of the subdifferential we have that the last scalar product is pointwise nonnegative. More precisely, combining (4.5), (4.6) and (4.7) we get pointwise

$$\begin{aligned}\langle \overline{\boldsymbol{\eta}}(x-y), \widetilde{\boldsymbol{\eta}}(x-z) - \widetilde{\boldsymbol{\eta}}(y-z)\rangle &= \frac{\overline{\eta}(x-y)}{|x-y|}\, \langle x-y, \widetilde{\boldsymbol{\eta}}(x-z) - \widetilde{\boldsymbol{\eta}}(y-z)\rangle \\ &\geq \frac{\partial^o w(|x-y|)}{|x-y|}\, \langle x-y, \widetilde{\boldsymbol{\eta}}(x-z) - \widetilde{\boldsymbol{\eta}}(y-z)\rangle \\ &= \langle \partial^o W(x-y), \widetilde{\boldsymbol{\eta}}(x-z) - \widetilde{\boldsymbol{\eta}}(y-z)\rangle \geq 0.\end{aligned}$$

Inserting this inequality in the representation of $\mathcal{I}_\mu$ the claim is proven.

Now, consider an arbitrary anti-symmetric selection $\boldsymbol{\eta}$ in ∂W. Using (4.9) twice and the symmetry (4.8) we get

$$\mathcal{I}_\mu(\boldsymbol{\eta},\boldsymbol{\eta}) \geq \mathcal{I}_\mu(\partial^o W,\boldsymbol{\eta}) = \mathcal{I}_\mu(\boldsymbol{\eta},\partial^o W) \geq \mathcal{I}_\mu(\partial^o W,\partial^o W).$$

Recalling (4.4), we have $\mathcal{I}_\mu(\boldsymbol{\eta},\boldsymbol{\eta}) = \|\boldsymbol{\eta} * \mu\|^2_{L^2(\mathbb{R}^d,\mu;\mathbb{R}^d)}$, thus the above inequality is equivalent to

$$\|\boldsymbol{\eta} * \mu\|^2_{L^2(\mathbb{R}^d,\mu;\mathbb{R}^d)} \geq \|\partial^o W * \mu\|^2_{L^2(\mathbb{R}^d,\mu;\mathbb{R}^d)}$$

for any anti-symmetric selection $\boldsymbol{\eta}$ in the subdifferential of ∂W. Therefore, $\partial^o W$ minimizes $\|\boldsymbol{\eta} * \mu\|^2_{L^2(\mathbb{R}^d,\mu;\mathbb{R}^d)}$ among all measurable anti-symmetric selections $\boldsymbol{\eta}$ in ∂W and then (4.3) holds. □

Example 4.6. The above result fails if we omit the convexity assumption. Indeed, let us consider a 1-dimensional example. Let

$$\widehat{W}(x) = \frac{1}{2}\,|x^2-1|.$$

Notice that this function is radial and -1-convex, and its subdifferential is

$$\partial\widehat{W}(x) = \begin{cases} x & \text{for } |x|>1, \\ -x & \text{for } |x|<1, \\ [-1,1] & \text{for } x=\pm 1. \end{cases}$$

Let us consider the measure $\mu = \frac{1}{3}\delta_{x_1} + \frac{1}{3}\delta_{x_2} + \frac{1}{3}\delta_{x_3}$. We have to minimize the quantity

$$\|\boldsymbol{\eta} * \mu\|_{L^2(\mathbb{R},\mu;\mathbb{R})} = \frac{1}{27}\sum_{j=1}^{3}\left|\sum_{i=1}^{3}\boldsymbol{\eta}(x_j-x_i)\right|^2$$

among all measurable antisymmetric selections $\boldsymbol{\eta}$ in $\partial\widehat{W}$.

If we let $x_1 = 1$, $x_2 = 0$, $x_3 = 3/4$, the only points where it is needed to select are ± 1, corresponding to $\pm(x_1 - x_2)$, hence expanding the sum above (using the antisymmetry) it is clear that we reduce to find the minimizer of

$$\min\{\boldsymbol{\eta}(x_1 - x_2)^2 + \boldsymbol{\eta}(x_1 - x_2)(\boldsymbol{\eta}(x_1 - x_3) - \boldsymbol{\eta}(x_2 - x_3)) : \boldsymbol{\eta}(x_1 - x_2) \in [-1, 1]\}.$$

Here $x_1-x_3 = 1/4$ (so that $\boldsymbol{\eta}(x_1-x_3) = -1/4$) and $x_2-x_3 = -3/4$ (so that $\boldsymbol{\eta}(x_2-x_3) = 3/4$). Then, letting $y = \boldsymbol{\eta}(x_1 - x_2)$, we are left with the problem $\min\{y^2 - y : y \in [-1, 1]\}$, whose solution is $y = 1/2$. This is different from the element of minimal norm in $\partial W(x_2 - x_1)$, which of course is 0.

We also point out that in this non convex case the choice of the selection is not independent from the measure μ. Indeed, if we change the value of x_3 to be, for instance, $-1/4$, we have $\boldsymbol{\eta}(x_1 - x_3) = 5/4$ and $\boldsymbol{\eta}(x_2 - x_3) = -1/4$, then we have to solve $\min\{y^2 + 3y/2 : y \in [-1, 1]\}$, and the solution is $y = \boldsymbol{\eta}(1) = -\boldsymbol{\eta}(-1) = -3/4$.

Example 4.7. The result of Theorem 4.5 fails if we omit the radial hypothesis on W. As a counterexample we provide a convex function $\mathbf{W}$ satisfying all the assumptions (1.2),(1.3), (1.4), (1.8) and a measure $\mu \in \mathscr{P}_2(\mathbb{R}^2)$ such that

$$\partial^o \mathcal{W}(\mu) \neq \partial^o W * \mu.$$

Let $\theta > 2$ and $W : \mathbb{R}^2 \to \mathbb{R}$ be the function defined, for $x = (u, v) \in \mathbb{R}^2$, by

$$W(u,v) = \begin{cases} \max\{|u|, |v|\} & \text{for} \quad \max\{|u|, |v|\} \leq 1, \\ \theta \max\{|u|, |v|\} + 1 - \theta, & \text{for} \quad \max\{|u|, |v|\} > 1. \end{cases}$$

The graph of W is a reversed pyramid with varying slopes, with vertex in the origin. W is symmetric and its level sets are squares centered in the origin. Then, we have

$$\nabla W(u,v) = \begin{cases} (1,0) & \text{for} \quad 0 < u < 1, \ -u < v < u, \\ (\theta,0) & \text{for} \quad u > 1, \ -u < v < u, \end{cases}$$

and ∇W is given by analogous expressions in the other three regions of the plane.
Let x_1, x_2, x_3 be points in $\mathbb{R}^2$ such that

$$x_2 - x_1 = (1, 1), \quad x_3 - x_2 = (-1/2 - \varepsilon, 1/2), \quad x_3 - x_1 = (1/2 - \varepsilon, 3/2).$$

Among these points, for small enough $\varepsilon > 0$, ∂W is not a singleton only at $x_2 - x_1$, and in particular it is the convex set K of $\mathbb{R}^2$ defined as

$$K = \{(u, v) \in \mathbb{R}^2 : \ u \geq 0, \ v \geq 0, \ 1 - u \leq v \leq \theta - u\}.$$

We let $\boldsymbol{\eta}_{ij}$, $i, j = 1, 2, 3$ denote the generic element of ∂W at $x_i - x_j$. We also let $\mu = \frac{1}{3}\delta_{x_1} + \frac{1}{3}\delta_{x_2} + \frac{1}{3}\delta_{x_3}$. In this particular case

$$\partial^o \mathcal{W}(\mu) = \operatorname{argmin} \left\{ \frac{1}{27} \sum_{j=1}^{3} \left| \sum_{i=1}^{3} \boldsymbol{\eta}_{ji} \right|^2 : \boldsymbol{\eta} \in \partial W, \boldsymbol{\eta}_{ji} = -\boldsymbol{\eta}_{ij} \right\}.$$

Since $\boldsymbol{\eta}_{13} = \nabla W(x_1 - x_3), \boldsymbol{\eta}_{23} = \nabla W(x_2 - x_3)$, we are left to minimize with respect to the unique variable $\boldsymbol{\eta}_{21}$, that is, the minimization problem above reduces to

$$\min_{\boldsymbol{\eta}_{21} \in K} \left(\frac{1}{27} |-\boldsymbol{\eta}_{21} + \boldsymbol{\eta}_{13}|^2 + \frac{1}{27} |\boldsymbol{\eta}_{23} + \boldsymbol{\eta}_{21}|^2 + \frac{1}{27} |\boldsymbol{\eta}_{32} + \boldsymbol{\eta}_{31}|^2 \right).$$

We have $\boldsymbol{\eta}_{13} = (0, -\theta)$ and $\boldsymbol{\eta}_{23} = (1, 0)$, and hence it is immediate to check the solution is the minimizer of

$$|\boldsymbol{\eta}_{21}|^2 + \langle \boldsymbol{\eta}_{21}, (1, \theta) \rangle. \tag{4.10}$$

If $\boldsymbol{\eta} = \partial^o W$ we have that $\boldsymbol{\eta}_{21}$ is the element of minimal norm in K, that is $(1/2, 1/2)$, and in this case the quantity above takes the value $1 + \theta/2$. But $\theta > 2$, so that the minimum value in K of the quadratic expression (4.10) is 2, attained in a different point, that is $\boldsymbol{\eta}_{21} = (1, 0)$.

Remark 4.8 (**Particle system**). As in [CDFLS], the well-posedness result in Theorem 2.4 for measure solutions allows to put in the same framework particle and continuum solutions. Assume that we are given N pointwise particles, each carrying a mass m_i, with $\sum_i m_i = 1$. Let $x_i(t)$ be the position in $\mathbb{R}^d$ of the i-th particle at time t. Let $\boldsymbol{\eta}^1 : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ and $\boldsymbol{\eta}^2 : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ be selections in the subdifferentials $\partial_1 \mathbf{W}$ and $\partial_2 \mathbf{W}$ respectively. We consider the system

$$\frac{dx_i}{dt} = \frac{1}{2} \sum_{j=1}^N m_j (\boldsymbol{\eta}^1(x_j, x_i) + \boldsymbol{\eta}^2(x_i, x_j)), \qquad i = 1, \ldots, N.$$

If the $t \mapsto x_i(t)$ are absolutely continuous curves, the empirical measure $\mu_t = \sum_{i=1}^N m_i \delta_{x_i(t)}$ solves the following PDE in the sense of distribution

$$\partial_t \mu_t - \frac{1}{2} \operatorname{div} \left(\left(\int_{\mathbb{R}^d} \boldsymbol{\eta}^1(\cdot, y) + \boldsymbol{\eta}^2(y, \cdot) \, d\mu_t(y) \right) \mu_t \right) = 0.$$

Now, if we choose $\boldsymbol{\eta}^1, \boldsymbol{\eta}^2$ as depending on t and realizing, for any t, the minimal selection in $\partial \mathcal{W}(\mu_t)$ (according to Theorem 3.3), we have correspondence between equation (2.7) and the ODE system above. If we are in the framework of assumption (1.8), the velocity vector field of the continuity equation is written as a convolution and the corresponding ODE system takes the form (1.14), where, for any t, $\boldsymbol{\eta}_t$ is the suitable measurable anti-symmetric selection in ∂W. If we are in the hypotheses of Theorem 4.5, then $\boldsymbol{\eta}_t = \partial^o W$ for any t (in particular, as previously observed, the selection does not depend on t). In this case, (1.14) reads

$$\frac{dx_i}{dt} = \sum_{j=1}^N m_j \partial^o w(|x_j - x_i|) \frac{x_j - x_i}{|x_j - x_i|}, \qquad i = 1, \ldots, N,$$

as usual w denoting the radial profile of W as assumption (1.9) holds.

We would like to remark that in the latter, radial and convex case, the finite-time collapse argument in [CDFLS] carries over. Indeed, it is enough to substitute ∇W therein with the new object $\partial^o W$. Let x_0 be the fixed center of mass (since $\partial^o W$ is anti-symmetric, the center of mass $x_0 := \sum_{j=1}^N m_j x_j$ does not move during the evolution). From such relation, following [CDFLS, Proposition 4.2], one obtains $\frac{d}{dt} R^2(t) \le -D R(t)$, where $R(t) := \max_{i \in \{1, \ldots, N\}} \{|x_i(t) - x_0|\}$ and $D := \inf_{r \in (0, +\infty)} \partial^o w(r)$. The solution then reaches the asymptotic state $\mu_\infty = \delta_{x_0}$ in finite time if $D > 0$. It is also possible to generalize the result to the case $D = 0$ under suitable assumptions on the behavior of w at the origin such that a Gronwall estimate can be achieved. For instance if we ask $\frac{\partial^o w(x)}{x}$ to be decreasing on some interval $(0, \varepsilon)$ we are done. For all the details and a more general discussion on this issue we refer to [CDFLS]. We only remark that if we have finite time collapse in the center of mass for a system of particles, by the general stability properties of gradient flow solutions we can also deduce the finite time collapse for general compactly supported initial data.

5. Examples of stationary states in the non convex case

In this section, we are dealing with stationary states to (1.5), that is, solutions μ to (1.5) which do not depend on time. In view of Theorem 2.4, they are characterized by $v = -\partial^o \mathcal{W}(\mu) = 0$ μ-a.e.

We have already seen in the case of particle collapse (see Remark 4.8) that, for convex nonnegative potentials, a single Dirac mass is the stationary state characterizing the long time asymptotics. It is easily seen that any Dirac mass minimizes the functional $\mathcal{W}$. So, in this case, by the translation invariance property, the only asymptotic solution is in fact the Dirac Delta at the center of mass of the initial datum. Dirac Deltas with all the mass concentrated at the center of mass are the only stationary states for purely attractive or purely repulsive potentials.

In the repulsive-attractive case, one can expect a much wider variety of stationary solutions. In [FR1, FR2] the authors show that the set of stationary states for short-range repulsive long-range attractive potentials even in 1D can be very large and complicated. Moreover, they give examples in which the stationary states are composed of a finite number of Dirac Deltas at points and others in which one has integrable compactly supported stationary solutions. They also show that having integrable or not stationary states depends on how strong the repulsion is at the origin. Numerical computations indicate that this is also the case in more dimensions [KSUB], and the dimensionality of the support dependence on the repulsion strength has been recently proved in any dimension in [BCLR2]. They show that the set of stable stationary states can be large and with complicated supports arising from instability modes of the uniform distribution on a sphere. Finally, [BCLR] contains a stability analysis for radial perturbations of the uniform distribution on a suitable sphere for general radial repulsive-attractive potentials. We will show a related example below.

In this section we are going to present some examples of stationary solutions based on the available characterization of the velocity vector field of the continuity equation. In particular, we will consider the (-1)-convex potentials in one dimension given by

$$\widetilde{W}(x) = \frac{1}{2}\,|x^2 - 1|^2, \qquad \widehat{W}(x) = \frac{1}{2}\,|x^2 - 1|.$$

Both cases correspond to attractive-repulsive potentials with the same behavior in the repulsive part at the origin. However, the change from repulsive to attractive in one case is smooth and in the other, it is only Lipschitz. In fact, in the two cases there many analogies, but also some different behaviors, as we are going to show with the next propositions. Let us remark that the potential $\widetilde{W}(x)$ has behavior larger than quadratic at infinity, and thus the theory developed in previous sections does not directly apply. In any case, we can reduce the growth of the potential outside a large ball not changing the discussion below about stationary states as soon as they have compact support. First of all, we search for stationary states made by a finite number of particles.

Proposition 5.1. *There exist two-particles stationary states for $\widetilde{W}$ and $\widehat{W}$.*

Proof. Let $\mu = m_1\delta_{x_1} + m_2\delta_{x_2}$, with $m_1 + m_2 = 1$. For the potential $\widetilde{W}$, we have to impose the condition

$$\frac{dx_i}{dt} = m_j \nabla\widetilde{W}(x_j - x_i) = 0, \tag{5.1}$$

for $i, j = 1, 2, i \neq j$. But $\nabla\widetilde{W}(x) = 2x(x^2 - 1)$. Hence it is enough to choose x_1, x_2 such that $|x_1 - x_2| = 1$, independently of the weights. Notice that we have infinitely many stationary states made by two Dirac Deltas. In the case of $\widehat{W}$, the subdifferential is

$$\partial\widehat{W}(x) = \begin{cases} x & \text{if } |x| > 1, \\ -x & \text{if } |x| < 1, \\ [-1, 1] & \text{if } |x| = 1. \end{cases} \tag{5.2}$$

In order to find a stationary state, we have to solve (5.1) with $\boldsymbol{\eta}$ in place of $\nabla\widetilde{W}$, where $\boldsymbol{\eta}$ is the suitable anti-symmetric selection in $\partial\widehat{W}$ realizing the minimal norm in $\partial\mathcal{W}(\mu)$ (see Corollary 4.2). Such selection is then found minimizing the quantity

$$\|\boldsymbol{\eta} * \mu\|^2_{L^2(\mathbb{R},\mu;\mathbb{R})} = \sum_{j=1}^{2} m_j \left| \sum_{i=1}^{2} m_i \boldsymbol{\eta}(x_j - x_1) \right|^2 = m_1 m_2 (m_1 + m_2)(\boldsymbol{\eta}(x_1 - x_2))^2$$

among the admissible selections $\boldsymbol{\eta}$ (we are using the anti-symmetry). Let again $|x_2 - x_1| = 1$, so that it is clear that the minimum above is zero, attained for $\boldsymbol{\eta}(1) = -\boldsymbol{\eta}(-1) = 0$. And this way, the two equations (5.1) are still satisfied. □

Proposition 5.2. *For both potentials $\widetilde{W}$ and $\widehat{W}$, there are no absolutely continuous stationary states in one dimension.*

Proof. Let us consider $\widehat{W}$. The argument is based on the fact that, if μ is absolutely continuous, it does not charge the points of non-differentiability of $\widehat{W}$. For a measure μ to be stationary, we have to verify that the corresponding velocity vector field vanishes. That is $\boldsymbol{\eta} * \mu = 0$, where $\boldsymbol{\eta}$ is the usual optimal selection in $\partial\widehat{W}$, as in Corollary 4.2. Suppose that μ is a stationary state and that $\mu = \rho\mathcal{L}^1$, for some $\rho \in L^1(\mathbb{R})$, then

$$\int_{\{|x-y|>1\}} (x - y)\rho(y)\, dy - \int_{\{|x-y|<1\}} (x - y)\rho(y)\, dy = 0.$$

By the translation invariance property, we can fix without loss of generality the center of mass, so we let $\int_{\mathbb{R}} y\rho(y)\, dy = 0$. We deduce

$$2x \int_{\{|x-y|>1\}} \rho(y)\, dy - x = 2 \int_{\{|x-y|>1\}} y\rho(y)\, dy,$$

hence

$$2x \left(\int_{-\infty}^{x-1} \rho(y)\, dy + \int_{x+1}^{+\infty} \rho(y)\, dy \right) - x = 2 \left(\int_{-\infty}^{x-1} y\rho(y)\, dy + \int_{x+1}^{+\infty} y\rho(y)\, dy \right).$$

Let us denote the term in the parenthesis in the left side by $\Theta(x)$ and let us take the derivative with respect to x. We have

$$2\Theta(x) + 2x(\rho(x-1) - \rho(x+1)) - 1 = 2((x-1)\rho(x-1) - (x+1)\rho(x+1)),$$

which yields $2\Theta(x) - 1 = -2(\rho(x-1) - \rho(x+1))$, that is $\Theta'(x) = -\Theta(x) + \frac{1}{2}$. We find

$$\Theta(x) = ke^{-x} + \frac{1}{2}, \quad k \in \mathbb{R},$$

then $\rho(x-1) - \rho(x+1) = -ke^{-x}$. But the integral of ρ is 1, so $k = 0$ and we are left with $\rho(x-1) = \rho(x+1)$. This is a contradiction, since ρ can not be periodic in this case. The proof for $\widetilde{W}$ is analogous, we omit the details. □

The following are more examples of stationary states

Example 5.3. There are stationary states for potential $\widehat{W}$ of the form

$$\mu = m_1\delta_{x_1} + m_2\delta_{x_2} + m_3\delta x_3,$$

with $m_1 + m_2 + m_3 = 1$. Indeed, we have to verify that

$$\begin{cases} \dfrac{dx_1}{dt} = m_2\boldsymbol{\eta}(x_2 - x_1) + m_3\boldsymbol{\eta}(x_3 - x_1) = 0 \\ \dfrac{dx_2}{dt} = m_1\boldsymbol{\eta}(x_1 - x_2) + m_3\boldsymbol{\eta}(x_3 - x_2) = 0 \\ \dfrac{dx_3}{dt} = m_1\boldsymbol{\eta}(x_1 - x_3) + m_2\boldsymbol{\eta}(x_2 - x_3) = 0, \end{cases} \tag{5.3}$$

where, as usual, $\boldsymbol{\eta}$ represents the anti-symmetric selection in the subdifferential (5.2) given by Corollary 4.2.

For instance, let us search for a solution in the following range

$$x_2 - x_1 = 1, \quad x_3 - x_1 > 1, \quad 0 < x_3 - x_2 < 1. \tag{5.4}$$

We begin searching for the right selection. As usual, we use the notation $\boldsymbol{\eta}_{ij} := \boldsymbol{\eta}(x_i - x_j)$. This is exactly the case treated in Example 4.6, and we need to select only at $\pm(x_2 - x_1)$. Taking the anti-symmetry into account, and recalling that the subdifferential of $\widehat{W}$ is (5.2) and that the relations (5.4) hold, there is

$$\begin{aligned} \|\boldsymbol{\eta} * \mu\|^2_{L^2(\mathbb{R},\mu;\mathbb{R})} &= \sum_{j=1}^{3} m_j \left| \sum_{i=1}^{3} m_i\boldsymbol{\eta}(x_j - x_1) \right|^2 \\ &= m_1(m_2^2\boldsymbol{\eta}_{12}^2 + 2m_2m_3(x_1 - x_3)\boldsymbol{\eta}_{12}) + m_2(m_1^2\boldsymbol{\eta}_{21}^2 + 2m_1m_3(x_3 - x_2)\boldsymbol{\eta}_{21}) + R_1 \\ &= m_1m_2(m_1 + m_2)\boldsymbol{\eta}_{12}^2 + 2m_1m_2m_3(x_1 + x_2 - 2x_3)\boldsymbol{\eta}_{12} + R_2, \end{aligned}$$

where the remainders R_1, R_2 do not depend on the value of $\boldsymbol{\eta}$ at ± 1, hence we only have to minimize, as in Example 4.6, with respect to $\boldsymbol{\eta}_{12}$ on the interval $[-1, 1]$. We have a quadratic function, so that if the vertex of the parabola is on the right of the interval $[-1, 1]$, then the minimizer is found for $\boldsymbol{\eta}_{12} = 1$, hence we get $\boldsymbol{\eta}(-1) = 1$ and then, by anti-symmetry, $\boldsymbol{\eta}(1) = -1$. Computing the vertex position, this condition is

$$\frac{m_3(x_3 - x_2) + m_3(x_3 - x_1)}{m_1 + m_2} \geq 1. \tag{5.5}$$

Hence, if such condition holds, making use of (5.2) the first two equations in (5.3) reduce to

$$\begin{cases} -m_2 + m_3(x_3 - x_1) = 0, \\ m_1 - m_3(x_3 - x_2) = 0. \end{cases} \tag{5.6}$$

If we define, for $0 < \alpha < \frac{1}{4}$,

$$m_1 = \frac{1}{4} - \alpha, \quad m_2 = \frac{1}{2}, \quad m_3 = \frac{1}{4} + \alpha, \quad x_1 = -1, \quad x_2 = 0, \quad x_3 = \frac{1-4\alpha}{1+4\alpha},$$

straightforward computations show that (5.3)-(5.4)-(5.5)-(5.6) are satisfied and that

$$\mu = \left(\frac{1}{4} - \alpha\right)\delta_{-1} + \frac{1}{2}\delta_0 + \left(\frac{1}{4} + \alpha\right)\delta_{\frac{1-4\alpha}{1+4\alpha}}$$

is a stationary state.

Remark 5.4. In the case of $\widetilde{W}$, it is very easy to construct an analogous example, solving system (5.3), where this time the actual gradient of $\widetilde{W}$ appears. In both cases, it seems clear that the procedure can be repeated for finding infinitely many stationary states with N Dirac masses for any $N > 3$.

We conclude with an example in two space dimensions. The reference functional is simply the radial version of $\widehat{W}$, still denoted by $\widehat{W}$, that is

$$\widehat{W}(x) = \frac{1}{2}||x|^2 - 1|, \quad x \in \mathbb{R}^2.$$

The potential is still (-1)-convex, and in this case

$$\partial\widehat{W}(x) = \begin{cases} x & \text{if } |x| > 1 \\ -x & \text{if } |x| \le 1 \\ [-1,1]x & \text{if } |x| = 1. \end{cases}$$

Example 5.5. Let σ_R denote the uniform measure on the circumference $\partial B_R(0)$, of radius R, centered in the origin. There exists $R > 0$ such that the measure σ_R is a stationary state for functional $\widehat{W}$.

Indeed, we can show that for any $x \in \partial B_R(0)$ there holds $\boldsymbol{\eta} * \sigma_R = 0$ for a suitable choice of the radius R, $\boldsymbol{\eta}$ being the optimal selection of Corollary 4.2. Explicitly, the convolution is

$$\int_{\{|x-y|>1\}} (x-y)\, d\sigma_R(y) - \int_{\{|x-y|\le 1\}} (x-y)\, d\sigma_R(y),$$

and the set of points $\{y : |x-y| = 1\}$, where one should select, is negligible. Fix x on the circle. We let $(\mathbf{e}_1, \mathbf{e}_2)$ be an orthogonal base in $\mathbb{R}^2$, where $\mathbf{e}_1$ is the direction of x, so that $x = R\mathbf{e}_1$. Hence we have to solve

$$R\mathbf{e}_1\sigma_R(\{|x-y|> 1\}) - R\mathbf{e}_1\sigma_R(\{|x-y|\le 1\}) - \int_{\{|x-y|>1\}} y\, d\sigma_R(y) + \int_{\{|x-y|\le 1\}} y\, d\sigma_R(y) = 0.$$

We write the integrals in polar coordinates with respect to the origin and the vector $\mathbf{e}_1$. In this system, we let $\alpha = \alpha(R)$ denote the (positive) angle corresponding to the intersection point between $\partial B_R(0)$ and the circle of radius 1 centered in x (see Figure 1 below). In particular

$$\sin\alpha(R) = \frac{\sqrt{R^2 - \frac{1}{4}}}{R^2}. \tag{5.7}$$

Since $y = R\cos\theta\mathbf{e}_1 + R\sin\theta\mathbf{e}_2$, it is immediately seen, by oddness of the sine function, that the equation in the direction of $\mathbf{e}_2$ is identically satisfied. In the direction of $\mathbf{e}_1$ we find

$$2R\int_\alpha^\pi R\, d\theta - 2R\int_0^\alpha R\, d\theta - 2R\int_\alpha^\pi \cos\theta R\, d\theta + 2R\int_0^\alpha \cos\theta R\, d\theta = 0.$$

Hence we have to solve

$$f(R) := \pi - 2\alpha(R) + 2\sin\alpha(R) = 0.$$

If $R < \frac{1}{2}$, the computation does not make sense, but indeed we can not have a stationary state for $R < \frac{1}{2}$, since in this case any point in $\partial B_R(0)$ has distance lower than 1 from x, so that for each of them the effect on x is a repulsion, and x tend to move far from the origin. Taking (5.7) into account, if $R = \frac{1}{2}$, we have $\alpha(R) = \pi$, hence the value of f at $\frac{1}{2}$ is $-\pi$. As R increases from $\frac{1}{2}$ to $+\infty$, the angle $\alpha(R)$ decreases from π to 0. Notice that the function

$$R \mapsto \frac{\sqrt{R^2 - \frac{1}{4}}}{R^2}$$

is increasing from $\frac{1}{2}$ to $\frac{\sqrt{2}}{2}$, where it has its maximum, and is decreasing in $(\frac{\sqrt{2}}{2}, +\infty)$. On the other hand, $\sin(x) - x$ is a decreasing function. Since $f(\sqrt{2}/2) = 2$, we conclude that f has only one zero, found in the interval $(1/2, \sqrt{2}/2)$. If R_0 is the zero, for $R = R_0$ the measure σ_R is stationary.

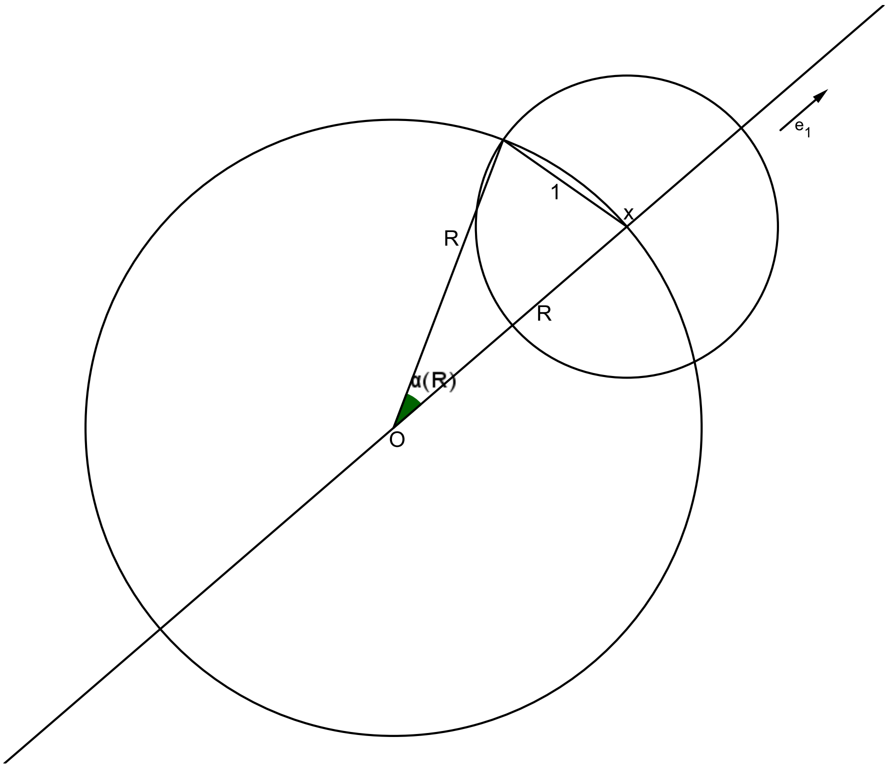

Figure 1. The construction of Example 5.5

6. Appendix: vector and plan subdifferential

Here we give a more complete overview about the Wasserstein subdifferential. In [AGS, §10.3], the theory is developed for functionals $\Phi : \mathscr{P}_2(\mathbb{R}^d) \to (-\infty, +\infty]$ such that

$$\Phi : \mathscr{P}_2(\mathbb{R}^d) \to (-\infty, +\infty] \text{ is proper and lower semicontinuous in } \mathscr{P}_2(\mathbb{R}^d) \tag{6.1}$$

and

$$(6.2)\quad \Phi(\cdot) + \frac{1}{2\tau}\, d_W^2(\cdot,\mu) \text{ admits minimizers for any small enough } \tau > 0 \text{ and } \mu \in \mathscr{P}_2(\mathbb{R}^d).$$

Indeed, after Proposition 3.6 we know that $\mathcal{W}$ satisfies these hypotheses. Hence, we are in the framework of [AGS, §10.3]. We will show how the results therein work for the case of $\mathcal{W}$.

First of all, we provide a characterization of the the Wasserstein subdifferential introduced in Definition 2.3.

Proposition 6.1. *$\boldsymbol{\xi} \in L^2(\mathbb{R}^d,\mu;\mathbb{R}^d)$ belongs to the Wasserstein subdifferential of $\mathcal{W}$ at $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ if and only if*

$$\mathcal{W}(\nu) - \mathcal{W}(\mu) \geq \int_{\mathbb{R}^d} \langle \boldsymbol{\xi}(x), y - x\rangle \, d\gamma(x,y) + o(\mathcal{C}(\mu,\nu;\gamma)) \tag{6.3}$$

as $\nu \to \mu$ in $\mathscr{P}_2(\mathbb{R}^d)$, for suitable optimal transport plans $\gamma \in \Gamma_o(\mu,\nu)$. Moreover, $\boldsymbol{\xi}$ is a strong subdifferential if and only if (6.3) *holds whenever $\nu \to \mu$ in $\mathscr{P}_2(\mathbb{R}^d)$ and $\Gamma(\mu,\nu) \ni \gamma \to (\mathbf{i},\mathbf{i})_\#\mu$ in $\mathscr{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$.*

Proof. It is clear that (6.3) holds if $\boldsymbol{\xi}$ is in the Wasserstein subdifferential of $\mathcal{W}$ at μ (and the same for strong versions). On the other hand, suppose that (6.3) holds. Let $\gamma \in \Gamma(\mu,\nu)$ and define the interpolating curve $\theta^\gamma(t) = ((1-t)\pi^1 + t\pi^2)_\#\gamma$ between μ and ν, so that $\theta^\gamma(0) = \mu$ and $\theta^\gamma(1) = \nu$. We take advantage of a property of Wasserstein constant speed geodesics, shown in [AGS, Lemma 7.2.1]: there exists γ^* in $\Gamma_o(\mu,\nu)$ such that $\Gamma_o(\mu,\theta^{\gamma^*}(t))$ contains a unique element for any $t \in [0,1)$, given by $\gamma_t := (\pi^1,(1-t)\pi^1 + t\pi^2)_\#\gamma^*$. Then, (6.3) can be applied in correspondence of γ_t and with $\theta^\gamma(t)$ in place of ν, and together with (2.4), it gives, for $t \to 0$,

$$\begin{aligned}\mathcal{W}(\nu) - \mathcal{W}(\mu) &\geq \frac{\mathcal{W}(\theta^{\gamma^*}(t)) - \mathcal{W}(\mu)}{t} + \frac{\lambda}{2}(1-t)\mathcal{C}^2(\mu,\nu;\gamma^*) \\ &\geq \int_{\mathbb{R}^d\times\mathbb{R}^d} \langle \boldsymbol{\xi}(x), y-x\rangle \, d\gamma^*(x,y) + \frac{\lambda}{2}(1-t)\mathcal{C}^2(\mu,\nu;\gamma^*) + \frac{1}{t}\, o(\mathcal{C}(\mu,\theta^{\gamma^*}(t);\gamma_t)).\end{aligned}$$

Passing to the limit as $t \to 0$, since $\mathcal{C}(\mu,\theta^{\gamma^*}(t);\gamma_t) = t\mathcal{C}(\mu,\nu;\gamma^*)$, we get (2.5) for the plan $\gamma^* \in \Gamma_o(\mu,\nu)$. By a similar argument one can recover the result for strong subdifferentials: indeed, one considers the generic plan $\gamma \in \Gamma(\mu,\nu)$ and makes use of the convexity of $\mathcal{W}$ along any interpolating curve $t \mapsto \theta^\gamma(t)$ (see Proposition 2.2). □

On the other hand, the general definition of subdifferential, given in [AGS, §10.3], is more technical. According to that notion, the subdifferential is in fact a plan $\beta \in \mathscr{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$, as in the following

Definition 6.2 (**Plan subdifferential**). *Let $\mu \in \mathscr{P}_2(\mathbb{R}^d)$. We say that $\beta \in \mathscr{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$ belongs to the extended subdifferential of $\mathcal{W}$ at μ if $\pi^1_\#\beta = \mu$ and, for any $\nu \in \mathscr{P}_2(\mathbb{R}^d)$, there holds*

$$\mathcal{W}(\nu) - \mathcal{W}(\mu) \geq \int_{\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d} \langle y, z-x\rangle \, d\boldsymbol{\mu}(x,y,z) + \frac{\lambda}{2}\int_{\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d} |z-x|^2 \, d\boldsymbol{\mu}(x,y,z) \tag{6.4}$$

for some $\boldsymbol{\mu} \in \boldsymbol{\Gamma}_o(\beta,\nu)$ (we say that β is a strong subdifferential if moreover the above inequality holds for any plan $\boldsymbol{\mu} \in \boldsymbol{\Gamma}(\beta,\nu)$). We write $\beta \in \boldsymbol{\partial}\mathcal{W}(\mu)$ (resp. $\beta \in \boldsymbol{\partial}_S\mathcal{W}(\mu)$). Here the elements *of* $\boldsymbol{\Gamma}(\beta,\mu)$ *are three-plans, that is, measures in* $\mathscr{P}(\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d)$*, such that* $\pi^{1,2}_{\#}\boldsymbol{\mu}=\beta$ *and* $\pi^3_{\#}\boldsymbol{\mu}=\nu$*. The definition of optimal plan in this case is*

$$\boldsymbol{\Gamma}_o(\beta,\nu):=\{\gamma\in\boldsymbol{\Gamma}(\beta,\nu):\pi^{1,3}\gamma\in\Gamma_o(\mu,\nu)\}.$$

Remark 6.3. Since $\mathcal{W}$ is convex along any linearly interpolating curve, as noticed in Proposition 2.2, from [AGS, Theorem 10.3.6] we learn that we can equivalently define the extended subdifferential of $\mathcal{W}$ by asking inequality (6.4) for any $\boldsymbol{\mu}\in\boldsymbol{\Gamma}_o(\beta,\nu)$.

Remark 6.4. We observe that if β is concentrated on the graph of a vector field $\boldsymbol{\xi}$, we have $\beta=(\mathbf{i},\boldsymbol{\xi})_{\#}\mu$ and in particular $y=\boldsymbol{\xi}(x)$ for $\boldsymbol{\mu}$-a.e. $(x,y,z)\in\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d$. In this case the definition reduces to (2.5).

We recall that, also for extended subdifferentials, there holds

$$|\partial\mathcal{W}|(\mu)=\min\left\{\left(\int_{\mathbb{R}^d}|y|^2\,d\pi^2_{\#}\beta(y)\right)^{1/2}:\beta\in\boldsymbol{\partial}\mathcal{W}(\mu)\right\}. \tag{6.5}$$

Moreover, the corresponding minimizer is unique. See [AGS, Theorem 10.3.11]. We denote it by $\boldsymbol{\partial}^o\mathcal{W}(\mu)$.

We have the following

Lemma 6.5. *Let* $\mu\in\mathscr{P}_2(\mathbb{R}^d)$*. The following assertions hold:*

$$\beta\in\boldsymbol{\partial}_S\mathcal{W}(\mu)\Rightarrow\bar\beta\in\partial_S\mathcal{W}(\mu)\qquad\text{and}\qquad\beta\in\boldsymbol{\partial}\mathcal{W}(\mu)\Rightarrow\bar\beta\in\partial\mathcal{W}(\mu), \tag{6.6}$$

where $\bar\beta$ *is the barycenter of the plan* β *(see* Definition 3.10*).*

Proof. We begin with the proof for strong subdifferentials. Let $\beta\in\boldsymbol{\partial}_S\mathcal{W}(\mu)$ and we write $\beta=\int_{\mathbb{R}^d}\beta_x\,d\mu(x)$. For any $\nu\in\mathscr{P}_2(\mathbb{R}^d)$ and $\gamma\in\Gamma(\mu,\nu)$ there holds

$$\begin{aligned}\int_{\mathbb{R}^d\times\mathbb{R}^d}\langle\bar\beta(x),z-x\rangle\,d\gamma(x,z)&=\int_{\mathbb{R}^d\times\mathbb{R}^d}\left\langle\int_{\mathbb{R}^d}y\,d\beta_x(y),z-x\right\rangle\,d\gamma(x,z)\\&=\int_{\mathbb{R}^d\times\mathbb{R}^d}\int_{\mathbb{R}^d}\langle y,z-x\rangle\,d\beta_x(y)\,d\gamma(x,z).\end{aligned}$$

Moreover, taking into account that $\int_{\mathbb{R}^d}d\beta_x(y)=1$ for μ-a.e. $x\in\mathbb{R}^d$, we have that

$$\int_{\mathbb{R}^d\times\mathbb{R}^d}|z-x|^2\,d\gamma(x,z)=\int_{\mathbb{R}^d\times\mathbb{R}^d}\int_{\mathbb{R}^d}|z-x|^2\,d\beta_x(y)\,d\gamma(x,z).$$

Let us define the three-plan $\boldsymbol{\mu}\in\mathscr{P}(\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d)$ by

$$\int_{\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d}\phi(x,y,z)d\boldsymbol{\mu}(x,y,z):=\int_{\mathbb{R}^d\times\mathbb{R}^d}\int_{\mathbb{R}^d}\phi(x,y,z)d\beta_x(y)\,d\gamma(x,z)$$

for all continuous and bounded functions $\phi:\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}$. Then, $\boldsymbol{\mu}$ belongs to $\boldsymbol{\Gamma}(\beta,\nu)$. Making use of (6.4) for this particular choice of $\boldsymbol{\mu}\in\boldsymbol{\Gamma}(\beta,\nu)$, we see that $\bar\beta$ satisfies inequality (2.5). Therefore, $\bar\beta\in\partial_S\mathcal{W}(\mu)$.

Recalling Remark 6.3, since $\gamma\in\Gamma_o(\mu,\nu)\Rightarrow\boldsymbol{\mu}\in\boldsymbol{\Gamma}_o(\beta,\nu)$, reasoning as done for strong subdifferentials the second implication in (6.6) follows. □

Corollary 6.6. *If* $\beta_o=\boldsymbol{\partial}^o\mathcal{W}(\mu)$*, then* $\beta_o=(\mathbf{i},\bar\beta_o)_{\#}\mu$ *and* $\bar\beta_o=\partial^o\mathcal{W}(\mu)$*.*

Proof. By (3.14) we have that

$$\int_{\mathbb{R}^d} |y|^2 \, d\pi^2_\# \beta(y) \geq \int_{\mathbb{R}^d} |\bar{\beta}(x)|^2 \, d\mu(x) \qquad \text{for any } \beta \in \boldsymbol{\partial}\mathcal{W}(\mu).$$

If β_o is the minimizer in (6.5), taking into account that $\bar{\beta}_o \in \partial\mathcal{W}(\mu)$, by uniqueness we have that $\beta_o = (\mathbf{i}, \bar{\beta}_o)_\# \mu$. By (2.6) and (6.5) we also obtain that $\bar{\beta}_o = \partial^o\mathcal{W}(\mu)$. □

Remark 6.7. Because of (6.6), under the same assumptions of Proposition 3.2 one sees that if $\beta \in \boldsymbol{\partial}_S\mathcal{W}(\mu)$, then its barycenter $\bar{\beta}$ is given by (3.3).

Remark 6.8. *We refer to* [BCDP] *for an example of a geodesically convex interaction functional $\mathcal{W}$ (associated to a non λ-convex potential $\mathbf{W}$) such that, for suitable choice of μ, $\boldsymbol{\partial}^o\mathcal{W}(\mu)$ is not concentrated on a graph. In such case $\partial\mathcal{W}(\mu)$ is empty.*

Remark 6.9. In the case of strong subdifferentials, the implication of Lemma 6.5 is true for any functional Φ satisfying the assumptions (6.1) and (6.2). It is shown in Lemma 10.3.4 and Remark 10.3.5 of [AGS] that, given $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, a minimizer μ_τ to $\Phi(\cdot) + \frac{1}{2\tau} d^2_W(\cdot, \mu)$ and a plan $\hat{\gamma}_\tau \in \Gamma_o(\mu_\tau, \mu)$, there holds

$$\gamma_\tau \in \boldsymbol{\partial}_S\Phi(\mu_\tau),$$

where γ_τ is the rescaled of $\hat{\gamma}_\tau$ (see Definition 3.12). Moreover, among these rescaled plans, there exists a plan whose barycenter belongs to $\partial_S\Phi(\mu)$. After Lemma 6.5, we may indeed infer that this holds true for the rescaled of any optimal plan in $\Gamma_o(\mu_\tau, \mu)$.

Eventually, we are ready for the proof of Proposition 3.13. We make use of the general convergence properties of rescaled plan subdifferentials shown in [AGS, Theorem 10.3.10], passing to barycenters by means of Lemma 6.5 and Corollary 6.6.

***Proof of* Proposition 3.13**. Let $\tau > 0$ be small enough. Once more, let $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, let μ_τ minimize $\mathcal{W}(\cdot) + \frac{1}{2\tau} d^2_W(\cdot, \mu)$ and let $\hat{\gamma}_\tau \in \Gamma_o(\mu_\tau, \mu)$. Moreover, let γ_τ be the rescaled of $\hat{\gamma}_\tau$ (as given by Definition 3.12). By [AGS, Lemma 10.3.4], $\gamma_\tau \in \boldsymbol{\partial}_S\mathcal{W}(\mu_\tau)$, and then by Lemma 6.5 we have $\bar{\gamma}_\tau \in \partial_S\mathcal{W}(\mu_\tau)$. By Theorem 3.1, $\boldsymbol{\partial}\mathcal{W}(\mu)$ is not empty, therefore we are in the hypotheses of [AGS, Theorem 10.3.10], which entails, taking into account also [AGS, Remark 10.3.14],

$$\lim_{\tau \to 0} \gamma_\tau = \boldsymbol{\partial}^o\mathcal{W}(\mu) \quad \text{in } \mathscr{P}_2(\mathbb{R}^d \times \mathbb{R}^d).$$

But Corollary 6.6 implies that $\boldsymbol{\partial}^o\mathcal{W}(\mu) = (\mathbf{i}, \partial^o\mathcal{W}(\mu))_\#\mu$. The convergence above then means that, as $\tau \to 0$,

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \phi(x,y) \, d(\gamma_\tau)_x(y) \, d\mu_\tau(x) \to \int_{\mathbb{R}^d \times \mathbb{R}^d} \phi(x,y) \, d(\mathbf{i}, \partial^o\mathcal{W}(\mu))_\#\mu(x,y) \tag{6.7}$$

for any continuous function $\phi : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ with at most quadratic growth at infinity, where $(\gamma_\tau)_x$ denotes the family of measures which disintegrates γ_τ with respect to μ_τ. Letting $\zeta \in C_0(\mathbb{R}^d; \mathbb{R}^d)$, and choosing $\phi(x,y) = \langle y, \zeta(x) \rangle$ in (6.7), we obtain the weak convergence in the sense of Definition 3.8 of $\bar{\gamma}_\tau$ to $\partial^o\mathcal{W}(\mu)$. On the other hand, using Jensen inequality as in (3.14) and (6.7) with $\phi(x,y) = |y|^2$ we obtain

$$\limsup_{\tau \to 0} \int_{\mathbb{R}^d} |\bar{\gamma}_\tau|^2 \, d\mu_\tau \leq \lim_{\tau \to 0} \int_{\mathbb{R}^d \times \mathbb{R}^d} |y|^2 \, d\gamma_\tau = \int_{\mathbb{R}^d \times \mathbb{R}^d} |y|^2 \, d(\mathbf{i}, \partial^o\mathcal{W}(\mu))_\#\mu = \int_{\mathbb{R}^d} |\partial^o\mathcal{W}(\mu)|^2 \, d\mu,$$

hence we also have the strong convergence in the sense of Definition 3.8. Notice that the second term in the last formula is $d^2_W(\mu_\tau, \mu)/\tau^2$, hence (3.16) holds. □

Acknowledgements. The authors would like to thank Giuseppe Savaré for several discussions about this work. JAC acknowledges support from the project MTM2011-27739-C04-02 DGI (Spain) and 2009-SGR-345 from AGAUR-Generalitat de Catalunya. JAC acknowledges support from the Royal Society through a Wolfson Research Merit Award. This work was supported by Engineering and Physical Sciences Research Council grant number EP/K008404/1. SL acknowledges support from Project nr. 25 of the 2007 Azioni Integrate Italia-Spagna. SL and EM have been partially supported by the INDAM-GNAMPA project 2011 "Measure solution of differential equations of drift-diffusion, interactions and of Cahn-Hilliard type". JAC and SL gratefully acknowledge the hospitality of the Centro de Ciencias Pedro Pascual de Benasque where this work was started. EM was supported by a postdoctoral scholarship of the Fondation Mathématique Jacques Hadamard, he acknowledges hospitality from Paris-sud University. EM also acknowledges the support from the project FP7-IDEAS-ERC-StG Grant #200497 (BioSMA).

References

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

[AGS] L. Ambrosio, N. Gigli, G. Savaré, *Gradient flows in metric spaces and in the spaces of probability measures*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel (2005).

[BCLR] D. Balagué, J.A. Carrillo, T. Laurent, G. Raoul, *Nonlocal interactions by repulsive-attractive potentials: radial ins/stability*, Physica D **260** (2013), 5–25.

[BCLR2] D. Balagué, J.A. Carrillo, T. Laurent, G. Raoul, *Dimensionality of Local Minimizers of the Interaction Energy*, Arch. Rat. Mech. Anal. **209** (2013), 1055–1088.

[BCP] D. Benedetto, E. Caglioti, M. Pulvirenti, *A kinetic equation for granular media*, RAIRO Model. Math. Anal. Numer. **31**, 5 (1997), 615–641. Erratum in M2AN Math. Model. Numer. Anal. **33**, 2 (1999), 439–441.

[BCCP] D. Benedetto, E. Caglioti, J.A. Carrillo, M. Pulvirenti *A non-Maxwellian steady distribution for one-dimensional granular media*, J. Statist. Phys. **91**, 5-6 (1998), 979–990.

[BCDP] G.A. Bonaschi, J.A. Carrillo, M. Di Francesco, M.A. Peletier *Equivalence of gradient flows and entropy solutions for singular nonlocal interaction equations in 1D*, arXiv:1310.4110 [math.AP]

[BCL] A. L. Bertozzi, J. A. Carrillo, T. Laurent, *Blow-up in multidimensional aggregation equations with mildly singular interaction kernels.* Nonlinearity **22** (2009), no. 3, 683–710.

[BL] A. L. Bertozzi, T. Laurent, *Finite-time blow-up of solutions of an aggregation equation in $\mathbb{R}^n$*, Comm. Math. Phys. **274** (2007), no. 3, 717–735.

[BLR] A. L. Bertozzi, T. Laurent, J. Rosado, *L^p theory for the multidimensional aggregation equation*, Comm. Pur. Appl. Math. **64** (2011), no. 1, 45–83.

[B] P. Billingsley, *Convergence of probability measures.* Second edition. Wiley Series in Probability and Statistics: Probability and Statistics. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York.

[CDFLS] J. A. Carrillo, M. Di Francesco, A. Figalli, T. Laurent, D. Slepcev, *Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations*, Duke Math. J. **156** (2011), 229–271.

[CDFLS2] J. A. Carrillo, M. Di Francesco, A. Figalli, T. Laurent, D. Slepcev, *Confinement in nonlocal interaction equations*, Nonlinear Analysis TMA **75** (2012), 550–558.

[CMV] J. A. Carrillo, R. J. McCann, C. Villani, *Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates*, Rev. Mat. Iberoamericana **19** (2003), no. 3, 971–1018.

[CMV2] J. A. Carrillo, R. J. McCann, C. Villani,*Contractions in the 2-Wasserstein length space and thermalization of granular media.* Arch. Ration. Mech. Anal. **179** (2006), no. 2, 217–263.

[CR] J. A. Carrillo, J. Rosado, *Uniqueness of Bounded Solutions to Aggregation Equations by Optimal Transport Methods,* Proceedings of the 5th European Congress of Mathematicians, 3–16, Eur. Math. Soc., Zurich, 2010.

[FR1] K. Fellner and G. Raoul, *Stable stationary states of non-local interaction equations*, Math. Models Methods Appl. Sci. **20** (2010), 2267–2291.

[FR2] K. Fellner and G. Raoul, *Stability of stationary states of non-local interaction equations*, Mathematical and Computer Modelling **53** (2011), 1436–1450.

[JKO] R. Jordan, D. Kinderlehrer and F. Otto: The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.* **29** (1998), 1–17.

[KSUB] T. Kolokonikov, H. Sun, D. Uminsky and A. Bertozzi, *Ring patterns and their stability in 2D kinematic particle aggregation models*, Physical Review E **84** (2011), 015203.

[LT] H. Li and G. Toscani, *Long–time asymptotics of kinetic models of granular flows*, Arch. Rat. Mech. Anal., 172 (2004), pp. 407–428.

[MY] S. McNamara, W. Young, *Kinetics of a one-dimensional granular medium*, Phys. Fluids A **5**, 1 (1993), 34–45.

[MEBS] A. Mogilner, L. Edelstein-Keshet, L. Bent, A. Spiros, *Mutual interactions, potentials, and individual distance in a social aggregation*, J. Math. Biol. **47** (2003), 353–389.

[O1] F. Otto, *Dynamics of Labyrinthine Pattern Formation in Magnetic Fluids: A Mean- Field Theory,* Arch.Rational Mech. Anal. **141** (1998), 63–103.

[O2] F. Otto: The geometry of dissipative evolution equations: the porous-medium equation. *Comm. PDE* **26** (2001), 101–174.

[T] G. Toscani, *One-dimensional kinetic models of granular flows.* M2AN Math. Model. Numer. Anal. **34** (2000), no. 6, 1277–1291.

[V] C. Villani, *Optimal transport, old and new.* Springer-Verlag (2008).

José A. Carrillo, Department of Mathematics, Imperial College London, South Kensington Campus, London SW7 2AZ, UK.

E-mail address: `carrillo@imperial.ac.uk`

Stefano Lisini, Dipartimento di Matematica "F. Casorati", Università degli Studi di Pavia, via Ferrata 1, 27100 Pavia, Italy.

E-mail address: `stefano.lisini@unipv.it`

Edoardo Mainini, Università degli Studi di Genova, Dipartimento di Ingegneria meccanica, energetica, gestionale e dei trasporti (DIME) - sezione MAT. P.le Kennedy 1, 16129 Genova, Italy

E-mail address: `edoardo.mainini@unipv.it`