CONVERGENCE OF REGULARIZED NONLOCAL INTERACTION ENERGIES

KATY CRAIG AND IHSAN TOPALOGLU

Abstract. Inspired by numerical studies of the aggregation equation, we study the effect of regularization on nonlocal interaction energies. We consider energies defined via a repulsive-attractive interaction kernel in power-law form, regularized by convolution with a mollifier. We prove that, with respect to the 2-Wasserstein metric, the regularized energies Γ-converge to the unregularized energy and minimizers converge to minimizers. We then apply our results to prove Γ-convergence of the gradient flows, when restricted to the space of measures with bounded density.

1. Introduction

We consider the nonlocal interaction energy

$$\mathsf{E}(\mu) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} K(x-y)\, d\mu(x)d\mu(y), \tag{1.1}$$

over the space $\mathcal{P}_2(\mathbb{R}^d)$ of probability measures with finite second moment, where the pairwise interaction kernel $K : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is a radially symmetric, locally integrable, and lower semicontinuous function. Depending on the choice of K, the asymptotic states of many physical and biological systems can be characterized as minimizers of this energy. Of particular interest are interaction kernels which are repulsive at short distances and attractive at long distances. Important examples of such repulsive-attractive kernels are Morse-type potentials

$$\begin{gathered} K(x) = C_r e^{-|x|/l_r} - C_a c^{-|x|/l_a} \\ \text{with } C_a/C_r < (l_r/l_a)^d,\ 0 < l_a < l_r, \text{ and } C_a, C_r > 0, \end{gathered} \tag{1.2}$$

and potentials in the power-law form

$$K(x) = |x|^q/q - |x|^p/p \quad \text{with} \quad -d < p < q, \tag{1.3}$$

which arise in models of granular media [9, 23, 24, 40], molecular self-assembly [32, 50, 60], biological swarming [10, 57], and the distributions of eigenvalues for Gaussian random matrices [26, 48].

Due to the competition between the repulsive and attractive terms, the minimization of these energies leads to complex equilibrium configurations [4, 11–14, 16, 31, 33–35, 38, 49, 55, 56]. In the case of repulsive-attractive power law kernels (1.3), local minimizers exhibit a variety

---

Date: March 16th, 2015.

1991 *Mathematics Subject Classification.* 49J45, 82B21, 82B05, 35R09, 45K05.

Key words and phrases. global minimizers, pair potentials, aggregation, vortex blob method, Wasserstein metric, gradient flow, Γ-convergence, Coulomb repulsion.

KC's research is supported by U.S. National Science Foundation grant DMS 1401867 and IT's research is supported by a Fields–Ontario Postdoctoral Fellowship.

of qualitatively different patterns, from solid rings to broken, rounded triangles [39], and the dimension of the minimizers' support can be characterized in terms of the strength of the repulsive forces [5].

In addition to qualitative properties of local minimizers, there has also been significant interest in the existence and uniqueness of global minimizers. (Since (1.1) is translation invariant, minimizers are only unique up to translation.) Existence was recently established for a range of kernels, including (1.3), by Simione, Slepčev, and the second author [52] and by Cañizo, Carrillo, and Patacchini [18]. On the other hand, characterizing which kernels have a unique global minimizer is essentially open, due to lack of convexity. In the particular case of Coulomb repulsion and quadratic attraction (when the energy is convex with respect to H^{-1}) the unique global minimizer is the indicator function on a ball (c.f. for example [20, 27]).

Alongside the static problem of characterizing the minimizers of the interaction energy, there has also been significant interest in the dynamic problem of understanding the behavior of systems as they evolve toward a local minimizer. Such systems arise in biological swarming [45, 46, 56], robotic swarming [28, 47], and assembly of viral capsid proteins [37], and they may be modeled as *gradient flows* of the energy with respect to the 2-*Wasserstein metric*. Formally, this corresponds to the nonlinear, nonlocal partial differential equation

$$\begin{cases} \rho_t + \nabla \cdot (v\rho) = 0 \quad \text{with} \quad v = -2\nabla K * \rho, \\ \rho(x,0) = \rho_0(x), \end{cases} \tag{1.4}$$

known as the *aggregation equation*. (We choose to put a factor of two in the velocity field instead of putting a factor of one half on the energy (1.1).) For semi-convex interaction kernels K, with up to a Lipschitz singularity, weak measures solutions to (1.4) exist for all time and are unique [1, 21]. If the kernels are sufficiently convex, solutions converge exponentially quickly to a unique global minimizer [23]. However, for nonconvex kernels with merely integrable singularities, much less is known about the evolution of measure solutions and their asymptotic behavior. In the particular case of Coulomb repulsion and quadratic attraction, solutions with bounded, continuous initial data converge to the unique global minimizer algebraically in time [14].

Due to the analytical difficulties repulsive-attractive kernels present, both theoretical and applied work is often complemented by numerical simulations. The most common method for simulating (1.4) is a particle method, in which solutions are approximated by a sum of Dirac masses. For repulsive-attractive power-law kernels (1.3) in the parameter regime $2-d < p \leqslant 2,\ q > 0$, Carrillo, Choi, and Hauray prove the convergence of the particle method to weak measure solutions of the aggregation equation [19]. In recent work, Bertozzi and the first author consider a modification of this method [29], analogous to classical vortex blob methods from fluid dynamics (c.f. [3, 7, 8] and references therein). By convolving the interaction kernel K with a mollifier, the authors extend particle methods to a wide range of interaction kernels, including the power-law kernels (1.3) for $2-d \leqslant p \leqslant q$, and obtain quantitative rates of convergence to classical solutions. This convergence result is limited to bounded time intervals, though numerical results indicate that regularizing the energy in this way may also be useful in studying asymptotic behavior (see Figure 1).

Inspired by these results, we consider a regularization of the nonlocal interaction energy analogous to Bertozzi and the first author's blob method. Specifically, given a smooth, rapidly

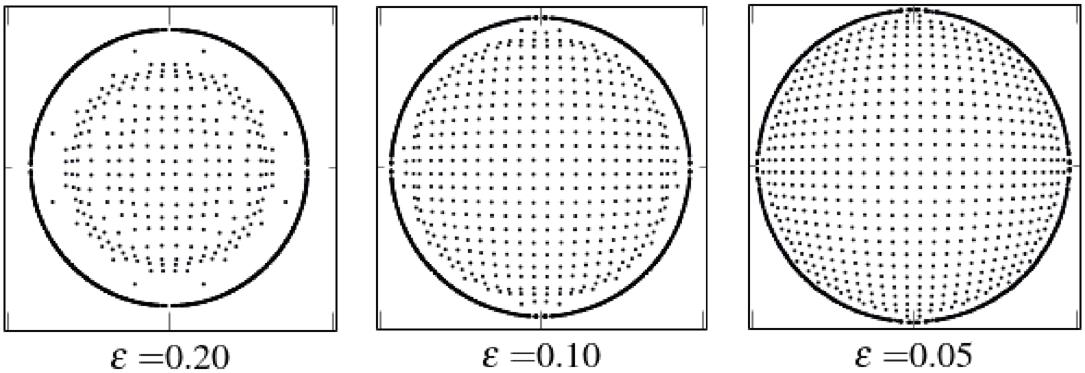

Figure 1. Numerical solutions of (1.4) at $t = 100$ for Coulomb repulsion and quadratic attraction in two dimensions. Solutions are generated by the blob method [29] with $\epsilon = 0.20, 0.10$, and 0.05 on the square $[-1,1] \times [-1,1]$, which is discretized on a 40x40 grid. The initial data is the function

$$\rho_0(x,y) = C(1 - x^2 - y^2)_+^2,$$

with C chosen so ρ_0 has mass one. As $\epsilon \to 0$, the support of the numerical solution spreads throughout the ball, suggesting that, in the limit, it converges to the indicator function on the ball, the unique steady state of the PDE.

decreasing mollifier φ, we define the regularized kernel

$$K_\epsilon := \varphi_\epsilon * K * \varphi_\epsilon,$$

with $\varphi_\epsilon(x) := \epsilon^{-d}\varphi(x/\epsilon)$, and consider the corresponding regularized interaction energy

$$\mathsf{E}_\epsilon(\mu) := \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} K_\epsilon(x-y)\, d\mu(x) d\mu(y). \tag{1.5}$$

We show that, for the repulsive-attractive power-law kernels (1.3) in the parameter regime $2-d \leqslant p < 0 < q \leqslant 2$, the regularized energies E_ϵ Γ-converge to E with respect to the 2-Wasserstein metric. Then, via a compactness argument, we show that, up to a subsequence, the minimizers of E_ϵ converge to a minimizer of E as $\epsilon \to 0$. This provides a method of approximating minimizers of E by minimizers of energies with superior convexity and regularity properties (see Remarks 2.7 and 2.8). It also imparts further theoretical justification for the success of the Bertozzi and the first author's numerical blob method.

We conclude by considering the convergence of the corresponding gradient flows as $\epsilon \to 0$. Indeed, using the abstract scheme of Serfaty [51], we show that gradient flows of E_ϵ that are bounded in $L^\infty(\mathbb{R}^d)$ Γ-converge to a generalized notion of gradient flow for E, that is to its *curve of maximal slope* on the metric space of probability measures with bounded density. This provides a link between the well-understood case of gradient flows of semi-convex energies E_ϵ and the curve of maximal slope of the unregularized energy E. It also provides a first step in understanding the connection between the gradient flows of unregularized and noncovex interaction energies E and the aggregation equation, via a singular perturbation approach (see Remark 4.1).

These results provide many directions for future work. A natural extension would be to consider the effect of simultaneously discretizing and regularizing the energy and study the distinguished limits as the discretization and regularization are removed. Another interesting direction would be to seek quantitative estimates on the rate of convergence of minimizers and gradient flows as a regularization is removed. Our present work strongly leverages compactness arguments, and we believe that a very different approach would be needed in order to provide quantitative information.

Our paper is organized as follows. In Section 2, we recall fundamental results on the 2-Wasserstein metric and the regularization of probability measures, and we use these to establish a monotonicity property of the regularized energies (1.5). In Section 3, we prove the Γ-convergence of E_ϵ to E and the corresponding convergence of minimizers. Finally, in Section 4, we introduce necessary background on curves of maximal slope and prove the Γ-convergence of these generalized gradient flows.

Throughout the paper we take $d \geqslant 3$ and $2-d \leqslant p < 0 < q \leqslant 2$. We require $q \leqslant 2$ so that the attractive term of the kernel K and its regularization are functions with at most quadratic growth, so their convolution with $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ is finitely valued (see Remark 2.9). We require $q > 0$ to obtain compactness of the sequence of minimizers corresponding to the regularized energies E_ϵ (see Remark 3.8). On the other hand, we take $2-d \leqslant p < 0$ in order to ensure monotonicity of the repulsive part of the regularized kernels with respect to ϵ, when the mollifier is the heat kernel (see Proposition 2.6).

Our results can extend to $d = 2$ by considering the energy E with the Coulomb repulsion given by the interaction kernel $\log(1/|x|)$ (corresponding to $p = 0$). When $d = 1$ and $p \geqslant 2-d$, the repulsive-attractive power-law kernels are continuous. In this case, Fellner and Raoul [33] consider regularization of the Coulomb repulsion and show that, in the presence of a confining potential, Dirac-type stationary states converge to a unique L^∞-stationary state.

Throughout we denote by $L^p(\mathbb{R}^d)$ the space of functions whose p-th power is integrable with respect to the Lebesgue measure on $\mathbb{R}^d$ and by $L^p(\mu)$ those functions whose p-th power is integrable with respect to the probability measure $\mu \in \mathcal{P}(\mathbb{R}^d)$. All constants which appear in the text are positive and may change from line to line. We express the dependence of constants to parameters by a subscript such as $C_{d,p}$ when necessary.

2. Preliminaries

2.1. **The 2-Wasserstein metric and regularization of probability measures.** We consider the energies E_ϵ and E over the space

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

of probability measures with finite second moment. We endow this space with the 2-Wasserstein metric, which we recall briefly now, along with some of its basic properties. For further background, we refer the reader to the books by Ambrosio, Gigli and Savaré [1] and Villani [59].

The 2-Wasserstein distance between $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ is

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

where $\mathcal{C}(\mu,\nu)$ is the set of transport plans between μ and ν,

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

Here π_1, π_2 denote the projections $\pi_1(x,y) = x$ and $\pi_2(x,y) = y$. The minimization problem (2.1) admits a solution, i.e., there exists an optimal transport plan $\gamma_0 \in \mathcal{C}_0(\mu,\nu)$ so that

$$d_W^2(\mu,\nu) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |x-y|^2 \, d\gamma_0(x,y).$$

Moreover, $(\mathcal{P}_2(\mathbb{R}^d), d_W)$ is a complete and separable metric space, and convergence in $(\mathcal{P}_2(\mathbb{R}^d), d_W)$ can be characterized as follows:

$$\begin{aligned} d_W(\mu_n,\mu) \to 0 \quad &\iff \quad \mu_n \to \mu \text{ weak-}* \text{ in } \mathcal{P}(\mathbb{R}^d) \text{ and } \textstyle\int_{\mathbb{R}^d} |x|^2 \, d\mu_n(x) \to \int_{\mathbb{R}^d} |x|^2 \, d\mu(x), \\ &\iff \quad \begin{array}{l} \int_{\mathbb{R}^d} f(x) \, d\mu_n(x) \to \int_{\mathbb{R}^d} f(x) \, d\mu(x), \\ \forall f \in C(\mathbb{R}^d) \text{ such that } |f(x)| \leqslant C(1+|x-x_0|^2). \end{array} \end{aligned}$$

We will refer to continuous functions satisfying $|f(x)| \leqslant C(1+|x-x_0|^2)$, for some $C > 0$ and $x_0 \in \mathbb{R}^d$, as functions with *at most quadratic growth.*

In addition to regularizing our energy functionals by convolution with a mollifier, we will also regularize our measures. We recall the definition of the convolution of a measure.

Definition 2.1. *For $\mu \in \mathcal{P}(\mathbb{R}^d)$ and $\varphi \in L^1(\mathbb{R}^d)$, $\mu * \varphi$ is defined by*

$$\int_{\mathbb{R}^d} f(x) \, d(\mu * \varphi)(x) = \int_{\mathbb{R}^d} f * \varphi(y) d\mu(y), \tag{2.2}$$

for all bounded measurable functions $f : \mathbb{R}^d \to \mathbb{R}$. Furthermore, if μ has finite second moment, then (2.2) holds for all continuous f with at most quadratic growth.

Remark 2.2. Note that $\mu * \varphi$ is absolutely continuous with respect to Lebesgue measure and $d(\mu * \varphi) = \mu * \varphi(x) \, dx$, where $\mu * \varphi(x) = \int_{\mathbb{R}^d} \varphi(x-y) \, d\mu(y)$. If, in addition, $\varphi(x) \in L^\infty(\mathbb{R}^d)$, then $\mu * \varphi(x) \in L^\infty(\mathbb{R}^d)$ and (2.2) holds for all $f \in L^1(\mathbb{R}^d)$.

We recall the following Lemma on the approximation of measures by convolution.

Lemma 2.3 (cf. [1, Lemma 7.1.10])**.** *Fix $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and a mollifier $\varphi \in C^\infty(\mathbb{R}^d)$ with finite second moment, $\varphi \geqslant 0$, and $\int_{\mathbb{R}^d} \varphi(x) dx = 1$. Then for $\varphi_\epsilon(x) := \epsilon^{-d}\varphi(x/\epsilon)$, $\mu * \varphi_\epsilon \in \mathcal{P}_2(\mathbb{R}^d)$ and*

$$d_W(\mu * \varphi_\epsilon, \mu) \leqslant \epsilon \left(\int_{\mathbb{R}^d} |x|^2 \varphi(x) \, dx \right)^{1/2}. \tag{2.3}$$

2.2. **Regularization of energies.** In what follows, we consider mollifiers which satisfy the following assumptions

(A1) $\varphi \in C^\infty(\mathbb{R}^d)$ and $\varphi(x) = \varphi(|x|) \geqslant 0$ for all $x \in \mathbb{R}^d$,

(A2) $\int_{\mathbb{R}^d} \varphi(x) \, dx = 1$,

(A3) $\varphi(x) \leqslant C|x|^{-l}$ for some $l \geqslant 2d+1$ and $C > 0$.

Remark 2.4 (Finite second moment of mollifiers). Any mollifier φ satisfying the assumptions (A1)–(A3) has finite second moment, i.e., $\int_{\mathbb{R}^d} |x|^2 \varphi(x) \, dx < +\infty$.

As a consequence of the previous results on regularization of measures in the Wasserstein metric, we have the following lemma which relates E_ϵ to E.

Lemma 2.5. *Given a mollifier φ satisfying (A1-A3), for all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$,*

$$\mathsf{E}_\epsilon(\mu) = \mathsf{E}(\mu * \varphi_\epsilon).$$

Proof. Define the right shift operator $\tau_y : f(x) \mapsto f(x-y)$. Note that for any f and g even,

$$\tau_y[f*g](x) = f*g(x-y) = \int_{\mathbb{R}^d} f(x-y-z)g(z)\,dz = \int_{\mathbb{R}^d} \tau_y[f](x-z)g(z)\,dz = \tau_y[f]*g(x).$$

Furthermore, since $f*g(x-y) = f*g(y-x)$, we also have $\tau_y[f*g](x) = \tau_x[f]*g(y)$. For the choice of powers in the regime $2-d \leqslant p < 0 < q \leqslant 2$, the interaction potential K is nonnegative. It is also radially symetric (hence, even) by definition.

Since $\varphi \in L^\infty(\mathbb{R}^d)$ and μ has finite second moment, (2.2) holds for all continuous f of at most quadratic growth and for all $f \in L^1(\mathbb{R}^d)$. As we may decompose K into the sum of a function of at most quadratic growth (away from the origin) and an integrable component (near the origin), the result then follows:

$$\begin{aligned}
\mathsf{E}_\epsilon(\mu) &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \varphi_\epsilon * K * \varphi_\epsilon(x-y)\,d\mu(x)d\mu(y) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \tau_y[\varphi_\epsilon * K]*\varphi_\epsilon(x)\,d\mu(x)d\mu(y),\\
&= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \varphi_\epsilon * \tau_y[K]*\varphi_\epsilon(x)\,d\mu(x)d\mu(y) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \varphi_\epsilon * \tau_y[K](x)\,d[\mu*\varphi_\epsilon](x)d\mu(y),\\
&= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \varphi_\epsilon * \tau_x[K](y)\,d[\mu*\varphi_\epsilon](x)d\mu(y) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} K(x-y)\,d[\mu*\varphi_\epsilon](y)d[\mu*\varphi_\epsilon](x),\\
&= \mathsf{E}(\mu*\varphi_\epsilon).
\end{aligned}$$

□

For general measures and mollifiers, there is no uniform relation between the size of $\mathsf{E}_\epsilon(\mu)$ compared to $\mathsf{E}(\mu)$. However, mollifying the interaction kernel via the heat kernel, we obtain monotonicity in ϵ of the repulsive part of the interaction energy. This generalizes a property used by Blanchet, Carlen, and Carrillo [17] for the Newtonian potential.

Let $K^a := \frac{1}{q}|x|^q$ denote the attractive part of the interaction kernel and $K^r := -\frac{1}{p}|x|^p$ the repulsive part. Similarly, let E^a denote the attractive part of the energy and E^r denote the repulsive part,

$$\mathsf{E}^a(\mu) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} K^a(x-y)\,d\mu(x)d\mu(y), \qquad \mathsf{E}^r(\mu) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} K^r(x-y)\,d\mu(x)d\mu(y). \tag{2.4}$$

Proposition 2.6 (Monotonicity of E^r_ϵ)**.** *Suppose the regularization mollifier ψ_ϵ is given by the heat kernel. Then for all $\epsilon_1 \geqslant \epsilon_2 > 0$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$*

$$\mathsf{E}^r_{\epsilon_1}(\mu) \leqslant \mathsf{E}^r_{\epsilon_2}(\mu).$$

Proof. Since $\psi_\epsilon(x) = \psi(\epsilon,x) = (4\pi\epsilon)^{-d/2}e^{-|x|^2/4\epsilon}$, $\partial_\epsilon\psi = \Delta_x\psi$ for all $\epsilon > 0$ and $x \in \mathbb{R}^d$. In the parameter regime $2-d \leqslant p < 0$, the Riesz potential

$$I_{d+p}(f)(x) = C_{d,p}\int_{\mathbb{R}^d} f(y)|x-y|^p\,dy = C_{d,p}pK^r * f$$

satisfies the identity

$$I_{d+p}(\Delta f) = \Delta I_{d+p}(f) = -I_{d+p-2}(f)$$

for any Schwartz class function f (see e.g. [53]). Moreover, $I_{d+p}(f) \geqslant 0$ for any $f \geqslant 0$.

Defining $K^r_\epsilon(x) = \psi_\epsilon * K^r * \psi_\epsilon$ we obtain

$$\begin{aligned}\frac{\partial}{\partial \epsilon} K^r_\epsilon(x) &= \psi_\epsilon * K^r * \left(\frac{\partial}{\partial \epsilon}\psi_\epsilon\right)\\ &= \psi_\epsilon * K^r * (\Delta \psi(\epsilon, x))\\ &= -C_{d+p}\, \psi_\epsilon * I_{p+d-2}(\psi_\epsilon) \leqslant 0\end{aligned}$$

since $\psi_\epsilon \geqslant 0$. Thus, K^r_ϵ is monotonically decreasing in ϵ. By definition of E^r (2.4), this gives the result. □

We close this section with brief remarks on the convexity and differentiability of the regularized energies.

Remark 2.7 (λ_ϵ-convexity of E_ϵ). Recall that a function $K : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is λ-convex if $K(x) - \frac{\lambda}{2}|x|^2$ is convex for some $\lambda \leqslant 0$. In particular, both the regularized potentials K_ϵ and their opposites $-K_\epsilon$ are λ_ϵ-convex with $\lambda_\epsilon = -C_\varphi \epsilon^{-d}$, where C_φ is a positive constant depending on φ and ϵ is sufficiently small. For an interaction energy of the form (1.1), λ-convexity of the kernel K ensures λ-convexity of the energy E with respect to the Wasserstein metric [21]. Thus, E_ϵ and $-\mathsf{E}_\epsilon$ are both λ_ϵ-convex.

Remark 2.8 (Differentiability of E_ϵ). Given a functional $\mathsf{F} : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R} \cup \{+\infty\}$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ so that $\mathsf{F}(\mu) < +\infty$, the *metric slope* of F at μ is

$$|\partial \mathsf{F}|(\mu) := \limsup_{\nu \to \mu} \frac{(\mathsf{F}(\mu) - \mathsf{F}(\nu))_+}{d_W(\mu, \nu)}.$$

For E_ϵ and $-\mathsf{E}_\epsilon$ the metric slope is well-defined (cf. [21, Proposition 2.2], [1, Lemma 10.1.5]) and

$$|\partial \mathsf{E}_\epsilon|(\mu) = 2\|\nabla K_\epsilon * \mu\|_{L^2(\mu)}.$$

Furthermore, for $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, we have $2\|\nabla K_\epsilon * \mu\|_{L^2(\mu)} \leqslant C_{\varphi,\mu}\epsilon^{1-d}$ for ϵ sufficiently small. To see this, let $0 \leqslant \eta(x) \leqslant 1$ be a smooth function satisfying $\eta(x) = 0$ if $|x| \leqslant 1$ and $\eta(x) = 1$ if $|x| \geqslant 2$. Using η, we decompose ∇K_ϵ into its singular and nonsingular components, $\nabla K_\epsilon = \nabla K_\epsilon(1-\eta) + \nabla K_\epsilon \eta = \nabla K^s_\epsilon + \nabla K^n_\epsilon$. By linearity of convolution and the triangle inequality for $L^2(\mu)$, it suffices to estimate $\|\nabla K^s_\epsilon * \mu\|_{L^2(\mu)}$ and $\|\nabla K^n_\epsilon * \mu\|_{L^2(\mu)}$. The former is bounded by $\|\nabla K^s_\epsilon * \mu\|_{L^\infty(d\mu)} \leqslant \sup_{x \in \mathbb{R}^d} |\nabla K^s_\epsilon(x)| \leqslant C_\varphi \epsilon^{1-d}$. To bound the latter, we apply Minkowski's integral inequality, the fact that ∇K^n_ϵ has at most linear growth, and the fact that μ has finite second moment to obtain

$$\|\nabla K^n_\epsilon * \mu\|_{L^2(\mu)} \leqslant \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |\nabla K^n_\epsilon(x-y)|^2 \, d\mu(y)\right)^{1/2} d\mu(x) \leqslant C_{\varphi,\mu}.$$

Remark 2.9 (Attraction powers $q > 2$). One possible way to extend our results to cover attraction powers $q > 2$ would be to consider the energy E over the space $\mathcal{P}_q(\mathbb{R}^d)$ of probability measures with finite moments up to order q, endowed with the q-Wasserstein distance

$$d_q(\mu, \nu) := \left(\min\left\{\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |x-y|^q \, d\gamma(x,y) \colon \gamma \in \mathcal{C}(\mu, \nu)\right\}\right)^{1/q},$$

where $\mathcal{C}(\mu,\nu)$ is the set of transport plans between μ and ν. Alternatively, to cover all $q>0$, one could consider the ∞-Wasserstein metric

$$d_\infty(\mu,\nu) := \inf_{\gamma\in\mathcal{C}(\mu,\nu)} \sup_{(x,y)\in \mathrm{supp}(\gamma)} |x-y|$$

over the space $\mathcal{P}_\infty(\mathbb{R}^d)$ of probability measures with finite moments of all orders. The space $(\mathcal{P}_\infty(\mathbb{R}^d), d_\infty)$ is a complete metric space [36] and has been used in the study of local minimizers for several variational problems, including nonlocal repulsive-attractive energies [5, 20, 44].

The main difficulty in extending our results to $\mathcal{P}_q(\mathbb{R}^d)$ or $\mathcal{P}_\infty(\mathbb{R}^d)$ lies in understanding the appropriate generalization of the λ-convexity to these distances and, in particular, the appropriate analogue of the HWI inequality, which plays a key role in the proof of Theorem 3.4. We leave this to future work.

3. Convergence of regularized energies and minimizers

We now turn to the proof that, up to a subsequence, minimizers of the regularized energies E_ϵ converge to a minimizer μ of E with respect to d_W. We establish this by first proving a Γ-convergence result for the sequence of energies E_ϵ and then obtaining compactness for any sequence $\{\mu_\epsilon\}_{\epsilon>0}$ when $\mathsf{E}_\epsilon(\mu_\epsilon)$ is uniformly bounded. For the latter, we use Lions' result on concentration compactness, which we recall for the readers' convenience.

Lemma 3.1 (Concentration compactness lemma for measures (cf. [41], [54, Section 4.3]))**.** *Let $\{\mu_n\}_{n\in\mathbb{N}}$ be a sequence of probability measures on $\mathbb{R}^d$. Then there exists a subsequence $\{\mu_{n_k}\}_{k\in\mathbb{N}}$ satisfying one of the three following possibilities:*

(i) (tightness up to translation) *There exists a sequence $\{y_k\}_{k\in\mathbb{N}}\subset\mathbb{R}^d$ such that for all $\epsilon>0$ there exists $R>0$ with the property that*

$$\int_{B_R(y_k)} d\mu_{n_k}(x) \geqslant 1-\epsilon \qquad \text{for all } k.$$

(ii) (vanishing) $\lim_{k\to\infty} \sup_{y\in\mathbb{R}^d} \int_{B_R(y)} d\mu_{n_k}(x) = 0$*, for all $R>0$;*

(iii) (dichotomy) *There exists $\alpha\in(0,1)$ such that for all $\epsilon>0$, there exist a number $R>0$ and a sequence $\{x_k\}_{k\in\mathbb{N}}\subset\mathbb{R}^d$ with the following property:*

Given any $R'>R$ there are nonnegative measures μ_k^1 and μ_k^2 such that

$$0\leqslant \mu_k^1+\mu_k^2 \leqslant \mu_{n_k}\,, \qquad \mathrm{supp}(\mu_k^1)\subset B_R(x_k), \qquad \mathrm{supp}(\mu_k^2)\subset \mathbb{R}^d\setminus B_{R'}(x_k)\,,$$

$$\limsup_{k\to\infty}\left(\left|\alpha-\int_{\mathbb{R}^d} d\mu_k^1(x)\right| + \left|(1-\alpha)-\int_{\mathbb{R}^d} d\mu_k^2(x)\right|\right)\leqslant \epsilon.$$

We begin our proof of the convergence of minimizers with the observation that minimizers exist for both the regularized and unregularized energies.

Proposition 3.2 (Existence of minimizers in $\mathcal{P}_2(\mathbb{R}^d)$)**.** *For all $2-d\leqslant p<0<q\leqslant 2$ and for $\epsilon>0$ sufficiently small, the energies E and E_ϵ defined by* (1.1) *and* (1.5) *via the interaction kernel* (1.3) *admit a minimizer in $\mathcal{P}_2(\mathbb{R}^d)$.*

Proof. First note that in the parameter regime $2-d \leqslant p < 0 < q \leqslant 2$ the interaction potential K, and its regularization K_ϵ are locally integrable and lower semicontinuous.

For K given by (1.3) and $x \neq 0$, we have that

$$\partial_{x_i} K(x) = x_i|x|^{q-2} - x_i|x|^{p-2} = x_i|x|^{q-2}(1-|x|^{p-q}).$$

Consequently, if $x_i > 1$, $\partial_{x_i} K(x)$ is nonnegative, and K is strictly increasing outside of the ball $|x| > 1$.

Now we show that K_ϵ is also strictly increasing outside some ball in $\mathbb{R}^d$ when $\epsilon > 0$ is sufficiently small. It suffices to show that if $x_i > 2$, then

$$|\partial_{x_i} K_\epsilon(x) - \partial_{x_i} K(x)| \leqslant C\epsilon|x|^{q-2} \tag{3.1}$$

for some constant $C > 0$. Once we have this, then for $x_i > 2$, we have $(1-|x|^{p-q}) > \tilde{C}_{p,q} > 0$ and

$$\begin{aligned}\partial_{x_i} K_\epsilon(x) &= \partial_{x_i} K(x) + \partial_{x_i} K_\epsilon(x) - \partial_{x_i} K(x) \geqslant x_i|x|^{q-2}(1-|x|^{p-q}) - C\epsilon|x|^{q-2}\\ &\geqslant 2|x|^{q-2}\tilde{C}_{p,q} - C\epsilon|x|^{q-2} = |x|^{q-2}(2\tilde{C}_{p,q} - C\epsilon).\end{aligned}$$

Choosing ϵ sufficiently small, this is nonnegative.

We now show the inequality (3.1). First, we use a cutoff function to rewrite $\partial_{x_i} K(x)$ as the sum of a compactly supported singular function $\partial_{x_i} K^s$ and a continuously differentiable function $\partial_{x_i} K^n$. Let $0 \leqslant \eta \leqslant 1$ be a smooth function satisfying $\eta(x) \equiv 0$ for $|x| < 1/4$ and $\eta(x) \equiv 1$ for $|x| > 1/2$. Write $\partial_{x_i} K = (1-\eta)\partial_{x_i} K + \eta\partial_{x_i} K =: \partial_{x_i} K^s + \partial_{x_i} K^n$. Then, for $\Phi_\epsilon(y) := \int_{\mathbb{R}^d} \varphi_\epsilon(y-z)\varphi_\epsilon(z)\,dz$,

$$\begin{aligned}&|\partial_{x_i} K_\epsilon(x) - \partial_{x_i} K(x)|\\ &\quad= \left|\int_{\mathbb{R}^d} \left(\partial_{x_i} K(x-y) - \partial_{x_i} K(x)\right)\Phi_\epsilon(y)\,dy\right|\\ &\quad\leqslant \left|\int_{\mathbb{R}^d} \left(\partial_{x_i} K^s(x-y) - \partial_{x_i} K^s(x)\right)\Phi_\epsilon(y)\,dy\right| + \left|\int_{\mathbb{R}^d} \left(\partial_{x_i} K^n(x-y) - \partial_{x_i} K^n(x)\right)\Phi_\epsilon(y)\,dy\right|\\ &\quad=: I_1 + I_2.\end{aligned}$$

The autocorrelation function Φ_ϵ satisfies $\Phi_\epsilon(y) = \epsilon^{-d}\Phi(y/\epsilon)$ for all $y \in \mathbb{R}^d$ and has the same decay property (A3) as φ_ϵ. To see this, suppose $|y| = 3R$ for some $R > 0$. Then for any $z \in \mathbb{R}^d$ we have $|z| > R$ or $|y-z| > R$; hence,

$$\begin{aligned}|\Phi_\epsilon(y)| &= \frac{1}{\epsilon^{2d}}\int_{\mathbb{R}^d} \varphi\left(\frac{y-z}{\epsilon}\right)\varphi\left(\frac{z}{\epsilon}\right)\,dz\\ &\leqslant \frac{1}{\epsilon^{2d}}\left(\int_{|z|>R} \varphi\left(\frac{y-z}{\epsilon}\right)\varphi\left(\frac{z}{\epsilon}\right)\,dz + \int_{|y-z|>R} \varphi\left(\frac{y-z}{\epsilon}\right)\varphi\left(\frac{z}{\epsilon}\right)\,dz\right)\\ &\leqslant C\epsilon^{l-2d}R^{-l} = C\epsilon^{l-2d}|y|^{-l}.\end{aligned}$$

First, we estimate I_1. If $x_i > 2$, then $|x| > 2$, so $\partial_{x_i} K^s(x) = 0$. Consequently,

$$I_1 = \left|\int_{|x-y|<1} \partial_{x_i} K^s(x-y)\Phi_\epsilon(y)\,dy\right|.$$

Since $|x-y| < 1$ and $|x| > 2$,

$$|y| \geqslant |x| - |y-x| \geqslant |x| - 1 \geqslant |x|/2.$$

Consequently,

$$|\Phi_\epsilon(y)| \leqslant C\epsilon^{l-2d}|y|^{-l} \leqslant C\,\epsilon\,|y|^{-l} \leqslant C\,2^l\,\epsilon\,|x|^{-l}$$

for $0<\epsilon<1$. Since $\partial_{x_i}K^s$ is locally integrable and $2-q<2d+1\leqslant l$, there exists $C>0$ so that for $|x|>2$,

$$I_1 \leqslant C\,\epsilon\,|x|^{-l} \leqslant C\,\epsilon\,|x|^{q-2}.$$

Now we estimate I_2. Since $\partial_{x_i}K^n$ is continuously differentiable,

$$\begin{aligned} I_2 = \left|\int_{\mathbb{R}^d}\int_0^1 \frac{d}{d\alpha}\partial_{x_i}K^n(x-\alpha y)\,d\alpha\,\Phi_\epsilon(y)\,dy\right| &= \left|\int_{\mathbb{R}^d}\int_0^1 \langle\nabla\partial_{x_i}K^n(x-\alpha y),-y\rangle\,d\alpha\,\Phi_\epsilon(y)\,dy\right| \\ \leqslant \int_{|y|\leqslant|x|/2}\int_0^1 |\nabla\partial_{x_i}K^n(x-\alpha y)|\,|y|\,d\alpha\,\Phi_\epsilon(y)\,dy &+ \int_{y>|x|/2}\int_0^1 |\nabla\partial_{x_i}K^n(x-\alpha y)|\,|y|\,d\alpha\,\Phi_\epsilon(y)\,dy. \end{aligned}$$

For $|y|\leqslant|x|/2$, $|x-\alpha y|\geqslant|x|/2>1$ for all $\alpha\in[0,1]$. Consequently, $\nabla\eta(x-\alpha y)\equiv 0$, and

$$|\nabla\partial_{x_i}K^n(x-\alpha y)| \leqslant |(\nabla\eta(x-\alpha y))\partial_{x_i}K(x-\alpha y)| + |\eta(x-\alpha y)\nabla\partial_{x_i}K(x-\alpha y)| \leqslant C|x|^{q-2}.$$

On the other hand, for all all $y\in\mathbb{R}^d$, $|\nabla\partial_{x_i}K^n(x-\alpha y)|\leqslant C$. Therefore,

$$\begin{aligned} I_2 &\leqslant C|x|^{q-2}\int_{|y|<|x|/2}|y|\,\Phi_\epsilon(y)\,dy + C\int_{|y|>|x|/2}|y|\,\Phi_\epsilon(y)\,dy \\ &= C\epsilon|x|^{q-2}\int_{|z|<|x|/(2\epsilon)}|z|\,\Phi(z)\,dz + C\epsilon\int_{|z|>|x|/(2\epsilon)}|z|\,\Phi(z)\,dz. \end{aligned}$$

The first integral is bounded by a constant since Φ satisfies decay property (A3), hence, has finite first moment. For the second integral, we use that $|z\Phi(z)|\leqslant C|z|^{-l+1}$ and $2-q<l-d-1$ to obtain

$$\int_{|z|>|x|/(2\epsilon)}|z|\,\Phi(z)\,dz \leqslant C\int_{|x|/(2\epsilon)}^{+\infty} r^{-l+1}r^{d-1}dr = C\left(\frac{|x|}{2\epsilon}\right)^{d-l+1} \leqslant C\epsilon|x|^{q-2}.$$

Therefore, (3.1) follows.

We conclude that both $K(x)$ and $K_\epsilon(x)$ are strictly increasing for $|x|>2$, and in particular, they approach infinity as $|x|\to\infty$. Thus, for both E and E_ϵ, there exists a minimizer in $\mathcal{P}(\mathbb{R}^d)$ that is compactly supported, hence belongs to $\mathcal{P}_2(\mathbb{R}^d)$ [52, Theorem 3.1] [18, Theorem 1.4]. ☐

Remark 3.3. Since $\mathcal{P}_2(\mathbb{R}^d)\subset\mathcal{P}(\mathbb{R}^d)$, the infimum of the energy E or E_ϵ over $\mathcal{P}(\mathbb{R}^d)$ is less than or equal to the infimum over $\mathcal{P}_2(\mathbb{R}^d)$. Due to the fact that minimizers have compact support [18, Lemma 2.10], the converse holds, as well. Hence,

$$\inf_{\mu\in\mathcal{P}(\mathbb{R}^d)}\mathsf{E}(\mu) = \inf_{\mu\in\mathcal{P}_2(\mathbb{R}^d)}\mathsf{E}(\mu) \quad \text{and} \quad \inf_{\mu\in\mathcal{P}(\mathbb{R}^d)}\mathsf{E}_\epsilon(\mu) = \inf_{\mu\in\mathcal{P}_2(\mathbb{R}^d)}\mathsf{E}_\epsilon(\mu)$$

for all $\epsilon>0$ sufficiently small.

Now we prove the regularized energies E_ϵ converge to E in the sense of Γ-convergence.

Theorem 3.4 (Γ-convergence of regularized energies)**.** *The sequence of regularized energies $\{\mathsf{E}_\epsilon\}_{\epsilon>0}$ defined by* (1.5) *Γ-converges to the energy E with respect to $(\mathcal{P}_2(\mathbb{R}^d),d_W)$. That is,*

(i) (Lower semicontinuity) *For any $\{\mu_\epsilon\}_{\epsilon>0}\subset\mathcal{P}_2(\mathbb{R}^d)$ and $\mu\in\mathcal{P}_2(\mathbb{R}^d)$ such that it holds $\lim_{\epsilon\to0}d_W(\mu_\epsilon,\mu)=0$, we have that*

$$\liminf_{\epsilon\to0}\mathsf{E}_\epsilon(\mu_\epsilon)\geqslant\mathsf{E}(\mu).$$

(ii) (Recovery sequence) *For any* $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ *there exists* $\{\nu_\epsilon\}_{\epsilon>0} \subset \mathcal{P}_2(\mathbb{R}^d)$ *such that*

$$\lim_{\epsilon\to 0} d_W(\nu_\epsilon, \mu) = 0 \text{ and } \lim_{\epsilon \to 0} \mathsf{E}_\epsilon(\nu_\epsilon) = \mathsf{E}(\mu).$$

Proof. We will prove this theorem in two steps.

Step 1 (Lower semicontinuity). By Lemma 2.5, we have $\mathsf{E}_\epsilon(\mu_\epsilon) = \mathsf{E}(\mu_\epsilon * \varphi_\epsilon)$. Furthermore,

$$d_W(\mu, \mu_\epsilon * \varphi_\epsilon) \leqslant d_W(\mu, \mu_\epsilon) + d_W(\mu_\epsilon, \mu_\epsilon * \varphi_\epsilon), \tag{3.2}$$

where the first term approaches zero by hypothesis and the second term approaches zero by Lemma 2.3.

As the interaction potential K is lower semicontinuous and bounded from below, the Portmanteau Theorem [58, Theorem 1.3.4] ensures the energy E is lower semicontinuous with respect to convergence in $\mathcal{P}_2(\mathbb{R}^d)$. Indeed, the Portmanteau Theorem states that the weak-* convergence of $\mu_\epsilon \times \mu_\epsilon$ to $\mu \times \mu$ is equivalent to the fact that

$$\liminf_{\epsilon\to 0} \iint_{\mathbb{R}^n\times\mathbb{R}^n} f(x-y)\, d(\mu_\epsilon \times \mu_\epsilon)(x,y) \geqslant \iint_{\mathbb{R}^n\times\mathbb{R}^n} f(x-y)\, d(\mu\times\mu)(x,y)$$

when f is lower-semicontinuous and bounded from below. Hence, we obtain

$$\liminf_{\epsilon\to 0} \mathsf{E}_\epsilon(\mu_\epsilon) = \liminf_{\epsilon \to 0} \mathsf{E}(\mu_\epsilon * \varphi_\epsilon) \geqslant \mathsf{E}(\mu).$$

Step 2 (Recovery sequence). Let $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ be arbitrary. We define a recovery sequence ν_ϵ for the measure μ by using the heat kernel. Let $\psi(x) = (4\pi)^{-d/2}\, e^{-|x|^2/4}$, and define $\psi_{\delta(\epsilon)}(x) := \delta(\epsilon)^{-d}\psi(x/\delta(\epsilon))$ where

$$\delta(\epsilon) := \epsilon^{1/2d}. \tag{3.3}$$

Clearly the function ψ satisfies the assumptions (A1)–(A3) in Subsection 2.2; hence, it is an admissible mollifier. Define the measure

$$\nu_\epsilon := \mu * \psi_{\delta(\epsilon)}.$$

Then, by Lemma 2.3, $\nu_\epsilon \in \mathcal{P}_2(\mathbb{R}^d)$ for all $\epsilon > 0$, and $d_W(\nu_\epsilon, \mu) \to 0$ as $\epsilon \to 0$.

We now show that $\lim_{\epsilon\to 0} \mathsf{E}_\epsilon(\nu_\epsilon) = \mathsf{E}(\mu)$. By Lemma 2.5,

$$\begin{aligned} |\mathsf{E}_\epsilon(\nu_\epsilon) - \mathsf{E}(\mu)| &= |\mathsf{E}(\mu * \psi_{\delta(\epsilon)} * \varphi_\epsilon) - \mathsf{E}(\mu)| = |\mathsf{E}_{\delta(\epsilon)}(\mu * \varphi_\epsilon) - \mathsf{E}(\mu)| \\ &\leqslant |\mathsf{E}_{\delta(\epsilon)}(\mu * \varphi_\epsilon) - \mathsf{E}_{\delta(\epsilon)}(\mu)| + |\mathsf{E}_{\delta(\epsilon)}(\mu) - \mathsf{E}(\mu)| =: I_1 + I_2. \end{aligned} \tag{3.4}$$

It suffices to show that $I_1, I_2 \to 0$ as $\epsilon \to 0$. We estimate I_2 first. As $\mathsf{E}_{\delta(\epsilon)}(\mu)$ is regularized using the heat kernel $\psi_{\delta(\epsilon)}$, Proposition 2.6 ensures that $\mathsf{E}^r_{\delta(\epsilon)}(\mu) \leqslant \mathsf{E}^r(\mu)$ for all $\epsilon > 0$. Hence,

$$\limsup_{\epsilon\to 0} \mathsf{E}_{\delta(\epsilon)}(\mu) \leqslant \mathsf{E}^r(\mu) + \limsup_{\epsilon\to 0} \mathsf{E}^a_{\delta(\epsilon)}(\mu).$$

Again, by Lemma 2.5, $\mathsf{E}^a_{\delta(\epsilon)}(\mu) = \mathsf{E}^a(\mu * \psi_{\delta(\epsilon)})$. Since $d_W(\mu * \psi_{\delta(\epsilon)}, \mu) \to 0$ as $\epsilon \to 0$ and the attractive interaction kernel K^a is a continuous function with at most quadratic growth,

$$\limsup_{\epsilon\to 0} \mathsf{E}^a_{\delta(\epsilon)}(\mu) = \limsup_{\epsilon \to 0} \mathsf{E}^a(\mu * \psi_{\delta(\epsilon)}) = \mathsf{E}^a(\mu).$$

Therefore

$$\limsup_{\epsilon\to 0} \mathsf{E}_{\delta(\epsilon)}(\mu) \leqslant \mathsf{E}(\mu).$$

On the other hand, by the lower semicontinuity of the energy E, as proved in the first step of the proof, we get that

$$\liminf_{\epsilon\to 0} \mathsf{E}_{\delta(\epsilon)}(\mu) = \liminf_{\epsilon\to 0} \mathsf{E}(\mu * \psi_{\delta(\epsilon)}) \geqslant \mathsf{E}(\mu);$$

hence, $I_2 = |\mathsf{E}_{\delta(\epsilon)}(\mu) - \mathsf{E}(\mu)| \to 0$ as $\epsilon \to 0$.

To estimate I_1, recall that by Remark 2.7 and 2.8 the regularized energies $\pm\mathsf{E}_{\delta(\epsilon)}$ are $\lambda_{\delta(\epsilon)}$-convex with $\lambda_{\delta(\epsilon)} = -C_\varphi \delta(\epsilon)^{-d}$ and have metric slope $\|\nabla K * \psi_{\delta(\epsilon)} * \mu\|_{L^2(\mu)}$. Therefore, they satisfy the following HWI-type inequality [1, Theorem 2.4.9],

$$|\mathsf{E}_{\delta(\epsilon)}(\mu * \varphi_\epsilon) - \mathsf{E}_{\delta(\epsilon)}(\mu)| \leqslant 2\|\nabla K * \psi_{\delta(\epsilon)} * \mu\|_{L^2(\mu)} d_W(\mu * \varphi_\epsilon, \mu) - \frac{\lambda_{\delta(\epsilon)}}{2} d_W^2(\mu * \varphi_\epsilon, \mu).$$

Consequently, for $\epsilon > 0$ sufficiently small, so that $d_W(\mu * \varphi_\epsilon, \mu) \leqslant 1$, we may use the bound $\|\nabla K * \psi_{\delta(\epsilon)} * \mu\|_{L^2(\mu)} \leqslant C_{\psi,\mu}\delta(\epsilon)^{1-d}$ from Remark 2.8 and the definition of $\delta(\epsilon)$ from equation (3.3) to obtain

$$I_1 = |\mathsf{E}_{\delta(\epsilon)}(\mu * \varphi_\epsilon) - \mathsf{E}_{\delta(\epsilon)}(\mu)| \leqslant C(\delta(\epsilon)^{1-d} + \delta(\epsilon)^{-d})\, d_W(\mu * \varphi_\epsilon, \mu) \leqslant C\epsilon\, \delta(\epsilon)^{-d} = C\epsilon^{1/2}$$

Therefore $I_1 \to 0$ as $\epsilon \to 0$, which completes the proof. □

Remark 3.5. Since the energy E is lower semicontinuous with respect to weak-* convergence of measures in $\mathcal{P}(\mathbb{R}^d)$ and convergence in $\mathcal{P}_2(\mathbb{R}^d)$ implies weak-* convergence in $\mathcal{P}(\mathbb{R}^d)$, the conclusion of Theorem 3.4(i) is also true with respect to weak-* convergence in $\mathcal{P}(\mathbb{R}^d)$.

As a result of the strong confining forces induced by the attractive part of the kernel K, we obtain the following compactness result for the sequence of energies E_ϵ in $\mathcal{P}(\mathbb{R}^d)$.

Proposition 3.6 (Compactness in $\mathcal{P}(\mathbb{R}^d)$)**.** *Let $\{\mu_\epsilon\}_{\epsilon>0} \subset \mathcal{P}(\mathbb{R}^d)$ be a sequence such that, for all $\epsilon > 0$ sufficiently small, $\mathsf{E}_\epsilon(\mu_\epsilon) \leqslant C$ for some constant $C > 0$. Then $\{\mu_\epsilon\}_{\epsilon>0}$ has a subsequence which is convergent with respect to the weak-* topology in $\mathcal{P}(\mathbb{R}^d)$.*

Proof. We will use Lemma 3.1 and argue by contradiction, as in [52]. Note that as in the proof of Proposition 3.2, $K(x)$ and $K_\epsilon(x) \to +\infty$ as $|x| \to \infty$ for $\epsilon > 0$ sufficiently small. This ensures that the energy E_ϵ exhibits long-range confinement. We will use this fact to eliminate "vanishing" and "dichotomy" possibilities of the sequence $\{\mu_\epsilon\}_{\epsilon>0}$.

Suppose a subsequence of $\{\mu_\epsilon\}_{\epsilon>0}$, which we still index by $\epsilon > 0$, "vanishes" in the sense of Lemma 3.1(ii). Then for any $\delta > 0$ and $R > 0$, there exists $\epsilon_0 > 0$ such that for all $\epsilon < \epsilon_0$ and $x \in \mathbb{R}^d$, we have

$$\mu_\epsilon(\mathbb{R}^d \setminus B_R(x)) \geqslant 1 - \delta.$$

Hence, for all $\epsilon < \epsilon_0$,

$$\iint_{|x-y|\geqslant R} d\mu_\epsilon(x) d\mu_\epsilon(y) = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d \setminus B_R(x)} d\mu_\epsilon(y) \right) d\mu_\epsilon(x) \geqslant 1 - \delta. \tag{3.5}$$

Since $K_\epsilon(x) \to +\infty$ as $|x| \to \infty$, given a constant $B > 0$, there exists $R > 0$ such that for all $|x| \geqslant \tilde{R}$, we have $K_\epsilon(x) \geqslant B$. Let $\delta < 1/2$, and choose ϵ_0 depending on δ and $\tilde{R}$ such that

(3.5) holds. Then, using (3.5) and the fact that $K_\epsilon \geqslant 0$ for all $\epsilon > 0$, we get that

$$\begin{aligned}\mathsf{E}_\epsilon(\mu_\epsilon) &= \iint_{|x-y|<\tilde{R}} K_\epsilon(x-y)\, d\mu_\epsilon(x) d\mu_\epsilon(y) + \iint_{|x-y|\geqslant\tilde{R}} K_\epsilon(x-y)\, d\mu_\epsilon(x) d\mu_\epsilon(y)\\ &\geqslant \iint_{|x-y|\geqslant\tilde{R}} K_\epsilon(x-y)\, d\mu_\epsilon(x) d\mu_\epsilon(y) \geqslant (1-\delta)B.\end{aligned}$$

This contradicts the uniform bound $\mathsf{E}_\epsilon(\mu_\epsilon) \leqslant C$ when B is sufficiently large; hence, "vanishing" does not occur.

Suppose that for a subsequence of $\{\mu_\epsilon\}_{\epsilon>0}$, which again we index by $\epsilon > 0$, "dichotomy" occurs. Then, since $K_\epsilon \geqslant 0$ for all $\epsilon > 0$,

$$\liminf_{\epsilon\to 0} \mathsf{E}_\epsilon(\mu_\epsilon) \geqslant \liminf_{\epsilon\to 0} \int_{\mathbb{R}^d\setminus B'_R(x_\epsilon)} \int_{B_R(x_\epsilon)} K_\epsilon(x-y)\, d\mu^1_\epsilon(x) d\mu^2_\epsilon(y) \geqslant \inf_{|x|\geqslant R'-R} K_\epsilon(x)\alpha(1-\alpha),$$

where $R > 0$, the sequence $\{x_\epsilon\}_{\epsilon>0}$, and the measures μ^1_ϵ and μ^2_ϵ are defined as in Lemma 3.1(iii), and $R' > R$ is arbitrary. Thus, again using the growth of K_ϵ, we get that

$$\liminf_{\epsilon\to 0} \mathsf{E}_\epsilon(\mu_\epsilon) \geqslant +\infty,$$

contradicting the uniform bound on $\mathsf{E}_\epsilon(\mu_\epsilon)$.

Therefore, Lemma 3.1 ensures the sequence $\{\mu_\epsilon\}_{\epsilon>0}$ is tight up to a translation by a sequence $\{y_\epsilon\}_{\epsilon>0} \subset \mathbb{R}^d$. However, the energies E_ϵ are translation invariant, so we can replace μ_ϵ by $\mu_\epsilon(\cdot + y_\epsilon)$ which yields a tight sequence. Therefore by Prokhorov's theorem (cf. [58, Theorem 1.3.9]), there exists a subsequence $\{\mu_\epsilon\}_{\epsilon>0}$, still indexed by $\epsilon > 0$, and a measure $\mu \in \mathcal{P}(\mathbb{R}^d)$ such that $\mu_\epsilon \to \mu$ with respect to weak-* convergence in $\mathcal{P}(\mathbb{R}^d)$ as $\epsilon \to 0$. □

Classically, if a sequence of Γ-convergent functionals also satisfies a compactness property, then, up to a subsequence, minimizers converge to a minimizer of the limiting functional. Though our compactness result, Proposition 3.6, is established in a weaker space ($\mathcal{P}(\mathbb{R}^d)$) than the topology in which the energies Γ-converge ($\mathcal{P}_2(\mathbb{R}^d)$), we still obtain the following corollary on the convergence of minimizers.

Corollary 3.7 (Convergence of minimizers)**.** *For any $\epsilon > 0$ sufficiently small let $\mu_\epsilon \in \mathcal{P}_2(\mathbb{R}^d)$ be a minimizer of the energy E_ϵ defined by* (1.5)*. Then there exists $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ such that up to a subsequence $\mu_\epsilon \to \mu$ in $\mathcal{P}_2(\mathbb{R}^d)$ as $\epsilon \to 0$, and μ minimizes the energy E over $\mathcal{P}_2(\mathbb{R}^d)$.*

Proof. Since the sequence $\{\mu_\epsilon\}_{\epsilon>0} \subset \mathcal{P}_2(\mathbb{R}^d)$ consists of minimizers of E_ϵ, for $\epsilon > 0$ sufficiently small, there exists a constant $C > 0$ so that $\mathsf{E}_\epsilon(\mu_\epsilon) \leqslant C$. Then by Proposition 3.6, there exists a measure $\mu \in \mathcal{P}(\mathbb{R}^d)$ such that a subsequence of $\{\mu_\epsilon\}_{\epsilon>0}$, which we still denote by μ_ϵ, weak-* converges to μ in $\mathcal{P}(\mathbb{R}^d)$.

To show that μ minimizes E over $\mathcal{P}_2(\mathbb{R}^d)$ we proceed in two steps. First, consider an arbitrary measure $\nu \in \mathcal{P}_2(\mathbb{R}^d)$. By Theorem 3.4 (ii), there exists a sequence $\{\nu_\epsilon\}_{\epsilon>0} \subset \mathcal{P}_2(\mathbb{R}^d)$ so that $d_W(\nu_\epsilon, \nu) \to 0$ as $\epsilon \to 0$ and

$$\lim_{\epsilon\to 0} \mathsf{E}(\nu_\epsilon) = \mathsf{E}(\nu).$$

By Theorem 3.4 (i), Remark 3.5, and the fact that the measures μ_ϵ are minimizers of E_ϵ over $\mathcal{P}_2(\mathbb{R}^d) \subset \mathcal{P}(\mathbb{R}^d)$, we obtain

$$\mathsf{E}(\mu) \leqslant \liminf_{\epsilon\to 0} \mathsf{E}_\epsilon(\mu_\epsilon) \leqslant \liminf_{\epsilon\to 0} \mathsf{E}_\epsilon(\nu_\epsilon) = \mathsf{E}(\nu). \tag{3.6}$$

Thus, $\mu \in \mathcal{P}(\mathbb{R}^d)$ has less energy than any other measure $\nu \in \mathcal{P}_2(\mathbb{R}^d)$, so by Remark 3.3, μ minimizes E over $\mathcal{P}(\mathbb{R}^d)$, as well. Consequently, [18, Lemma 2.10] ensures μ is compactly supported; hence it is in $\mathcal{P}_2(\mathbb{R}^d)$.

Finally, we show that in fact μ_ϵ converges to μ in $\mathcal{P}_2(\mathbb{R}^d)$. In [18, Lemma 2.10] the authors show that

$$\operatorname{diam}(\operatorname{supp} \mu_\epsilon) \leqslant M_\epsilon$$

where $M_\epsilon := \sqrt{d}(4r_\epsilon + (\lceil 1/m_\epsilon \rceil - 1)(4r_\epsilon + 2R_\epsilon))$. Here

$$m_\epsilon := \frac{C - \mathsf{E}_\epsilon(\mu_\epsilon)}{C - K_\epsilon^{\min}}$$

with $C > 0$ so that $\mathsf{E}_\epsilon(\mu_\epsilon) \leqslant C$ for $\epsilon > 0$ sufficiently small. $K_\epsilon^{\min}$ denotes the absolute minimum value of the interaction kernel K_ϵ, r_ϵ is chosen such that $K_\epsilon(x) \geqslant C$ for all $|x| \geqslant r_\epsilon$, and R_ϵ is the radius after which the kernel K_ϵ is strictly increasing. We denote the corresponding quantities for the unregularized energy E by removing the subscript ϵ.

We show that M_ϵ, the upper bound on the diameter of the support of μ_ϵ, is bounded by some $\tilde{M} > 0$ for all $\epsilon > 0$ sufficiently small. We have $\limsup_{\epsilon \to 0} -\mathsf{E}_\epsilon(\mu_\epsilon) \leqslant -\mathsf{E}(\mu)$ and $K_\epsilon \to K$ uniformly on compact sets away from the origin. Thus, $m_\epsilon \to m$ as $\epsilon \to 0$, so $\lceil 1/m_\epsilon \rceil \leqslant \lceil 1/m \rceil + 1$ for $\epsilon > 0$ sufficiently small.

As shown in the proof of Proposition 3.2, K_ϵ is strictly increasing for $|x| \geqslant 2$, so we may take $R_\epsilon = 2$. Likewise, by estimate (3.1), there exists $r > 2$ so that $K_\epsilon(x) \geqslant C$ for $|x| > r$. Therefore

$$M_\epsilon \leqslant \tilde{M} := \sqrt{d}(4r + (\lceil 1/m \rceil)(4r + 4))$$

for all $\epsilon > 0$ sufficiently small.

This shows that the diameter of the support of μ_ϵ is uniformly bounded by $\tilde{M}$. Consequently, the second moments of the sequence μ_ϵ are uniformly integrable

$$\lim_{k \to +\infty} \int_{\{|x|^2 \geqslant k\}} |x|^2 \, d\mu_\epsilon = \lim_{k \to +\infty} \int_{\{\sqrt{k} \leqslant |x| \leqslant \tilde{M}\}} |x|^2 \, d\mu_\epsilon = 0$$

Since μ_ϵ converges to μ with respect to the weak-* topology on $\mathcal{P}(\mathbb{R}^d)$ and μ_ϵ has uniformly integrable second moments, $\mu_\epsilon \to \mu$ in $\mathcal{P}_2(\mathbb{R}^d)$ with respect to d_W [1, Proposition 7.1.5]. □

Remark 3.8 (Negative attraction power). The condition $q > 0$ is a sufficient condition for Propositions 3.2 and 3.6. If $q < 0$, both K and K_ϵ converge to zero as $|x| \to \infty$. In [52], the authors characterized the existence of minimizers for these types of potentials using the existence of a measure for which the energy is negative. This is clearly true for E and E_ϵ, as one can consider the characteristic function of a sufficiently large ball $B_R(0)$. However, we still require $q > 0$ for the compactness result in Proposition 3.6, and we believe that it is a necessary condition, as well.

4. Convergence of gradient flows

In this section, we apply our result on the Γ-convergence of the regularized energies to show that if the Wasserstein gradient flows of E_ϵ are bounded in $L^\infty(\mathbb{R}^d)$, they converge to a suitable metric space generalization of the gradient flow of E, known as a *curve of maximal slope*.

As previously mentioned, without regularization E is not convex (or λ-convex for $\lambda < 0$). Consequently, it falls outside the scope of much of the existing theory on well-posedness of Wasserstein gradient flows [1, 21]. However, the regularized interaction energies E_ϵ are

$\lambda_\epsilon = C_\varphi \epsilon^{-d}$ convex, so that for fixed $\epsilon > 0$, their Wasserstein gradient flows exist and are unique [21]. Thus, by showing these gradient flows Γ-converge to the curve of maximal slope for the unregularized energy, we provide a link between the well-understood case of convex gradient flow and emerging results on the Wasserstein gradient flow of non-convex energies [2, 15, 22, 23, 25, 30, 42, 43].

We restrict our attention to the space of probability measures with bounded density

$$\mathcal{P}_{2,R}(\mathbb{R}^d) := \{\mu \in \mathcal{P}_{2,ac}(\mathbb{R}^d) \colon \|\mu\|_{L^\infty(\mathbb{R}^d)} \leqslant R\}, \tag{4.1}$$

for any $R > 0$, due to the fact that, though E is not convex (or λ-convex for $\lambda < 0$), once it is restricted to $\mathcal{P}_{2,R}(\mathbb{R}^d)$, it possesses a generalized notion of convexity known as *ω-convexity* [22, 23]. This notion has been used, explicitly and implicitly, in many previous works on non-convex gradient flow. For our purposes, ω-convexity allows us to define a notion of upper gradient for E, which is an essential component in defining its curve of maximal slope.

Both from the perspective of energy minimization and the dynamics of the aggregation equation (1.4), this restriction is quite natural when $p = 2 - d$. Indeed, several results in the literature support the fact that both the energy minimization and evolution take place in the space of bounded functions. For quadratic attraction (i.e., for $q = 2$), the global minimizer of E is the characteristic function on a ball [20, 27]. Likewise, from the dynamics point of view, if the initial data belongs to $\mathcal{P}_2(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ and has compact support, then it remains bounded for all time and converges to the characteristic function on the unit ball [14]. For more general attraction powers ($0 < q \leqslant 2$), all compactly supported local minimizers are bounded [20] and smooth, compactly supported classical solutions of the aggregation equation remain bounded for all time [6, Lemma 1].

For $p \neq 2 - d$, the restriction to measures with bounded density is perhaps less natural, since the minimizers may concentrate mass on sets of measure zero and classical solutions to the aggregation equation can approach these steady states asymptotically [4]. However, since our proof regarding the Γ-convergence of gradient flows works for any choice of parameters $2 - d \leqslant p < 0 < q \leqslant 2$, we choose to include this regime for the sake completeness.

Remark 4.1 (Gradient flow vs. aggregation equation). While the aggregation equation (1.4) does inform our choice of the space of measures with bounded density (4.1), the relationship between weak solutions of the aggregation equation and the Wasserstein gradient flow of the interaction energy is purely formal for the unregularized (hence nonconvex) interaction energy. There is hope this relationship can be made rigorous following the approach of Ambrosio and Serfaty [2], but we leave this for future work.

For the regularized energy, if $\mu(t) \in \mathcal{P}_2(\mathbb{R}^d)$ is a weak solution of the aggregation equation (1.4), then it is also a Wasserstein gradient flow of E_ϵ [1, Corollary 11.1.8]. Below, we consider the *curve of maximal slope* of E_ϵ on the metric space of measures with bounded density $(\mathcal{P}_{2,R}(\mathbb{R}^d), d_W)$, and if $\mu(t) \in \mathcal{P}_2(\mathbb{R}^d)$ is a weak solution of the aggregation equation with $\|\mu(t)\|_\infty \leqslant R$, then it is a curve of maximal slope for the weak upper gradient $g_\epsilon(\mu) := 2\|\nabla K_\epsilon * \mu\|_{L^2(\mu)}$ on $(\mathcal{P}_{2,R}(\mathbb{R}^d), d_W)$ [1, Theorem 11.1.3]. The reverse implication is in general false, due to the height constraint imposed by $\mathcal{P}_{2,R}(\mathbb{R}^d)$.

4.1. **Curves of maximal slope and Γ-convergence.** We now briefly recall the notion of curves of maximal slope on a compete metric space $(\mathcal{S}, d)$. We refer the reader to the book by Ambrosio, Gigli, and Savaré [1, Chapter 1] for further details. A curve $u(t) : (a, b) \to \mathcal{S}$ is *absolutely continuous* if there exists $m \in L^2(a,b)$ so that

$$d(u(t),u(s)) \leqslant \int_s^t m(r)\,dr \text{ for all } a < s \leqslant t < b. \tag{4.2}$$

For any absolutely continuous curve, the limit

$$|u'(t)| = \lim_{s\to t} \frac{d(u(s),u(t))}{|s-t|}$$

exists for a.e. $t \in (a,b)$. Furthermore $m(t) := |u'(t)| \in L^2(a,b)$ satisfies (4.2) and for any $m \in L^2(a,b)$ satisfying (4.2), we have

$$|u'(t)| \leqslant m(t) \text{ for a.e. } t \in (a,b).$$

Given a functional $\mathsf{F} : \mathcal{S} \to (-\infty,+\infty]$ that is *proper*, i.e., $D(\mathsf{F}) = \{u \in \mathcal{S} : \mathsf{F}(u) < +\infty\} \neq \emptyset$, its upper gradient is a generalization of the modulus of the gradient from Euclidean space. Specifically, $g : \mathcal{S} \to [0,+\infty]$ is a *strong upper gradient* for F if for every absolutely continuous curve $u(t) : (a,b) \to \mathcal{S}$ the function $g \circ u$ is measurable and

$$|\mathsf{F}(u(t) - \mathsf{F}(u(s))| \leqslant \int_s^t g(u(r))|u'|(r)\,dr \text{ for all } a < s \leqslant t < b. \tag{4.3}$$

One example of a strong upper gradient is given by the metric slope

$$|\partial \mathsf{F}|(u) := \limsup_{v\to u} \frac{(\mathsf{F}(u) - \mathsf{F}(v))_+}{d(u,v)},$$

when F is a λ-convex and lower semicontinuous functional [1, Corollary 2.4.10].

Finally, we recall the definition of a curve a maximal slope. A locally absolutely continuous curve $u : (a,b) \to \mathcal{S}$ is a *curve of maximal slope* for F with respect to the strong upper gradient g if there exists a non-increasing function ϕ so that $\phi(t) = \mathsf{F} \circ u(t)$ for a.e. $t \in (a,b)$ and

$$\phi'(t) \leqslant -\frac{1}{2}|u'|^2(t) - \frac{1}{2}g^2(u(t)) \text{ for a.e. } t \in (a,b). \tag{4.4}$$

For all $\epsilon > 0$, [1, Corollary 2.4.12] ensures that if $\mu \in D(\mathsf{E}_\epsilon)$, then there exists a curve of maximal slope $\mu_\epsilon(t)\colon (0,+\infty) \to \mathcal{P}_{2,R}(\mathbb{R}^d)$ for E_ϵ with respect to the strong upper gradient $g_\epsilon(\nu) = 2||\nabla K_\epsilon * \nu\|_{L^2(d\nu)}$ satisfying $\mu_\epsilon(0) = \mu$.

With these definitions in hand, we now recall a general result of Serfaty on the Γ-convergence of gradient flows on a metric space.

Theorem 4.2 (cf. [51, Theorem 2])**.** *Let F_ϵ and F be functionals defined on metric spaces $(\mathcal{S}_\epsilon, d_\epsilon)$ and $(\mathcal{S}, d)$ with strong upper gradients g_ϵ and g, respectively. Suppose the following criteria hold:*

(i) *(Γ-liminf convergence) There is a notion of convergence S of $u_\epsilon \in \mathcal{S}_\epsilon$ to $u \in \mathcal{S}$ so that*

$$u_\epsilon \xrightharpoonup{S} u \quad \textit{implies} \quad \liminf_{\epsilon\to 0} \mathsf{F}_\epsilon(u_\epsilon) \geqslant \mathsf{F}(u).$$

(ii) *(Lower bound on metric derivatives) If $u_\epsilon(t) \xrightharpoonup{S} u(t)$ for $t \in (0,T)$, then for $s \in [0,T)$,*

$$\liminf_{\epsilon\to 0} \int_0^s |u_\epsilon'|^2_{d_\epsilon}(t)\,dt \geqslant \int_0^s |u'|^2_d(t)\,dt.$$

(iii) *(Lower bound on slopes) If $u_\epsilon \xrightharpoonup{S} u$, then $\liminf_{\epsilon\to 0} g_\epsilon(u_\epsilon) \geqslant g(u)$.*

If $u_\epsilon(t)$ is a curve of maximal slope on $(0,T)$ for F_ϵ with respect to g_ϵ satisfying

$$u_\epsilon(t) \xrightharpoonup{S} u(t) \text{ for } t \in (0,T) \quad \text{and} \quad \lim_{\epsilon\to 0} \mathsf{F}_\epsilon(u_\epsilon(0)) = \mathsf{F}(u(0)),$$

then $u(t)$ is a curve of maximal slope for F with respect to g and

$$\lim_{\epsilon\to 0} \mathsf{F}_\epsilon(u_\epsilon(t)) = \mathsf{F}(u(t)) \quad \text{for all} \;\; t \in [0,T),$$
$$g_\epsilon(u_\epsilon) \to g(u) \text{ and } |u_\epsilon'|_{d_\epsilon} \to |u'|_d \text{ in } L^2_{\mathrm{loc}}(0,T).$$

Remark 4.3. Although the metric d induces a natural topology on $\mathcal{S}$, the above result admits a notion of convergence S that can be induced by a weaker topology σ on $\mathcal{S}$. Indeed, $u_\epsilon \xrightharpoonup{S} u$ means that $\pi_\epsilon(u_\epsilon) \xrightharpoonup{\sigma} u$ for some map $\pi_\epsilon \colon \mathcal{S}_\epsilon \to \mathcal{S}$ (see [51] and [1, Remark 2.0.5] for details).

4.2. **Γ-convergence of the curves of maximal slope for the regularized energies.** Serfaty's general result on the Γ-convergence of curves of maximal slope provides a powerful general framework. In practice, it can be challenging to verify conditions $(i)-(iii)$. However, for the regularized interaction energies E_ϵ, these conditions follow from our previous results and the following HWI-type inequality from the first author's work with Kim and Yao, which allows us to compute the strong upper gradients.

Lemma 4.4 (c.f. [30])**.** *Suppose $\mu, \nu \in \mathcal{P}_{2,R}(\mathbb{R}^d)$ and E is an interaction energy induced by a repulsive-attractive kernel of the form (1.3). Define the function*

$$\omega(x) := \begin{cases} x|\log x| & \text{if } 0 \leqslant x \leqslant e^{-1-\sqrt{2}} \\ \sqrt{x^2 + 2(1+\sqrt{2})e^{-1-\sqrt{2}}x} & \text{if } x > e^{-1-\sqrt{2}} \end{cases}.$$

Then there exists $C_{R,d,p,q} > 0$ so that

$$|\mathsf{E}(\mu) - \mathsf{E}(\nu)| \leqslant 2\|\nabla K * \mu\|_{L^2(\mu)} d_W(\mu,\nu) + C_{R,d,p,q}\,\omega\big(d_W^2(\mu,\nu)\big).$$

With this, we show that the curves of maximal slope for the regularized energies Γ-converge to the curve of maximal slope of the unregularized energy.

Theorem 4.5. *Let E be the interaction energy with repulsive-attractive kernel (1.3), with $2-d \leqslant p < 0 < q \leqslant 2$, and let E_ϵ be the corresponding regularized energy (1.5).*

Suppose $\mu_\epsilon(t) \colon (0,T) \to \mathcal{P}_{2,R}(\mathbb{R}^d)$ is a curve of maximal slope of E_ϵ for the strong upper gradient $g_\epsilon(\mu) := 2\|\nabla K_\epsilon * \mu\|_{L^2(\mu)}$ on the metric space $(P_{2,R}(\mathbb{R}^d), d_W)$. Suppose also that $\mu_\epsilon(0)$ is well-prepared, in the sense that for some $\mu(0) \in \mathcal{P}_{2,R}(\mathbb{R}^d)$,

$$\mu_\epsilon(0) \to \mu(0) \text{ weak-* in } \mathcal{P}(\mathbb{R}^d) \text{ and } \lim_{\epsilon\to 0} \mathsf{E}_\epsilon(\mu_\epsilon(0)) = \mathsf{E}(\mu(0)).$$

Then for all $t \in (0,T)$, $\mu_\epsilon(t)$ has a weak- convergent subsequence $\mu_\epsilon(t) \xrightarrow{\epsilon\to 0} \mu(t)$, and $\mu(t) \colon (0,T) \to \mathcal{P}_{2,R}(\mathbb{R}^d)$ is a curve of maximal slope of E for $g(\mu) := 2\|\nabla K * \mu\|_{L^2(\mu)}$. Furthermore, as $\epsilon \to 0$,*

$$\mathsf{E}_\epsilon(\mu_\epsilon(t)) \to \mathsf{E}(\mu(t)) \text{ for all } \;\; t \in [0,T),$$
$$2\|\nabla K_\epsilon * \mu_\epsilon\|_{L^2(\mu_\epsilon)} \to 2\|\nabla K * \mu\|_{L^2(\mu)} \text{ in } L^2_{\mathrm{loc}}(0,T),$$
$$\text{and} \;\; |\mu_\epsilon'|_{d_W} \to |\mu'|_{d_W} \text{ in } L^2_{\mathrm{loc}}(0,T).$$

Remark 4.6. For any $\mu(0) \in \mathcal{P}_{2,R}(\mathbb{R}^d)$, there exists $\mu_\epsilon(0) \in \mathcal{P}_{2,R}$ satisfying the conditions of the theorem: simply define $\mu_\epsilon(0)$ by convolving $\mu(0)$ with the heat kernel, as in Theorem 3.4(ii).

Proof. Let $(\mathcal{S}_\epsilon, d_\epsilon) = (\mathcal{S}, d) = (\mathcal{P}_{2,R}(\mathbb{R}^d), d_W)$. Note that

$$\mathcal{P}_{2,R}(\mathbb{R}^d) = \{\mu \in \mathcal{P}_2(\mathbb{R}^d) \colon \|\mu\|_{L^\infty(\mathbb{R}^d)} \leqslant R\}$$

is closed with respect to d_W, thus $\mathcal{P}_{2,R}(\mathbb{R}^d)$ is a complete metric space. Furthermore, any d_W bounded set of $\mathcal{P}_{2,R}(\mathbb{R}^d)$ is relatively compact with respect to weak-* convergence in $\mathcal{P}(\mathbb{R}^d)$, and its limit points lie in $\mathcal{P}_{2,R}(\mathbb{R}^d)$. Given $\mu_\epsilon \in \mathcal{S}_\epsilon$, $\mu \in \mathcal{S}$ we say $\mu_\epsilon \xrightarrow{S} \mu$ if the μ_ϵ converges with respect to the weak-* convergence in $\mathcal{P}(\mathbb{R}^d)$.

We now define the strong upper gradients. For all $\epsilon > 0$, Remark (2.8) ensures that $g_\epsilon(\nu) = 2\|\nabla K_\epsilon * \nu\|_{L^2(\nu)}$ is a strong upper gradient of E_ϵ on $\mathcal{P}_2(\mathbb{R}^d)$, hence on $\mathcal{P}_{2,R}(\mathbb{R}^d)$. Next, we show that $g(\nu) = 2\|\nabla K * \nu\|_{L^2(\nu)}$ is a strong upper gradient of E. Throughout, we use the fact that if $\mu \in L^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$, then $K * \mu(x)$ is continuously differentiable and $\nabla(K * \mu) = (\nabla K) * \mu$.

Suppose $\nu(t) : (a,b) \to \mathcal{P}_{2,R}(\mathbb{R}^d)$ is an absolutely continuous curve. The function $t \mapsto g(\nu(t))$ is measurable, since it is given by the composition of measurable functions. By Lemma 4.4,

$$|E(\nu(t)) - E(\nu(s))| \leqslant 2\|\nabla K * \nu(t)\|_{L^2(\nu(t))} d_W(\nu(t), \nu(s)) + C_{R,d,p,q}\,\omega\big(d_W^2(\nu(t), \nu(s))\big). \tag{4.5}$$

As in Remark 2.8, we estimate $\|\nabla K * \nu(t)\|_{L^2(\nu(t))}$ by breaking ∇K into its singular and nonsingular parts,

$$\begin{aligned}
\|\nabla K * \nu(t)\|_{L^2(\nu(t))} &\leqslant \|\nabla K^n * \nu(t)\|_{L^2(\nu(t))} + \|\nabla K^s * \nu(t)\|_{L^2(\nu(t))} \\
&\leqslant \left(\int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} \nabla K^n(x-y)\, d\nu(y,t) \right|^2 d\nu(x,t) \right)^{1/2} + \sqrt{R} \|\nabla K^s * \nu(t)\|_{L^2(\mathbb{R}^d)} \\
&\leqslant \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |\nabla K^n(x-y)|^2 \, d\nu(x,t) \right)^{1/2} d\nu(y,t) + \sqrt{R} \|\nabla K^s\|_{L^1(\mathbb{R}^d)} \|\nu(t)\|_{L^2(\mathbb{R}^d)}.
\end{aligned}$$

Since ∇K^n has at most linear growth, the first term is bounded by the second moments of $\nu(t)$, which are uniformly bounded for $t \in (a,b)$ since $\{\nu(t)\}_{t \in (a,b)}$ is bounded with respect to d_W. The second term is uniformly bounded since $\|\nu(t)\|_{L^1} = 1$ and $\|\nu(t)\|_{L^\infty} \leqslant R$.

By the absolute continuity of $\nu(t)$, $d_W(\nu(t), \nu(s)) \leqslant \int_s^t |\nu'|(r)\, dr$ for $|\nu'|(r) \in L^2(a,b)$. Therefore (4.5) ensures $t \mapsto E(\nu(t))$ is absolutely continuous and

$$\left| \frac{d}{dt} E(\nu(t)) \right| \leqslant 2 \|\nabla K * \nu(t)\|_{L^2(\nu(t))} |\nu'|(t),$$

hence

$$|E(\nu(t)) - E(\nu(s))| = \left| \int_s^t \frac{d}{dr} E(\nu(r))\, dr \right| \leqslant \int_s^t 2\|\nabla K * \nu(r)\|_{L^2(\nu(r))} |\nu'|(r)\, dr.$$

We conclude that $g(\nu) = 2\|\nabla K * \nu\|_{L^2(\nu)}$ is a strong upper gradient of E.

Now we show that for all $t \in (0,T)$, $\mu_\epsilon(t) \xrightarrow{\epsilon \to 0} \mu(t) \in \mathcal{P}_{2,R}(\mathbb{R}^d)$, with respect to weak-* convergence of probability measures. Since $\mu_\epsilon(t) \colon [0,\infty) \to \mathcal{P}_2(\mathbb{R}^d)$ is a curve of maximal slope for the strong upper gradient $g_\epsilon(\mu) = \|\nabla K_\epsilon * \mu\|_{L^2(\mu)}$, applying Cauchy's inequality to (4.3) and comparing with (4.4), we obtain (see [1, Remark 1.3.3])

$$\mathsf{E}_\epsilon(\mu_\epsilon(s)) - \mathsf{E}_\epsilon(\mu_\epsilon(t)) = \int_s^t |\mu_\epsilon'|^2(r)dr. \tag{4.6}$$

for all $0 \leqslant s \leqslant t < \infty$. By definition of the metric slope and the fact that $\mathsf{E}_\epsilon(\mu_\epsilon(t)) \geqslant 0$,

$$d_W^2(\mu_\epsilon(0), \mu_\epsilon(t)) \leqslant \left(\int_0^t |\mu_\epsilon'|(r)\, dr \right)^2 \leqslant t \int_0^t |\mu_\epsilon'|^2(t)\, dt \leqslant T\, \mathsf{E}_\epsilon(\mu_\epsilon(0))$$

for all $t \in (0,T)$. Since $\lim_{\epsilon\to 0} \mathsf{E}_\epsilon(\mu_\epsilon(0)) = \mathsf{E}(\mu(0))$, the right hand side is uniformly bounded for ϵ sufficiently small. Therefore, $\{\mu_\epsilon(t)\}_{\epsilon>0}$ is uniformly bounded in $\mathcal{P}_{2,R}(\mathbb{R}^d)$, and, up to a subsequence, $\mu_\epsilon(t) \xrightarrow{\epsilon\to 0} \mu(t) \in \mathcal{P}_{2,R}(\mathbb{R}^d)$ with respect to weak-*convergence in $\mathcal{P}(\mathbb{R}^d)$.

It remains to verify criteria $(i) - (iii)$ of Theorem (4.2) to conclude that $\mu(t)$ is a curve of maximal slope of E for $g(\mu)$, and the corresponding energies, strong upper gradients and metric slopes converge as $\epsilon \to 0$.

Criterion (i) follows immediately from Theorem 3.4, part (i) and Remark 3.5. To prove (ii), we assume without loss of generality that there exists $0 \leqslant C < +\infty$ so that

$$C = \liminf_{\epsilon\to 0} \int_0^s |\mu_\epsilon'|^2(t)\, dt.$$

Choose a subsequence $|\tilde{\mu}_\epsilon'|(t)$ so that $\lim_{\epsilon\to 0} \int_0^s |\tilde{\mu}_\epsilon'|^2(t)\, dt = C$. Then $|\tilde{\mu}_\epsilon'|(t)$ is bounded in $L^2(0,s)$ so, up to a further subsequence, it is weakly convergent to some $v(t) \in L^2(0,s)$. Consequently, for any $0 \leqslant s_0 \leqslant s_1 \leqslant s$,

$$\lim_{\epsilon\to 0} \int_{s_0}^{s_1} |\tilde{\mu}_\epsilon|(t)\, dt = \int_{s_0}^{s_1} v(t)\, dt.$$

By taking limits in the definition of the metric derivative and using the lower-semicontinuity of d_W with respect to weak-* convergence,

$$d_W(\mu_\epsilon(s_0), \mu_\epsilon(s_1)) \leqslant \int_{s_0}^{s_1} |\mu_\epsilon'|(t)\, dt \quad \text{yields} \quad d_W(\mu(s_0), \mu(s_1)) \leqslant \int_{s_0}^{s_1} v(t)\, dt.$$

By [1, Theorem 1.1.2], this implies that $|\mu'|(t) \leqslant v(t)$ for a.e. $t \in (0,s)$. Thus, by the lower semicontinuity of the $L^2(0,s)$ norm with respect to weak convergence,

$$\liminf_{\epsilon\to 0} \int_0^s |\mu_\epsilon'|^2(t)\, dt = \lim_{\epsilon\to 0} \int_0^s |\tilde{\mu}_\epsilon'|^2(t)\, dt \geqslant \int_0^s v(t)^2\, dt \geqslant \int_0^s |\mu'|^2(t)\, dt.$$

Finally, we turn to (iii). We assume without loss of generality that

$$C = \liminf_{\epsilon\to 0} g_\epsilon(\mu_\epsilon),$$

for some $0 \leqslant C < +\infty$. Choose a subsequence $g_\epsilon(\mu_\epsilon)$ so that $\lim_{\epsilon\to 0} g_\epsilon(\mu_\epsilon) = C$. Since $\|\mu_\epsilon(t)\|_{L^\infty(\mathbb{R}^d)} \leqslant R$ for all $\epsilon > 0$ and all $t \in (0,T)$, there exists a further subsequence so that μ_ϵ converges to some limit ν in the weak-* topology of $L^\infty(\mathbb{R}^d)$. Since μ_ϵ also converges to μ with respect to the weak-* topology of $\mathcal{P}(\mathbb{R}^d)$, for all $f \in C_c^\infty(\mathbb{R}^d)$,

$$\int_{\mathbb{R}^d} f(x)\nu(x)\, dx = \lim_{\epsilon\to 0} \int_{\mathbb{R}^d} f(x)\mu_\epsilon(x)\, dx = \int_{\mathbb{R}^d} f(x)\mu(x)\, dx;$$

hence, $\nu = \mu$.

By [1, Theorem 5.4.4] it suffices to show

$$\lim_{\epsilon\to 0}\int_{\mathbb{R}^d} f(x)(\nabla K_\epsilon * \mu_\epsilon)(x)\,d\mu_\epsilon(x) = \int_{\mathbb{R}^d} f(x)(\nabla K * \mu)(x)\,d\mu(x) \quad \text{for all } f \in C_c^\infty(\mathbb{R}^d).$$

Then using the convexity and the lower semicontinuity of the function $|\cdot|^2$ along with the weak-* convergence of μ_ϵ yields the result.

Following a similar argument as in Lemma 2.5 and defining $\tilde{\mu}_\epsilon = \varphi_\epsilon * \mu_\epsilon * \varphi_\epsilon$, we have

$$\begin{aligned}
&\left|\int_{\mathbb{R}^d} f(x)(\nabla K_\epsilon * \mu_\epsilon)(x)\,d\mu_\epsilon(x) - \int_{\mathbb{R}^d} f(x)(\nabla K * \mu)(x)\,d\mu(x)\right| \\
&\qquad = \left|\int_{\mathbb{R}^d} f(x)(\nabla K * \tilde{\mu}_\epsilon)(x)\,d\mu_\epsilon(x) - \int_{\mathbb{R}^d} f(x)(\nabla K * \mu)(x)\,d\mu(x)\right| \\
&\qquad \leqslant \left|\int_{\mathbb{R}^d} f(x)[(\nabla K * \tilde{\mu}_\epsilon)(x) - (\nabla K * \mu)(x)]\,d\mu_\epsilon(x)\right| \\
&\qquad\qquad + \left|\int_{\mathbb{R}^d} f(x)(\nabla K * \mu)(x)\,d\mu_\epsilon(x) - \int_{\mathbb{R}^d} f(x)(\nabla K * \mu)(x)\,d\mu(x)\right| \\
&\qquad =: A_\epsilon + B_\epsilon.
\end{aligned}$$

Since $\mu_\epsilon \to \mu$ weak-* in $\mathcal{P}(\mathbb{R}^d)$ and $f(x)(\nabla K * \mu)(x)$ is bounded and continuous, $\lim_{\epsilon\to 0} B_\epsilon = 0$.

It remains to show $\lim_{\epsilon\to 0} A_\epsilon = 0$. First, note that $\tilde{\mu}_\epsilon \to \mu$ in the weak-* topology of $L^\infty(\mathbb{R}^d)$ as $\epsilon \to 0$. Indeed, for any $f \in L^1(\mathbb{R}^d)$, we have that

$$\begin{aligned}
\left|\int_{\mathbb{R}^d} f(x)\,d\tilde{\mu}_\epsilon(x) - \int_{\mathbb{R}^d} f(x)\,d\mu(x)\right| &= \left|\int_{\mathbb{R}^d} \varphi_\epsilon * f * \varphi_\epsilon(x)\,d\mu_\epsilon(x) - \int_{\mathbb{R}^d} f(x)\,d\mu(x)\right| \\
&\leqslant R\|\varphi_\epsilon * f * \varphi_\epsilon - f\|_{L^1} + \left|\int_{\mathbb{R}^d} f(x)\,d\mu_\epsilon(x) - \int_{\mathbb{R}^d} f(x)\,d\mu(x)\right|,
\end{aligned}$$

where both terms approach zero as $\epsilon \to 0$. Returning to A_ϵ,

$$A_\epsilon \leqslant \int_{\mathbb{R}^d} |f(x)(\nabla K * (\tilde{\mu}_\epsilon - \mu))(x)|\ |\mu_\epsilon(x)|\ dx \leqslant R\,\|f\|_{L^\infty(\mathbb{R}^d)} \int_{\operatorname{supp} f} |\nabla K * (\tilde{\mu}_\epsilon - \mu)(x)|\,dx.$$

Since the integrand has at most linear growth, it is bounded on the compact set $\operatorname{supp} f$, and we may apply the dominated convergence theorem, provided the integrand converges pointwise.

When $0 < q \leqslant 1$, ∇K is the sum of a bounded continuous function and an integrable function, and since $\tilde{\mu}_\epsilon \xrightarrow{\epsilon\to 0} \mu$ in both weak-* probability and weak-* $L^\infty(\mathbb{R}^d)$, the integrand converges for each x. On the other hand, when $1 < q \leqslant 2$, it suffices to show that ∇K is the sum of a continuous function, which is uniformly integrable with respect to $\tilde{\mu}_\epsilon$, and an integrable function [1, Lemma 5.1.7]. In particular, is enough to show that $|x|^{q-1}$ is uniformly integrable with respect to $\tilde{\mu}_\epsilon$. Since $|x| \geqslant k$ implies that $|x|/k \geqslant 1$ we have that

$$\lim_{k\to\infty}\int_{|x|\geqslant k} |x|^{q-1}\,d\tilde{\mu}_\epsilon \leqslant \lim_{k\to\infty}\frac{1}{k}\int_{|x|\geqslant k} |x|^q\,d\tilde{\mu}_\epsilon \leqslant \lim_{k\to\infty}\frac{1}{k}\mathsf{E}(\tilde{\mu}_\epsilon) = \lim_{k\to\infty}\frac{1}{k}\mathsf{E}_\epsilon(\mu_\epsilon).$$

Since $\mu_\epsilon(t)$ is a curve of maximal slope for E_ϵ, (4.6) gives $\mathsf{E}_\epsilon(\mu_\epsilon(t)) \leqslant \mathsf{E}_\epsilon(\mu_\epsilon(0))$ for all $t \in (0,T)$. The well-preparedness of the initial data gives $\mathsf{E}_\epsilon(\mu_\epsilon(0)) < C$, hence the result follows. □

Acknowledgments. This project was initiated during the Thematic Program on Variational Problems in Physics, Economics and Geometry at the Fields Institute for Research in Mathematical Sciences. The authors would like to thank the organizers and the Institute for their hospitality and support during their visits. They also would like to thank Andrea Bertozzi, Eric Carlen, Andres Contreras, Inwon Kim, and Xavier Lamy for useful discussions.

References

[1] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.

[2] L. Ambrosio and S. Serfaty. A gradient flow approach to an evolution problem arising in superconductivity. *Comm. Pure Appl. Math.*, 61(11):1495–1539, 2008.

[3] C. Anderson and C. Greengard. On vortex methods. *SIAM J. Numer. Anal.*, 22(3):413–440, 1985.

[4] D. Balagué, J. Carrillo, T. Laurent, and G. Raoul. Nonlocal interactions by repulsive-attractive potentials: Radial ins/stability. *Phys. D.*, 260:5–25, 2013.

[5] D. Balagué, J. A. Carrillo, T. Laurent, and G. Raoul. Dimensionality of local minimizers of the interaction energy. *Arch. Ration. Mech. Anal.*, 209(3):1055–1088, 2013.

[6] D. Balagué, J.A. Carrillo, and Y. Yao. Confinement for repulsive-attractive kernels. *Discrete Contin. Dyn. Syst. Ser. B*, 19(5):1227–1248, 2014.

[7] J.T. Beale and A. Majda. Vortex methods. I. Convergence in three dimensions. *Math. Comp.*, 39(159):1–27, 1982.

[8] J.T. Beale and A. Majda. Vortex methods. II. Higher order accuracy in two and three dimensions. *Math. Comp.*, 39(159):29–52, 1982.

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

[10] A.J. Bernoff and C.M. Topaz. A primer of swarm equilibria. *SIAM J. Appl. Dyn. Syst.*, 10(1):212–250, 2011.

[11] A.L. Bertozzi and J. Brandman. Finite-time blow-up of L^∞-weak solutions of an aggregation equation. *Commun. Math. Sci.*, 8(1):45–65, 2010.

[12] A.L. Bertozzi, J.A. Carrillo, and T. Laurent. Blow-up in multidimensional aggregation equations with mildly singular interaction kernels. *Nonlinearity*, 22(3):683–710, 2009.

[13] A.L. Bertozzi, J.B. Garnett, and T. Laurent. Characterization of radially symmetric finite time blowup in multidimensional aggregation equations. *SIAM J. Math. Anal.*, 44(2):651–681, 2012.

[14] A.L. Bertozzi, T. Laurent, and F. Léger. Aggregation and spreading via the Newtonian potential: the dynamics of patch solutions. *Math. Models Methods Appl. Sci.*, 22(suppl. 1):1140005, 39, 2012.

[15] A.L. Bertozzi, T. Laurent, and J. Rosado. L^p theory for the multidimensional aggregation equation. *Comm. Pure Appl. Math.*, 64(1):45–83, 2011.

[16] A.L. Bertozzi, H. Sun, T. Kolokolnikov, D. Uminsky, and J. Von Brecht. Ring patterns and their bifurcations in a nonlocal model of biological swarms. *Preprint*.

[17] A. Blanchet, E.A. Carlen, and J.A. Carrillo. Functional inequalities, thick tails and asymptotics for the critical mass Patlak-Keller-Segel model. *J. Funct. Anal.*, 262(5):2142–2230, 2012.

[18] J.A. Cañizo, J.A. Carrillo, and F.S. Patacchini. Existence of compactly supported global minimizers for the interaction energy. to appear.

[19] J.A. Carrillo, Y.-P. Choi, and M. Hauray. The derivation of swarming models: mean-field limit and Wasserstein distances. *Collective Dynamics from Bacteria to Crowds: An Excursion Through Modeling, Analysis and Simulation Series, CISM International Centre for Mechanical Sciences*, 553:1–46, 2014.

[20] J.A. Carrillo, M.G. Delgadino, and A. Mellet. Regularity of local minimizers of the interaction energy via obstacle problems. Preprint, 2014.

[21] J.A. Carrillo, M. Di Francesco, A. Figalli, T. Laurent, and D. Slepčev. Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations. *Duke Math. J.*, 156(2):229–271, 2011.

[22] J.A. Carrillo, S. Lisini, and E. Mainini. Uniqueness for Keller-Segel-type chemotaxis models. *Discrete Contin. Dyn. Syst.*, 34(4):1319–1338, 2014.

[23] 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. Mat. Iberoamericana*, 19(3):971–1018, 2003.
[24] J.A. Carrillo, R.J. McCann, and C. Villani. Contractions in the 2-Wasserstein length space and thermalization of granular media. *Arch. Ration. Mech. Anal.*, 179(2):217–263, 2006.
[25] J.A. Carrillo and J. Rosado. Uniqueness of bounded solutions to aggregation equations by optimal transport methods. In *European Congress of Mathematics*, pages 3–16. Eur. Math. Soc., Zürich, 2010.
[26] D. Chafaï, N. Gozlan, and P.-A. Zitt. First-order global asymptotics for confined particles with singular pair repulsion. *Ann. Appl. Probab.*, 24(6):2371–2413, 2014.
[27] R. Choksi, R.C. Fetecau, and I. Topaloglu. On minimizers of interaction functionals with competing attractive and repulsive potentials. *Ann. Inst. H. Poincaré Anal. Non Linéaire*. to appear.
[28] 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*, pages 2292–2299, 2007.
[29] K. Craig and A. Bertozzi. A blob method for the aggregation equation. *Math. Comp.* to appear.
[30] K. Craig, I. C. Kim, and Y. Yao. The Wasserstein gradient flow of the height constrained aggregation equation. in progress, 2015.
[31] H. Dong. The aggregation equation with power-law kernels: ill-posedness, mass concentration and similarity solutions. *Comm. Math. Phys.*, 304(3):649–664, 2011.
[32] J.P.K. Doye, D.J. Wales, and R.S. Berry. The effect of the range of the potential on the structures of clusters. *J. Chem. Phys.*, 103:4234–4249, 1995.
[33] K. Fellner and G. Raoul. Stable stationary states of non-local interaction equations. *Math. Models Methods Appl. Sci.*, 20(12):2267–2291, 2010.
[34] R.C. Fetecau and Y. Huang. Equilibria of biological aggregations with nonlocal repulsive-attractive interactions. *Phys. D*, 260:49–64, 2013.
[35] R.C. Fetecau, Y. Huang, and T. Kolokolnikov. Swarm dynamics and equilibria for a nonlocal aggregation model. *Nonlinearity*, 24(10):2681–2716, 2011.
[36] C.R. Givens and R.M. Shortt. A class of Wasserstein metrics for probability distributions. *Michigan Math. J.*, 31(2):231–240, 1984.
[37] M.F. Hagan and D. Chandler. Dynamic pathways for viral capsid assembly. *Biophysical Journal*, 91:42–54, 2006.
[38] Y. Huang and A.L. Bertozzi. Self-similar blowup solutions to an aggregation equation in $\mathbf{R}^n$. *SIAM J. Appl. Math.*, 70(7):2582–2603, 2010.
[39] T. Kolokolnikov, H. Sun, D. Uminsky, and A.L. Bertozzi. Stability of ring patterns arising from two-dimensional particle interactions. *Phys. Rev. E*, 84(1), 2011.
[40] H. Li and G. Toscani. Long-time asymptotics of kinetic models of granular flows. *Arch. Ration. Mech. Anal.*, 172(3):407–428, 2004.
[41] P.-L. Lions. The concentration-compactness principle in the calculus of variations. The locally compact case. I. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 1(2):109–145, 1984.
[42] E. Mainini. A global uniqueness result for an evolution problem arising in superconductivity. *Boll. Unione Mat. Ital. (9)*, 2(2):509–528, 2009.
[43] E. Mainini. Well-posedness for a mean field model of Ginzburg-Landau vortices with opposite degrees. *NoDEA Nonlinear Differential Equations Appl.*, 19(2):133–158, 2012.
[44] R.J. McCann. Stable rotating binary stars and fluid in a tube. *Houston J. Math.*, 32(2):603–631, 2006.
[45] A. Mogilner and L. Edelstein-Keshet. A non-local model for a swarm. *J. Math. Biol.*, 38(6):534–570, 1999.
[46] A. Mogilner, L. Edelstein-Keshet, L. Bent, and A. Spiros. Mutual interactions, potentials, and individual distance in a social aggregation. *J. Math. Biol.*, 47(4):353–389, 2003.
[47] L. Perea, G. Gómez, and P. Elosegui. Extension of the Cucker–Smale control law to space flight formations. *AIAA J. of Guidance, Control, and Dynamics*, 32:527–537, 2009.
[48] D. Petz and F. Hiai. Logarithmic energy as an entropy functional. In *Advances in differential equations and mathematical physics (Atlanta, GA, 1997)*, volume 217 of *Contemp. Math.*, pages 205–221. Amer. Math. Soc., Providence, RI, 1998.
[49] F. Poupaud. Diagonal defect measures, adhesion dynamics and Euler equation. *Methods Appl. Anal.*, 9(4):533–561, 2002.

[50] M.C. Rechtsman, F.H. Stillinger, and S. Torquato. Optimized interactions for targeted self- assembly: application to a honeycomb lattice. *Phys. Rev. Lett.*, 95(22), 2005.
[51] S. Serfaty. Gamma-convergence of gradient flows on Hilbert and metric spaces and applications. *Discrete Contin. Dyn. Syst.*, 31(4):1427–1451, 2011.
[52] R. Simione, D. Slepčev, and I. Topaloglu. Existence of ground states of nonlocal-interaction energies. *J. Stat. Phys.* to appear.
[53] E.M. Stein. *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.
[54] M. Struwe. *Variational Methods: Applications to Nonlinear Partial Differential Equations and Hamiltonian Systems*. Springer-Verlag, Berlin, third edition, 2000.
[55] H. Sun, D. Uminsky, and A.L. Bertozzi. Stability and clustering of self-similar solutions of aggregation equations. *J. Math. Phys.*, 53(11):115610, 18, 2012.
[56] C.M. Topaz and A.L. Bertozzi. Swarming patterns in a two-dimensional kinematic model for biological groups. *SIAM J. Appl. Math.*, 65(1):152–174, 2004.
[57] C.M. Topaz, A.L. Bertozzi, and M.A. Lewis. A nonlocal continuum model for biological aggregation. *Bull. Math. Biol.*, 68(7):1601–1623, 2006.
[58] A.W. van der Vaart and J. Wellner. *Weak Convergence and Empirical Processes: With Applications to Statistics*. Springer Series in Statistics. Springer, 1996.
[59] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.
[60] D.J. Wales. Energy landscapes of clusters bound by short-ranged potentials. *Chem. Eur. J. Chem. Phys.*, 11:2491–2494, 2010.

Department of Mathematics, University of California, Los Angeles, CA
E-mail address: kcraig@math.ucla.edu

Department of Mathematics and Statistics, McMaster University, Hamilton, ON
E-mail address: itopalog@math.mcmaster.ca