GLOBAL-IN-TIME WEAK MEASURE SOLUTIONS, FINITE-TIME AGGREGATION AND CONFINEMENT FOR NONLOCAL INTERACTION EQUATIONS

J. A. CARRILLO[1], M. DIFRANCESCO[2], A. FIGALLI[3], T. LAURENT[4] AND D. SLEPČEV[5]

ABSTRACT. In this paper, we provide a well-posedness theory for weak measure solutions of the Cauchy problem for a family of nonlocal interaction equations. These equations are continuum models for interacting particle systems with attractive/repulsive pairwise interaction potentials. The main phenomenon of interest is that, even with smooth initial data, the solutions can concentrate mass in finite time. We develop an existence theory that enables one to go beyond the blow-up time in classical norms and allows for solutions to form atomic parts of the measure in finite time. The weak measure solutions are shown to be unique and exist globally in time. Moreover, in the case of sufficiently attractive potentials, we show the finite time total collapse of the solution onto a single point, for compactly supported initial measures. Finally, we give conditions on compensation between the attraction at large distances and local repulsion of the potentials to have global-in-time confined systems for compactly supported initial data. Our approach is based on the theory of gradient flows in the space of probability measures endowed with the Wasserstein metric. In addition to classical tools, we exploit the stability of the flow with respect to the transportation distance to greatly simplify many problems by reducing them to questions about particle approximations.

Keywords: well-posedness for measure solutions, gradient flows, optimal transport, nonlocal interactions, finite time blow-up, particle approximation, confinement.

AMS Classification: 35B40, 45K05, 49K20, 92DXX.

1. INTRODUCTION

We consider a mass distribution of particles, $\mu \geq 0$, interacting under a continuous *interaction potential*, W. The associated *interaction energy* is defined as

$$\mathcal{W}[\mu] := \frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} W(x - y) \, d\mu(x) \, d\mu(y). \qquad (1.1)$$

Our paper is devoted to the class of continuity equations of the form

$$\frac{\partial \mu}{\partial t} = \operatorname{div} \left[\left(\nabla \frac{\delta \mathcal{W}}{\delta \mu} \right) \mu \right] = \operatorname{div} \left[(\nabla W * \mu) \mu \right] \qquad x \in \mathbb{R}^d, t > 0. \qquad (1.2)$$

The equation is typically coupled with an initial datum

$$\mu(0) = \mu_0. \qquad (1.3)$$

The velocity field in the continuity equation, $-(\nabla W * \mu)(t, x)$, represents the combined contributions, at the point x, of the interaction through the potential W with particles at all other points.

The choice of W depends on the phenomenon studied. For instance in population dynamics, one is interested in the description of the evolution of a density of individuals. Very often the interaction between two individuals only depends on the distance between them. This suggests a choice of W as a *radial* function, i.e. $W(x) = w(|x|)$. Moreover, a choice of w such that $w'(r) > 0$ corresponds to an *attractive* force among the particles (or individuals), whereas $w'(r) < 0$ models a *repulsive* force.

Equation (1.2) arises in several applications in physics and biology. Simplified inelastic interaction models for granular media were considered in [4, 18] with $W = |x|^3/3$ and [41, 26] with $W = |x|^\alpha$, $\alpha > 1$. Such models usually lead to convex attractive potentials.

Mathematical modelling of the collective behavior of individuals, such as swarming, has also been treated by continuum models steaming from discrete particle models [31, 14, 39, 32, 40, 15, 34, 13, 21, 22, 16, 17]. Typical examples of interaction potentials appearing in these works are the attractive Morse potential $W(x) = -e^{-|x|}$, attractive-repulsive Morse potentials $W(x) = -C_a e^{-|x|/\ell_a} + C_r e^{-|x|/\ell_r}$, $W(x) = -e^{-|x|^2}$, $W(x) = -C_a e^{-|x|^2/\ell_a} + C_r e^{-|x|^2/\ell_r}$, or W being the characteristic function of a set in $\mathbb{R}^d$. A major issue is the possibility of a *finite time blow-up* of initially regular solutions, which occurs when w is attractive enough near $r = 0$. In particular, the solution can aggregate (collapse) part (or all) of its mass to a point in finite time. Blow-up producing potentials feature a suitable singularity in their second derivative at $r = 0$. Typically, the potential is of the form $W(x) \approx |x|^{1+\alpha}$ with $0 \leq \alpha < 1$, see [25, 7, 6, 5] in case of the Lipschitz singularity. Related questions with diffusion added to the system have been tackled in [9, 27, 28, 29].

Finally, another source of models with interaction potential appear in the modelling of cell movement by chemotaxis. In fact, the classical Patlak-Keller-Segel [36, 24] system, see [12, 10, 11], corresponds to the choice of the Newtonian potential in $\mathbb{R}^2$ as interaction, $W = \frac{1}{2\pi} \log |x|$, with linear diffusion. In the case without diffusion, a notion of weak measure solutions was introduced in [37] for which the author proved global-in-time existence, although uniqueness is lacking.

Given a continuous potential W, thanks to the structure of (1.2), we can assume without loss of generality that the following basic assumption holds:

(NL0) W is continuous, $W(x) = W(-x)$, and $W(0) = 0$.

We will say that W is a *locally attractive potential* if it verifies **(NL0)** and:

(NL1) W is λ–convex for a certain $\lambda \leq 0$, i.e. $W(x) - \frac{\lambda}{2}|x|^2$ is convex.
(NL2) There exists a constant $C > 0$ such that

$$W(z) \leq C(1 + |z|^2), \qquad \text{for all} \ \ z \in \mathbb{R}^d.$$

(NL3) $W \in C^1(\mathbb{R}^d \setminus \{0\})$.
(NL4) W has local minimum at $x = 0$.

We will say that the potential is a *pointy locally attractive potential* if it satisfies **(NL0)**-**(NL4)** and it has a Lipschitz singularity at the origin. In case the potential is continuously differentiable at the origin, we will speak about a *smooth potential*.

Remark 1.1. Assumptions **(NL0)**-**(NL1)** imply that

$$W(x) \geq \frac{\lambda}{2}|x|^2, \tag{1.4}$$

since $0 \in \partial W(0)$ and $W(0) = 0$. Hypotheses **(NL1)**-**(NL3)** imply a growth control on the gradient of W. More precisely, using the convexity of $x \mapsto V(x) := W(x) - \frac{\lambda}{2}|x|^2$ and the quadratic growth of $W(x)$, there exists $K > 0$ such that

$$\nabla V(x) \cdot p \le V(x+p) - V(x) \le K(1 + |x|^2 + |p|^2)$$

for any $x \ne 0$. Now, taking the supremum among all vectors p such that $|p| = \max\{|x|, 1\}$, we get $|\nabla V(x)| \le K(2 + 2|x|)$ from which

$$|\nabla W(x)| \le 2K + (2K + |\lambda|)\,|x|. \tag{1.5}$$

Let us also remark that **(NL1)** together with **(NL3)** imply that if the potential is not differentiable at the origin, then it has at most a Lipschitz singularity at the origin. Examples of locally attractive potentials neither pointy nor smooth are the ones with a local behavior at the origin like $|x|^{1+\alpha}$, with $0 < \alpha < 1$.

The first problem we treat in this paper is to give a well-posedness theory of weak measure solutions in the case of locally attractive potentials. Due to the possible concentration of solutions in a finite time, one has to allow for a concept of weak solution in a (nonnegative) measure sense. Our work fills in an important gap in the present studies of the equation. Simplistically speaking: [2, 3] provide a good theory for weak measure solutions for potentials which do not produce blow-up in finite time. On the other hand in the works that study potentials that do produce blow-up [25, 7, 6, 5] the notion of the solution breaks down at the blow-up time.

Before discussing the main results of this work, we introduce the concept of weak measure solution to (1.2). A natural way to introduce a concept of weak measure solution is to work in the space $\mathcal{P}(\mathbb{R}^d)$ of probability measures on $\mathbb{R}^d$. Since the class of equation described here does not feature mass–threshold phenomena, we normalize the mass to 1 without loss of generality. Following the approach developed in [2, 3], we shall consider weak measure solutions which additionally belong to the metric space

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

of probability measures with finite second moment, endowed with the 2–$Wasserstein$ distance d_W; see the next section.

Definition 1.2. *A locally absolutely continuous curve $\mu : [0, +\infty) \ni t \mapsto \mathcal{P}_2(\mathbb{R}^d)$ is said to be a* weak measure solution *to* (1.2) *with initial datum $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$ if and only if $\partial^0 W * \mu \in L^1_{loc}((0, +\infty); L^2(\mu(t)))$ and*

$$\int_0^{+\infty} \int_{\mathbb{R}^d} \frac{\partial \varphi}{\partial t}(x, t)\, d\mu(t)(x)\, dt + \int_{\mathbb{R}^d} \varphi(x, 0)\, d\mu_0(x) =$$

$$\int_0^{+\infty} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \nabla \varphi(x, t) \cdot \partial^0 W(x - y)\, d\mu(t)(x)\, d\mu(t)(y)\, dt,$$

for all test functions $\varphi \in C_c^\infty([0, +\infty) \times \mathbb{R}^d)$.

In this definition, $\partial^0 W(x)$ denotes the element of minimal norm in the subdifferential of W at x. In particular, thanks to **(NL3)** and **(NL4)**, we show

$$(\partial^0 W * \mu)(x) = \int_{y \ne x} \nabla W(x - y)\, d\mu(y).$$

Here, the absolute continuity of the curve of measures means that its metric derivative is integrable, see next section. Let us point out that the nonlinear term is well-defined due to the integrability of the velocity field against the weak measure solution.

The main idea to construct weak measure solutions to (1.2) is to use the interpretation of these equations as gradient flows in the space $\mathcal{P}_2(\mathbb{R}^d)$ of the interaction potential functional (1.1) with respect to the transport distance d_W. Such an interpretation turns out to be extremely well-adapted to proving uniqueness and stability results for *gradient flow solutions* compared to other strategies. This basic intuitive idea, introduced in [35] for the porous medium equation and generalized to a wide class of equations in [19], was made completely rigorous for a large class of equations in [2, 3] including some particular instances of (1.2). For gradient flow solutions we are able to obtain the existence, uniqueness and d_W-stability. Let us point out that gradient flow solutions are eventually shown to be equivalent to weak measure solutions.

Let us remark that the well-posedness theory of gradient flow solutions in the space of probability measures is developed in [2, 3] in case the potential W is smooth and λ-convex. Here, we mainly focus on generalizing this theory to allow Lipschitz singularities at the origin. In the case of locally attractive potentials, the technical point to deal with is the characterization of the subdifferential and its element of minimal norm. Moreover, we generalize this gradient flow theory allowing a negative quadratic behaviour at infinity. This fact introduces certain technical difficulties at the level of coercivity and lower semi-continuity of the functional defining the variational scheme. The well-posedness theory of gradient-flow solutions is the goal of Section 2.

One of the key properties of the constructed solutions is the stability with respect to d_W: given two gradient flow solutions μ_t^1 and μ_t^2,

$$d_W(\mu_t^1, \mu_t^2) \leq e^{-\lambda t} d_W(\mu_0^1, \mu_0^2)$$

for all $t \geq 0$. If $\lambda > 0$ the above estimate still holds provided that the initial measures have the same center of mass. Let us point out that the center of mass is preserved along the evolution. If the potential is smooth the stability estimate follows from [2, 19]. This stability result is not only useful for showing uniqueness but it is mainly a tool for approximating general solutions by particle ones. In fact, the previous estimate can be considered as a proof of the convergence of the continuous particle method for this equation on bounded time intervals. This is very much in the spirit of early works in the convergence of particle approximations to Vlasov-type equations in kinetic theory [33, 38].

Section 3 is devoted to show qualitative properties of the approximate solutions obtained by the variational scheme as in [23]. More precisely, we prove that particles remain particles at the level of a discrete variational scheme, provided the time step is small enough. In particular, this shows that the gradient flow solution starting from a finite number of particles remains at any time a finite number of particles, whose positions are determined by an ODE system. Although the fact that this construction give the solutions for finite number of particles can be directly checked on the solution concept, it is quite interesting to prove it directly at the variational scheme level, as it shows its suitability as a numerical scheme.

Section 4 is devoted to the question of finite-time blow-up of solutions. For a radially symmetric attractive potential, i.e. $W(x) = w(|x|)$, $w'(r) > 0$ for $r > 0$, the number

$$T(\varepsilon_1) := \int_0^{\varepsilon_1} \frac{dr}{w'(r)}, \qquad \varepsilon_1 > 0 \tag{1.6}$$

can be thought as the time it takes for a particle obeying the ODE $\dot{X} = -\nabla W(X)$ to reach the origin if it start at a distance ε_1 from the origin. This number quantifies the attractive strength of the potential: the smaller $T(\varepsilon_1)$ is, the more attractive the potential is. It was shown in [6, 7, 8] that if $T(\varepsilon_1) = +\infty$ for some (or equivalently for all) $\varepsilon_1 > 0$, then solutions of (1.2) starting with initial data in L^p will stay in L^p for all time, whereas if $T(\varepsilon_1) < +\infty$ for some $\varepsilon_1 > 0$, then compactly supported solutions will leave L^p in finite time (this result holds in the class of potentials which does not oscillate pathologically around the origin). In Section 4, thanks to our developed existence theory, we are able to obtain further understanding of the nature of the blow-up: loosely speaking, we prove that if the potential is attractive enough (i.e. $T(\varepsilon_1) < +\infty$ for some $\varepsilon_1 > 0$) then solutions of (1.2) starting with measure initial data will concentrate to a single Delta Dirac in finite time. We refer to this phenomena as finite time total collapse.

We will say that W is a *locally attractive non-Osgood potential* if in addition to **(NL0)-(NL4)**, it satisfies the finite time blow-up condition:

(NL-FTBU) W is radial, i.e. $W(x) = w(|x|)$, $W \in C^2(\mathbb{R}^d \setminus \{0\})$ with $w'(r) > 0$ for $r > 0$ and satisfying the following monotonicity condition: either **(a)** $w'(0^+) > 0$, or **(b)** $w'(0^+) = 0$ with $w''(r)$ monotone decreasing on an interval $(0, \varepsilon_0)$. Moreover, the potential satisfies the integrability condition

$$\int_0^{\varepsilon_1} \frac{1}{w'(r)} dr < +\infty, \qquad \text{for some } \varepsilon_1 > 0. \tag{1.7}$$

Let us point out that the condition of monotonicity of $w''(r)$ is not too restrictive. It is actually automatically satisfied by any potential who satisfies (1.7) and whose second derivative doesn't oscillate badly at the origin, as in [6, 7] (more comments on this assumption are done in Section 4). Examples of this type of potentials are the ones having a local behavior at the origin like $w'(r) \simeq r^\alpha$ with $0 \leq \alpha < 1$ or $w'(r) \simeq r \log^2 r$.

The proof is based on showing a finite-time total collapse result for the particles approximation independent of the number of particles possibly depending on the initial support. This fact, together with the convergence of the particle approximation, leads to the finite-time aggregation onto a single particle with the total mass of the system. This is the main technical novelty of our approach to blow-up.

Let us remark that the blow-up of the solution in L^p norms will in general happen before the total aggregation/collapse onto a single point. The transition from the first L^∞-blow-up to the total collapse can be very complicated. For instance one could have multiple points of aggregation onto Dirac deltas interacting between them and with smooth parts of the measure in a challenging evolution before the total aggregation onto a single point. This is also explained in Section 4. This complex behavior was already encountered in [37] in the case of the chemotaxis model without diffusion, but his notion of solution lacks of uniqueness and stability. Many problems on the details of the blowup in (1.2) and the interaction of delta masses with surrounding absolutely-continuous-measure part remain.

The last section is devoted to proving confinement of solutions for attractive/repulsive potentials which are radial and increasing outside a ball, that is for ones that satisfy

(NL-RAD) There exists $R_a \geq 0$ such that $W(x) = w(|x|)$ for $|x| \geq R_a$, and $w'(r) \geq 0$ for $r > R_a$.

We say that a potential is *strongly-attractive-at-infinity* if in addition to **(NL0)-(NL4)** and **(NL-RAD)** it satisfies the strong confinement condition:

(NL-CONF-strong)

$$\liminf_{r \to +\infty} w'(r) > 8\sqrt{2}\, C_W,$$

where

$$C_W := \begin{cases} 0 & \text{if } R_a = 0 \\ \sup_{x \in B(0,R_a) \setminus \{0\}} |\nabla W(x)| & \text{if } R_a > 0 \end{cases} \tag{1.8}$$

¿From (1.5) follows that C_W is finite.

We say that a potential is *weakly-attractive-at-infinity* if in addition to **(NL0)-(NL4)** and **(NL-RAD)** it satisfies the weak confinement condition:

(NL-CONF-weak)

$$\lim_{r \to +\infty} w'(r)\sqrt{r} = +\infty.$$

We prove estimates on the evolution of the radius of the support of solutions. For potentials satisfying **(NL-CONF-strong)** these estimates guarantee that if the radius of the support is large then it must be decreasing. We then refine the argument to show that even if only **(NL-CONF-weak)** is satisfied, the radius must remain uniformly bounded in time. To showcase the robustness of the notion of the solution we use two different techniques: for the first result we use the JKO approximation scheme, while for the second one we use the particle approximations.

2. The Jordan–Kinderlehrer–Otto (JKO) scheme

In this section we develop the existence theory for measure–valued solutions in the sense of Definition 1.2 by following the set up developed in [2]. A natural choice of a space of measures where to develop such a theory is the space $\mathcal{P}_2(\mathbb{R}^d)$ endowed with the *Wasserstein* distance

$$d_W(\mu,\nu) := \left[\min \left\{ \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |x-y|^2 d\gamma(x,y) : \; \gamma \in \Gamma(\mu,\nu) \right\} \right]^{1/2}, \tag{2.1}$$

where the set $\Gamma(\mu,\nu)$ of *transport plans* between μ and ν is defined by

$$\Gamma(\mu,\nu) := \left\{ \gamma \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d) : (\pi_1)_{\#}\gamma = \mu \text{ and } (\pi_2)_{\#}\gamma = \nu \right\}$$

with $\pi_1(x,y) = x$ and $\pi_2(x,y) = y$, that is,

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \phi(x) d\gamma = \int_{\mathbb{R}^d} \phi(x) d\mu, \quad \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \phi(y) d\gamma = \int_{\mathbb{R}^d} \phi(y) d\nu, \quad \text{for all } \phi \in C_b(\mathbb{R}^d).$$

The space $(\mathcal{P}_2(\mathbb{R}^d), d_W)$ is a complete metric space [42, 2]. The standard theory of optimal transportation [2, 42] provides the existence of an optimal transport plan for variational

problem (2.1), i.e. there exists $\gamma_o \in \Gamma(\mu, \nu)$ such that

$$d_W^2(\mu, \nu) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |x - y|^2 d\gamma_o(x, y). \tag{2.2}$$

The set of all the *optimal plans* γ_o satisfying (2.2) is denoted by $\Gamma_o(\mu, \nu)$.

We recall that the *interaction energy* $\mathcal{W} : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is defined as follows:

$$\mathcal{W}[\mu] := \frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} W(x - y) d\mu(x) d\mu(y). \tag{2.3}$$

Note that the definition of $\mathcal{W}$ is well-posed due to assumptions **(NL1)-(NL2)**, which provide suitable control of the integral at infinity. Also, the continuity of W ensures the well-posedness of $\mathcal{W}$ on singular measures.

Following [2], we shall first address the problem of the existence of a *curve of maximal slope* for the functional $\mathcal{W}$. For this purpose, let us introduce some definitions. The *slope of $\mathcal{W}$* is defined as:

$$|\partial \mathcal{W}|[\mu] := \limsup_{\nu \to \mu} \frac{(\mathcal{W}[\mu] - \mathcal{W}[\nu])^+}{d_W(\mu, \nu)}, \tag{2.4}$$

where $u^+ := \max\{u, 0\}$. Given an absolutely continuous curve $[0, T] \ni t \mapsto \mu(t) \in \mathcal{P}_2(\mathbb{R}^d)$, its *metric derivative* is:

$$|\mu'|(t) := \limsup_{s \to t} \frac{d_W(\mu(s), \mu(t))}{|s - t|}. \tag{2.5}$$

Finally, we recall the definition of a curve of maximal slope for the functional $\mathcal{W}$. With the notation in [2], such a notion is referred to as a "curve of maximal slope with respect to $|\partial \mathcal{W}|$".

Definition 2.1. *A locally absolutely continuous curve $[0, T] \ni t \mapsto \mu(t) \in \mathcal{P}_2(\mathbb{R}^d)$ is a curve of maximal slope for the functional $\mathcal{W}$ if $t \mapsto \mathcal{W}[\mu(t)]$ is an absolutely continuous function, and the following inequality hold for every $0 \le s \le t \le T$:*

$$\frac{1}{2} \int_s^t |\mu'|^2(r) \, dr + \frac{1}{2} \int_s^t |\partial \mathcal{W}|^2[\mu(r)] \, dr \le \mathcal{W}[\mu(s)] - \mathcal{W}[\mu(t)]. \tag{2.6}$$

The notion of solutions provided in Definition 2.1 is purely metric (see [2, Part I]). We shall improve this notion of solution (in the spirit of [2, Part II]) to a solution in the "gradient flow" sense in Subsection 2.3.

The inequality (2.6), which defines the notion of a curve of maximal slope, is better understood after providing a representation formula for the slope $|\partial \mathcal{W}|$ in terms of an integral norm of a vector field involving the "gradient" of W, or rather its minimal subdifferential $\partial^0 W$. Moreover, the metric derivative $|\mu'|$ should be interpreted in a "length space" sense, which accounts for the metric space $\mathcal{P}_2(\mathbb{R}^d)$ being endowed with a kind of Riemannian structure, first introduced in [35] and then proven rigorously in [2]. For the sake of clarity, let us briefly recall this framework (see [2, Chapter 8] for further details).

Given a measure $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, the tangent space $Tan_\mu \mathcal{P}_2(\mathbb{R}^d)$ to $\mathcal{P}_2(\mathbb{R}^d)$ in μ is the closed vector subspace of $L^2(\mu)$ given by

$$Tan_\mu \mathcal{P}_2(\mathbb{R}^d) := \overline{\{\nabla \phi \ : \ \phi \in C_c^\infty(\mathbb{R}^d)\}}^{L^2(\mu)}.$$

Moreover, given an absolutely continuous curve $t \mapsto \mu(t) \in \mathcal{P}_2(\mathbb{R}^d)$, the "tangent vectors" to $\mu(t)$ can be identified as elements of the set of vector fields $v(t)$ solving the continuity equation

$$\partial_t \mu(t) + \nabla \cdot (v(t)\mu(t)) = 0 \tag{2.7}$$

in the sense of distributions. Among all the possible velocity fields $v(t)$ solving (2.7), as a consequence of [2, Theorem 8.3.1], there is one with minimal $L^2(\mu(t))$ norm, equal to the metric derivative of $\mu(t)$. Therefore, we have the following representation formula for $|\mu'|(t)$:

$$|\mu'|(t) = \min \Big\{ \ \|v(t)\|_{L^2(\mu(t))} : v(t) \text{ solves } \partial_t \mu(t) + \nabla \cdot (v(t)\mu(t)) = 0$$

$$\text{on } \mathbb{R}^d \times (0, T) \ \text{ in the sense of distributions} \Big\}.$$

We recall here the inequality

$$\limsup_{\varepsilon \searrow 0} \frac{d_W((id + \varepsilon\xi)_\# \mu, \mu)}{\varepsilon} \leq \|\xi\|_{L^2(d\mu)} \tag{2.8}$$

which follows easily from the definition of d_W.

As for the slope $|\partial \mathcal{W}|$ of the functional $\mathcal{W}$ (similarly to the classical subdifferential calculus in Hilbert spaces), it can be written as

$$|\partial \mathcal{W}|(\mu) = \min \left\{ \|w\|_{L^2(\mu)} : \ w \in \partial \mathcal{W}(\mu) \right\},$$

where $\partial \mathcal{W}(\mu)$ is the (possibly multivalued) *subdifferential* of $\mathcal{W}$ at the measure μ. The definition of subdifferential of a functional $\mathcal{W}$ on $\mathcal{P}_2(\mathbb{R}^d)$ in the general case is pretty involved (see [2, Definition 10.3.1]) and we shall not need to recall it here. In the next subsection, we follow the approach of [2] to characterize the (unique) element of the *minimal subdifferential of* $\mathcal{W}$ denoted by $\partial^0 \mathcal{W}(\mu)$.

2.1. Subdifferential of $\mathcal{W}$.
Given W a locally attractive potential, i.e. satisfying **(NL0)-(NL4)**, let $\partial W(x)$ be the (possibly multivalued) subdifferential of W at the point x, namely the convex set

$$\partial W(x) := \left\{ \kappa \in \mathbb{R}^d : \ W(y) - W(x) \geq \kappa \cdot (y - x) + o(|x - y|), \text{ for all } y \in \mathbb{R}^d \right\}.$$

Denoting by $\partial^0 W(x)$ the (unique) element of $\partial W(x)$ with minimal norm, due to the assumptions **(NL3)-(NL4)** and **(NL0)**, $\partial^0 W(x) = \nabla W(x)$ for $x \neq 0$ and $\partial^0 W(0) = 0$.

A vector field $\mathbf{w} \in L^2(\mu)$ is said to be an element of the subdifferential of $\mathcal{W}$ at μ, and we write $\mathbf{w} \in \partial \mathcal{W}[\mu]$, if

$$\mathcal{W}[\nu] - \mathcal{W}[\mu] \geq \int_{\mathbb{R}^d \times \mathbb{R}^d} \mathbf{w}(x) \cdot (y - x) \, d\gamma_o(x, y) + o(d_W(\nu, \mu)) \tag{2.9}$$

for all $\gamma_o \in \Gamma_o(\mu, \nu)$. In principle, according to [2, Definition 10.3.1], the elements of $\partial W(x)$ are *plans* γ in the set $\mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$ such that $(\pi_1)_\# \gamma = \mu$. If a plan $\gamma \in \partial \mathcal{W}[\mu]$ is concentrated on the graph of a vector field $\mathbf{w} \in L^2(\mu)$, then [2, Definition 10.3.1] reduces to (2.9). By following the approach of [2, Sections 10.3 and 10.4], it is easy to see that the (unique) element with minimal norm of $\partial \mathcal{W}[\mu]$ is concentrated on the graph of a vector field. Following [2], we call this element the *minimal subdifferential* of $\mathcal{W}$ at μ, and we denote it by $\partial^0 \mathcal{W}[\mu]$. The following characterization of the subdifferential is obtained in

[2, Theorem 10.4.11] for smooth potentials, and here we generalize it to locally attractive potentials:

Proposition 2.2. *Given a locally attractive potential, the vector field*

$$\kappa(x) := (\partial^0 W * \mu)(x) = \int_{y \neq x} \nabla W(x - y) \, d\mu(y)$$

is the unique element of the minimal subdifferential of $\mathcal{W}$, i.e. $\partial^0 W * \mu = \partial^0 \mathcal{W}[\mu]$.

Proof. We divide the proof into two steps.
Step 1: $\kappa(x) \in \partial \mathcal{W}[\mu]$. We have to show that

$$\mathcal{W}[\nu] - \mathcal{W}[\mu] \geq \int_{\mathbb{R}^d \times \mathbb{R}^d} \kappa(x) \cdot (y - x) \, d\gamma(x, y) + o(d_W(\nu, \mu))$$

for all $\gamma \in \Gamma_o(\mu, \nu)$. Thanks to the λ-convexity of W, it suffices to prove that

$$\liminf_{t \to 0} \frac{\mathcal{W}[(1 - t)\pi_1 + t\pi_2)_\# \gamma] - \mathcal{W}[\mu]}{t} \geq \int_{\mathbb{R}^d \times \mathbb{R}^d} \kappa(x) \cdot (y - x) \, d\gamma(x, y). \qquad (2.10)$$

To see this, we observe that the λ–convexity of W implies that the function

$$t \mapsto f(t) := \frac{W(ty + (1 - t)x) - W(x)}{t} - \frac{\lambda}{2} t |x - y|^2 \qquad (2.11)$$

is nondecreasing in t. Therefore, by writing $f(1) \geq \liminf_{t \searrow 0} f(t)$, integrating in $d\gamma(x, y)$ and using the monotone convergence theorem, we easily recover

$$\mathcal{W}[\nu] - \mathcal{W}[\mu] \geq \liminf_{t \to 0} \frac{\mathcal{W}[(1 - t)\pi_1 + t\pi_2)_\# \gamma] - \mathcal{W}[\mu]}{t} + \frac{\lambda}{2} d_W^2(\nu, \mu).$$

We now prove (2.10). Let us compute

$$\mathcal{W}[(1 - t)\pi_1 + t\pi_2)_\# \gamma] - \mathcal{W}[\mu]$$
$$= \frac{1}{2} \int_{\mathbb{R}^d \times \mathbb{R}^d} \int_{\mathbb{R}^d \times \mathbb{R}^d} [W(t(y_2 - y_1) + (1 - t)(x_2 - x_1)) - W(x_2 - x_1)] \, d\gamma(x_1, y_1) \, d\gamma(x_2, y_2)$$
$$= \frac{1}{2} \int_{x_1 \neq x_2} [W(t(y_2 - y_1) + (1 - t)(x_2 - x_1)) - W(x_2 - x_1)] \, d\gamma(x_1, y_1) \, d\gamma(x_2, y_2)$$
$$+ \frac{1}{2} \int_{x_1 = x_2} W(t(y_2 - y_1)) \, d\gamma(x_1, y_1) \, d\gamma(x_2, y_2). \qquad (2.12)$$

We decompose μ as $\mu^a + \mu^r$, where $\mu^a = \sum_i m_i \delta_{a_i}$ denotes the atomic part of μ. Disintegrating γ with respect to μ as $\gamma_x(dy) \otimes \mu(dx)$, we get

$$\sum_i m_i \int_{\mathbb{R}^d} |y - a_i|^2 \, d\gamma_{a_i}(y) = \int_{\mathbb{R}^d \times \mathbb{R}^d} |y - x|^2 \, d\gamma_x(y) \, d\mu^a(x)$$

$$\leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |y - x|^2 \, d\gamma(x, y) = d_W^2(\mu, \nu).$$

Therefore, observing that $0 \leq m_i \leq 1$ and using the fact that $\delta_{a_i}(dx_1) \otimes \mu^r(dx_2)$ and $\mu^r(dx_1) \otimes \mu^r(dx_2)$ give no mass to the diagonal set $\{x_1 = x_2\}$, we get

$$
\int_{x_1=x_2} W(t(y_2 - y_1)) \, d\gamma(x_1, y_1) \, d\gamma(x_2, y_2) = \sum_i m_i^2 \int_{x_1=x_2} W(t(y_2 - y_1)) \, d\gamma_{a_i}(y_1) \, d\gamma_{a_i}(y_2)
$$

$$
\geq -\frac{\lambda}{2} t^2 \sum_i m_i^2 \int_{\mathbb{R}^d \times \mathbb{R}^d} |y_2 - y_1|^2 \, d\gamma_{a_i}(y_1) \, d\gamma_{a_i}(y_2)
$$

$$
\geq -\lambda t^2 \sum_i m_i \int_{\mathbb{R}^d \times \mathbb{R}^d} \left[|y_2 - a_i|^2 + |y_1 - a_i|^2 \right] \, d\gamma_{a_i}(y_1) \, d\gamma_{a_i}(y_2)
$$

$$
\geq -2\lambda t^2 d_W^2(\mu, \nu).
$$

which implies that in taking the limit as $t \to 0$ in (2.12) we can neglect the term coming from the integral over the set $\{x_1 = x_2\}$. As for the other term in (2.12), we write it as follows:

$$
\lim_{t \to 0} \int_{x_1 \neq x_2} \frac{W(t(y_2 - y_1) + (1 - t)(x_2 - x_1)) - W(x_2 - x_1)}{2t} \, d\gamma(x_1, y_1) \, d\gamma(x_2, y_2)
$$

$$
= \lim_{t \to 0} \int_{x_1 \neq x_2} \left[\frac{W(t(y_2 - y_1) + (1 - t)(x_2 - x_1)) - W(x_2 - x_1)}{2t} \right.
$$

$$
\left. - \frac{\lambda t}{4} |x_2 - x_1 - y_2 + y_1|^2 \right] \, d\gamma(x_1, y_1) \, d\gamma(x_2, y_2)
$$

$$
+ \lim_{t \to 0} \int_{x_1 \neq x_2} \frac{\lambda t}{4} |x_2 - x_1 - y_2 + y_1|^2 \, d\gamma(x_1, y_1) \, d\gamma(x_2, y_2). \tag{2.13}
$$

Now, since W is locally Lipschitz with quadratic growth at infinity and has no singularity out of the origin, in the first integral in the right-hand side of (2.13) we can pass the limit into the integral because of the monotonicity of the function $f(t)$ defined in (2.11) and the monotone convergence limit. The second term in the right-hand side of (2.13) equals zero since $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$. Therefore, using the fact that ∇W is odd, we have

$$
\lim_{t \to 0} \frac{1}{2} \int_{x_1 \neq x_2} \frac{W(t(y_2 - y_1) + (1 - t)(x_2 - x_1)) - W(x_2 - x_1)}{t} \, d\gamma(x_1, y_1) \, d\gamma(x_2, y_2)
$$

$$
= \frac{1}{2} \int_{x_1 \neq x_2} \nabla W(x_2 - x_1) \cdot [(y_2 - x_2) - (y_1 - x_1)] \, d\gamma(x_1, y_1) \, d\gamma(x_2, y_2)
$$

$$
= \int_{x_1 \neq x_2} \nabla W(x_2 - x_1) \cdot (y_2 - x_2) \, d\gamma(x_1, y_1) \, d\gamma(x_2, y_2)
$$

$$
= \int_{\mathbb{R}^d \times \mathbb{R}^d} \kappa(x) \cdot (y - x) \, d\gamma(x, y),
$$

and this gives the desired result.

Step 2: w is the element of minimal norm of $\partial \mathcal{W}[\mu]$. We closely follows the argument in [2, Theorem 10.4.11]. Fix a vector field $\xi \in C_c^\infty(\mathbb{R}^d, \mathbb{R}^d)$. Observing that

$W(x - z + t(\xi(x) - \xi(z))) = W(x - z) = 0$ when $x = z$, we get

$$\lim_{t \to 0} \frac{\mathcal{W}[(id + t\xi)_{\#}\mu] - \mathcal{W}[\mu]}{t}$$

$$= \lim_{t \to 0} \frac{1}{2} \int_{\mathbb{R}^d \times \mathbb{R}^d} \frac{W((x - z) + t(\xi(x) - \xi(z))) - W(x - z)}{t} \, d\mu(x) \, d\mu(z)$$

$$= \lim_{t \to 0} \frac{1}{2} \int_{x \neq z} \frac{W((x - z) + t(\xi(x) - \xi(z))) - W(x - z)}{t} \, d\mu(x) \, d\mu(z)$$

$$= \frac{1}{2} \int_{x \neq z} \nabla W(x - z) \cdot (\xi(x) - \xi(z)) \, d\mu(x) \, d\mu(z)$$

$$= \int_{\mathbb{R}^d} \kappa(x) \cdot \xi(x) \, d\mu(x). \tag{2.14}$$

Hence, since the definition of slope (2.4) easily implies

$$\liminf_{t \searrow 0} \frac{\mathcal{W}[(id + t\xi)_{\#}\mu] - \mathcal{W}[\mu]}{d_W((id + t\xi)_{\#}\mu, \mu)} \geq -|\partial \mathcal{W}|(\mu),$$

we can use (2.8) and (2.14) to get

$$\int_{\mathbb{R}^d} \kappa(x) \cdot \xi(x) \, d\mu(x) \geq -|\partial \mathcal{W}|(\mu) \liminf_{t \to 0} \frac{d_W((id + t\xi)_{\#}\mu, \mu)}{t} \geq -|\partial \mathcal{W}|(\mu) \|\xi\|_{L^2(\mu)}.$$

Changing ξ with $-\xi$ gives

$$\left| \int_{\mathbb{R}^d} \kappa(x) \cdot \xi(x) \, d\mu(x) \right| \leq |\partial \mathcal{W}|(\mu) \|\xi\|_{L^2(\mu)},$$

so that by the arbitrariness of ξ we get $\|\kappa\|_{L^2(\mu)} \leq |\partial \mathcal{W}|(\mu)$, and therefore κ is the (unique) element of minimal norm. $\qquad\square$

2.2. Well-posedness and convergence of the scheme.

The approach of [2] in proving the existence of a curve of maximal slope for a functional on $\mathcal{P}_2$ is based on a variational version of the implicit Euler scheme, sometimes referred to as the *Jordan–Kinderlehrer–Otto (JKO) scheme* or *minimizing movement scheme* [23, 1, 2]. Given an initial measure $\mu_0 \in \mathcal{P}_2$ and time-step $\tau > 0$, we consider a sequence μ_k^τ recursively defined by $\mu_0^\tau = \mu_0$ and

$$\mu_{k+1}^\tau \in \arg\min_{\mu \in \mathcal{P}_2} \left\{ \mathcal{W}[\mu] + \frac{1}{2\tau} d_W^2(\mu_k^\tau, \mu) \right\}, \tag{2.15}$$

for all $k \in \mathbb{N}$.

We shall address here the well-posedness of the definition (2.15) and the convergence of μ_k^τ as $\tau \to 0$ (after a suitable interpolation) to a limit which satisfies Definition 2.1. Such a problem has been widely studied for smooth convex potentials in [2], where convergence of the discrete scheme to a suitable limit is shown. However, allowing for $W(x)$ behaving like $-C|x|^2$ as $|x| \to +\infty$ and for a pointy singularity at $x = 0$ requires some improvements of the arguments in [2, Part I], as we shall see below. Indeed let us point out that, for $W(x)$ behaving like $-C|x|^2$, the functional $\mathcal{W}[\mu]$ is upper (and not lower!) semicontinuous with respect to the convergence in d_W.

For the sake of clarity, we shall recall all the main steps of the JKO scheme developed in [2] in the particular case of a functional given by a pure nonlocal interaction energy. We shall perform this task also for another reason, namely to relax the set of assumptions

(NL0)-(NL4) in order to admit $|\nabla W|$ to be possibly unbounded at the origin (see Remark 2.11).

We start by showing that the minimization problem (2.15) admits at least one solution, which in our situation is not a trivial issue. To this aim, we prove a technical lemma which will be also useful in the sequel.

Lemma 2.3 (Weak lower semi–continuity of the penalized interaction energy). *Suppose W satisfies* **(NL0)-(NL3)**. *Then, there exists $\tau_0 > 0$ depending only on W such that, for a fixed $\overline{\mu} \in \mathcal{P}_2(\mathbb{R}^d)$, the penalized interaction energy functional*

$$\mathcal{P}_2(\mathbb{R}^d) \ni \mu \mapsto \mathcal{W}[\mu] + \frac{1}{2\tau}d_W^2(\mu, \overline{\mu})$$

is lower semi–continuous with respect to the narrow topology of $\mathcal{P}(\mathbb{R}^d)$ for all $0 < \tau < \tau_0$.

Proof. Let $\{\mu_n\}_n \subset \mathcal{P}_2(\mathbb{R}^d)$ such that $\lim_{n\to+\infty}\mu_n = \mu_\infty$ narrowly. We have to prove that

$$\liminf_{n\to+\infty}\left[\mathcal{W}[\mu_n] + \frac{1}{2\tau}d_W^2(\mu_n, \overline{\mu})\right] \geq \mathcal{W}[\mu_\infty] + \frac{1}{2\tau}d_W^2(\mu_\infty, \overline{\mu}). \tag{2.16}$$

Let us first suppose that $\overline{\mu}$ has compact support contained in the sphere $B(0, R) = \{|x| \leq R\}$. From the estimate in Remark 1.1, we have

$$W(x - y) \geq \frac{\lambda}{2}(1 + |x - y|^2) \geq \frac{\lambda}{2}(1 + 2|x|^2 + 2|y|^2),$$

which implies that

$$h(x, y) := W(x - y) - \frac{\lambda}{2}(1 + 2|x|^2 + 2|y|^2)$$

is a nonnegative continuous function. Therefore,

$$
\begin{aligned}
\mathcal{W}[\mu_n] + \frac{1}{2\tau}d_W^2(\mu_n, \overline{\mu}) &= \frac{\lambda}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}(1 + 2|x|^2 + 2|y|^2)\,d\mu_n(x)\,d\mu_n(y) \\
&\quad + \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}h(x, y)\,d\mu_n(x)\,d\mu_n(y) + \frac{1}{2\tau}d_W^2(\mu_n, \overline{\mu}) \\
&= \frac{1}{2\tau}d_W^2(\mu_n, \overline{\mu}) + \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}h(x, y)\,d\mu_n(x)\,d\mu_n(y) \\
&\quad \frac{\lambda}{2} + 2\lambda\int_{\mathbb{R}^d}|x|^2\,d\mu_n(x).
\end{aligned}
\tag{2.17}
$$

Since $h \geq 0$, we easily get

$$\liminf_{n\to+\infty}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}h(x, y)d\mu_n(x)d\mu_n(y) \geq \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}h(x, y)d\mu_\infty(x)d\mu_\infty(y). \tag{2.18}$$

Now, let $\gamma_n \in \Gamma_o(\overline{\mu}, \mu_n)$. Then,

$$
\begin{aligned}
\frac{1}{2\tau}d_W^2(\mu_n, \overline{\mu}) + 2\lambda\int_{\mathbb{R}^d}|x|^2 d\mu_n(x) &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\left(\frac{1}{2\tau}|x - y|^2 + 2\lambda|y|^2\right)d\gamma_n(x, y) \\
&= \int\int_{|x-y|\leq 4R}\left(\frac{1}{2\tau}|x - y|^2 + 2\lambda|y|^2\right)d\gamma_n(x, y) \tag{2.19} \\
&\quad + \int\int_{|x-y|>4R}\left(\frac{1}{2\tau}|x - y|^2 + 2\lambda|y|^2\right)d\gamma_n(x, y). \tag{2.20}
\end{aligned}
$$

As for the term in (2.19), we first note the compactness of the domain of integration since $\text{supp}(\bar{\mu}) \subset B(0, R)$ and thus $\text{supp}(\gamma_n) \subset B(0, R) \times \mathbb{R}^d$ and the integration is reduced to the set where $x \in B(0, R)$ and $|x - y| \leq 4R$. Stability of optimal transportation plans (see [43, Theorem 5.20]) implies that there exists a subsequence, that we may assume to be the whole sequence, such that γ_n converges narrowly to an optimal plan $\gamma_\infty \in \Gamma_o(\mu_\infty, \bar{\mu})$. As a consequence, we easily obtain

$$\lim_{n \to +\infty} \int \int_{|x-y| \leq 4R} \left(\frac{1}{2\tau} |x-y|^2 + 2\lambda |y|^2 \right) d\gamma_n(x, y)$$
$$= \int \int_{|x-y| \leq 4R} \left(\frac{1}{2\tau} |x-y|^2 + 2\lambda |y|^2 \right) d\gamma_\infty(x, y) \qquad (2.21)$$

where γ_∞ is the narrow limit of γ_n as $n \to +\infty$.

As for the integrand in the term (2.20), we observe that, since $|x| \leq R$ and $|y| \geq |x-y| - |x| \geq 4R - R = 3R$,

$$\frac{|x-y|}{|y|} \geq \frac{|y| - |x|}{|y|} = 1 - \frac{|x|}{|y|} \geq 1 - \frac{R}{3R} = \frac{2}{3}.$$

Hence, on the set $\{|x-y| \geq 4R\}$, we have

$$\frac{1}{2\tau} |x-y|^2 + 2\lambda |y|^2 \geq \left(\frac{1}{3\tau} + 2\lambda \right) |y|^2,$$

and the right-hand side above is nonnegative for τ small enough. Therefore, as for (2.18), we easily get

$$\liminf_{n \to +\infty} \int \int_{|x-y| > 4R} \left(\frac{1}{2\tau} |x-y|^2 + 2\lambda |y|^2 \right) d\gamma_n(x, y)$$
$$\geq \int \int_{|x-y| > 4R} \left(\frac{1}{2\tau} |x-y|^2 + 2\lambda |y|^2 \right) d\gamma_\infty(x, y), \qquad (2.22)$$

where once again, $\gamma_\infty \in \Gamma_o(\mu_\infty, \bar{\mu})$. By combining together (2.17), (2.18), (2.21) and (2.22), and by using that $\gamma_\infty \in \Gamma_o(\mu_\infty, \bar{\mu})$, we obtain the desired assertion (2.16).

The proof for a general $\bar{\mu} \in \mathcal{P}_2(\mathbb{R}^d)$ easily follows by approximating $\bar{\mu}$ with respect to d_W by a sequence of compactly supported probability measures. $\square$

Remark 2.4. We observe that the optimality of the plans γ_n is never actually needed in the previous proof.

Next, we prove the solvability of the minimization problem (2.15).

Proposition 2.5 (Existence of minimizers). *Suppose W satisfies* **(NL0)-(NL3)**. *Then, there exists $\tau_0 > 0$ depending only on W such that, for all $0 < \tau < \tau_0$ and for a given $\bar{\mu} \in \mathcal{P}_2(\mathbb{R}^d)$, there is $\mu_\infty \in \mathcal{P}_2(\mathbb{R}^d)$ such that*

$$\mathcal{W}[\mu_\infty] + \frac{1}{2\tau} d_W^2(\bar{\mu}, \mu_\infty) = \min_{\mu \in \mathcal{P}_2(\mathbb{R}^d)} \left\{ \mathcal{W}[\mu] + \frac{1}{2\tau} d_W^2(\bar{\mu}, \mu) \right\}.$$

Proof. STEP 1: COMPACTNESS. Let us fix a measure $\bar{\mu} \in \mathcal{P}_2(\mathbb{R}^d)$ and a time step $\tau > 0$, and consider a minimizing sequence $\mu_n \in \mathcal{P}_2(\mathbb{R}^d)$, i.e.

$$\inf_{\mu \in \mathcal{P}_2(\mathbb{R}^d)} \left\{ \mathcal{W}[\mu] + \frac{1}{2\tau} d_W^2(\mu, \bar{\mu}) \right\} = \lim_{n \to +\infty} \left\{ \mathcal{W}[\mu_n] + \frac{1}{2\tau} d_W^2(\mu_n, \bar{\mu}) \right\}.$$

Since μ_n is a minimizing sequence, we have

$$\mathcal{W}[\mu_n] + \frac{1}{2\tau}d_W^2(\mu_n, \overline{\mu}) \leq C_1 \qquad (2.23)$$

for some constant C_1. Then, the lower estimate of W in Remark 1.1 implies

$$
\begin{aligned}
\frac{1}{2\tau}d_W^2(\mu_n, \overline{\mu}) + \mathcal{W}[\mu_n] &\geq \frac{1}{2\tau}d_W^2(\mu_n, \overline{\mu}) + \frac{\lambda}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}(1+|x-y|^2)d\mu_n(x)d\mu_n(y) \\
&\geq \frac{1}{2\tau}d_W^2(\mu_n, \overline{\mu}) + \frac{\lambda}{2} + 2\lambda\int_{\mathbb{R}^d}|x|^2 d\mu_n(x) \\
&\geq \frac{1}{2\tau}d_W^2(\mu_n, \overline{\mu}) + \frac{\lambda}{2} + 2\lambda\int_{\mathbb{R}^d}|x|^2 d\overline{\mu}(x) + 2\lambda d_W^2(\mu_n, \overline{\mu}) \quad (2.24)
\end{aligned}
$$

and the above with τ small enough together with (2.23) implies that $d_W^2(\mu_n, \overline{\mu})$ is uniformly bounded with respect to n. Prokhorov's compactness theorem then implies that the sequence $\{\mu_n\}_n$ is tight.

STEP 2: COERCIVITY. We need to prove that

$$\liminf_{n\to+\infty}\left[\mathcal{W}[\mu_n] + \frac{1}{2\tau}d_W^2(\mu_n, \overline{\mu})\right] \geq C_0\, d_W^2(\mu_n, \overline{\mu}) - C_1$$

for some positive constant C_0, C_1 independent on n. This follows easily from (2.24) for τ small enough.

STEP 3: PASSING TO THE LIMIT BY LOWER SEMI–CONTINUITY. This is a consequence of Lemma 2.3. $\square$

Next we have to establish that the family $\{\mu_k^\tau\}_{\tau\in(0,\tau_0)}$ (up to a suitable interpolation) converges narrowly to a certain limit. This task can be performed exactly as described in [2, Chapters 2, 3]. For the sake of clarity, we recall here the result in [2] stating the convergence of the JKO scheme. The proof can be found in [2, Proposition 2.2.3]. First, we introduce the piecewise constant interpolation

$$
\begin{aligned}
\mu^\tau(0) &:= \mu_0 \\
\mu^\tau(t) &:= \mu_k^\tau \qquad \text{if} \quad t\in((k-1)\tau, k\tau], \quad k\geq 1.
\end{aligned}
$$

Proposition 2.6 (Compactness in the JKO scheme [2]). *Suppose W satisfies* **(NL0)-(NL4)**. *There exists a sequence $\tau_n \searrow 0$, and a limit curve $\mu \in AC_{loc}\left([0,+\infty); \mathcal{P}_2(\mathbb{R}^d)\right)$, such that*

$$\mu^n(t) := \mu^{\tau_n}(t) \to \mu(t), \qquad narrowly \ as \quad n\to+\infty$$

for all $t\in[0,+\infty)$.

Let us first remark that the previous theorem holds true for λ-convex potentials with at most quadratic growth that are smooth everywhere in the sense of replacing **(NL4)** by $W\in C^1(\mathbb{R}^d)$, fact that will be used later in Section 5.

According to the notation recalled in [2, Definition 2.0.6], the above proposition states that the set of minimizing movements for $\mathcal{W}$ starting from μ_0 is not empty. The last step of the procedure proposed in [2] is to check that the limit curve provided by Proposition 2.6 is a curve of maximal slope for W according to definition 2.1.

Let $\tilde{\mu}^n(t)$ denote the *De Giorgi variational interpolation* (see [2, Section 3.2]). Then, from [2, Lemma 3.2.2] we have the energy inequality

$$\mathcal{W}[\mu_0] \geq \frac{1}{2}\int_0^T \|v^n(t)\|_{L^2(\mu^n(t))}^2\,dt + \frac{1}{2}\int_0^T |\partial\mathcal{W}|(\tilde{\mu}^n(t))^2\,dt + \mathcal{W}[\mu^n(T)] \qquad (2.25)$$

for all $T > 0$, where $v^n(t)$ denotes the minimal velocity field associated to the curve $t \mapsto \mu^n(t)$, namely $v^n(t)$ is the minimizing velocity field in definition given in (2.5). Up to a subsequence, both $\mu^n(t)$ and $\tilde{\mu}^n(t)$ narrowly converge to the same limit curve $\mu(t)$ on $[0,+\infty)$ provided by Proposition 2.6. The following lemma is needed to suitably pass to the limit the slope term in (2.25).

Lemma 2.7 (Lower semicontinuity of the slope).

$$\liminf_{n\to+\infty}\int_0^T |\partial\mathcal{W}|(\tilde{\mu}^n(t))dt \geq \int_0^T |\partial\mathcal{W}|(\tilde{\mu}(t))dt.$$

Proof. By using the representation formula proven in Proposition 2.2, we have to prove that

$$\liminf_{n\to+\infty}\int_0^T \int_{\mathbb{R}^d} |\kappa^n(x,t)|^2 d\mu^n(t)(x)dt \geq \int_0^T \int_{\mathbb{R}^d} |\kappa(x,t)|^2 d\mu(t)(x)dt,$$

where

$$\kappa^n(x,t) := \partial^0 W * \mu^n(x,t), \qquad \kappa(x,t) := \partial^0 W * \mu(x,t).$$

As a byproduct of [2, Theorem 5.4.4] on the measure space $X := \mathbb{R}^d \times [0,T]$ with the family of measures $\mu^n \otimes dt$, we get the desired assertion once we prove that

$$\int_0^T \int_{\mathbb{R}^d} |\kappa^n(x,t)|^2 d\mu^n(x)dt \quad \text{is uniformly bounded} \qquad (2.26)$$

and that κ^n converges *weakly* to κ, i.e. that for any vector field $\phi \in C_c^\infty(\mathbb{R}^d \times [0,T];\mathbb{R}^d)$

$$\int_0^T \int_{\mathbb{R}^d} \phi(x,t)\cdot\kappa^n(x,t)d\mu^n(t)(x) \to \int_0^T \int_{\mathbb{R}^d} \phi(x,t)\cdot\kappa(x,t)d\mu(t)(x) \qquad (2.27)$$

as $n \to +\infty$. We start by showing (2.27). The term on the left-hand side is given by

$$\int_0^T \int_{\mathbb{R}^d} \phi(x,t)\cdot\kappa^n(x,t)\,d\mu^n(t)(x) = \int_0^T \int_{x\neq y} \phi(x,t)\cdot\nabla W(x-y)\,d\mu^n(t)(y)\,d\mu^n(t)(x)\,dt$$

$$= \frac{1}{2}\int_0^T \int_{x\neq y} (\phi(x,t)-\phi(y,t))\cdot\nabla W(x-y)\,d\mu^n(t)(y)\,d\mu^n(t)(x)dt.$$

By [2, Lemma 3.2.2], the sequence μ_n has uniformly bounded second moments. Therefore, thanks to the linear growth control on the gradient of W in (1.5), the function $(\phi(x,t) - \phi(y,t))\cdot\nabla W(x-y)$ is uniformly integrable with respect to $\mu^n(t)\otimes\mu^n(t)\otimes dt$, and we easily recover (2.27) by weak convergence arguments.

Thanks to (1.4) and the fact that the sequence μ^n has uniformly bounded second moments, $\mathcal{W}[\mu^n]$ is bounded from below. This fact together with (2.25) trivially implies (2.26). $\qquad\square$

We are now ready to complete the proof of the existence of a solution to (1.2)–(1.3) in the sense of Definition 2.1.

Theorem 2.8 (Existence of curves of maximal slope). *Let W satisfy the assumptions* **(NL0)-(NL4)**. *Then, there exists at least one curve of maximal slope for the functional* $\mathcal{W}$, *i.e. there exists at least one curve* $\mu \in AC_{loc}\left([0, +\infty); \mathcal{P}_2(\mathbb{R}^d)\right)$ *such that the energy inequality*

$$\mathcal{W}[\mu_0] \geq \frac{1}{2} \int_0^T \|v(t)\|_{L^2(\mu(t))}^2 \, dt$$
$$+ \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} \left| \int_{x \neq y} \nabla W(x - y) d\mu(t)(y) \right|^2 d\mu(t)(x) \, dt + \mathcal{W}[\mu(T)], \quad (2.28)$$

is satisfied, where $v(t) \in L^2(\mu(t))$ *is the minimal velocity field associated to* μ.

Proof. We want to prove that the curve $\mu(t)$ provided by Proposition 2.6 satisfies the desired condition. As a consequence of (2.25) and of Lemma 2.7, if we show that

$$\liminf_{n \to \infty} \frac{1}{2} \int_0^T \|v^n(t)\|_{L^2(\mu^n(t))}^2 \, dt + \mathcal{W}[\mu^n(T)] \geq \frac{1}{2} \int_0^T \|v(t)\|_{L^2(\mu^n(t))}^2 \, dt + \mathcal{W}[\mu(T)], \quad (2.29)$$

all the remaining parts of the proof of the convergence of the scheme to a solution goes through like in the case when $\mathcal{W}$ is lower semicontinuous with respect to the narrow topology, see [2, Chapter 3].

To prove the inequality (2.29), having in mind the representation formula (2.7) linking μ^n and v^n, we regularize the solutions of $\partial_t \mu^n(t) + \text{div}(v^n(t)\mu^n(t)) = 0$ and $\partial_t \mu(t) + \text{div}(v(t)\mu(t)) = 0$ as follows:

$$v^{n,\varepsilon}(t) := \frac{(v^n(t)\mu^n(t)) * \eta_\varepsilon}{\mu^n(t) * \eta_\varepsilon}, \qquad \mu^{n,\varepsilon}(t) := \mu^n(t) * \eta_\varepsilon,$$
$$v^\varepsilon(t) := \frac{(v(t)\mu(t)) * \eta_\varepsilon}{\mu(t) * \eta_\varepsilon}, \qquad \mu^\varepsilon(t) := \mu(t) * \eta_\varepsilon,$$

where $\eta_\varepsilon = \frac{1}{\varepsilon^d} \eta\left(\frac{\cdot}{\varepsilon}\right) \in C^\infty(\mathbb{R}^d)$ is a smooth convolution kernel with support the whole $\mathbb{R}^d$, say a gaussian. Applying [2, Proposition 8.1.8] we deduce that the measures $\mu^{n,\varepsilon}(t)$, $\mu^\varepsilon(t)$ are given by the formula $\mu^{n,\varepsilon}(t) = (X^{n,\varepsilon}(t))_{\#}\mu_0$ and $\mu^\varepsilon(t) = (X^\varepsilon(t))_{\#}\mu_0$, where $X^{n,\varepsilon}(t)$ and $X^\varepsilon(t)$ denote the flows of $v^{n,\varepsilon}(t)$ and $v^\varepsilon(t)$ respectively, more precisely

$$\frac{d}{dt} X^{n,\varepsilon}(t, x) = v^{n,\varepsilon}(t, X^{n,\varepsilon}(t, x)), \qquad X^{n,\varepsilon}(0, x) = x,$$
$$\frac{d}{dt} X^\varepsilon(t, x) = v^\varepsilon(t, X^\varepsilon(t, x)), \qquad X^\varepsilon(0, x) = x.$$

We now define the transport map from $\mu^\varepsilon(T)$ to $\mu^{n,\varepsilon}(T)$ as $T_n^\varepsilon := X^{n,\varepsilon}(T) \circ (X^\varepsilon(T))^{-1}$. We have

$$d_W^2(\mu^\varepsilon(T), \mu^{n,\varepsilon}(T)) \leq \int_{\mathbb{R}^d} |T_n^\varepsilon(x) - x|^2 \, d\mu^\varepsilon(T)(x)$$

$$= \int_{\mathbb{R}^d} |X^{n,\varepsilon}(T) \circ (X^\varepsilon(T))^{-1}(x) - (X^\varepsilon(T))^{-1}(x) + (X^\varepsilon(T))^{-1}(x) - x|^2 \, d\mu(T)(x)$$

$$= \int_{\mathbb{R}^d} \left| \int_0^T \left[v^{n,\varepsilon}(t, X^{n,\varepsilon}(t) \circ (X^\varepsilon(T))^{-1}(x)) - v^\varepsilon(t, X^\varepsilon(t) \circ (X^\varepsilon(T))^{-1}(x)) \right] dt \right|^2 d\mu^\varepsilon(T)(x)$$

$$= \int_{\mathbb{R}^d} \left| \int_0^T \left[v^{n,\varepsilon}(t, X^{n,\varepsilon}(t, x)) - v^\varepsilon(t, X^\varepsilon(t, x)) \right] dt \right|^2 d\mu_0^\varepsilon(x)$$

By Hölder's inequality and expanding the squares, we get

$$d_W^2(\mu^\varepsilon(T), \mu^{n,\varepsilon}(T)) \leq T \int_{\mathbb{R}^d} \int_0^T \left| v^{n,\varepsilon}(t, X^{n,\varepsilon}(t, x)) - v^\varepsilon(t, X^\varepsilon(t, x)) \right|^2 dt \, d\mu_0^\varepsilon(x)$$

$$\leq T \int_{\mathbb{R}^d} \int_0^T |v^{n,\varepsilon}(t, x)|^2 \, d\mu^{n,\varepsilon}(t)(x) \, dt + T \int_{\mathbb{R}^d} \int_0^T |v^\varepsilon(t, x)|^2 \, d\mu^\varepsilon(t)(x) \, dt$$

$$- 2T \int_{\mathbb{R}^d} \int_0^T v^{n,\varepsilon}(t, X^{n,\varepsilon}(t, x)) \cdot v^\varepsilon(t, X^\varepsilon(t, x)) \, dt \, d\mu_0(x). \tag{2.30}$$

Thanks to [2, Lemma 8.1.10] we have

$$\int_{\mathbb{R}^d} \int_0^T |v^{n,\varepsilon}(t, x)|^2 \, d\mu^{n,\varepsilon}(t)(x) \, dt \leq \int_{\mathbb{R}^d} \int_0^T |v^n(t, x)|^2 \, d\mu^n(t)(x) \, dt \qquad \forall \, \varepsilon > 0. \tag{2.31}$$

Moreover, thanks to the weak convergence of $(\mu^n(t), v^n(t)\mu^n(t))$ to $(\mu(t), v(t)\mu(t))$, which is a consequence of the linear growth control of the gradient of W in (1.5) and the fact that $\mu^{n,\varepsilon}(t)$ and $\mu^\varepsilon(t)$ are uniformly (in $n \in \mathbb{N}$) bounded away from zero on compact sets of $\mathbb{R}^d$, we deduce that

$$v^{n,\varepsilon}(t) \to v^\varepsilon(t) \quad \text{in} \quad L^1([0, T], C_{loc}^\infty(\mathbb{R}^d)). \tag{2.32}$$

Indeed,

$$D^\alpha[v^{n,\varepsilon} - v^\varepsilon] = \frac{D^\alpha \eta^\varepsilon * (v^n \mu^n)}{\mu^{n,\varepsilon}} - \frac{D^\alpha \eta^\varepsilon * (v\mu)}{\mu^\varepsilon}$$

$$= D^\alpha \eta^\varepsilon * (v^n \mu^n) \left(\frac{\mu^\varepsilon - \mu^{n,\varepsilon}}{\mu^\varepsilon \mu^{n,\varepsilon}} \right) + \frac{1}{\mu^\varepsilon} D^\alpha \eta^\varepsilon * (v\mu - v^n \mu^n)$$

and v^n is uniformly bounded in $L^2(\mu^n)$ with respect to n. Since the flows $X^{n,\varepsilon}(t)$ and $X^\varepsilon(t)$ are globally defined (see for instance [2, Proposition 8.1.8]), (2.32) easily implies that for any $t \in [0, T]$

$$X^{n,\varepsilon}(t) \to X^\varepsilon(t) \quad \text{locally uniformly on compact subsets of } \mathbb{R}^d. \tag{2.33}$$

This fact, together with the fact that $v^{n,\varepsilon}(t, X^{n,\varepsilon}(t))$ are uniformly bounded in $L^2(\mu_0 \otimes dt)$ thanks to (2.31), implies that

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} \int_0^T v^{n,\varepsilon}(t, X^{n,\varepsilon}(t,x)) \cdot v^\varepsilon(t, X^\varepsilon(t,x)) \, dt \, d\mu_0(x)$$

$$= \int_{\mathbb{R}^d} \int_0^T |v^\varepsilon(t, X^\varepsilon(t,x))|^2 \, dt \, d\mu_0(x) = \int_{\mathbb{R}^d} \int_0^T |v^\varepsilon(t,x)|^2 \, d\mu_0(x) \, dt. \qquad (2.34)$$

To prove (2.34), split the integral on the left-hand side as follows

$$\int_{\mathbb{R}^d} \int_0^T v^{n,\varepsilon}(t, X^{n,\varepsilon}(t,x)) \cdot v^\varepsilon(t, X^\varepsilon(t,x)) \, dt \, d\mu_0(x)$$

$$= \int_{|x|>R} \int_0^T v^{n,\varepsilon}(t, X^{n,\varepsilon}(t,x)) \cdot v^\varepsilon(t, X^\varepsilon(t,x)) \, dt \, d\mu_0(x)$$

$$+ \int_{|x|\le R} \int_0^T v^{n,\varepsilon}(t, X^{n,\varepsilon}(t,x)) \cdot v^\varepsilon(t, X^\varepsilon(t,x)) \, dt \, d\mu_0(x) =: I_1 + I_2.$$

Now, thanks to (2.31) and the fact that v^n is uniformly bounded in $L^2(\mu^n)$ with respect to n, we can estimate

$$I_1^2 \le \int_{\mathbb{R}^d} \int_0^T |v^{n,\varepsilon}(t, X^{n,\varepsilon}(t,x))|^2 \, dt \, d\mu_0(x) \int_{|x|>R} \int_0^T |v^\varepsilon(t, X^\varepsilon(t,x))|^2 \, dt \, d\mu_0(x)$$

$$\le C \int_{|x|>R} \int_0^T |v^\varepsilon(t, X^\varepsilon(t,x))|^2 \, dt \, d\mu_0(x)$$

for some constant C independent on n. Hence, one can choose R large enough such that $|I_1| < \eta$ for an arbitrarily small $\eta > 0$. On the other hand, (2.32) and (2.33) imply

$$I_2 \to \int_{|x|<R} \int_0^T |v^\varepsilon(t,x)|^2 \, d\mu_0(x) \, dt$$

as $n \to +\infty$, and (2.34) follows by letting $R \to +\infty$.

Therefore, by combining (2.34) with (2.30) and (2.31) we obtain

$$\liminf_{n\to\infty} d_W^2(\mu^\varepsilon(T), \mu^{n,\varepsilon}(T)) + 2T\mathcal{W}[\mu^n(T)] \qquad (2.35)$$

$$\le \liminf_{n\to\infty} T \left[\int_{\mathbb{R}^d} \int_0^T |v^n(t,x)|^2 \, d\mu^n(t)(x) \, dt - \int_{\mathbb{R}^d} \int_0^T |v^\varepsilon(t,x)|^2 \, d\mu^\varepsilon(t)(x) \, dt + 2\mathcal{W}[\mu^n(T)] \right].$$

We now claim that there exists a constant $C_0 > 0$, depending only on the convolution kernel η, such that for any $\mu \in \mathcal{P}(\mathbb{R}^d)$

$$d_W^2(\mu, \mu * \eta_\varepsilon) \le C_0 \varepsilon^2. \qquad (2.36)$$

Indeed is suffices to consider the transport plan $\gamma^\varepsilon \in \Gamma(\mu, \mu * \eta_\varepsilon)$ defined has

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} f(x,y) \, d\gamma^\varepsilon(x,y) := \int_{\mathbb{R}^d \times \mathbb{R}^d} f(x,y) \eta_\varepsilon(y-x) \, dy \, d\mu(x) \qquad \forall f \in C_b(\mathbb{R}^d \times \mathbb{R}^d),$$

to get that

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} |y-x|^2 \, d\gamma^\varepsilon(x,y) = \int_{\mathbb{R}^d} |z|^2 \eta_\varepsilon(z) \, dz = \varepsilon^2 \int_{\mathbb{R}^d} |z|^2 \eta(z) \, dz,$$

which proves (2.36). We finally observe that

$$\liminf_{\varepsilon\to 0}\int_{\mathbb{R}^d}\int_0^T |v^\varepsilon(t,x)|^2\,d\mu^\varepsilon(t)(x)\,dt \geq \int_{\mathbb{R}^d}\int_0^T |v(t,x)|^2\,d\mu(t)(x)\,dt \qquad (2.37)$$

(actually using (2.31) one could prove that the above liminf is a limit, and equality holds). Combining (2.35) with (2.36) we obtain

$$\liminf_{n\to\infty} d_W^2(\mu(T),\mu^n(T)) + 2T\mathcal{W}[\mu^n(T)] \leq \liminf_{n\to\infty} T\left[\int_{\mathbb{R}^d}\int_0^T |v^n(t,x)|^2\,d\mu^n(t)\,dt\right.$$

$$\left. -\int_{\mathbb{R}^d}\int_0^T |v^\varepsilon(t,x)|^2\,d\mu^\varepsilon(t)(x)\,dt + 2\mathcal{W}[\mu^n(T)]\right] + O(\varepsilon),$$

so, that letting $\varepsilon\to 0$, thanks to (2.37) we finally get

$$\liminf_{n\to\infty} d_W^2(\mu(T),\mu^n(T)) + 2T\mathcal{W}[\mu^n(T)] \qquad (2.38)$$

$$\leq \liminf_{n\to\infty} T\left[2\mathcal{W}[\mu^n(T)] + \int_{\mathbb{R}^d}\int_0^T |v^n(t,x)|^2\,d\mu^n(t)(x)\,dt - \int_{\mathbb{R}^d}\int_0^T |v(t,x)|^2\,d\mu(t)(x)\,dt\right].$$

Then, in view of Lemma 2.3, we deduce that

$$\liminf_{n\to\infty} d_W^2(\mu(T),\mu^n(T)) + 2T\mathcal{W}[\mu^n(T)] \geq 2T\mathcal{W}[\mu(T)] \qquad (2.39)$$

holds for T small enough. From (2.39) combined with (2.38) we obtain that (2.29) holds provided T is sufficiently small (but independent on the initial datum μ_0), and this allows to prove the existence of a curve of maximal slope on a small time interval $[0,T]$. Iterating now the construction via minimizing movements on $[T,2T]$, $[2T,3T]$ and so on, and adding the energy inequalities (2.28) on each time interval, we get the desired result. $\qquad\square$

Remark 2.9 (λ-Generalised convexity)**.** Let us emphasize that our functional is not only λ-convex with respect to Wasserstein geodesics, but also with respect to generalized geodesics (see [2, Definition 9.2.4]). It follows directly from [2, Remark 9.2.6]. Exploiting this fact, the existence of solutions for the discrete scheme and the convergence of the scheme follow from the theory developed in [2]. On the other hand, we believe that our strategy to show the convergence of the discrete scheme is of interest in itself, being much more flexible and since it may be applied to more general situations where λ-convexity fails. This is why we chose to prove existence of curves of maximal slope in this way.

Remark 2.10 (The ODE system)**.** It is straightforward to verify that C^1 solutions of the ODE system (for the time intervals that such exist)

$$\dot{x}_i = -\sum_{j\neq i} m_j \nabla W(x_i - x_j), \qquad i = 1,\ldots,N \qquad (2.40)$$

with $m_i > 0$ for $i = 1,\ldots,N$ and $\sum_i m_i = 1$, are solutions, in the distributional sense, as in Definition 1.2, of the PDE

$$\frac{\partial\mu}{\partial t} = \mathrm{div}[(\nabla W * \mu)\mu],$$

with $\mu(t) = \sum_{i=1}^N m_i \delta_{x_i(t)}$. Conversely if $\mu(t)$ of this form solves the PDE, and x_i are C^1 for $i = 1,\ldots,N$ then x_i solve the ODE system.

The question is what happens if the particles collide: can the solutions of the PDE still be represented by an ODE? To give an affirmative answer to this question we amend the ODE system (2.40). We consider absolutely continuous solutions of

$$\dot{x}_i = - \sum_{j \in C(i)} m_j \nabla W(x_i - x_j), \qquad i = 1, \dots, N \tag{2.41}$$

$$\text{where } C(i) := \{j \in \{1, \dots, N\} \; : \; j \neq i, \, x_j(t) \neq x_i(t)\} \tag{2.42}$$

More precisely we consider the solutions of the associated integral equation. If $C(i)$ is empty, then all particles have collapsed to a single particle. We then define the right hand side to be zero, that is we define the sum over empty set of indexes to be zero. The right hand side of this ODE system is bounded and Lipschitz-continuous in space on short time intervals. Thus the ODE system has a unique Lipschitz-continuous solutions on short time intervals. The estimate (1.5) then implies that the solutions are global-in-time. Note that the solutions are Lipschitz (in time) on bounded time intervals. Also note that collisions of particles can occur, but that we do not relabel the particles when they collide. Since the number of particles is N there exist $0 \leq k \leq N - 1$ times $0 =: T_0 < T_1 < T_2 < \cdots < T_k < \infty =: T_{k+1}$ at which collisions occur. Note that $\mu(t) = \sum_{i=1}^{N} m_i \delta_{x_i(t)}$ is a solution of the PDE on the time intervals $[T_l, T_{l+1})$. Furthermore, the Lipschitz continuity of x_i implies that μ is an absolutely continuous curve in $\mathcal{P}_2(\mathbb{R}^d)$. It is then straightforward to verify that μ is a weak solution according to Definition 1.2. Since the solution to the PDE is unique the converse claim also holds.

Let us mention that while above we did not relabel the particles after collisions, at times it is useful to do so. That is on time intervals $[T_l, T_{l+1})$ the ODE system (2.41) is equivalent to

$$\frac{d\tilde{x}_i}{dt} = -\sum_{j \neq i} \tilde{m}_j \nabla W(\tilde{x}_l - \tilde{x}_j), \qquad i = 1, \dots, N_l \tag{2.43}$$

where N_l is the number of distinct particles on the time interval $[T_l, T_{l+1})$, and $\tilde{x}_j$, $\tilde{m}_j$ are their locations and masses, respectively.

Remark 2.11 (Existence of minimizing movements when ∇W is unbounded)**.** We remark here that the construction of the JKO scheme up to the proof of the Proposition 2.6 can be performed even in case ∇W has a singular behavior such as $W(x) = |x|^\alpha$ for $\alpha \in (0, 1)$ (although in this case we are not able to characterize the subdifferential). Therefore, one can easily prove that there exist at least one minimizing movement for such a kind of functional. Note that the case $\alpha = 0$ is critical, since one recovers the logarithmic kernel $W(x) = \log|x|$ as $\alpha \to 0$, for which it is an open problem how to define unique global-in-time weak measure solutions for all initial masses, see [37].

2.3. **Gradient Flow Solutions.** In this subsection, we will show the existence of global-in-time weak measure solutions for (1.2) for locally attractive potentials as a consequence of the general abstract theorems proved in [2]. In fact, using that the potential is λ-convex by **(NL1)**, Lemma 2.3 and the existence of minimizers in Proposition 2.5, we meet the hypotheses of [2, Theorem 11.1.3]. This abstract theorem shows that curves of maximal slope are equivalent under certain hypotheses to gradient flows. As a direct consequence of the existence of curves of maximal slope in Theorem 2.8, we can assert the following result. Let us remark that Proposition 2.2 has played a key role in the argument leading

to Theorem 2.8 in two ways: allowing to show the lower semicontinuity of the slope to get the energy inequalities, and in order to identify the limiting velocity field.

Theorem 2.12 (Existence of the Gradient Flow). *Given any $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$, then there exists a gradient flow solution, i.e. a curve $\mu \in AC_{loc}([0, \infty); \mathcal{P}_2(\mathbb{R}^d))$ satisfying*

$$\frac{\partial \mu(t)}{\partial t} + \operatorname{div}(v(t)\mu(t)) = 0 \ \ in \ \mathcal{D}'([0, \infty) \times \mathbb{R}^d),$$

$$v(t) = -\partial^0 \mathcal{W}[\mu(t)] = -\partial^0 W * \mu(t),$$

$$\|v(t)\|_{L^2(\mu(t))} = |\mu'|(t) \quad a.e. \ t > 0,$$

with $\mu(0) = \mu_0$. Moreover, the energy identity

$$\int_a^b \int_{\mathbb{R}^d} |v(t, x)|^2 \, d\mu(t)(x) \, dt + \mathcal{W}[\mu(b)] = \mathcal{W}[\mu(a)]$$

holds for all $0 \leq a \leq b < \infty$.

To summarize, the notions of curves of maximal slope and gradient flow solutions are equivalent and they imply the notion of weak measure solutions in the sense of Definition 1.2. Furthermore, absolute continuity of the curve of weak measure solutions and the characterization of the subdifferential imply that weak measure solutions are gradient flow solutions, see [2, Sections 8.3 and 8.4]. Thus, the three notions of solution are equivalent. The λ-geodesic convexity of the functional plays a crucial role for the uniqueness of gradient flow solutions. The following result follows readily from [2, Theorem 11.1.4].

Theorem 2.13 (d_W-Contraction). *Given two gradient flow solutions $\mu^1(t)$ and $\mu^2(t)$ in the sense of the theorem above, we have*

$$d_W(\mu^1(t), \mu^2(t)) \leq e^{-\lambda t} d_W(\mu_0^1, \mu_0^2)$$

for all $t \geq 0$. In particular, we have a unique gradient flow solution for any given $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$. Moreover, the gradient flow solution is characterized by a system of evolution variational inequalities:

$$\frac{1}{2}\frac{d}{dt} d_W^2(\mu(t), \sigma) + \frac{\lambda}{2} d_W^2(\mu(t), \sigma) \leq \mathcal{W}[\sigma] - \mathcal{W}[\mu(t)] \quad a.e. \ t > 0,$$

for all $\sigma \in \mathcal{P}_2(\mathbb{R}^d)$.

With this we have completed the existence, uniqueness and stability for gradient flow solutions for locally attractive potentials and smooth potentials with less restrictive growth conditions at infinity than in [2].

Remark 2.14 (Case $\lambda > 0$). All the theory and results obtained in this section can be applied in the case of uniform convex non-smooth potentials, i.e. allowing for $\lambda > 0$ in **(NL1)**, provided we restrict ourselves to measures with equal initial center of mass. This relies on the fact that the interaction potential in this case is λ-geodesically convex in the set of probability measures with fixed center of mass [30, 18, 42, 2].

Remark 2.15 (Comparison with classical PDE arguments). Let us observe that a more classical strategy to construct weak measure solutions is based on approximating the initial datum by atomic measures, i.e. showing the convergence of the particle method. More precisely, one exploits the existence of solutions for the discrete particle system in Remark 2.10 and the stability result in Theorem 2.13 to show convergence of the discrete

approximating solutions to a limit curve. In this way, everything reduces to prove that the limit curve is a weak measure solution to (1.2), which is however not completely trivial, and would require some work. Moreover, it is not clear how to show directly that the weak measure solutions constructed in this way are both gradient flow solutions and curves of maximal slope, and that they satisfy the energy identity. This kind of strategy is well-known in kinetic theory, see for instance [33, 38].

3. Particle measures in the JKO scheme

In this section, we show that the JKO scheme preserves the atomic part of the initial datum for all times, provided the time step is small enough. In particular, if we start with N-particles measure, it remains so, possibly with less particles, for all times. As a consequence, this immediately identifies the limit solution of the JKO scheme in this particular case. Moreover, it shows the well-posedness of a particle numerical scheme for solving numerically (1.2). More comments on this will be given below. Throughout this section, we allow λ to be positive, in which case our statements are stronger.

Given $\overline{\mu} \in \mathcal{P}_2(\mathbb{R}^d)$ let, for $\tau > 0$,

$$F_\tau[\mu] := \mathcal{W}[\mu] + \frac{1}{2\tau} d_W^2(\overline{\mu}, \mu). \tag{3.1}$$

Let us denote $u^- := \max\{0, -u\}$. We show that during a sufficiently small step of the JKO scheme, the mass contained in a particle remains concentrated, regardless of what the rest of the state looks like.

Definition 3.1 (atomization). *Given* $\mu \in \mathcal{P}_1(\mathbb{R}^d)$, μ^* *stands for the point mass located at the center of mass of* μ, *i.e:*

$$\mu^* := \delta_z \quad where \quad z = \int_{\mathbb{R}^d} x \, d\mu(x)$$

We say that μ^* *is the atomization of* μ.

Theorem 3.2. *Assume* W *satisfies* **(NL0)-(NL1)**. *Let* $\overline{\mu} = m\delta_a + \mu_r \in \mathcal{P}_2(\mathbb{R}^d)$ *with* $0 < m \leq 1$ *and* $\delta_a \perp \mu_r$. *Given any* $\tau > 0$ *such that* $\tau\lambda^- < 1$, *let*

$$\mu \in \operatorname{argmin}_{\nu \in \mathcal{P}_2(\mathbb{R}^d)} F_\tau[\nu], \tag{3.2}$$

and denote by π *an optimal transportation plan between* $\overline{\mu}$ *and* μ. *Let us define*

$$\mu_1(E) := \frac{1}{m} \pi(\{a\} \times E), \tag{3.3}$$

for any Borel set E. *Then* $\mu_1 = \mu_1^*$. *In particular,*

$$\mu = m\delta_z + \mu_s$$

for some $z \in \mathbb{R}^d$ *and* μ_s *a nonnegative measure.*

To rephrase the statement of the theorem in plain language: Any optimal transportation plan from the present state $\overline{\mu}$ to a minimizer of the JKO step μ carries all the mass from the particle at a to another point z. Thus the updated state has a particle at z, whose mass is at least the same as the one of the particle which was in a.

In case the measure $\overline{\mu}$ is a sum of N particles, by applying Theorem 3.2 to each particle, we easily conclude that μ is still a sum of particles, possibly less than N.

Corollary 3.3 (Particles remain particles). *Assume W satisfies* **(NL0)-(NL1)**. *Let* $\overline{\mu} = \sum_{i=1}^{N} m_i \delta_{x_i}$, *where* $x_1, \ldots, x_N$ *are distinct points in* $\mathbb{R}^d$, $\sum_{i=1}^{N} m_i = 1$ *and* $m_i \in (0,1)$. *Given any* $\tau > 0$ *such that* $\tau \lambda^- < 1$, *let*

$$\mu \in \operatorname{argmin}_{\nu \in \mathcal{P}_2(\mathbb{R}^d)} F_\tau[\nu].$$

Then there exist $y_1, \ldots, y_N \in \mathbb{R}^d$, *not necessarily distinct, such that* $\mu = \sum_{i=1}^{N} m_i \delta_{y_i}$.

To prove Theorem 3.2, given a minimizer μ of the JKO step, we show that

$$\mu_{new} := m\mu_1^* + (\mu - m\mu_1) \tag{3.4}$$

decreases the JKO functional:

$$F_\tau[\mu_{new}] < F_\tau[\mu], \qquad \text{if } \mu_1 \neq \mu_1^*. \tag{3.5}$$

This implies $\mu = \mu_{new}$. To prove (3.5) we examine what effect does atomizing μ_1 have on the terms in the JKO functional: the energy and the Wasserstein distance. We show that, as expected, atomizing decreases the Wasserstein distance. On the other hand atomizing can increase the interaction energy, but only if λ is negative. The key observation is that in each of the terms the change is controlled by the variance of μ_1. Taking the time step small enough allows us to conclude.

Lemma 3.4. *Assume* $\mathcal{W}$ *satisfies* **(NL0)-(NL1)**. *Let* $\nu_1, \nu_2 \in \mathcal{P}_2(\mathbb{R}^d)$ *and* $\nu = m_1\nu_1 + m_2\nu_2$ *with* $0 \leq m_1 \leq 1$ *and* $m_2 = 1 - m_1$. *Let* $\nu_{new} := m_1\nu_1^* + m_2\nu_2$. *Then*

$$\mathcal{W}[\nu] - \mathcal{W}[\nu_{new}] \geq \frac{\lambda}{2} m_1 Var(\nu_1)$$

Proof. Introduce the symmetric bilinear form

$$B(\eta_1, \eta_2) = \frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} W(x - y) \, d\eta_1(x) \, d\eta_2(y)$$

so that

$$\mathcal{W}[\nu] - \mathcal{W}[\nu_{new}] = B(m_1\nu_1 + m_2\nu_2, m_1\nu_1 + m_2\nu_2) - B(m_1\nu_1^* + m_2\nu_2, m_1\nu_1^* + m_2\nu_2)$$

$$= 2m_1m_2 B(\nu_1, \nu_2) + m_1^2 B(\nu_1, \nu_1) - 2m_1m_2 B(\nu_1^*, \nu_2) - m_1^2 B(\nu_1^*, \nu_1^*)$$

$$= m_1m_2 \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} [W(x - y) - W(z_1 - y)] \, d\nu_1(x) \, d\nu_2(y)$$

$$+ \frac{m_1^2}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} [W(x - y) - W(z_1 - z_1)] \, d\nu_1(x) \, d\nu_1(y),$$

where z_1 is the center of mass of ν_1. By the λ-convexity assumption **(NL1)**, for each $p \in \mathbb{R}^d$ and $r_p \in \partial W(p) \neq \emptyset$ the inequality

$$W(q) \geq W(p) + r_p \cdot (q - p) + \frac{\lambda}{2}|q - p|^2$$

holds for all $q \in \mathbb{R}^d$. So for $y \in \mathbb{R}^d$ there exist $r_{z_1-y} \in \partial W(z_1 - y)$ and $r_0 \in \partial W(0)$ such that

$$W(x - y) - W(z_1 - y) \geq r_{z_1-y} \cdot (x - z_1) + \frac{\lambda}{2}|x - z_1|^2$$

$$W(x - y) - W(0) \geq r_0 \cdot (x - y) + \frac{\lambda}{2}|x - y|^2$$

holds for all $x, y \in \mathbb{R}^d$. Using that $\int_{\mathbb{R}^d}(x - z_1) \, d\nu_1(x) = 0$, we obtain

$$\mathcal{W}[\nu] - \mathcal{W}[\nu_{new}] \geq m_1 m_2 \frac{\lambda}{2} Var(\nu_1) + \frac{m_1^2}{2} \frac{\lambda}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |(x - z_1) - (y - z_1)|^2 \, d\nu_1(x) \, d\nu_1(y)$$

$$= m_1 m_2 \frac{\lambda}{2} Var(\nu_1) + m_1^2 \frac{\lambda}{2} Var(\nu_1)$$

$$= \frac{m_1 \lambda}{2} Var(\nu_1),$$

concluding the proof. $\qquad \square$

Lemma 3.5. *Let $\overline{\mu}$ given as in Theorem 3.2. Given any $\nu \in \mathcal{P}_2(\mathbb{R}^d)$, let π be the optimal plan between $\overline{\mu}$ and ν and ν_1 defined by (3.3). Then, being $\nu_{new} := m\nu_1^* + (\nu - m\nu_1)$,*

$$d_W^2(\overline{\mu}, \nu) - d_W^2(\overline{\mu}, \nu_{new}) \geq m \, \mathrm{Var}(\nu_1),$$

Proof. Let z be the center of mass of ν_1 (so that $\nu_1^* = \delta_z$). Denote by π_1 the restriction of π to $\{a\} \times \mathbb{R}^d$, and $\pi_2 := \pi - \pi_1$. Let $\pi_{new} := m\delta_{(a,z)} + \pi_2$. Note that π_{new} is a transportation plan between $\overline{\mu}$ and ν_{new}. Therefore,

$$d_W^2(\overline{\mu}, \nu) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |x - y|^2 \, d(\pi_1 + \pi_2)$$

$$= m \int_{\mathbb{R}^d} |y - a|^2 \, d\nu_1(y) + \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |x - y|^2 \, d\pi_2$$

$$= m \int_{\mathbb{R}^d} \left[|y - z|^2 + 2(y - z) \cdot (z - a) + |z - a|^2 \right] d\nu_1(y) + \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |x - y|^2 \, d\pi_2$$

$$= m \int_{\mathbb{R}^d} |y - z|^2 \, d\nu_1(y) + \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |x - y|^2 \, d(m\delta_{(a,z)} + \pi_2)$$

$$\geq m Var(\nu_1) + d_W^2(\overline{\mu}, \nu_{new}),$$

as desired. $\qquad \square$

Proof of Theorem 3.2. Assume that the claim does not hold, and consider μ_{new} defined by (3.4). Then, Lemmas 3.4 and 3.5 imply that

$$F_\tau[\mu] - F_\tau[\mu_{new}] = \mathcal{W}[\mu] - \mathcal{W}[\mu_{new}] + \frac{1}{2\tau} \left(d_W^2(\overline{\mu}, \mu) - d_W^2(\overline{\mu}, \mu_{new}) \right)$$

$$\geq \frac{m}{2} \left(\lambda + \frac{1}{\tau} \right) Var(\mu_1) > 0,$$

contradicting the minimality of μ. $\qquad \square$

Remark 3.6. The above property of minimizers in each step of the JKO scheme carries over to the limiting solution, thanks to the convergence of the JKO scheme towards curves of maximal slope and gradient flow solutions of Section 2, see Theorem 2.8. Therefore, solutions corresponding to initial data with a finite number of particles plus an orthogonal part remain so for all times, with a possibly decreasing number of particles in time, see also Proposition 4.6. Moreover, combining Corollary 3.3 with the convergence of the JKO scheme in Theorem 2.8 allows to recover Remark 2.10, i.e. the correspondence between solutions of the ODE system (2.40) and gradient flow solutions of (1.2) with atomic initial measures.

4. Finite-time Total Collapse and Multiple Collapse by Stability

In this section, we focus on studying the large-time asymptotics of locally attractive non-Osgood potential, i.e. potentials satisfying assumptions **(NL0)-(NL3)** and **(NL-FTBU)**. Note that the hypothesys **(NL-FTBU)** implies **(NL4)** since the origin becomes a global minimum.

We start by discussing the monotonicity assumption in **(NL-FTBU)**. If the potential satisfies $w'(0^+) > 0$, i.e. it has a Lipschitz singularity at the origin, nearby particles move towards each other with a relative speed of about $2w'(0^+)$, and thus we expect the concentration in finite time. In case **(NL-FTBUb)**, thanks to the non-Osgood condition, we do expect again concentration in finite time. In fact, in the case of a single particle subject to the potential $W(x)$, one easily checks that the particle reaches the origin in finite time. As we show in Theorem 4.4, compactly supported measures do collapse completely in finite time.

Let us start by showing that the monotonicity assumption in the hypotheses **(NL-FTBUb)** gives local information about the potential behaviour.

Lemma 4.1 (Local behavior at the singularity). *Given $\omega \in C^2((0, +\infty))$ such that $\omega'(0^+) = 0$ with $\omega''(r)$ monotone decreasing on an interval $(0, \varepsilon_0)$, then $\frac{\omega'(r)}{r}$ is decreasing in $(0, \varepsilon_0)$.*

Proof. Since ω'' is monotone decreasing on $(0, \varepsilon_0)$, for all $r \in (0, \varepsilon_0)$ there exists $s \in (0, r)$ such that $\frac{\omega'(r)}{r} = \omega''(s) \geq \omega''(r)$. This implies

$$\frac{d}{dr}\left(\frac{\omega'(r)}{r}\right) = \frac{1}{r}\left(\omega''(r) - \frac{\omega'(r)}{r}\right) \leq 0.$$

Therefore $\omega'(r)/r$ decreases in this neighborhood of 0. $\qquad\square$

Remark 4.2 (No Oscillation Condition on the potential). Let us point out that the condition of $\omega''(r)$ being monotone decreasing on an interval $(0, \varepsilon_0)$ is not too restrictive. Actually, the non-Osgood assumption, i.e. $1/\omega'$ integrable at 0, together with plain monotonicity of $\omega''(r)$ on an interval around 0, implies that ω'' is monotone decreasing on some interval around 0 (the only possibility to violate this condition would be that the second derivative oscillate wildly at 0). Let us give a quick proof. By the non-Osgood condition, we easily have

$$\lim_{\varepsilon \to 0^+} \sup_{r \leq \varepsilon} \frac{\omega'(r)}{r} = +\infty \tag{4.1}$$

Now, we use it to deduce that there exists a sequence $r_n \to 0^+$ such that $\omega'(r_n)/r_n \to +\infty$. Since by assumption $\omega'(0^+) = 0^+$, we know that $\omega'(r_n)/r_n = \omega''(s_n) \to +\infty$ for some sequence $0 < s_n \leq r_n$. But since ω'' is monotone, this implies that $\lim_{r \to 0^+} \omega''(r) = +\infty$ and that ω'' is decreasing around 0.

Let us start by showing the finite total collapse in the case of finite number of particles.

Proposition 4.3 (Finite Time Particles Collapse). *Assume W satisfies* **(NL0)-(NL3)** *and* **(NL-FTBU)**. *Given the initial datum $\mu_0 = \sum_{i=1}^N m_i \delta_{x_i^0}$ with center of mass*

$$x_c := \sum_{i=1}^N m_i x_i^0,$$

let $\mu(t)$ denote the unique gradient flow solution with $\mu(0) = \mu_0$. Set R_0 to be the largest distance from the initial particles to the center of mass:

$$R_0 := \max_{i=1,\ldots,N} |x_i^0 - x_c|.$$

Then there exists $T^* > 0$, depending only on R_0 but not on the number of particles, such that $\mu(t) = \delta_{x_c}$ for $t \geq T^*$.

Proof.   Let us define the curves $t \mapsto x_i(t)$, $i = 1,\ldots,N$ as the solution of the ODE system discussed in Remark 2.10:

$$\dot{x}_i = -\sum_{j \in C(i)} m_j \nabla W(x_i - x_j), \qquad i = 1,\ldots,N$$

where $C(i) = \{j \in \{1,\ldots,N\} \ : \ j \neq i, \ x_j(t) \neq x_i(t)\}$. Recall also that we define the sum over empty set of indexes to be zero. Then, $\mu(t) = \sum_{i=1}^N m_i \delta_{x_i(t)}$, where possibly $x_i(t) = x_j(t)$ for some $i \neq j$.

Our claim is equivalent to saying that there exists $T^* > 0$ such that $x_i(t) = x_c$ for all $t \geq T^*$ and $i = 1,\ldots,N$. Note that, due to assumption **(NL0)** the center of mass of the particles is preserved in time for the solutions of the ODE system. Since the system is translation invariant, we can assume that $x_c = 0$ without loss of generality.

We define the Lipschitz function $R(t)$ to be the distance of the furthest particle from the center of mass:

$$R(t) := \max_{i=1,\ldots,N} |x_i(t)|.$$

Recall that x_i are Lipschitz in time, and are C^1 for all but finitely many *collision times* $0 =: T_0 < T_1 < T_2 < \cdots < T_l < T_{l+1} := +\infty$.

We first compute a differential inequality for the function $R(t)$. Due to assumption **(NL-FTBU)**, for all $t \geq 0$ and all $i = 1,\ldots,N$

$$\frac{d^+ x_i}{dt} := \lim_{h \to 0^+} \frac{x_i(t+h) - x_i(t)}{h} = -\sum_{j \in C(i)} m_j \nabla W(x_i - x_j)$$

$$= -\sum_{j \in C(i)} m_j \frac{x_i - x_j}{|x_i - x_j|} w'(|x_i - x_j|). \qquad (4.2)$$

While it would have been sufficient to deal with the derivative $\frac{dx_i}{dt}$ which exists a.e., we wanted to highlight the fact that the right-hand derivative exists for all times, including the collision times. Using (4.2), we have

$$\frac{d^+}{dt} R^2(t) = \max_{\{i \, : \, x_i(t) = R(t)\}} \frac{d^+}{dt} |x_i|^2$$

$$= \max_{\{i \, : \, x_i(t) = R(t)\}} -2 \sum_{j \in C(i)} m_j \frac{(x_i - x_j) \cdot x_i}{|x_i - x_j|} w'(|x_i - x_j|). \qquad (4.3)$$

Note that since R is Lipschitz, $\frac{d}{dt} R^2$ exists a.e. and is equal to $\frac{d^+}{dt} R^2$. Observe that for any i as above, $(x_i - x_j) \cdot x_i \geq 0$ since all other particles are inside $\overline{B}(0, R(t))$. Using again assumption **(NL-FTBU)**, we have $w'(|x_i - x_j|) > 0$ and thus $\frac{d^+}{dt} R(t) \leq 0$, from which we deduce that $R(t) \leq R_0$ for all $t \geq 0$. Let us distinguish two cases:

Case (a): $w'(0^+) > 0$. Let us define

$$D := \min_{r \in [0, 2R_0]} w'(r) > 0.$$

By coming back to (4.3), we deduce that for all $t \geq 0$

$$\frac{d^+}{dt} R(t)^2 \leq \max_{\{i \,:\, x_i(t) = R(t)\}} -2D \sum_{j \in C(i)} m_j \frac{(x_i - x_j) \cdot x_i}{|x_i - x_j|}$$

$$\leq \max_{\{i \,:\, x_i(t) = R(t)\}} -\frac{D}{R(t)} \sum_{j \neq i} m_j (x_i - x_j) \cdot x_i$$

since $|x_i - x_j| \leq 2R(t)$ for $j \neq i$ and $(x_i - x_j) \cdot x_i \geq 0$. It is easy to check, using the unit total mass of the measure and that the center of mass is zero, that for any i as above

$$\sum_{j \neq i} m_j (x_i - x_j) \cdot x_i = R(t)^2.$$

Hence $\frac{d^+}{dt} R(t) \leq -D$. We conclude that the claim of the proposition holds with $T^* = R_0 / D$.

Case (b): $w'(0^+) = 0$ together with the other assumptions in **(NL-FTBUb)**. From Lemma 4.1 we deduce that there exists $\varepsilon_0 > 0$ such that $w'(r)/r$ is decreasing in $(0, \varepsilon_0)$. Without of loss of generality, we can assume that $\varepsilon_0 < \varepsilon_1$, with ε_1 as in **(NL-FTBUb)**.

Let us first show that $R(t)$ must reach values less than $\varepsilon_0/2$ in finite time. Since $R(t)$ is decreasing, it suffices to consider the case $R_0 > \varepsilon_0/2$. Fix any time such that $R(t) \geq \varepsilon_0/2$. Coming back to (4.3), we distinguish for any i such that $R(t) = |x_i(t)|$, two sets of particles: I, where $|x_i - x_j| \leq \varepsilon_0/2$, and II, where $|x_i - x_j| > \varepsilon_0/2$. For indexes in the set I we can use that $w'(r)/r$ is decreasing, while to handle the set II we define

$$D := \min_{r \in [\varepsilon_0/2, 2R_0]} w'(r) > 0.$$

Using again $|x_i - x_j| \leq 2R(t)$ for $j \neq i$ and $(x_i - x_j) \cdot x_i \geq 0$, we can write

$$\frac{d^+}{dt} R^2(t) \leq \max_{\{i \,:\, x_i(t) = R(t)\}} \left\{ -2 \frac{w'(\varepsilon_0)}{\varepsilon_0} \sum_{(I)} m_j (x_i - x_j) \cdot x_i - \frac{D}{R(t)} \sum_{(II)} m_j (x_i - x_j) \cdot x_i \right\}.$$

Thanks to $R(t) \geq \varepsilon_0/2$ and $w'(\varepsilon_0) \geq D$, we can finally conclude that

$$\frac{d^+}{dt} R^2(t) \leq \max_{\{i \,:\, x_i(t) = R(t)\}} -\frac{D}{R(t)} \sum_{j \neq i} m_j (x_i - x_j) \cdot x_i = -DR(t)$$

for all times such that $R(t) \geq \varepsilon_0/2$. Thus, there exists a time τ such that $R(t) \leq \varepsilon_0/2$ for $t \geq \tau$.

We now refine the above argument for $t \geq \tau$ using that the distance between any two particles satisfies $|x_i - x_j| \leq 2R(t) \leq \varepsilon_0$. Since $w'(r)/r$ is decreasing on $(0, \varepsilon_0)$ we deduce that for times $t \geq \tau$

$$\frac{d^+}{dt} R(t)^2 \leq \max_{\{i \,:\, x_i(t) = R(t)\}} -\frac{w'(2R(t))}{R(t)} \sum_{j \neq i} m_j (x_i - x_j) \cdot x_i = -w'(2R(t))R(t),$$

so that $\frac{d}{dt}R(t) \le -w'(2R(t))/2$ for almost all $t \ge \tau$. Using the non-Osgood condition, i.e. the integrability of $1/w'(r)$ at the origin, we conclude that $R(t) = 0$ for a certain T^* completely determined by the inequality $\frac{d}{dt}R(t) \le -w'(2R(t))/2$.

Let us remark that this proof shows that the time of total collapse of the particles to their center of mass does not depend either on the number of particles or their masses, but only on R_0. $\qquad\square$

Making use of the stability result, the convergence of the particle method, and the total collapse for finite number of particles, we deduce the second main result of this work.

Theorem 4.4 (Finite Time Total Collapse). *Assume W satisfies* **(NL0)-(NL3)** *and* **(NL-FTBU)**. *Let $\mu(t)$ denote the unique gradient flow solution starting from the probability measure μ_0 with center of mass*

$$x_c := \int_{\mathbb{R}^d} x \, d\mu_0,$$

supported in $\overline{B}(x_c, R_0)$. Then there exists T^*, depending only on R_0, such that $\mu(t) = \delta_{x_c}$ for all $t \ge T^*$.

Proof. As in the previous proposition, we can assume $x_c = 0$. Given any compactly supported measure μ_0 in $B(0, R_0)$ and any $\eta > 0$, we can find a number of particles $N = N(\eta)$, a set of positions $\{x_1^0, \dots, x_N^0\} \subset B(0, R_0)$, and masses $\{m_1, \dots, m_N\}$, such that

$$d_W \left(\mu_0, \sum_{i=1}^N m_i \delta_{x_i^0} \right) \le \eta.$$

Let us denote by $\mu_\eta(t)$ the particle solution associated to the initial datum $\mu_\eta(0) = \sum_{i=1}^N m_i \delta_{x_i^0}$.

By Proposition 4.3, there exists a time T^* independent of N such that $\mu_\eta(t) = \delta_0$ for $t \ge T^*$. Hence, by the stability result in Theorem 2.13 we obtain

$$d_W(\mu(T^*), \delta_0) = d_W(\mu(T^*), \mu_\eta(T^*)) \le e^{-\lambda T^*} d_W \left(\mu_0, \sum_{i=1}^N m_i \delta_{x_i^0} \right) \le e^{-\lambda T^*} \eta.$$

By the arbitrariness of η, we conclude that $\mu(t) = \delta_0$ for all $t \ge T^*$ as desired. $\qquad\square$

Remark 4.5 (Finite Time Total Collapse and Tail Behavior). The previous result can be generalized for measures which are not compactly supported by the following procedure. Let us consider the case in which $c_0 := \inf_{[0,+\infty)} w' = w'(0^+) > 0$. Then the proof of case (a) in Proposition 4.3 shows that, if μ_0 is supported in $B(x_c, R_0)$, then $\mu(t) = \delta_{x_c}$ for $t \ge R_0/c_0$. From this fact and the stability estimate, it is not difficult to show that for any initial datum μ_0 decaying more than exponentially at infinity (say a gaussian), $\mu(t)$ converges exponentially fast to δ_{x_c} in infinite time. Indeed, if

$$\mu_{0,R} := \frac{\mu_0 \lfloor_{B(x_c, R)}}{\mu_0(B(x_c, R))},$$

then one easily gets $d_W(\mu_0, \mu_{0,R}) \lesssim e^{-\overline{C}R}$ for any $\overline{C} > 0$, and their centers of mass x_c and $x_{c,R}$ are exponentially close too. Hence, if $\mu_R(t)$ denotes the solution starting from

$\mu_{0,R}$, then $\mu_R(R/c_0) = \delta_{x_{c,R}}$. Therefore, choosing $\overline{C} > 2|\lambda|/c_0$, we get

$$d_W(\mu(t), \delta_{x_c}) \leq d_W(\mu(t), \mu_{c_0t}(t)) + |x_c - x_{c,c_0t}| \lesssim e^{-\lambda t} d_W(\mu_0, \mu_{0,c_0t}) + e^{-\overline{C}c_0t} \lesssim e^{-\overline{C}c_0t},$$

as desired. As expected the tail behavior of the initial measure has to be fast enough to compensate the exponential growing bound in the stability when $\lambda < 0$. On the other hand, if $\lambda \geq 0$ then we do not need any assumption on the initial datum to prove convergence in infinite time, although having estimates on the tails allows to prove better rates of convergence.

The aim of the following proposition is to show that, if we start with a measure which has some atomic part, then the atoms can only increase their mass. We present a proof based on particle approximations, an alternative approach is using the JKO-scheme, via Theorem 3.2.

Proposition 4.6 (Dirac Deltas can only increase mass). *Let $\mu(t)$ denote the unique solution starting from the probability measure $\sum_{i=1}^N m_i \delta_{x_i^0} + \nu_0$, and define the curves $t \mapsto x_i(t)$, $i = 1, \ldots, N$, as the solution of the ODE*

$$\dot{x}_i(t) = -(\partial^0 W * \mu(t))(x_i(t)).$$

Then $\mu(t) \geq \sum_{i=1}^N m_i \delta_{x_i(t)}$ for all $t \geq 0$, with possibly $x_i(t) = x_j(t)$ for some $t > 0$, $i \neq j$.

Proof. This result is again an application of the result in the case of a finite number of particles, combined with the stability of solutions. Let us approximate ν_0 with a sequence $\nu_0^k = m \sum_{j=1}^k \frac{1}{k} \delta_{y_j}$, with $m := \nu_0(\mathbb{R}^d)$. Then the unique solution starting from $\sum_{i=1}^N m_i \delta_{x_i^0} + \nu_0^k$ is given by

$$\mu^k(t) = \sum_{i=1}^N m_i \delta_{x_i^k(t)} + \sum_{j=1}^k \frac{m}{k} \delta_{y_j^k(t)},$$

where $x_i^k(t)$ and $y_j^k(t)$ solve the ODE system

$$\dot{x}_i^k = - \sum_{l \neq i, \, x_l^k \neq x_i^k} m_j \nabla W(x_l^k - x_i^k) - \sum_{l, \, y_l^k \neq x_i^k} \frac{m}{k} \nabla W(y_l^k - x_i^k), \qquad i = 1, \ldots, N,$$

$$\dot{y}_j^k(t) = - \sum_{i, \, x_i^k \neq y_l^k} m_i \nabla W(x_i^k - y_j^k) - \sum_{l \neq j, \, y_l^k \neq y_j^k} \frac{m}{k} \nabla W(y_l^k - y_j^k), \qquad j = 1, \ldots, k.$$

This gives in particular

$$\mu^k(t) \geq \sum_{i=1}^k m_i \delta_{x_i^k(t)} \tag{4.4}$$

as measures, since the particles coming from the "discretization" of ν_0 can only join the fixed particles x_i but they will not split them. We now observe that the curves $t \mapsto x_i^k(t)$ are uniformly Lipschitz (locally in time). Indeed to obtain a bound on the velocity $\partial^0 W * \mu$, by (1.5) it suffices to show that the second moments of the measures $\mu^k(t)$ are uniformly bounded, locally in time. To check this, we use as test function $|x|^2$

for a general gradient flow solution $\mu(t)$ of (1.2), and exploiting the λ-convexity of W we get

$$\frac{d}{dt} \int_{\mathbb{R}^d} |x|^2 \, d\mu(t)(x) = -2 \int_{\mathbb{R}^d} x \cdot (\partial^0 W * \mu(t)) \, d\mu(t)(x)$$

$$= -\int_{x \neq y} (x - y) \cdot (\nabla W(x) - \nabla W(y)) \, d\mu(t)(y) \, d\mu(t)(x)$$

$$\leq -\lambda \int_{x \neq y} |x - y|^2 \, d\mu(t)(y) \, d\mu(t)(x) \leq 4|\lambda| \int_{\mathbb{R}^d} |x|^2 \, d\mu(t)(x).$$

Therefore, using the stability of solutions and Ascoli-Arzelà Theorem, up to a subsequence each curve $t \mapsto x_i^k(t)$ converges locally uniformly to a limit curve $t \mapsto x_i(t)$ which satisfies

$$\dot{x}_i(t) = -\int_{\mathbb{R}^d} \partial^0 W(y - x_i(t)) \, d\mu_t(x), \qquad i = 1, \ldots, N.$$

Taking the limit in the inequality (4.4) we get the desired result. $\qquad \square$

Finally, let us show that the blow up of the L^∞ norm of a solution to (1.2) may occur at a time strictly less than the time of total collapse. In order to produce such a phenomenon, we shall work again with the ODE system (2.40) and then we argue by approximation. We first show a simple argument in a particular situation. We introduce some notation following Proposition 4.3. Let us define the curves $t \mapsto x_i(t)$, $i = 1, \ldots, 2N$ as the solution of the ODE system

$$\frac{dx_i}{dt} = -\sum_{j \in C(i)} \frac{1}{2N} \nabla W(x_i - x_j), \qquad x_i(0) = x_i^0, \qquad i = 1, \ldots, 2N,$$

so that $\mu(t) = \sum_{i=1}^{2N} \frac{1}{2N} \delta_{x_i(t)}$ is a gradient flow solution to (1.2). We define $x_{c_1}(t)$ and $x_{c_2}(t)$ to be the center of masses of the first N and the last N particles respectively. Let us consider the functions

$$R_1(t) := \max_{i=1,\ldots,N} |x_i(t) - x_{c_1}(t)| \qquad \text{and} \qquad R_2(t) := \max_{i=N+1,\ldots,2N} |x_i(t) - x_{c_2}(t)|.$$

Finally, we denote by $\mu_1(t)$ and $\mu_2(t)$ the measures $\sum_{i=1}^{N} \frac{1}{2N} \delta_{x_i(t)}$ and $\sum_{i=N+1}^{2N} \frac{1}{2N} \delta_{x_i(t)}$ respectively.

Proposition 4.7 (Multiple Collapse). *Assume the potential W satisfies* **(NL0)-(NL3)**, **(NL-FTBUa)**, *and* $\lim_{x \to +\infty} w'(x) = 0$. *There exist* $r_0, d_0, T_0, T_1 > 0$ *such that if* $\max\{R_1(0), R_2(0)\} \leq r_0$ *and* $|x_{c_1}(0) - x_{c_2}(0)| \geq d_0$, *then*

$$\mu_1(t) = \delta_{x_{c_1}(t)} \neq \mu_2(t) = \delta_{x_{c_2}(t)} \quad \text{for all } T_0 \leq t < T_1.$$

Proof. The ODE system satisfied by the particles is given by

$$\frac{dx_i}{dt} = -\sum_{j \in C(i)} \frac{1}{2N} \frac{x_i - x_j}{|x_i - x_j|} w'(|x_i - x_j|), \qquad i = 1, \ldots, 2N.$$

We distinguish two sets of particles: (I) the set of the first N particles and (II) the set of last N particles. Arguing as in Proposition 4.3, we obtain

$$\frac{d^+}{dt} R_1^2(t) = \max_{\{i \, : \, |x_i(t) - x_{c_1}(t)| = R_1(t)\}} -\sum_{j \in C(i)} \frac{1}{N} \frac{(x_i - x_j) \cdot (x_i - x_{c_1})}{|x_i - x_j|} w'(|x_i - x_j|) .$$

with $(x_i - x_j) \cdot (x_i - x_{c_1}) \geq 0$ for $j = 1, \ldots, N$. Fix d_0 large enough such that $|w'(r)| \leq \frac{1}{4} w'(0^+)$ for $r \geq d_0/2$. Then, as long as $\max\{R_1(t), R_2(t)\} \leq \frac{1}{8} d_0$ and $|x_{c_1}(t) - x_{c_2}(t)| \geq \frac{3}{4} d_0$, we have that for some i for which $|x_i(t) - x_{c_1}(t)| = R_1(t)$

$$\frac{d^+}{dt} R_1^2(t) \leq -\frac{w'(0^+)}{2N \, R_1(t)} \underbrace{\sum (x_i - x_j) \cdot (x_i - x_{c_1})}_{(I)} + \frac{|x_i - x_{c_1}|}{N} \underbrace{\sum |w'(|x_i - x_j|)|}_{(II)}$$

$$\leq -\frac{w'(0^+)}{2} R_1(t) + \frac{w'(0^+)}{4} R_1(t) = -\frac{w'(0^+)}{4} R_1(t), \tag{4.5}$$

where we used

$$\sum_{(I)} (x_i - x_j) \cdot (x_i - x_{c_1}) = N R_1(t)^2.$$

By continuity in time of solutions, there exists $t_* > 0$ small enough such that $|x_{c_1}(t) - x_{c_2}(t)| \geq \frac{3}{4} d_0$ is satisfied for $0 \leq t \leq t_*$. Choosing $r_0 \leq \min\{\frac{1}{8} d_0, \frac{w'(0^+)}{8} t_*\}$ and using (4.5), we ensure that $\max\{R_1(t), R_2(t)\} \leq \frac{1}{8} d_0$ in $0 \leq t \leq t_*$ and $R_1(t_*) = 0$. Analogously, we have that $R_2(t_*) = 0$. Then, it is clear by continuity in time that we can choose $T_0 \leq t_* < T_1$ such that the statement holds. $\qquad\square$

By a more refined analysis, one could produce an analogous result in case the potential W satisfies **(NL0)-(NL4)**, **(NL-FTBUb)**, and $\lim_{x \to +\infty} w'(x) = 0$. For instance, one can explicitly construct examples of particle configurations with special symmetries where one can check by tedious computations the multiple collapse phenomena. As a consequence, we obtain the following result

Corollary 4.8. *Assume the potential W satisfies* **(NL0)-(NL3)**, **(NL-FTBU)**, *and* $\lim_{x \to +\infty} w'(x) = 0$. *Then, there exists a nonnegative function $\rho_0 \in C_c^\infty(\mathbb{R}^d)$ with unit mass and there two curves $x_{c_k}(t)$, $k = 1, 2$, and $0 < T_0 < T_1$ such that the gradient flow solution associated with the initial datum $\rho_0 \, dx$ satisfies*

$$\mu(t) = \frac{1}{2} \delta_{x_{c_1}(t)} + \frac{1}{2} \delta_{x_{c_2}(t)} \quad and \quad x_{c_1}(t) \neq x_{c_2}(t)$$

for all $t \in (T_0, T_1)$.

It is clear from the previous proof that this two particle collapse can be generalized to multiple collapse situations with as many particle collapses as we want and choosing the time ordering of their collapses in any desired manner.

5. Confinement

We prove two results on confinement for potentials that satisfy **(NL-RAD)**. The first one gives an estimate on the evolution of the radius of the support of solutions starting from compactly supported data. For potentials that are strongly attractive at infinity, that is the ones satisfying **(NL-CONF-strong)**, it implies that if the radius of the support is large, then it is decreasing. Heuristically, we can estimate the outward velocity of the particles that are the furthest from the center of mass. In particular for potentials satisfying **(NL-CONF-strong)**, if the particles are far enough from the center of mass, then the attractive effects of the particles that are on the "opposite side" of the center of mass overwhelm the repulsive effects of the nearby particles. For sufficiently attractive potentials, with no repulsive part, this leads to an upper bound on the time of collapse.

For the potentials that are weakly-attractive at infinity, that is for ones which satisfy **(NL-CONF-weak)**, we show that the support of solutions stays uniformly bounded for all times, with an explicit bound.

For variety, we prove the confinement statements using two different approaches. For the first set of results, we use the JKO scheme to approximate the solution, and establish control on the support at each step of the JKO scheme. For the weak confinement we approximate the solution by atomic measures. For such an approximation the equation reduces to the ODE system discussed in Remark 2.10. The heart of the matter is proving confinement for the ODE approximation, with bounds independent of the number of particles.

We define for $r \geq R_a$

$$\sigma(r) := \inf_{s \geq r} w'(s). \tag{5.1}$$

Lemma 5.1. *Assume that W satisfies* **(NL0)-(NL3)** *and* **(NL-RAD)**. *Let $r_0 > 2R_a$ and $\alpha = \frac{\sigma(r_0)}{8\sqrt{2}} - C_W$. Then, given a compactly supported probability measure $\overline{\mu}$ with the center of mass at x_0 such that for some $\tau > 0$*

$$\operatorname{supp} \overline{\mu} \subset \overline{B}(x_0, r_0 + \alpha\tau),$$

if $\mu \in \operatorname{argmin}_{\nu \in \mathcal{P}_2(\mathbb{R}^d)} F_\tau[\nu]$ then

$$\operatorname{supp} \mu \subset \overline{B}(x_0, r_0).$$

Here C_W is defined by (1.8) and F_τ is the JKO functional (3.1). Note that α may be negative, and the only restriction on τ is that $r_0 + \alpha\tau > 0$, since otherwise the statement has no content.

Proof. As we did previously, we can assume without loss of generality that $x_0 = 0$. Since $\mathcal{W}$ is translation invariant, μ has center of mass 0 as well. We divide the proof in two steps

Step 1: The support does not grow faster than at the maximal repulsive speed C_W. More precisely, we first show that $\operatorname{supp} \mu \subset \overline{B}(0, r_0 + (\alpha + C_W)\tau)$. We do this by constructing a measure μ_{new} for which $F_\tau(\mu_{new}) < F_\tau(\mu)$ unless the support of μ is contained in the appropriate ball above.

Consider Π_r, the projection onto the ball $\overline{B}(0, r)$,

$$\bar{x} = \Pi_r(x) := \begin{cases} x & \text{if } |x| \leq r \\ r \frac{x}{|x|} & \text{if } |x| \geq r \end{cases}$$

Consider $r \geq r_0$ and define $\mu_{new} := \Pi_{r\#}\mu$. Let π be an optimal transportation plan between $\overline{\mu}$ and μ. Consider $\pi_{new} := (I \times \Pi_r)_{\#}\pi$. Note that π_{new} is a transportation plan between $\overline{\mu}$ and μ_{new}. Thus

$$d_W^2(\overline{\mu}, \mu_{new}) \leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 \, d\pi_{new}(x, y) = \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - \bar{y}|^2 \, d\pi(x, y). \tag{5.2}$$

where the notation $\bar{y} = \Pi_r(y)$ was used for simplicity. We claim that for all $x \in B(0, r_0 + \alpha\tau)$ and all $y \in \mathbb{R}^d$

$$|x - \bar{y}|^2 \leq |x - y|^2 + 2(\alpha\tau + r_0 - r)|y - \bar{y}|.$$

The proof relies only on elementary geometric discussion of the cases: $r_0 + \alpha\tau < r$ and $r_0 + \alpha\tau > r$ with subcases $r_0 + \alpha\tau < |y|$, $r < |y| \leq r_0 + \alpha\tau$, and $|y| \leq r$. While the proof is

straightforward, it is not short, due to the need to discuss the four cases. Thus we chose to leave the details to the reader. The above inequalities imply

$$d_W^2(\overline{\mu}, \mu_{new}) \le \int_{\mathbb{R}^d \times \mathbb{R}^d} \left[|x - y|^2 + 2(\alpha\tau + r_0 - r)|y - \bar{y}| \right] \, d\pi(x, y)$$

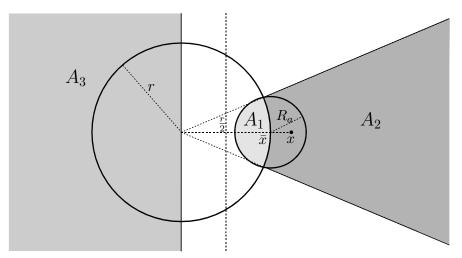$$= d_W^2(\overline{\mu}, \mu) + 2(\alpha\tau + r_0 - r) \int_{|y| > r} |y - \bar{y}| \, d\mu(y).$$

(5.3)

FIGURE 1. Relevant regions for $|x| > r$.

To estimate the change in the interaction energy we estimate the contributions from the "attractive" and "repulsive" parts:

$$\mathcal{W}[\mu_{new}] - \mathcal{W}[\mu] = \frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} W(\tilde{x} - \tilde{y}) \, d\mu_{new}(\tilde{x}) \, d\mu_{new}(\tilde{y})$$

$$- \frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} W(x - y) \, d\mu(y) \, d\mu(x)$$

$$= \int_{|x| > r} \int_{|y| \le r} [W(\bar{x} - y) - W(x - y)] \, d\mu(y) \, d\mu(x)$$

$$+ \frac{1}{2} \int_{|x| > r} \int_{|y| > r} [W(\bar{x} - \bar{y}) - W(x - y)] \, d\mu(y) \, d\mu(x)$$

$$\le \int_{|x| > r} \int_{A_1(x)} [W(\bar{x} - y) - W(x - y)] \, d\mu(y) \, d\mu(x)$$
$$\left. + \frac{1}{2} \int_{|x| > r} \int_{A_2(x)} [W(\bar{x} - \bar{y}) - W(x - y)] \, d\mu(y) \, d\mu(x) \right\} I$$

$$\left. + \frac{1}{2} \int_{|x| > r} \int_{A_3(x)} [W(\bar{x} - \bar{y}) - W(x - y)] \, d\mu(y) \, d\mu(x) \right\} II$$

$$= I + II \le I$$

(5.4)

where $A_1(x) := \{y : |y| \le r, |\bar{x} - \bar{y}| < R_a\}$, $A_2(x) := \{y : |y| > r, |\bar{x} - \bar{y}| < R_a\}$, and $A_3(x) := \{y : x \cdot y < 0\}$ are illustrated on Figure 1. Note that on the region we omitted

above $R_a \leq |\bar{x} - \bar{y}| \leq |x - y|$, so the contribution is non-positive, which allowed us to omit it.

We first estimate I. To do so observe that for $y \in A_2(x)$ if $|x - y| \leq R_a$ then by definition of C_W and (1.8),

$$W(\bar{x} - \bar{y}) - W(x - y) \leq C_W |\bar{x} - \bar{y} - (x - y)| \leq C_W(|x - \bar{x}| + |y - \bar{y}|).$$

Since W is radial and increasing outside of $B(0, R_a)$, when $y \in A_2(x)$, and $|x - y| > R_a$ then

$$W(\bar{x} - \bar{y}) - W(x - y) \leq W(\bar{x} - \bar{y}) - W(\Pi_{R_a}(x - y)) \leq C_W |\bar{x} - \bar{y} - \Pi_{R_a}(x - y)|$$
$$\leq C_W |\bar{x} - \bar{y} - (x - y)| \leq C_W(|x - \bar{x}| + |y - \bar{y}|)$$

where the inequality between the first and the second line follows from the fact that $\bar{x} - \bar{y} \in B(0, R_a)$ and $x - y \notin B(0, R_a)$. The estimate also holds for $y \in A_1(x)$ only that then $y = \bar{y}$. Therefore

$$
\begin{aligned}
I &\leq \int_{|x|>r} \int_{A_1(x)} C_W |\bar{x} - x| \, d\mu(y) \, d\mu(x) \\
&\quad + \frac{1}{2} \int_{|x|>r} \int_{A_2(x)} [C_W(|x - \bar{x}| + |y - \bar{y}|)] \, d\mu(y) \, d\mu(x) \\
&= C_W \int_{|x|>r} |x - \bar{x}| \, \mu(\{y \,:\, |\bar{x} - \bar{y}| < R_a\}) \, d\mu(x) \\
&\leq C_W \int_{|x|>r} |x - \bar{x}| \, \mu\left(\left\{y \,:\, y \cdot \frac{x}{|x|} \geq \frac{r}{2}\right\}\right) \, d\mu(x)
\end{aligned}
\tag{5.5}
$$

where in the last inequality (which we only need for the refined argument) we used that since $r \geq r_0 \geq 2R_a$, if $|\bar{x} - \bar{y}| < R_a$ then $y \cdot \frac{x}{|x|} \geq r - R_a \geq \frac{r}{2}$.

To obtain the first estimate on the radius of support of μ it is enough to note that the largest outward velocity of a particle is C_W. To obtain the appropriate statement at the level of the JKO scheme, we combine the above estimates (5.4), (5.5) and (5.3) and use that μ is a minimizer:

$$
\begin{aligned}
0 \leq F_\tau[\mu_{new}] - F_\tau[\mu] &\leq I + \frac{\alpha\tau + r_0 - r}{\tau} \int_{|x|>r} |x - \bar{x}| \, d\mu(x) \\
&\leq \left(C_W + \alpha - \frac{r - r_0}{\tau}\right) \int_{|x|>r} |x - \bar{x}| d\mu(x).
\end{aligned}
$$

Thus $\int_{|x|>r} |x - \bar{x}| d\mu(x) = 0$ if $r > r_0 + (C_W + \alpha)\tau$, which implies

$$\operatorname{supp}\mu \subset \overline{B}(0, r_0 + (C_W + \alpha)\tau). \tag{5.6}$$

Step 2: Compensation between repulsion and attraction. Since the support of μ is bounded, we can consider $\bar{r} := \inf\{s > 0 : \operatorname{supp}\mu \subseteq \overline{B}(0, s)\}$. Note that $\operatorname{supp}\mu \subseteq \overline{B}(0, \bar{r})$. Suppose that the claim of the Lemma does not hold. Then $\bar{r} > r_0$. Let $r = \frac{1}{2}(r_0 + \bar{r})$. Note that $\operatorname{supp}\mu \not\subset \overline{B}(0, r)$ but that $\operatorname{supp}\mu \subset B(0, 2r)$.

To refine the argument we need to estimate the effects of attraction, that is the term II in (5.4). We utilize the following inequality: If $y \cdot x < 0$ then, since the angle between

$x - \bar{x}$ and $\bar{y} - \bar{x}$ is greater than $3\pi/4$, $(x - \bar{x}) \cdot (\bar{x} - \bar{y}) \geq \frac{1}{\sqrt{2}} |x - \bar{x}| \, |\bar{x} - \bar{y}|$. Therefore

$$|x - \bar{y}|^2 = |(x - \bar{x}) + (\bar{x} - \bar{y})|^2$$

$$\geq |x - \bar{x}|^2 + 2\frac{1}{\sqrt{2}} |x - \bar{x}| \, |\bar{x} - \bar{y}| + |\bar{x} - \bar{y}|^2$$

$$\geq \left(\frac{1}{\sqrt{2}} |x - \bar{x}| + |\bar{x} - \bar{y}| \right)^2$$

Consequently

$$|x - y| - |\bar{x} - \bar{y}| \geq |x - \bar{y}| - |\bar{x} - \bar{y}| \geq \frac{1}{\sqrt{2}} |x - \bar{x}|.$$

We now estimate

$$II \leq -\frac{1}{2} \int_{|x|>r} \int_{y \cdot x < 0} \sigma(r)(|x - y| - |\bar{x} - \bar{y}|) \, d\mu(y) \, d\mu(x)$$

$$\leq -\frac{1}{2\sqrt{2}} \int_{|x|>r} \sigma(r)|x - \bar{x}|\mu(\{y \ : \ x \cdot y < 0\}) \, d\mu(x).$$

(5.7)

We remark that if τ is small, more precisely when $0 < \tau < \frac{r_0}{M(C_W + \alpha)}$ for some $M > 0$ large, then, due to (5.6), $\sigma(r)$ above can be replaced by $\sigma_M(r) := \min_{s \in [r, 2(1+1/M)r]} w'(s)$.

We now use that the center of mass of μ is 0 in order to show that the measure of the set which is "repulsing" the particle at $|x| > r$ is bounded by a multiple of the measure of the set which is "attracting". In particular, $\int_{\mathbb{R}^d} x \cdot y \, d\mu(y) = 0$ implies

$$- \int_{y \cdot x < 0} x \cdot y \, d\mu(y) \geq \int_{y \cdot x \geq \frac{r}{2}|x|} x \cdot y \, d\mu(y).$$

Thanks to this estimate, using that $\text{supp} \, \mu \subset \overline{B}(0, 2r)$ and $2R_a < r$, we obtain that for $|x| > r$

$$2r \, \mu(\{y \ : \ x \cdot y < 0\}) \geq \frac{r}{2} \mu \left(\left\{ y \ : \ y \cdot x \geq \frac{r}{2}|x| \right\} \right)$$

Combining this with estimates (5.4), (5.5), and (5.7) yields

$$\mathcal{W}[\mu_{new}] - \mathcal{W}[\mu] \leq \int_{|x|>r} \left(C_W - \frac{1}{8\sqrt{2}} \sigma(r) \right) |x - \bar{x}| \, \mu \left(\left\{ y \ : \ y \cdot x \geq \frac{r}{2}|x| \right\} \right) d\mu(x)$$

$$\leq \int_{|x|>r} |x - \bar{x}| \left(C_W - \frac{1}{8\sqrt{2}} \sigma(r) \right) d\mu(x).$$

The minimality of μ, the above inequality and (5.3) imply

$$0 \leq F_\tau[\mu_{new}] - F_\tau[\mu] \leq \left(C_W - \frac{1}{8\sqrt{2}} \sigma(r) + \alpha \right) \int_{|x|>r} |x - \bar{x}| d\mu(x).$$

Since $C_W - \frac{1}{8\sqrt{2}} \sigma(r) + \alpha < 0$ we conclude that $\text{supp} \, \mu \subset \overline{B}(0, r)$. Contradiction. $\square$

We use the lemma to obtain detailed upper bound on the radius of support of a solution.

Corollary 5.2. *Assume W satisfies* **(NL0)-(NL3)**, **(NL-RAD)**, *and either W is C^1 near 0 or* **(NL4)** *holds. Let $r_0 \geq 2R_a$ and define $\tilde{\sigma}(r) := \inf_{r \leq s \leq 2r} w'(s)$ for $r \geq R_a$.*

Then, given a compactly supported probability measure μ_0 with the center of mass at x_0 such that

$$\operatorname{supp} \mu_0 \subset \overline{B}(x_0, r_0),$$

the solution $\mu(t)$ to (1.2) satisfies the following estimates:

(i) *Let $r(t)$ be the solution of the ODE*

$$\begin{cases} \frac{dr}{dt} = -\frac{\tilde{\sigma}(r)}{8\sqrt{2}} + C_W \\ r(0) = r_0, \end{cases}$$

and define $\bar{r}(t) := \max\{r(t), 2R_a\}$. Then for all $t \geq 0$

$$\operatorname{supp} \mu(t) \subseteq \overline{B}(x_0, \bar{r}(t)).$$

(ii) *Special case - attractive, radial potentials: Assume further that W is radial, and such that $w(r)$ is an increasing function. Then $C_W = 0$ and*

$$\operatorname{supp} \mu(t) \subseteq \overline{B}(x_0, r(t)^+)$$

for all $t \geq 0$.

In both cases, if W satisfies the assumption (NL-CON-strong) the support stays uniformly bounded in time.

We remark that this corollary complements the results on collapse of Section 4. Furthermore analyzing the ODE $\frac{dr}{dt} = -c\tilde{\sigma}(r)$ can be used to show the finite time collapse for potentials satisfying the condition (NL-FTBU). On the other hand, the particle approximation that was used in proving the results of Section 4, can be used to prove this lemma as well.

Proof. Part (i) of the corollary with $\sigma(r)$ in place of $\tilde{\sigma}(r)$ follows easily from Lemma 5.1 as it gives the rate α, at which the support of the minimizers of the JKO scheme is changing. To conclude, we only need to recall that the minimizers of the JKO scheme converge narrowly towards $\mu(t)$, and note that uniform bounds on the support of measures are preserved in the limit. Part (ii), is just a special case, since $R_a = 0$.

We need to justify that $\tilde{\sigma}$ can be used instead of σ. This follows from the fact that, thanks to the remark below (5.7), for small τ we can improve the estimates on attractive part in Lemma 5.1 replacing σ by σ_M. Since $M > 0$ is arbitrary, letting $M \to +\infty$ we obtain the estimate with $\tilde{\sigma}$, as desired.

Note that if the assumption (NL-CON-strong) is satisfied then $\frac{dr}{dt} \leq 0$ whenever r is large enough, more precisely in $r \geq \inf\{r > 2R_a : \sigma(r) > 8\sqrt{2}C_W\}$. Thus the support stays uniformly bounded. $\qquad\square$

For potentials satisfying (NL-CONF-weak) we define for $r \geq R_a$

$$\theta(r) := \inf_{s \geq r} w'(s)\sqrt{s}.$$

Lemma 5.3 (Weak confinement for weakly attractive-at-infinity potentials). *Assume that W satisfies the conditions (NL0)-(NL3) and (NL-CONF-weak). Then, for every $R > 0$, there exists $\overline{R} \geq R$, depending only on R, C_W and W (via θ), such that the following holds: Let $x_i(t)$ be the solution of the ODE system considered in Remark 2.10*

$$\dot{x}_i = - \sum_{j \in C(i)} m_j \nabla W(x_j - x_i), \qquad i = 1, \ldots, N,$$

with $m_j > 0$, $\sum_j m_j = 1$, and $\sum_j m_j x_j(0) = 0$. If $|x_i(0)| \leq R$ for $i = 1, \ldots, N$ then

$$|x_i(t)| \leq \overline{R} \qquad \text{for all } t > 0 \text{ and } i = 1, \ldots, N.$$

Proof. The idea of the proof is as follows: Note that there are no direct energetic obstacles to prevent the support of the solution becoming large. That is the boundedness of the energy does not prevent a particle from traveling far, as long as its mass is small. However it turns out that for even a small particle to go far from the center of mass, there must exist a large mass nearby. That is for the small particle to go far, there must be particles of relatively large total mass which are "pushing" it out. However the existence of a large mass far from the center of mass does violate the fact that the energy is dissipated.

By adding a constant to W, we can assume that $W(x) \geq 0$ for all $x \in \mathbb{R}^d$. Let $\overline{R}$ be such that

$$\overline{R} > 4R_a, \ \overline{R} > R, \ \text{ and } \theta\left(\frac{\overline{R}}{8}\right) > 12\sqrt{2C_W \|W\|_{L^\infty(B(0,2R))}}. \tag{5.8}$$

Let us observe that for any $r > 2R_a$

$$w(r) \geq \frac{\sqrt{r}}{2}\theta\left(\frac{r}{2}\right). \tag{5.9}$$

This follows by noting that $w'(s) \geq \theta(r/2)/\sqrt{r}$ for all $s \in (r/2, r)$, integrating from $r/2$ to r, and using that $w(r/2) \geq 0$.

Assume that the claim of the lemma does not hold. Since x_i are C^1 for all but finitely many times, there exists $\tilde{R} \geq \overline{R}$ for which the inequalities (5.8) still hold and such that at the first time a particle reaches distance $\tilde{R}$ from the origin all of the particle trajectories are differentiable. Let t_1 be the first time that at which particle reaches the distance $\tilde{R}$ from the origin. Consider the *relabeled* ODE system (2.43) near time t_1. For notational simplicity, we keep the symbols x_i and m_i for particle positions and masses. We can assume that $|x_1(t_1)| = \tilde{R}$. Note that $\dot{x}_1(t_1) \cdot x_1(t_1) \geq 0$. Therefore

$$-\sum_{j \geq 2} m_j \nabla W(x_1(t) - x_j(t)) \cdot x_1(t) \geq 0. \tag{5.10}$$

We can assume that $x_1(t)/|x_1(t)| = e_1$. Let J_R be the set of indexes of particles that at time t_1 are repulsing x_1, that is $J_R := \{j \ : \ \nabla W(x_1(t_1) - x_j(t_1)) \cdot e_1 < 0\}$, $J_A := \{2, \ldots, N(t_1)\} \backslash J_R$, and $J_a = \{j \ : \ x_j(t_1) \cdot e_1 \leq \tilde{R}/2\}$ with J_a^c its complementary set of indices. We notice that, since $\tilde{R} > 2R_a$, $J_a \subseteq J_A$. Using that at time t_1 all particles are contained in $\overline{B}(0, \tilde{R})$, it follows that

$$\nabla W(x_1 - x_j) \cdot e_1 \geq \frac{1}{2}w'(|x_j - x_1|) \ \text{ for all } j \in J_a.$$

Furthermore, since 0 is the center of mass,

$$m_a := \sum_{j \in J_a} m_j = 1 - \sum_{j \in J_a^c} m_j =: 1 - m_a^c \geq \frac{1}{3}.$$

The argument is analogous to one in the proof of Lemma 5.1: Since the center of mass is zero, using the definition of J_a and that all particles lie inside $\overline{B}(0, \tilde{R})$, we get

$$\frac{\tilde{R}}{2}m_a^c \leq \sum_{j \in J_a^c} m_j \left(x_j(t_1) \cdot e_1\right) = -\sum_{j \in J_a} m_j \left(x_j(t_1) \cdot e_1\right) \leq \tilde{R} m_a = \tilde{R}(1 - m_a^c),$$

from which $m_a^c \leq \frac{2}{3}$. Let $m_R := \sum_{j \in J_R} m_j$. From (5.10) it follows that

$$\sum_{j \in J_R} -m_j \nabla W(x_1 - x_j) \cdot e_1 \geq \sum_{j \in J_a} m_j \nabla W(x_1 - x_j) \cdot e_1 \geq \sum_{j \in J_a} m_j \frac{1}{2} w'(|x_1 - x_j|),$$

so that

$$m_R C_W \geq \frac{1}{12\sqrt{2\tilde{R}}} \theta \left(\frac{\tilde{R}}{2} \right),$$

which implies a lower bound on m_R.

Note that energy $\mathcal{W}[\mu(0)]$ is bounded by $\frac{1}{2} \|W\|_{L^\infty(B(0,2R))}$. One the other hand, using that $\tilde{R} > 4R_a$,

$$2\mathcal{W}[\mu(t_1)] \geq \sum_{j \in J_R} \sum_{k \in J_a} m_j m_k W(x_j - x_k) \geq \frac{m_R}{3} \inf_{r \geq \tilde{R}/4} w(r).$$

Using that $\mathcal{W}[\mu(0)] \geq \mathcal{W}[\mu(t_1)]$ and (5.9), we conclude that

$$\|W\|_{L^\infty(B(0,2R))} \geq \frac{m_R}{3} \inf_{r \geq \tilde{R}/4} w(r)$$

$$\geq \frac{1}{3} \left[\frac{1}{12\sqrt{2\tilde{R}}\,C_W} \theta \left(\frac{\tilde{R}}{2} \right) \right] \left[\frac{\sqrt{\tilde{R}}}{4} \theta \left(\frac{\tilde{R}}{8} \right) \right] \geq \frac{1}{144\sqrt{2}\,C_W} \theta \left(\frac{\tilde{R}}{8} \right)^2,$$

which contradicts (5.8). □

Arguing by approximation as in the proof of Theorem 4.4, we immediately obtain the following:

Corollary 5.4. *Assume W satisfies (NL0)-(NL3), (NL-CONF-weak), and either (NL4) or W is C^1 near 0. Then, given a compactly supported probability measure μ_0 with center of mass at x_0 such that $\mathrm{supp}\,\overline{\mu} \subset \overline{B}(x_0, R)$, there exists $\overline{R} \geq R$, depending only on R, C_W and W, such that the solution $\mu(t)$ to (1.2) satisfies*

$$\mathrm{supp}\,\mu(t) \subset \overline{B}(x_0, \overline{R}) \qquad \text{for all } t \geq 0.$$

Remark 5.5. We remark that the claim of the above corollary does not hold if W does not grow at infinity (for instance, consider any W positive in $B(0,1)$ and zero outside). We also point out that the assumptions **(NL4)** or W is C^1 near 0 are needed here to be able to apply the well-posedness theory developed either in Section 2 or in [2] leading to the d_W-stability.

We conclude the paper by making the following conjecture: The result of the corollary above holds if the condition **(NL-CONF-weak)** is replaced by the assumption that w is increasing on $[R, \infty)$ for some R and

$$\lim_{r \to +\infty} w(r) = +\infty.$$

References

[1] M. AGUEH, *Existence of solutions to degenerate parabolic equations via the Monge-Kantorovich theory*, Adv. Differential Equations., 10, No 3 (2005), pp. 309-360.

[2] L. AMBROSIO, N. GIGLI AND G. SAVARÉ, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Mathematics, Birkhäuser, 2005.

[3] L. AMBROSIO AND G. SAVARÉ, *Gradient Flows of Probability Measures* in Handbook of Evolution Equations Dafermos, C. and Feireisl, E. (ed.); Vol. 3, Elsevier (2006), pp. 1-136.

[4] D. BENEDETTO, E. CAGLIOTI, M. PULVIRENTI, *A kinetic equation for granular media*, RAIRO Modél. Math. Anal. Numér., 31, (1997), 615–641.

[5] A.L. BERTOZZI AND J. BRANDMAN, *Finite-time blow-up of L^∞-weak solutions of an aggregation equation*, to appear in Comm. Math. Sci.

[6] A.L. BERTOZZI, J.A. CARRILLO AND T. LAURENT, *Blowup in multidimensional aggregation equations with mildly singular interaction kernels*, Nonlinearity, 22 (2009), pp. 683-710.

[7] A.L. BERTOZZI AND T. LAURENT, *Finite-time Blow-up of Solutions of an Aggregation Equation in $\mathbb{R}^n$*, Comm. Math. Phys., 274 (2007), pp. 717-735.

[8] A.L. BERTOZZI, T. LAURENT AND J. ROSADO, *L^p-theory for the aggregation equation*, preprint (2008).

[9] P. BILER, G. KARCH, P. LAURENÇOT, *Blowup of solutions to a diffusive aggregation model*, preprint (2009).

[10] A. BLANCHET, V. CALVEZ AND J.A. CARRILLO, *Convergence of the mass-transport steepest descent scheme for the subcritical Patlak-Keller-Segel model*, SIAM J. Numer. Anal., 46 (2008), pp. 691–721.

[11] A. BLANCHET, J.A. CARRILLO, AND N. MASMOUDI, *Infinite Time Aggregation for the Critical Patlak-Keller-Segel model in $\mathbb{R}^2$*, Comm. Pure Appl. Math., 61 (2008), pp. 1449–1481.

[12] A. BLANCHET, J. DOLBEAULT, AND B. PERTHAME, *Two-dimensional Keller-Segel model: optimal critical mass and qualitative properties of the solutions*, Electron. J. Differential Equations, (2006), No. 44, 32 pp. (electronic).

[13] M. BODNAR AND J.J.L. VELÁZQUEZ, *An integro-differential equation arising as a limit of individual cell-based models*, J. Differential Equations 222, (2006), pp. 341–380.

[14] S. BOI, V. CAPASSO AND D. MORALE, *Modeling the aggregative behavior of ants of the species Polyergus rufescens*, Spatial heterogeneity in ecological models (Alcalá de Henares, 1998), Nonlinear Anal. Real World Appl., 1 (2000), pp. 163–176.

[15] M. BURGER, V. CAPASSO AND D. MORALE, *On an aggregation model with long and short range interactions*, Nonlinear Analysis. Real World Applications. An International Multidisciplinary Journal, 8 (2007), pp. 939–958.

[16] M. BURGER AND M. DI FRANCESCO, *Large time behavior of nonlocal aggregation models with nonlinear diffusion*, Networks and Heterogeneous Media, 3, (2008), 749 - 785.

[17] J.A. CARRILLO, M.R. D'ORSOGNA AND V. PANFEROV, *Double milling in self-propelled swarms from kinetic theory*, preprint UAB (2008).

[18] J.A. CARRILLO, R.J. MCCANN AND C. VILLANI, *Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates*, Rev. Matemática Iberoamericana, 19 (2003), pp. 1-48.

[19] J.A. CARRILLO, R.J. MCCANN AND C. VILLANI, *Contractions in the 2-Wasserstein length space and thermalization of granular media*, Arch. Rat. Mech. Anal., 179 (2006), pp. 217–263.

[20] J.A. CARRILLO AND J. ROSADO, *Uniqueness of Bounded Solutions to Aggregation Equations by Optimal Transport Methods*, preprint UAB (2008).

[21] Y.-L. CHUANG, Y.R. HUANG, M.R. D'ORSOGNA AND A.L. BERTOZZI, *Multi-vehicle flocking: scalability of cooperative control algorithms using pairwise potentials*, IEEE International Conference on Robotics and Automation, 2007, pp. 2292-2299.

[22] F. Golse, *The mean-field limit for the dynamics of large particle systems*, Journées "Équations aux Dérivées Partielles", Exp. No. IX, 47, Univ. Nantes, Nantes, 2003.

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

[24] E.F. Keller and L.A. Segel, *Initiation of slide mold aggregation viewed as an instability*, J. Theor. Biol., 26 (1970).

[25] T. Laurent, *Local and Global Existence for an Aggregation Equation*, Comm. Partial Differential Equations, 32 (2007), pp. 1941-1964.

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

[27] D. Li and J. Rodrigo, *Finite-time singularities of an aggregation equation in $\mathbb{R}^n$ with fractional dissipation,* to appear in Comm. Math. Phys.

[28] D. Li and J. Rodrigo, *Refined blowup criteria and nonsymmetric blowup of an aggregation equation,* to appear in Advances in Mathematics.

[29] D. Li and X. Zhang, *On a nonlocal aggregation model with nonlinear diffusion,* preprint (2008).

[30] R.J. McCann, *A convexity principle for interacting gases*, Adv. Math. 128, (1997), pp. 153–179.

[31] A. Mogilner and L. Edelstein-Keshet, *A non-local model for a swarm*, J. Math. Bio. 38, (1999), pp. 534–570.

[32] D. Morale, V. Capasso and K. Oelschläger, *An interacting particle system modelling aggregation behavior: from individuals to populations*, J. Math. Biol., 50 (2005), pp. 49–66.

[33] H. Neunzert, *The Vlasov equation as a limit of Hamiltonian classical mechanical systems of interacting particles*, Trans. Fluid Dynamics, 18 (1977), pp. 663?678.

[34] A. Okubo, S. Levin, *Diffusion and Ecological Problems: Modern Perspectives*, Springer, Berlin, 2002.

[35] F. Otto, *The geometry of dissipative evolution equations: the porous medium equation*, Comm. Partial Differential Equations, 26 (2001) pp. 101–174.

[36] C.S. Patlak, *Random walk with persistence and external bias*, Bull. Math. Biophys., 15 (1953), pp. 311–338.

[37] F. Poupaud, *Diagonal defect measures, adhesion dynamics and Euler equation*, Methods Appl. Anal., 9 (2002), pp. 533–561.

[38] H. Spohn, *Kinetic equations from Hamiltonian dynamics: Markovian limits*, Rev. Modern Phys., 52 (1980), pp. 569–615.

[39] C.M. Topaz and A.L. Bertozzi, *Swarming patterns in a two-dimensional kinematic model for biological groups*, SIAM J. Appl. Math., 65 (2004), pp. 152–174.

[40] C.M. Topaz, A.L. Bertozzi, and M.A. Lewis, *A nonlocal continuum model for biological aggregation*, Bulletin of Mathematical Biology, 68(7), pp. 1601-1623, 2006.

[41] G. Toscani, *One-dimensional kinetic models of granular flows*, RAIRO Modél. Math. Anal. Numér., 34, 6 (2000), pp. 1277–1291.

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

[43] C. Villani, *Optimal transport, old and new*, Grundlehren der mathematischen Wissenschaften 338, Springer 2009.

[1] ICREA and Departament de Matemàtiques, Universitat Autònoma de Barcelona, E-08193 Bellaterra, Spain. E-mail: `carrillo@mat.uab.es`.

[2] Sezione di Matematica per l'Ingegneria, Dipartimento di Matematica Pura ed Applicata, Università di L'Aquila, Piazzale E. Pontieri 2, Monteluco di Roio, 67040 L'Aquila, Italy. E-mail: `difrance@univaq.it`.

[3] Centre de Mathématiques Laurent Schwartz - Ecole Polytechnique, UMR 7640, 91128 Palaiseau, France. E-mail: `figalli@math.polytechnique.fr`.

[4] Department of Mathematics, University of California - Los Angeles, Los Angeles, CA 90095, USA. E-mail: `laurent@math.ucla.edu`.

⁵ Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA. E-mail: slepcev@math.cmu.edu.