

# The derivation of Swarming models: Mean-Field Limit and Wasserstein distances

José Antonio Carrillo<sup>†</sup>, Young-Pil Choi<sup>†</sup>, and Maxime Hauray<sup>‡</sup>

<sup>†</sup> Department of Mathematics, Imperial College London,  
SW7 2AZ London, United Kingdom

<sup>‡</sup> Centre de Mathématiques et Informatique, Université d'Aix-Marseille,  
Technopôle Château-Gombert, Marseille, France

**Abstract** These notes are devoted to a summary on the mean-field limit of large ensembles of interacting particles with applications in swarming models. We first make a summary of the kinetic models derived as continuum versions of second order models for swarming. We focus on the question of passing from the discrete to the continuum model in the Dobrushin framework. We show how to use related techniques from fluid mechanics equations applied to first order models for swarming, also called the aggregation equation. We give qualitative bounds on the approximation of initial data by particles to obtain the mean-field limit for radial singular (at the origin) potentials up to the Newtonian singularity. We also show the propagation of chaos for more restricted set of singular potentials.

## 1 Introduction

In the last years, we have seen the development of a great deal of different models in the biology, applied mathematics, and physics literature to describe the collective behavior of individuals. Here, individuals may mean animals (insects, fish, birds,...), bacteria, and even robots. Most of these models involve the nonlocal character of the interaction as a basic modelling pillar, see for instance Camazine, Deneubourg, Franks, Sneyd, Theraulaz, and Bonabeau (2003); Couzin, Krause, Franks, and Levin (2005); Li, Luke-man, and Edelstein-Keshet (2008); Vicsek, Czirok, Ben-Jacob, Cohen, and Shochet (1995). In fact, one of largest source of collective behavior models comes from control engineering. There, the aim is to produce a suitable control of the movement of small squads of robots in order to perform unmanned vehicle operations, for instance Perea, Gómez, and Elosequi (2009). Even, these ideas have been proposed to model crowd motion, including

more “intelligent” particles deciding their movement based on optimization of certain quantities: time to exit from a room or a stadium, for instance Burger, Markowich, and Pietschmann (2011).

Either in social or in biological sciences, these models encounter many interesting features such as the spontaneous formation of different pattern behaviors. When we talk about patterns, we do not mean static patterns like in the study of crystals but rather dynamic patterns leading to the collective motion of the individual ensemble. For instance, two of the main collective motion patterns studied in different models are the flock and the milling behavior, see D’Orsogna, Chuang, Bertozzi, and Chayes (2006); Carrillo, D’Orsogna, and Panferov (2009); Cañizo, Carrillo, and Rosado (2010); Carrillo, Klar, Martin, and Tiwari (2010); Carrillo, Panferov, and Martin (2013). In the flock pattern, individuals achieve a consensus on the direction or orientation towards some objective, producing as a consequence a particular spatial shape showing their preferred comfort structure. This kind of swiftly moving flocks have been reported in many species although the most spectacular or bucolic ones are the bird flocks, starlings for instance. In the mill pattern, individuals arrange into a kind of vortex like motion around some point. This particular moving pattern has been observed in fish schools. Hundreds of movies can be easily accessed through internet search showing them.

There are many reasons one can argue, why such a large number of individuals react to external stimuli producing these macroscopic patterns without seemingly the presence of a leader in the swarm. Hydrodynamic enhancement, predators avoidance, social interactions, spawning survival rate, and many others have been proposed to explain this behavior in different species, see Parrish, and Edelstein-Keshet (1999).

One of the main question in describing this behavior by mathematical models is how to include the interaction between individuals. In any case, there is a consensus that the modelling starts from particle-like models as in statistical physics. These particle models are also called Individual Based Models (IBMs) in the community. They are usually formed by a set of differential equations of Newton type (called 2nd order models) or by kinematic equations where the inertia terms are neglected (called first order models). Essentially, by admitting that the inertia term is negligible, we assume that individuals can adjust to the velocity field instantaneously, an approximation valid when their speed is not too large. In any case, these first order models were proposed in the literature derived in a phenomenological manner Mogilner, Edelstein-Keshet, Bent, and Spiros (2003); Mogilner and Edelstein-Keshet (1999); Parrish, and Edelstein-Keshet (1999); Topaz and Bertozzi (2004); Topaz, Bertozzi, and Lewis (2006); Eftimie, de Vries, and

Lewis (2007). The literature on first and second order models for swarming has increased exponentially fast in the last few years. Many of these models find also their origin in social sciences, where consensus or opinion formation was also described in similar grounds. Another typical ingredient in these models is some kind of noise leading to systems of SDEs. In this work, we will not discuss how to incorporate noise in these models, we refer to Bolley, Cañizo, and Carrillo (2011) and the references therein.

Most of these models are based on discrete approaches incorporating certain effects that we like to call the “first principles” of swarming. These first principles are based on modelling the “sociological behavior” of animals with very simple rules such as the social tendency to produce grouping (attraction/aggregation), the inherent minimal space they need to move without problems and feel comfortably inside the group (repulsion/collisional avoidance) and the mimetic adaptation or synchronization to a group (orientation/alignment). Even if these minimal models contain very basic rules, the patterns observed in their simulation and their complex asymptotic behavior are already very challenging from the mathematical viewpoint. The 3-zone models including attraction, repulsion, and alignment effects are classical in fish modelling Aoki (1982); Huth and Wissel (1992) for instance. Based on them, one can incorporate many other effects to render more realistic the outputs of the simulations and the models, see Barbaro, Taylor, Trethewey, Youseff, and Birnir (2009) for fish schools or Hemelrijk and Hildenbrandt (2008) for birds flocks. We also refer to the reader to the recent review Carrillo, Fornasier, Toscani, and Vecil (2010) about the kinetic modelling of swarming.

To the eyes of a kinetic theorist or a statistical physicist, studying such systems of ODEs when the number of individuals get large is doomed to failure. Dynamical system approaches are quite useful but they typically have huge problems to describe large systems of particles. A classical approach to attack the problem is to pass to a continuous description of the system. This means to go from particle descriptions to kinetic descriptions where the unknown is the particle density distribution in position-velocity (phase) space for 2nd order models or in position space for 1st order models.

Going from particle to continuum descriptions is one of the most classical problems in kinetic theory. It is at the basis of the derivation of the mother and father kinetic equations, namely: the Vlasov and the Boltzmann equations. A rigorous derivation of the Boltzmann equation from the Newtonian dynamics has only been given for short times (of the order of the average time of first collision), see Lanford (1974) Gallagher, St-Raymond, and Texier (2012). In that case, interactions between the particles are modelled by short-range potentials leading to collision kernels. The question of

the derivation of the Boltzmann equation from particles with jump processes was also raised and solved by Kac (1956), and further results are given in the recent important work by Mischler and Mouhot (2013). The derivation of the Vlasov equation is well understood only for regular or not too singular potentials Braun and Hepp (1977); Neunzert (1984); Dobrushin (1979); Hauray and Jabin (2012). In fact, a full derivation of the Vlasov-Poisson system in 3D is also lacking. The problem of passing to the limit from particle to continuum models like the Vlasov equation is called the mean-field limit. This name just comes from the fact that the resulting equation is a kind of averaged version of the interaction between the large number of individuals. Moreover, the resulting equation gives the typical behavior of one isolated individual among all the others since they are assumed to be completely indistinguishable.

Finally, there are other famous mean-field limit equations, such as the Euler and the Navier-Stokes equations for incompressible fluids, see Marchioro and Pulvirenti (1994); Majda and Bertozzi (2002). It has been extensively used for numerical purposes that both equations in the 2D incompressible case can be derived from particle approximations, called vortex point approximations. The convergence in the viscous case has been rigorously proved for very general initial data Osada (1985); Fournier, Hauray, and Mischler (2012). In the non-viscous case Schochet (1996) proves that particle approximations converge towards solutions of the Euler equation, but they may not converge to the good solution because of the lack of uniqueness in the Euler equation, see De Lellis and L. Székelyhidi (2009). However, in the case where the initial particles are equally spaced on a grid to approximate a smooth solution of the Euler equation, the convergence was shown in Goodman, Hou, and Lowengrub (1990). These vortex methods have been proven to be convergent and estimates of the error committed have been obtained in recent works using optimal transport techniques Hauray (2009) but not for the real Euler equation in 2D.

The aim of this work is to show in detail a particular example of the mean field limit in the case of first order models not covered in the previous literature. Nevertheless, we will first discuss some of these issues for 2nd order models summarizing results in Cañizo, Carrillo, and Rosado (2011); Bolley, Cañizo, and Carrillo (2011). We will also discuss that the spatial shape of the main patterns: flock and mills, are given by stationary solutions of the 1st order models. This gives another reason from a more conceptual mathematical viewpoint of reducing to 1st order models. Section 3 will be devoted to obtain the mean field limit to the so-called aggregation equation for singular potentials recovering some of the models studied in Bertozzi, Carrillo, and Laurent (2009); Bertozzi, Laurent, and Rosado (2010). Here,

the idea is to assume that we have solutions of the model in better functional spaces due to the singularity of the potential, but we have to pay in terms of conditions on the initial distribution of particles (how they are distributed) in such a way that the particle solution converges to the continuum solution of the aggregation equation as  $N \rightarrow \infty$ . We will make use of similar arguments to Hauray (2009) to show the mean-field limit for first order swarming models with singular potentials up to the Newtonian singularity. In Section 4, we study a local existence of a unique  $L^p$ -solution for the aggregation equation. This complements the well-posedness theory in Bertozzi, Laurent, and Rosado (2010). Finally, Section 5 is devoted to show the propagation of chaos property for the aggregation equation. This property is very important from the physical relevance of the kinetic and aggregation models, since it states that one can derive the mean-field equations under quite generic randomly generated initial location of the particles. We are only able to show it for a more restricted set of singular potentials with respect to the mean-field limit.

## 2 The Dobrushin approach

### 2.1 Some Individual Based Models

As we described in the introduction, the modelling in swarming starts by introducing some particle models, IBMs in the jargon of this community, incorporating some of the basic effects: repulsion, attraction, and alignment. Let us discuss briefly some of these models, starting with the ones that have recently attracted more attention due to their simplicity while having a rich mathematical structure and pattern formation. One of these models was introduced by the UCLA group in D’Orsogna, Chuang, Bertozzi, and Chayes (2006) and it consists in Newton’s like equations where all the effect of repulsion and attraction is encoded via a pairwise potential  $W : \mathbb{R}^d \rightarrow \mathbb{R}$ . A popular choice for the interaction potential  $W$  is the Morse potential given by

$$W(x) = -C_A e^{-|x|/\ell_A} + C_R e^{-|x|/\ell_R}, \quad (2.1)$$

where  $C_A, C_R$  and  $\ell_A, \ell_R$  are the strengths and the typical lengths of attraction and repulsion, respectively. They are chosen for having biologically reasonable potentials with  $C = C_R/C_A > 1$  and  $\ell_R/\ell_A < 1$ , see Carrillo, Panferov, and Martin (2013) for other nice choices of the interaction potentials and a deeper discussion on the issue of biologically relevant interaction potentials. Apart from this, the other effect included is the tendency of the particles to travel asymptotically at a fixed speed as in Levine, Rappel, and Cohen (2000). Consequently, a term producing a balance between

self-propulsion and friction is introduced imposing an asymptotic speed to the particles (if other effects are ignored), but it does not influence the orientation vector. The resulting ODE system reads as:

$$\begin{cases} \frac{dx_i}{dt} = v_i, & (i = 1, \dots, N), \\ \frac{dv_i}{dt} = (\alpha - \beta |v_i|^2)v_i - \frac{1}{N} \sum_{j \neq i} \nabla W(|x_i - x_j|), & (i = 1, \dots, N). \end{cases}$$

where  $\alpha, \beta$  are nonnegative parameters, determining the asymptotic speed of particles given by  $\sqrt{\alpha/\beta}$ . Here, the potential has been scaled depending on the mass of each particle as in Carrillo, D'Orsogna, and Panferov (2009) and in such a way that the effect of the potential per particle diminishes while the energy is of constant order as the number of particles  $N$  diverges. This scaling is the so-called mean-field scaling, see the introduction of Bodnar and Velazquez (2012) for a nice discussion of the different scalings in first order models.

Another popular IBM including only the alignment effect is the so-called Cucker and Smale (2007) model. Each individual in the swarm changes its velocity vector based on the other individuals by adjusting/averaging their relative velocity with all the others. This averaging is weighted in such a way that closer individuals have more influence than further ones. For a system with  $N$  individuals the Cucker-Smale model reads as

$$\begin{cases} \frac{dx_i}{dt} = v_i, \\ \frac{dv_i}{dt} = \frac{1}{N} \sum_{j=1}^N w_{ij} (v_j - v_i), \end{cases}$$

with the *communication rate*  $w(x)$  given by:

$$w_{ij} = w(x_i - x_j) = \frac{1}{(1 + |x_i - x_j|^2)^\gamma},$$

for some  $\gamma \geq 0$ .

Associated to the above models, one can formally write the expected Vlasov-like kinetic equations as  $N \rightarrow \infty$ , see for instance Carrillo, D'Orsogna, and Panferov (2009), leading to

$$\partial_t f + v \cdot \nabla_x f - (\nabla W * \rho) \cdot \nabla_v f + \operatorname{div}_v((\alpha - \beta |v|^2)v f) = 0, \quad (2.2)$$

where  $\rho$  represents the macroscopic *density* of  $f$ :

$$\rho(t, x) := \int_{\mathbb{R}^d} f(t, x, v) dv \quad \text{for } t \geq 0, \quad x \in \mathbb{R}^d.$$

The Cucker-Smale particle model leads to the following kinetic equation:

$$\frac{\partial f}{\partial t} + v \cdot \nabla_x f = \nabla_v \cdot [\xi[f] f], \quad (2.3)$$

where  $\xi[f](x, v, t) = (H * f)(x, v, t)$ , with  $H(x, v) = w(x)v$  and  $*$  standing for the convolution in both position and velocity ( $x$  and  $v$ ). We refer to Cucker and Smale (2007); Ha and Tadmor (2008); Ha and Liu (2009); Carrillo, Fornasier, Rosado, and Toscani (2010) for further discussion about this model and qualitative properties.

Moreover, quite general models incorporating the three effects previously discussed with additional ingredients, such as vision cones or topological interactions, have been considered in Carrillo, Fornasier, Toscani, and Vecil (2010); Li, Lukeman, and Edelstein-Keshet (2008); Agueh, Illner, and Richardson (2011); Albi and Pareschi (2013); Haskovec (2013). In particular in Li, Lukeman, and Edelstein-Keshet (2008), they consider that the  $N$  individuals follow the system:

$$\begin{cases} \frac{dx_i}{dt} = v_i, \\ \frac{dv_i}{dt} = F_i^A + F_i^I, \end{cases} \quad (2.4)$$

where  $F_i^A$  is the self-propulsion generated by the  $i$ th-individual, while  $F_i^I$  is due to interaction with the others. The interaction with other individuals can be generally modeled as:

$$F_i^I = F_i^{I,x} + F_i^{I,v} = \sum_{j=1}^N g_{\pm}(|x_i - x_j|) \frac{x_j - x_i}{|x_i - x_j|} + \sum_{j=1}^N h_{\pm}(|v_i - v_j|) \frac{v_j - v_i}{|v_i - v_j|}.$$

Here,  $g_+$  and  $h_+$  ( $g_-$  and  $h_-$ ) are chosen when the influence comes from the front (behind), i.e., if  $(x_j - x_i) \cdot v_i > 0$  ( $< 0$ ); choosing  $g_+ \neq g_-$  and  $h_+ \neq h_-$  means that the forces from particles in front and those from particles behind are different. The sign of the functions  $g_{\pm}(r)$  encodes the short-range repulsion and long-range attraction for particles in front of (+) and behind (-) the  $i$ th-particle. Similarly,  $h_+ > 0$  ( $< 0$ ) implies that the velocity-dependent force makes the velocity of particle  $i$  get closer to (away from) that of particle  $j$ .

Some of these models, for instance Agueh, Illner, and Richardson (2011); Albi and Pareschi (2013); Haskovec (2013), include sharp boundaries for the vision cone or for the interaction with the nearest neighbors. As we shall see later, these are typical situations in which the mean-field limit for general

measures will not work. By sharp boundaries we mean that the functions involved in the kernels such as  $w(x)$ ,  $g_{\pm}$ , or  $h_{\pm}$  are given by characteristic functions on sets depending on the location/velocity of the agent.

## 2.2 Basic tools in transport distances

In this subsection, we present several definitions of Wasserstein distances and their properties.

**Definition 2.1.** (Wasserstein  $p$ -distance) Let  $\rho_1, \rho_2$  be two Borel probability measures on  $\mathbb{R}^d$ . Then the Euclidean Wasserstein distance of order  $1 \leq p < \infty$  between  $\rho_1$  and  $\rho_2$  is defined as

$$d_p(\rho_1, \rho_2) := \inf_{\gamma} \left( \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p d\gamma(x, y) \right)^{1/p},$$

and, for  $p = \infty$  (this is the limiting case, as  $p \rightarrow \infty$ ),

$$d_{\infty}(\rho_1, \rho_2) := \inf_{\gamma} \left( \gamma - \sup_{(x, y) \in \mathbb{R}^d \times \mathbb{R}^d} |x - y| \right),$$

where the infimum runs over all transference plans, i.e., all probability measures  $\gamma$  on  $\mathbb{R}^d \times \mathbb{R}^d$  with marginals  $\rho_1$  and  $\rho_2$  respectively,

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} \phi(x) d\gamma(x, y) = \int_{\mathbb{R}^d} \phi(x) \rho_1(x) dx,$$

and

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} \phi(y) d\gamma(x, y) = \int_{\mathbb{R}^d} \phi(y) \rho_2(y) dy,$$

for all  $\phi \in \mathcal{C}_b(\mathbb{R}^d)$ .

We also remind the definition of the push-forward of a measure by a mapping in order to give the relation between Wasserstein distances and optimal transportation.

**Definition 2.2.** Let  $\rho_1$  be a Borel measure on  $\mathbb{R}^d$  and  $\mathcal{T} : \mathbb{R}^d \rightarrow \mathbb{R}^d$  be a measurable mapping. Then the push-forward of  $\rho_1$  by  $\mathcal{T}$  is the measure  $\rho_2$  defined by

$$\rho_2(B) = \rho_1(\mathcal{T}^{-1}(B)) \quad \text{for } B \subset \mathbb{R}^d,$$

and denoted as  $\rho_2 = \mathcal{T} \# \rho_1$ .



The set of probability measures with bounded moments of order  $p$ , denoted by  $\mathcal{P}_p(\mathbb{R}^d)$ ,  $1 \leq p < \infty$ , is a complete metric space endowed with the  $p$ -Wasserstein distance  $d_p$ , see Villani (2003). We refer to Givens and Shortt (1984); McCann (2006) for more details in the case of the  $d_\infty$  distance.

**Remark 2.3.** The definition of  $\rho_2 = \mathcal{T}\#\rho_1$  is equivalent to

$$\int_{\mathbb{R}^d} \phi(x) d\rho_2(x) = \int_{\mathbb{R}^d} \phi(\mathcal{T}(x)) d\rho_1(x),$$

for all  $\phi \in \mathcal{C}_b(\mathbb{R}^d)$ . Given a probability measure with bounded  $p$ -th moment  $\rho_0$ , consider two measurable mappings  $X_1, X_2 : \mathbb{R}^d \rightarrow \mathbb{R}^d$ , then the following inequality holds.

$$d_p^p(X_1\#\rho_0, X_2\#\rho_0) \leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^p d\gamma(x, y) = \int_{\mathbb{R}^d} |X_1(x) - X_2(x)|^p d\rho_0(x).$$

Here, we used as transference plan  $\gamma = (X_1 \times X_2)\#\rho_0$  in Definition 2.1.

### 2.3 A quick review of the classical Dobrushin result

Under smoothness assumptions on the ingredient functions of the swarming models, one can use adaptations of the classical result of Dobrushin (1979) to obtain what is called the mean-field limit equation for general particle approximations of any initial measure. These arguments are classical in kinetic theory and were also introduced in Braun and Hepp (1977); Neunzert (1984), making use of the bounded Lipschitz distance, and reviewed in Spohn (1991); Villani (2002), see also Sznitman (1991); Méléard (1996) for the case with noise. The bounded Lipschitz distance or dual  $W^{1,\infty}$ -norm is equivalent to the Wasserstein distance  $d_1$  for compactly supported measures. This strategy works as soon as the velocity field defining the characteristics of the model is a bounded and globally Lipschitz function whose dependence on the measure itself is Lipschitz continuous in the  $d_1$  sense. These ideas were improved to allow for locally Lipschitz velocity fields for compactly supported initial measures in Cañizo, Carrillo, and Rosado (2011) and for suitable decay conditions at infinity and with noise in Bolley, Cañizo, and Carrillo (2011). With these techniques one can include quite general kinetic models for swarming in this well-posedness theory.

Let us introduce some notation for this section:  $\mathcal{A} = \mathcal{P}_c(\mathbb{R}^d \times \mathbb{R}^d)$  denotes the subset of  $\mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$  consisting of measures of compact support in  $\mathbb{R}^d \times \mathbb{R}^d$ . On the other hand, we consider the set of functions  $\mathcal{B} := \text{Lip}_{loc}(\mathbb{R}^d \times \mathbb{R}^d)$ , which in particular are locally Lipschitz with respect to  $(x, v)$ .  $B_R$  will denote the ball centered at 0 of radius  $R$  in  $\mathbb{R} \times \mathbb{R}$ .

Let us consider general operators from measures to vector fields,  $\mathcal{H}[\cdot] : \mathcal{A} \rightarrow \mathcal{B}$ , satisfying the following hypotheses: for any  $R_0 > 0$  and  $f, g \in \mathcal{A}$  such that  $\text{supp } f \cup \text{supp } g \subseteq B_{R_0}$ , there exists some ball  $B_R \subset \mathbb{R}^d \times \mathbb{R}^d$  and a constant  $C = C(R, R_0) > 0$ , such that

$$\|\mathcal{H}[f] - \mathcal{H}[g]\|_{L^\infty(B_R)} \leq C d_1(f, g), \quad (2.5)$$

$$\text{Lip}_R(\mathcal{H}[f]) \leq C, \quad \|\mathcal{H}[f]\|_{L^\infty(B_R)} \leq C. \quad (2.6)$$

Here,  $\text{Lip}_R(\cdot)$  denotes the Lipschitz constant of a function in  $B_R$ .

Given  $f \in \mathcal{C}([0, T], \mathcal{P}_c(B_{R_0}))$ , and for any initial condition  $(X^0, V^0) \in \mathbb{R}^d \times \mathbb{R}^d$ , the following system of ordinary differential equations has a unique locally defined solution thanks to conditions (2.5)

$$\frac{d}{dt} X = V, \quad X(0) = X^0 \quad (2.7a)$$

$$\frac{d}{dt} V = \mathcal{H}[f(t)](X, V), \quad V(0) = V^0. \quad (2.7b)$$

We will additionally require that the solutions to that system are global. Of course, this is a requirement that has to be checked for every particular model. We prefer to give a general condition which reduces the problem of existence and stability to the simpler one of existence of the ODEs. Under the above conditions, the existence and uniqueness of associated transport equation

$$\partial_t f + v \cdot \nabla_x f - \nabla_v \cdot [\mathcal{H}[f]f] = 0. \quad (2.8)$$

was obtained in Cañizo, Carrillo, and Rosado (2011) to which we refer for full details. In Cañizo, Carrillo, and Rosado (2011), the interactions  $\mathcal{H}[f] = (\alpha - \beta|v|^2)v - \nabla W * \rho$  and  $\mathcal{H}[f] = H * f$  corresponding to (2.2) and (2.3), respectively, and

$$\mathcal{H}[f] = F_A(x, v) + G(x) * \rho + H(x, v) * f,$$

with  $F_A$ ,  $G$  and  $H$  given functions satisfying suitable hypotheses, such that the kinetic equation (2.8) corresponds to the model (2.4) are investigated.

**Theorem 2.4.** *Given an operator  $\mathcal{H}[\cdot] : \mathcal{A} \rightarrow \mathcal{B}$  satisfying Hypotheses (2.5) and (2.6) for which the characteristics (2.7a)-(2.7b) are globally well-defined, and  $f_0$  a measure on  $\mathbb{R}^d \times \mathbb{R}^d$  with compact support. There exists a solution  $f$  on  $[0, +\infty)$  to equation (2.8) with initial condition  $f_0$ . In addition,*

$$f \in \mathcal{C}([0, +\infty); \mathcal{P}_c(\mathbb{R}^d \times \mathbb{R}^d)) \quad (2.9)$$

and there is some increasing function  $R = R(T)$  such that for all  $T > 0$ ,

$$\text{supp } f_t \subseteq B_{R(T)} \subseteq \mathbb{R}^d \times \mathbb{R}^d \quad \text{for all } t \in [0, T]. \quad (2.10)$$

This solution is unique among the family of solutions satisfying (2.9) and (2.10). Moreover, given any other initial data  $g_0 \in \mathcal{P}_c(\mathbb{R}^d \times \mathbb{R}^d)$  and  $g$  its corresponding solution, then there exists a strictly increasing function  $r(t) : [0, \infty) \rightarrow \mathbb{R}_0^+$  with  $r(0) = 1$  depending only on  $\mathcal{H}$  and the size of the support of  $f_0$  and  $g_0$ , such that

$$d_1(f_t, g_t) \leq r(t) d_1(f_0, g_0), \quad t \geq 0.$$

The stability theorem 2.4 gives in particular a rigorous derivation of the kinetic equation (2.8) from the large particle limit of the system of ordinary differential equations. This is the exact statement of the mean-field limit for general measures as initial data. Let us consider the system of ordinary differential equations:

$$\dot{x}_i = v_i, \quad i = 1, \dots, N, \quad (2.11a)$$

$$\dot{v}_i = \sum_{j \neq i} m_j \mathcal{H}[f^N(t)](x_i, v_i), \quad i = 1, \dots, N. \quad (2.11b)$$

where  $m_1, \dots, m_N \geq 0$  and  $\sum_i m_i = 1$  and  $f^N$  is defined next. Under the conditions of Theorem 2.4, we first notice that if  $x_i, v_i : [0, T] \rightarrow \mathbb{R}^d$ , for  $i = 1, \dots, N$ , are a solution to the system (2.11), then the function  $f^N : [0, T] \rightarrow \mathcal{P}_c(\mathbb{R}^d \times \mathbb{R}^d)$  given by

$$f_t^N := \sum_{i=1}^N m_i \delta_{(x_i(t), v_i(t))} \quad (2.12)$$

is the solution to (2.8) with initial condition

$$f_0^N = \sum_{i=1}^N m_i \delta_{(x_i(0), v_i(0))}. \quad (2.13)$$

In fact, the solution (2.12) is called the empirical measure associated to the system of ODEs (2.11). We finally write the full statement of the mean-field limit in the Dobrushin strategy.

**Corollary 2.5.** *Given  $f_0 \in \mathcal{P}_c(\mathbb{R}^d \times \mathbb{R}^d)$  and  $\mathcal{H}[f]$  satisfying the conditions of Theorem 2.4, take a sequence of  $f_0^N$  of measures of the form (2.13) (with  $m_i, x_i(0)$  and  $v_i(0)$  possibly varying with  $N$ ), in such a way that*

$$\lim_{N \rightarrow \infty} d_1(f_0^N, f_0) = 0.$$

Consider  $f_t^N$  the empirical measure associated to the solution of the system (2.11) with initial conditions  $x_i(0), v_i(0)$ . Then,

$$\lim_{N \rightarrow \infty} d_1(f_t^N, f_t) = 0, \quad (2.14)$$

for all  $t \geq 0$ , where  $f = f(t, x, v)$  is the unique measure solution to eq. (2.8) with initial data  $f_0$ .

This section can be directly applied to the models recently introduced in Agueh, Illner, and Richardson (2011) to account for vision cones and braking/acceleration of individuals and those in Albi and Pareschi (2013); Haskovec (2013) to include topological (nearest neighbours) interactions once the parameter functions are smoothed out to avoid sharp boundaries.

## 2.4 First-order models: Aggregation Equation

Summarizing the previous subsection, under suitable smoothness of the parameters involved in the swarming models, the empirical measures are solutions themselves of the Vlasov-like kinetic equation (2.8). Thus, a stability result in  $d_1$  with respect to the initial data is enough to conclude the mean-field limit. Let us consider one of the particular examples in subsection 2.1, the model introduced in D'Orsogna, Chuang, Bertozzi, and Chayes (2006) with the Morse potential (2.1). This potential does not satisfy the smoothness assumption in Theorem 2.4. In principle, one cannot expect to have a mean-field result for general measures as initial data and for general approximations by particles. In fact, we do not have a well-posedness theory for such initial data in those cases. However, one can develop well-posedness theories in better functional spaces, say  $L^1 \cap L^p(\mathbb{R}^d \times \mathbb{R}^d)$  for the initial data and then impose suitable conditions to the distribution of the approximated particles initially to be able to conclude the mean-field limit (2.14). This is the strategy that have been followed in Hauray and Jabin (2012) for the classical Vlasov equation and in Hauray (2009) for Euler-like equations in fluid mechanics.

In the next sections, the objective is to show this strategy applied to a simpler swarming model than the ones showed above. We will showcase these tools in the case of the so-called aggregation equation. Let us assume that we have just particles interacting through the pairwise potential  $W(x)$ . Assuming that the variations of the velocity and speed are much smaller than spatial variations, see Mogilner and Edelstein-Keshet (1999), then one can neglect the inertia term in Newton's equation to deduce that

$$\frac{dX_i}{dt} = - \sum_{j \neq i} \nabla W(X_i - X_j) \text{ in the } N \rightarrow \infty \text{ limit} \Rightarrow \begin{cases} \frac{\partial \rho}{\partial t} + \operatorname{div}(\rho u) = 0 \\ u = -\nabla W * \rho \end{cases} .$$

Another reason to study this first order equation is that the stationary states of the first order model determine the spatial shape of the flock solutions to the second order models, see Carrillo, Panferov, and Martin (2013). Let us remark that one could apply the Dobrushin strategy to the aggregation equation for  $\mathcal{C}^2(\mathbb{R}^d)$  smooth potential with at most quadratic growth at infinity by following the same argument as in Theorem (2.4). This argument was detailed in a nice summer school notes in Golse (2003).

### 3 Mean-Field Limit for the Aggregation Equation

Now, we analyse the mean-field limit of the first order model for swarming introduced in the previous section. More precisely, we will study sufficient conditions on the initial distribution of particles for the convergence of a particle system towards the aggregation equation. This model consists of the continuity equation for the probability density of individuals  $\rho(x, t)$  at position  $x \in \mathbb{R}^d$  and time  $t > 0$  given by:

$$\begin{cases} \partial_t \rho + \nabla \cdot (\rho u) = 0, & t > 0, \quad x \in \mathbb{R}^d, \\ u(t, x) := -\nabla W * \rho, & t > 0, \quad x \in \mathbb{R}^d, \\ \rho(0, x) := \rho^0(x), & x \in \mathbb{R}^d, \end{cases} \quad (3.15)$$

where  $u(x, t)$  is velocity field non-locally computed in terms of the density of individuals.

As an approximation by particles of the aggregation equation (3.15), we consider the following ODE system:

$$\begin{cases} \dot{X}_i(t) = -\sum_{j \neq i} m_j \nabla W(X_i(t) - X_j(t)), \\ X_i(0) = X_i^0, \quad i = 1, \dots, N. \end{cases} \quad (3.16)$$

Here,  $\{X_i\}_{i=1}^N$  and  $\{m_i\}_{i=1}^N$  are the positions and weights of  $i$ -th particles, respectively. We define the associated empirical distribution  $\mu_N(t)$  as

$$\mu_N(t) = \sum_{i=1}^N m_i \delta_{X_i(t)}, \quad \sum_{i=1}^N m_i = \int_{\mathbb{R}^d} \rho_0(x) dx = 1, \quad (3.17)$$

with  $m_i > 0$ ,  $i = 1, \dots, N$ . As long as two particles (or more) do not collide, and if we set  $\nabla W(0) = 0$  (arbitrarily if there is a singularity), then  $\mu_N$  satisfies (3.15) in the sense of distributions, i.e.,  $\mu_N(t)$  and  $\rho(t)$  satisfy the same equation. In this framework, the convergence:

*“ $\mu_N^0 \rightharpoonup \rho^0$  weakly-\* as measures  $\implies \mu_N(t) \rightharpoonup \rho(t)$  weakly-\* as measures for small time or for every time?”*

is a natural question. If the answer is yes, we say that the continuity equation (3.15) is the mean-field limit of the particle approximation (3.16). In other words, we can say that the continuum nonlocal equation (3.15) has been rigorously derived from particle systems.

Because of the singularity in the interaction force, the natural transport distance to use is the one induced by the  $d_\infty$ -topology. Remark that this distance also allows to understand linearized stability of particle systems around singular steady state measures with a ring shape in first order aggregation models, see Balagué, Carrillo, Laurent, and Raoul (2013b); Kolokonikov, Sun, Uminsky, and Bertozzi (2011). Actually, a local perturbation of the dynamical system (3.16) keeping the number of particles fixed is obtained by transporting the particle to other locations nearby. One could even allow for splitting of the mass into different particles, but all of them located in a local neighborhood of the unperturbed particle positions. Certainly, sending a small portion of mass very far away from the location of one particle is not a  $d_\infty$ -perturbation of the atomic measure but it is a  $d_p$  small perturbation for all  $1 \leq p < \infty$ . These ideas have also recently been used in Balagué, Carrillo, Laurent, and Raoul (2013b) to study local minimizers of the energy functional associated to (3.15).

Another issue to cope with is that we are dealing with particle systems whose characteristics may lead to collisions in finite time. Therefore, we will be able to obtain meaningful results only on intervals in which collisions are avoided (although in some particular cases we can allow collisions).

We next introduce several notations that are used throughout the rest of this work to compare the distance between a solution  $\rho(t)$  of the continuum aggregation equation (3.15) and the empirical measure  $\mu_N(t)$  defined by (3.17) associated to a solution  $\{X_i\}_{i=1}^N$  of the particle system (3.16). The main two quantities appearing in this comparison are the  $d_\infty$ -distance between  $\rho(t)$  and  $\mu_N(t)$ , and the minimum inter-particle distance:

$$\eta(t) := d_\infty(\mu_N(t), \rho(t)), \quad \eta_m(t) := \min_{1 \leq i \neq j \leq N} (|X_i(t) - X_j(t)|), \quad (3.18)$$

with  $\eta^0 := \eta(0)$  and  $\eta_m^0 := \eta_m(0)$ . Our strategy does not take advantage, as we do not know how, of the repulsive or attractive character of the potentials, being the proof equal for both cases.

A theory of well-posedness for measure solutions have been obtained for the aggregation equation (3.15) allowing collision of particles in finite time in Carrillo, Di Francesco, Figalli, Laurent, and Slepčev (2011); Carrillo, Di Francesco, Figalli, Laurent, and Slepčev (2012). In these works, the potential is assumed to be smooth except at the origin, where the allowed singularity cannot be worse than Lipschitz and the potential has to be  $\lambda$ -convex, see Carrillo, Di Francesco, Figalli, Laurent, and Slepčev (2011) for

details. This convexity allows for attractive at the origin potentials, but not repulsive, with local behaviors of the form  $|x|^b$  with  $1 \leq b < 2$ . In these works, the essential tools that allow to get the mean-field limit for more singular potentials that quadratic are based on gradient flows in the Wasserstein distance  $d_2$  sense as in Ambrosio, Gigli, and Savaré (2005). The additional dissipation in the system of the natural Liapunov functional given by the total interaction energy is crucial to get the mean field limit for general measures for a potential behaving locally at 0 like  $W(x) \simeq |x|$ , for instance for the attractive Morse potential  $W(x) = 1 - e^{-|x|}$ .

In this work, we want to allow for more singular potentials at the origin as in Bertozzi, Carrillo, and Laurent (2009); Bertozzi, Laurent, and Rosado (2010), and thus we need to work with solutions in better functional spaces. More precisely, we will work with solutions of the aggregation equation (3.15) in  $L^\infty(0, T; (L^1 \cap L^p)(\mathbb{R}^d))$  with  $1 \leq p \leq \infty$  to be determined depending on the singularity of the potential. We will use the notation

$$\|\rho\|_{(L^1 \cap L^p)(\mathbb{R}^d)} := \|\rho\|_1 + \|\rho\|_p, \quad \|\rho\| := \|\rho\|_{L^\infty(0, T; (L^1 \cap L^p)(\mathbb{R}^d))},$$

where  $\|\rho\|_p$  denotes the  $L^p(\mathbb{R}^d)$ -norm of  $\rho$ ,  $1 \leq p \leq \infty$ .

In order to make sense of solutions to (3.15), we need the following assumptions on the interaction potential: we first fix  $W(0) = 0$  by definition, even if  $W$  is singular at the origin, and

$$|\nabla W(x)| \leq \frac{C}{|x|^\alpha}, \quad \text{and} \quad |D^2 W(x)| \leq \frac{C}{|x|^{1+\alpha}}, \quad \forall x \in \mathbb{R}^d \setminus \{0\}, \quad (3.19)$$

for  $-1 \leq \alpha < d-1$ . Note that due to the assumptions on  $W$ , we can always find  $1 < p < \infty$  such that  $(\alpha+1)p' < d$ , and thus  $\nabla W$  belongs to  $\mathcal{W}_{loc}^{1, p'}(\mathbb{R}^d)$ .

Our results also apply with minor modifications for interaction potentials of the form  $W := W_1 + W_2$ , with  $W_1$  satisfying assumptions (3.19), and  $\nabla W_2$  being a global Lipschitz function, or even more general satisfying a one-sided Lipschitz (or convexity) condition  $y \cdot D^2 W_2(x)y \leq C|y|^2$  for all  $y \in \mathbb{R}^d$ . This last generalization is important because it is satisfied if  $W_2 = c|x|^a$ , ( $0 \leq a \leq 2$ ) with  $c$  positive. So that any repulsive-attractive potential  $W$ , see Balagué, Carrillo, Laurent, and Raoul (2013a,b) for a definition, such that  $W(x) \simeq -|x|^b/b$  locally at  $x$  near the origin, satisfies assumptions (3.19) locally with  $\alpha = 1 - b$ . Therefore, our mean-field limit results apply to locally repulsive potentials with exponent range  $2 - d < b < a \leq 2$  and without much restriction on the attractive part at  $+\infty$ , i.e.,  $a > 0$ . We will discuss further on localizing assumptions (3.19) at the end of this section. Finally, we cannot apply our techniques to the Newtonian singularity Bertozzi, Laurent, and Léger (2012) being the limiting case of

our strategy as it was the case for the Euler-like models in fluid mechanics studied in Hauray (2009).

We next summarize the results on the existence and uniqueness of solutions to the aggregation equation (3.15). For the local well-posedness of solutions to equation (3.15), we refer to Bertozzi and Laurent (2007); Bertozzi, Carrillo, and Laurent (2009); Bertozzi, Laurent, and Rosado (2010); Laurent (2007). In particular, unique solutions for the system (3.15) were obtained in Bertozzi, Laurent, and Rosado (2010) with second moment bounded initial data. More precisely, Bertozzi et al. (Bertozzi, Laurent, and Rosado, 2010, Theorem 1.1) showed that if  $\nabla W \in \mathcal{W}^{1,p'}(\mathbb{R}^d)$  and  $\rho^0 \in L^p(\mathbb{R}^d) \cap \mathcal{P}_2(\mathbb{R}^d)$ , then there exists  $T^* > 0$  and a unique nonnegative solution to (3.15) satisfying

$$\rho \in \mathcal{C}([0, T^*], (L^1 \cap L^p)(\mathbb{R}^d)) \cap \mathcal{C}^1([0, T^*], \mathcal{W}^{-1,p}(\mathbb{R}^d)).$$

Unfortunately, one can not directly apply those results for potentials satisfying assumptions (3.19). We will complement the results in Bertozzi, Laurent, and Rosado (2010) to show the local existence of a unique solution to the system (3.15) with the interaction potential function  $W$  satisfying (3.19) in Section 4. We prefer to postpone the well-posedness theory in order to emphasize the mean-field limit result contained in the following theorem, whose proof follows the strategy in Hauray (2009).

**Theorem 3.1.** *Suppose the kernel  $W$  satisfies (3.19), and let  $\rho$  be a solution to the system (3.15) up to time  $T > 0$ , such that  $\rho \in L^\infty(0, T; (L^1 \cap L^p)(\mathbb{R}^d)) \cap \mathcal{C}([0, T], \mathcal{P}_1(\mathbb{R}^d))$ , with initial data  $\rho^0 \in (\mathcal{P}_1 \cap L^p)(\mathbb{R}^d)$ ,  $0 \leq \alpha < -1 + d/p'$ , and  $1 < p \leq \infty$ . Furthermore, we assume  $\mu_N^0$  converges to  $\rho^0$  for the distance  $d_\infty$  as the number of particles  $N$  goes to infinity, i.e.,*

$$d_\infty(\mu_N^0, \rho^0) \rightarrow 0 \quad \text{as } N \rightarrow \infty,$$

and that the initial quantities  $\eta^0, \eta_m^0$  satisfy

$$\lim_{N \rightarrow \infty} \frac{(\eta^0)^{d/p'}}{(\eta_m^0)^{1+\alpha}} = 0. \quad (3.20)$$

Then, for  $N$  large enough the particle system (3.16) is well-defined up to time  $T$ , in the sense that there is no collision between particles before that time, and moreover

$$\mu_N(t) \rightharpoonup \rho(t) \quad \text{weakly-}^* \text{ as measures as } N \rightarrow \infty, \quad \text{for all } t \in [0, T].$$

**Remark 3.2.** Let us first discuss the assumptions on the initial data in Theorem 3.1. The mean-field limit is valid for particular approximations



$\mu_N^0$  of  $\rho^0$ , that is, for well chosen particle approximations of the initial data. In fact, a procedure to construct initial atomic measures approximating the initial condition in the sense of (3.20) is the following: define a regular mesh of size  $\varepsilon$  and approximate  $\rho^0$  by a sum of Dirac masses  $\mu_N^0$  located at the center of the cells such that the mass at each particle is exactly equals to the mass of  $\rho^0$  contained in the associated cell. In that case, we have  $\eta^0 \sim \varepsilon$  and  $\eta_m^0 \sim \varepsilon$  (for the last condition we need that the mesh has some regularity). In that case, the assumption (3.20) is automatically fulfilled since  $(1 + \alpha)p' < d$ . Notice that no bound on the masses  $m_i$  of the particles is required.

*Proof of Theorem 3.1.* The proof of Theorem 3.1 is divided into three steps:

- In Step A, we estimate the growth of the  $d_\infty$  Wasserstein distance between the continuum and the discrete solutions  $\eta$  that involves  $\eta$  itself and  $\eta_m$  in the form:

$$\frac{d\eta}{dt} \leq C\eta\|\rho\| \left(1 + \eta^{d/p'} \eta_m^{-(1+\alpha)}\right). \quad (3.21)$$

- In Step B, we estimate the decay of the minimum inter-particle distance  $\eta_m$ , which also involves the terms  $\eta$  and  $\eta_m$  in the form:

$$\frac{d\eta_m}{dt} \geq -C\eta_m\|\rho\| \left(1 + \eta^{d/p'} \eta_m^{-(1+\alpha)}\right). \quad (3.22)$$

- In Step C, under the assumption of the initial approximation (3.20), we combine (3.21) and (3.22) to conclude the desired result.

**Step A.-** We first introduce the flows generated by the two velocity fields:  $u(x, t) = -\nabla W * \rho$  and  $u_N := -\nabla W * \mu_N$ . Let us remark that the convolution in the definition of  $u_N$  is just a notation for the right-hand side of (3.16) since the convolution of a Dirac Delta with a (possibly) singular potential is not well-defined. These flows  $\Psi_N, \Psi : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  are defined as solutions of

$$\begin{cases} \frac{d}{dt}(\Psi(t; s, x)) = u(t; s, \Psi(t; s, x)), \\ \Psi(s; s, x) = x, \end{cases} \quad (3.23)$$

for all  $s, t \in [0, T]$ , and

$$\begin{cases} \frac{d}{dt}(\Psi_N(t; s, x)) = u_N(t; s, \Psi_N(t; s, x)), \\ \Psi_N(s; s, x) = x, \end{cases} \quad (3.24)$$

for all  $s, t \in [0, T_0^N]$ . Notice that the solution  $X_i(t)$  to the system (3.16) is well-defined and continuous by the Cauchy-Lipschitz theorem as long as there is no collision between particles. Since  $\eta_m^0 > 0$ , there exists  $T_0^N > 0$  such that  $\eta_m(t) > 0$  for  $t \in [0, T_0^N]$  by continuity. Then the flow map  $\Psi_N(t; s, x)$  solution to (3.24) is well-defined for  $t, s \in [0, T_0^N]$ . Now, let us check that the flow for the solution associated to the continuum equation in (3.23) is well-defined. Assumptions (3.19) imply that

$$|\nabla W(x) - \nabla W(y)| \leq \frac{2|x - y|}{\min(|x|, |y|)^{\alpha+1}}. \quad (3.25)$$

One can see this by integrating along a straight line joining  $x$  and  $y$  but avoiding the singularity using a small circle if needed, see Hauray (2009). The estimate (3.25) implies that the velocity field is Lipschitz continuous with respect to the spatial variable. Actually, one can estimate it as

$$\begin{aligned} |u(t, x) - u(t, y)| &\leq \int_{\mathbb{R}^d} |\nabla W(x - z) - \nabla W(y - z)| \rho(t, z) dz \\ &\leq 2|x - y| \int_{\mathbb{R}^d} \frac{1}{\min(|x - z|, |y - z|)^{\alpha+1}} \rho(t, z) dz \\ &\leq 4|x - y| \sup_{x \in \mathbb{R}^d} \int_{\mathbb{R}^d} \frac{1}{|x - z|^{\alpha+1}} \rho(t, z) dz. \end{aligned}$$

Now, splitting the last integral into the near- and far-field sets  $\mathcal{A} := \{z : |x - z| \geq 1\}$  and  $\mathcal{B} := \mathbb{R}^d - \mathcal{A}$  and estimating the two terms, we deduce

$$\begin{aligned} \int_{\mathbb{R}^d} \frac{1}{|x - z|^{\alpha+1}} \rho(t, z) dz &\leq \|\rho(t)\|_1 + \left( \int_{\mathcal{B}} \frac{1}{|x - y|^{(1+\alpha)p'}} dy \right)^{1/p'} \|\rho(t)\|_p \\ &\leq C\|\rho\|, \end{aligned} \quad (3.26)$$

for all  $x \in \mathbb{R}^d$  due to the assumption  $(1+\alpha)p' < d$ . Putting together previous inequalities, we get the desired Lipschitz continuity of the velocity field with respect to  $x$ , which is moreover uniform in time. A similar estimate using (3.19) shows that the velocity field is bounded, and then the flow  $\Psi$  in (3.23) is well-defined. Our first aim is to find an expansion of the velocity of the  $d_\infty$  Wasserstein distance. The idea is similar to the evolution of the euclidean Wasserstein distance in Carrillo, McCann, and Villani (2003, 2006); Otto (2001). Fixed  $0 \leq t_0 < \min(T, T_0^N)$  and choose an optimal transport map for  $d_\infty$  denoted by  $\mathcal{T}^0$  between  $\rho(t_0)$  and  $\mu_N(t_0)$ ;  $\mu_N(t_0) = \mathcal{T}^0 \# \rho(t_0)$ . It is known that such an optimal transport map exists when  $\rho(t_0)$  is absolutely continuous with respect to the Lebesgue measure Champion, Pascale, and Juutinen (2008). Then it follows from Theorem 4.1 that  $\rho(t) =$

$\Psi(t; t_0, \cdot) \# \rho(t_0)$  and obviously  $\mu_N(t) = \Psi_N(t; t_0, \cdot) \# \mu_N(t_0)$  for  $t \geq t_0$ . We also notice that for  $t \geq t_0$

$$\mathcal{T}^t \# \rho(t) = \mu_N(t), \quad \text{where} \quad \mathcal{T}^t = \Psi_N(t; t_0, \cdot) \circ \mathcal{T}^0 \circ \Psi(t_0; t, \cdot).$$

By Definition 2.1 of the  $d_p$  Wasserstein distance, we get

$$d_p^p(\mu_N(t), \rho(t)) \leq \int_{\mathbb{R}^d} |\Psi(t; t_0, x) - \Psi_N(t; t_0, \mathcal{T}^0(x))|^p \rho(t_0, x) dx.$$

In the case of  $p = \infty$ , we obtain

$$\eta(t) = d_\infty(\mu_N(t), \rho(t)) \leq \|\Psi(t; t_0, \cdot) - \Psi_N(t; t_0, \cdot) \circ \mathcal{T}^0\|_\infty.$$

We notice that

$$\left. \frac{d}{dt} (\Psi_N(t; t_0, \mathcal{T}^0(x)) - \Psi(t; t_0, x)) \right|_{t=t_0} = u_N(t_0, \mathcal{T}^0(x)) - u(t_0, x).$$

Thus, writing the integral form, dividing by  $t - t_0$ , and taking the limit  $t \rightarrow t_0^+$  we easily get

$$\left. \frac{d}{dt} \|\Psi_N(t; t_0, \cdot) \circ \mathcal{T}^0 - \Psi(t; t_0, \cdot)\|_\infty \right|_{t=t_0^+} \leq \|u_N(t_0, \cdot) \circ \mathcal{T}^0 - u(t_0, \cdot)\|_\infty. \quad (3.27)$$

We now note that

$$\begin{aligned} u_N(t_0, \mathcal{T}^0(x)) - u(t_0, x) &= - \int_{\mathbb{R}^d} \nabla W(\mathcal{T}^0(x) - y) d\mu_N(t_0, y) + \int_{\mathbb{R}^d} \nabla W(x - y) \rho(t_0, y) dy \\ &= - \int_{\mathbb{R}^d} (\nabla W(\mathcal{T}^0(x) - \mathcal{T}^0(y)) - \nabla W(x - y)) \rho(t_0, y) dy. \end{aligned}$$

For notational simplicity, we omit the time dependency on  $t_0$  in the next few computations. This yields that (3.27) can be rewritten as

$$\frac{d^+ \eta}{dt} \leq C \sup_{x \in \mathbb{R}^d} \int_{\mathbb{R}^d} |\nabla W(\mathcal{T}(x) - \mathcal{T}(y)) - \nabla W(x - y)| \rho(y) dy. \quad (3.28)$$

We decompose the integral on  $\mathbb{R}^d$  into the near- and the far-field parts as  $\mathcal{A} := \{z : |x - z| \geq 4\eta\}$  and  $\mathcal{B} := \mathbb{R}^d - \mathcal{A}$  as

$$\begin{aligned} \int_{\mathbb{R}^d} |\nabla W(\mathcal{T}(x) - \mathcal{T}(y)) - \nabla W(x - y)| \rho(y) dy &= \int_{\mathcal{A}} \cdots + \int_{\mathcal{B}} \cdots \\ &:= \mathcal{I}_1 + \mathcal{I}_2. \end{aligned} \quad (3.29)$$

For the estimate in the set  $\mathcal{A}$ , we use

$$|\mathcal{T}(x) - \mathcal{T}(y)| \geq |x - y| - |\mathcal{T}(x) - x| - |\mathcal{T}(y) - y| \geq |x - y| - 2\eta \geq \frac{|x - y|}{2}$$

together with (3.25) and (3.26) to obtain

$$\begin{aligned} \mathcal{I}_1 &\leq \int_{\mathcal{A}} \frac{2(|x - \mathcal{T}(x)| + |y - \mathcal{T}(y)|)}{\min(|x - y|, |\mathcal{T}(x) - \mathcal{T}(y)|)^{\alpha+1}} \rho(y) dy \\ &\leq 4\eta \int_{\mathcal{A}} \left( \frac{1}{|x - y|^{\alpha+1}} + \frac{2^{\alpha+1}}{|x - y|^{\alpha+1}} \right) \rho(y) dy \leq C\eta \int_{\mathcal{A}} \frac{1}{|x - y|^{\alpha+1}} \rho(y) dy \\ &\leq C\eta \int_{\mathbb{R}^d} \frac{1}{|x - y|^{\alpha+1}} \rho(y) dy \leq C\eta \|\rho\|. \end{aligned} \quad (3.30)$$

For the second part  $\mathcal{I}_2$ , we estimate separately each term using (3.19) to deduce

$$\begin{aligned} \mathcal{I}_2 &\leq \int_{\mathcal{B}} \frac{\rho(y)}{|x - y|^\alpha} dy + \int_{\mathcal{B}} \frac{\rho(y)}{\eta_m^\alpha} dy \\ &\leq \left( \int_{\mathcal{B}} \frac{1}{|x - y|^{\alpha p'}} dy \right)^{1/p'} \|\rho\|_p + \frac{1}{\eta_m^\alpha} \left( \int_{\mathcal{B}} 1 dy \right)^{1/p'} \|\rho\|_p \\ &\leq C(\eta^{d/p' - \alpha} + \eta^{d/p'} \eta_m^{-\alpha}) \|\rho\|_p \leq C(\eta^{d/p' - \alpha} + \eta^{d/p'} \eta_m^{-\alpha}) \|\rho\|. \end{aligned} \quad (3.31)$$

Notice that  $|\mathcal{T}(x) - \mathcal{T}(y)| \geq \eta_m$  by definition of the minimum inter-particle distance (3.18) as soon as  $\mathcal{T}(x) \neq \mathcal{T}(y)$ ,  $\nabla W(\mathcal{T}(x) - \mathcal{T}(y)) = 0$  otherwise.

Finally, we choose two indices  $i, j$  so that  $|X_i - X_j| = \eta_m$ , then we observe that the middle point between  $X_i$  and  $X_j$  has to be transported by  $\mathcal{T}$  to either  $X_i$  or  $X_j$ , and thus  $\eta_m \leq 2\eta$ . Hence by combining (3.28)-(3.31) and being  $t_0$  arbitrary in  $[0, \min(T, T_0^N))$ , we have

$$\frac{d^+ \eta}{dt} \leq C\eta \|\rho\| \left( 1 + \eta^{d/p' - 1} \eta_m^{-\alpha} \right) \leq C\eta \|\rho\| \left( 1 + \eta^{d/p'} \eta_m^{-(1+\alpha)} \right), \quad (3.32)$$

for all  $t \in [0, \min(T, T_0^N))$ .

**Step B.-** We now focus on showing the lower bound estimate of  $\eta_m$  to make the system (3.32) closed. We again choose two indices  $i, j$  so that  $|X_i - X_j| = \eta_m$ . Neglecting the time dependency to simplify the notation,

we get

$$\begin{aligned}
\frac{d}{dt}|X_i - X_j| &\geq -|u_N(X_i) - u_N(X_j)| \\
&\geq - \int_{\mathbb{R}^d} |\nabla W(X_i - y) - \nabla W(X_j - y)| d\mu_N(y) \\
&= - \int_{\mathbb{R}^d} |\nabla W(X_i - \mathcal{T}(y)) - \nabla W(X_j - \mathcal{T}(y))| \rho(y) dy,
\end{aligned}$$

where  $\mathcal{T}$  is the optimal map satisfying  $\mu_N(t) = \mathcal{T}\#\rho(t)$ , for each  $t \in [0, \min(T, T_0^N))$ . Similar to (3.29), we split in near- and far-field parts the domain  $\mathbb{R}^d$  as  $\mathcal{A} := \{y : |X_i - y| \geq 2\eta \text{ and } |X_j - y| \geq 2\eta\}$  and  $\mathcal{B} := \mathbb{R}^d - \mathcal{A}$ . We can again use (3.25) to deduce

$$\begin{aligned}
&\int_{\mathcal{A}} |\nabla W(X_i - \mathcal{T}(y)) - \nabla W(X_j - \mathcal{T}(y))| \rho(y) dy \tag{3.33} \\
&\leq \int_{\mathcal{A}} \frac{2|X_i - X_j|}{\min(|X_i - \mathcal{T}(y)|, |X_j - \mathcal{T}(y)|)^{\alpha+1}} \rho(y) dy \\
&\leq 2^{2+\alpha}|X_i - X_j| \int_{\mathcal{A}} \left( \frac{1}{|X_i - y|^{\alpha+1}} + \frac{1}{|X_j - y|^{\alpha+1}} \right) \rho(y) dy \leq C\eta_m \|\rho\|,
\end{aligned}$$

where we used that  $|X_i - \mathcal{T}(y)| \geq |X_i - y| - \eta \geq \frac{1}{2}|X_i - y|$  and similarly for  $X_j$  together with (3.26). For the integral over  $\mathcal{B}$ , we use that as soon as  $X_i \neq \mathcal{T}(y)$ , then we obtain from (3.19) that

$$|\nabla W(X_i - \mathcal{T}(y))| \leq \frac{1}{|X_j - \mathcal{T}(y)|^\alpha} \leq \frac{1}{\eta_m^\alpha},$$

and  $\nabla W(X_i - \mathcal{T}(y)) = 0$  otherwise, and similarly for  $X_j$ . A simple Hölder computation as in (3.26) implies that

$$\int_{\mathcal{B}} \rho(y) dy \leq C\eta^{d/p'} \|\rho\|,$$

from which we infer that

$$\int_{\mathcal{B}} |\nabla W(X_i - \mathcal{T}(y)) - \nabla W(X_j - \mathcal{T}(y))| \rho(y) dy \leq C\eta^{d/p'} \eta_m^{-\alpha} \|\rho\|. \tag{3.34}$$

Putting together (3.33) and (3.34), we finally conclude that

$$\frac{d\eta_m}{dt} \geq -C\eta_m \|\rho\| \left( 1 + \eta^{d/p'} \eta_m^{-(1+\alpha)} \right), \tag{3.35}$$

for all  $t \in [0, \min(T, T_0^N))$ .

**Step C.-** Until now, we have proved from (3.32) and (3.35) that

$$\begin{cases} \frac{d^+\eta}{dt} \leq C\eta\|\rho\| \left(1 + \eta^{d/p'} \eta_m^{-(1+\alpha)}\right), \\ \frac{d\eta_m}{dt} \geq -C\eta_m\|\rho\| \left(1 + \eta^{d/p'} \eta_m^{-(1+\alpha)}\right), \end{cases} \quad (3.36)$$

for  $t \in [0, \min(T, T_0^N))$ . We first notice from (3.36) that if  $\eta^{d/p'} \eta_m^{-(1+\alpha)} \leq 1$ , then

$$\eta(t) \leq \eta^0 e^{2\|\rho\|t} \quad \text{and} \quad \eta_m(t) \geq \eta_m^0 e^{-2\|\rho\|t} \quad t \in [0, \min(T, T_0^N)). \quad (3.37)$$

We now show that (3.37) holds for time  $t \in [0, T]$  when  $N$  goes to infinity, in other words that  $T < T_0^N$  when  $N$  is sufficiently large. For this, we set

$$f(t) := \frac{\eta(t)}{\eta^0}, \quad g(t) := \frac{\eta_m(t)}{\eta_m^0} \quad \text{and} \quad \xi_N := (\eta^0)^{d/p'} (\eta_m^0)^{-(1+\alpha)}.$$

Note that  $\xi_N$  depends on the number of particles  $N$  as in (3.18). It yields

$$\begin{cases} \frac{d^+f}{dt} \leq C\|\rho\| f \left(1 + \xi_N f^{d/p'} g^{-(1+\alpha)}\right), \\ \frac{dg}{dt} \geq -C\|\rho\| g \left(1 + \xi_N f^{d/p'} g^{-(1+\alpha)}\right). \end{cases}$$

Since  $f(0) = g(0) = 1$  and  $\xi_N \rightarrow 0$  as  $N$  goes to infinity, we obtain that there exists a positive constant  $T_*^N (\leq T_0^N)$  such that

$$\xi_N f^{d/p'} g^{-(1+\alpha)} \leq 1 \quad \text{for} \quad t \in [0, T_*^N],$$

for sufficiently large  $N$ . Then it follows from (3.37) that

$$f(t) \leq e^{2\|\rho\|t} \quad \text{and} \quad g(t) \geq e^{-2\|\rho\|t}.$$

This yields  $\xi_N f^{d/p'} g^{-(1+\alpha)} \leq \xi_N e^{2(d/p' + (1+\alpha))\|\rho\|t}$ , that is,

$$\xi_N f^{d/p'} g^{-(1+\alpha)} \leq 1 \quad \text{holds for} \quad t \leq -\frac{\ln(\xi_N)}{2(d/p' + (1+\alpha))\|\rho\|},$$

so that

$$-\frac{\ln(\xi_N)}{2(d/p' + (1+\alpha))\|\rho\|} \leq T_*^N.$$

On the other hand, our assumption for the initial data (3.20) implies

$$\liminf_{N \rightarrow \infty} T_*^N \geq \lim_{N \rightarrow \infty} -\frac{\ln(\xi_N)}{2(d/p' + (1 + \alpha))\|\rho\|} = \infty,$$

and thus for  $N$  large enough,  $T < T_*^N < T_0^N$ . This completes the proof.  $\square$

**Remark 3.3.** One can use almost the same argument with the above to obtain an stability estimate in  $d_\infty$ : let  $\rho_1$  and  $\rho_2$  be solutions given by Theorem 4.1 to the system (3.15) satisfying (3.19), then we have

$$\frac{d}{dt} d_\infty(\rho_1(t), \rho_2(t)) \leq C \max(\|\rho_1\|, \|\rho_2\|) d_\infty(\rho_1(t), \rho_2(t)).$$

In fact, the estimate of mean field limit in Theorem 3.1 holds for  $-1 \leq \alpha < 0$  without any condition on  $\eta^0$  and  $\eta_m^0$ . This is coherent with the results in Carrillo, Di Francesco, Figalli, Laurent, and Slepčev (2011) in which the mean field limit is obtained for all measure initial data without restriction in the way initial data are approximated by Dirac masses at least for attractive potentials.

**Corollary 3.4.** *Suppose the interaction potential  $W$  satisfies (3.19) with  $-1 \leq \alpha < 0$ , and let  $\rho$  be a solution to the system (3.15) such that  $\rho \in L^\infty(0, T; (L^1 \cap L^p)(\mathbb{R}^d)) \cap \mathcal{C}([0, T], \mathcal{P}_1(\mathbb{R}^d))$ . Suppose that*

$$d_\infty(\mu_N^0, \rho^0) \rightarrow 0 \quad \text{as } N \rightarrow \infty.$$

*Then for any solution of the ODE system (3.16) the associated empirical distributions  $\mu_N(t)$  converge toward  $\rho(t)$  uniformly in time:*

$$\sup_{t \in [0, T]} d_\infty(\mu_N(t), \rho(t)) \rightarrow 0 \quad \text{as } N \rightarrow \infty.$$

**Remark 3.5.** It is remarkable that even if we do not have uniqueness of solution of (3.16) under assumption (3.19) with  $-1 \leq \alpha < 0$ , we get the mean field limit without restriction. If one collision occurs, then uniqueness may be lost, but the existence of solution is still guaranteed. Thus Corollary 3.4 is interesting because it is valid for density solutions to (3.15) even if collisions occur and uniqueness is lost at the particle level.

*Proof of Corollary 3.4.* We first notice that the existence of solutions to the ODE system (3.16) is guaranteed thanks to Cauchy-Peano-Arzela theorem

since  $\alpha$  is strictly negative with (3.19) implies that  $\nabla W$  is continuous. One can use the same arguments as in the proof of Theorem 3.1 to find

$$\begin{aligned} \frac{d^+\eta}{dt} &\leq C \sup_{x \in \mathbb{R}^d} \left( \int_{\mathcal{A}} + \int_{\mathcal{B}} \right) |\nabla W(\mathcal{T}(x) - \mathcal{T}(y)) - \nabla W(x - y)| \rho(y) dy \\ &:= \mathcal{K}_1 + \mathcal{K}_2, \end{aligned}$$

where the same notation for the sets  $\mathcal{A}$  and  $\mathcal{B}$  is used and the time dependency has been avoided for simplicity. Using (3.30) we estimate  $\mathcal{K}_1$  by  $C\eta\|\rho\|$ . To estimate  $\mathcal{K}_2$ , we use that  $\alpha < 0$  to get

$$|\nabla W(\mathcal{T}(x) - \mathcal{T}(y)) - \nabla W(x - y)| \leq \frac{C}{\eta^\alpha} + \frac{C}{|x - y|^\alpha},$$

and to obtain by Hölder's inequality that

$$\begin{aligned} \mathcal{K}_2 &\leq C \int_{\mathcal{B}} \frac{\rho(y)}{|x - y|^\alpha} dy + \frac{C}{\eta^\alpha} \int_{\mathcal{B}} \rho(y) dy \leq C\eta^{d/p' - \alpha} \|\rho\|_p + C\eta^{d/p'} \eta^{-\alpha} \|\rho\|_p \\ &\leq C\eta^{d/p' - \alpha} \|\rho\|. \end{aligned}$$

Hence, we have

$$\frac{d^+\eta}{dt} \leq C\eta\|\rho\| \left( 1 + \eta^{d/p' - \alpha - 1} \right),$$

and this yields for sufficiently large  $N$

$$\eta(t) \leq \left( (\eta^0)^{1 - (d/p' - \alpha)} e^{-C\|\rho\|(d/p' - \alpha - 1)t} + e^{-C\|\rho\|(d/p' - \alpha - 1)t} - 1 \right)^{-\frac{1}{d/p' - \alpha - 1}},$$

for all  $t \in [0, T]$ . Note that  $d/p' - \alpha - 1 > 0$  and then, the right hand side of previous estimate goes to zero as  $N$  goes to infinity. This completes the proof.  $\square$

We next show that there is no collision between particles when the initial quantities  $\eta^0$  and  $\eta_m^0$  in (3.18) satisfy

$$\lim_{N \rightarrow \infty} \frac{(\eta^0)^{d/p' - \alpha}}{\eta_m^0} = 0. \quad (3.38)$$

Note that the same strategy as in Remark 3.2 allows us to find suitable approximations for the initial data satisfying (3.38).



**Corollary 3.6.** *Under the assumptions of Corollary 3.4 with  $-1 \leq \alpha < 0$ , if we further assume that  $\eta^0, \eta_m^0$  satisfy (3.38). Then we have that for  $N$  large enough, the particle system (3.16) is uniquely well-defined till time  $T$  in the sense that there is no collision between particles before that time, and the convergence*

$$\sup_{t \in [0, T]} d_\infty(\mu_N(t), \rho(t)) \rightarrow 0 \quad \text{as } N \rightarrow \infty,$$

holds.

*Proof.* The proof of Corollary 3.4 shows that for sufficiently large  $N$

$$\eta \leq \left( (\eta^0)^{1-(d/p'-\alpha)} e^{-C\|\rho\|(d/p'-\alpha-1)t} + e^{-C\|\rho\|(d/p'-\alpha-1)t} - 1 \right)^{-\frac{1}{d/p'-\alpha-1}}.$$

For the estimate of  $\eta_m$ , one can obtain from the proof of Theorem 3.1 that

$$\frac{d\eta_m}{dt} \geq -C\eta_m\|\rho\| \left( 1 + \eta^{d/p'-\alpha}\eta_m^{-1} \right) \quad \text{for all } t \in [0, \min(T, T_0^N)],$$

where  $T_0^N$  denotes the first collision time between particles. Then we conclude the desired result employing the same arguments in Step C of Theorem 3.1 using (3.38).  $\square$

As a corollary of Theorem 3.1, we consider interaction potentials under weaker assumptions than (3.19): there exists  $R > 0$  such that  $W$  satisfies

$$|\nabla W(x)| \leq \frac{C}{|x|^\alpha}, \quad \text{and} \quad |D^2W(x)| \leq \frac{C}{|x|^{1+\alpha}}, \quad \forall x \in B(0, R), \quad (3.39)$$

where  $B(0, R) := \{x \in \mathbb{R}^d : |x| < R\}$ . Then one can assume that the initial data  $\rho^0$  has compact support, and show that the local solution  $\rho(t)$  has compact support on a small time interval  $[0, T]$ . This is possible since characteristics are locally in time well defined and the velocity is uniformly bounded under the assumptions (3.39) initially. This argument was made rigorous under stricter assumptions on the local behaviour of the interaction potential but allowing growth of the potential at infinity in Balagué and Carrillo (2012). Thus, one can cut-off the potential outside a large ball in such a way that the solution is unaffected but the potential satisfies the global assumption  $\nabla W \in \mathcal{W}^{1,p'}(\mathbb{R}^d)$  entering the well-posedness theory in Bertozzi, Laurent, and Rosado (2010) or satisfying (3.19) allowing for the application of Theorem 4.1. Concerning the interaction potential  $W$  satisfying (3.39), the same results of convergence in Theorem 3.1 and Corollary 3.6 can be obtained. We leave the details to the reader.

## 4 Local existence and uniqueness of $L^p$ -solutions

In this section, we provide a local existence and uniqueness result of weak solutions in  $L^p$ -spaces to the system (3.15) under the assumptions (3.19).

As we mentioned before, we can not directly apply the arguments in Bertozzi, Laurent, and Rosado (2010) for the potentials satisfying (3.19). Of course, we can overcome these difficulties using the property of compact supports on the initial data  $\rho^0$  (see the paragraph below Corollary 3.6). However, we use the arguments of dividing near- and far-field parts of the interaction potential function  $W$  to establish the local existence of a unique  $L^p$ -solution to the continuity aggregation equation (3.15).

**Theorem 4.1.** *Assume that  $W$  satisfies the condition (3.19), for some  $0 \leq \alpha < \frac{d}{p} - 1$ , and that  $\rho^0 \in \mathcal{P}_1(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$ ,  $1 < p \leq \infty$ . Then there exists a time  $T > 0$ , depending only on  $\|\rho^0\|_p$  and  $\alpha$ , and a unique nonnegative solution to (3.15) satisfying  $\rho \in L^\infty(0, T; L^1 \cap L^p(\mathbb{R}^d)) \cap \mathcal{C}([0, T], \mathcal{P}_1(\mathbb{R}^d))$ . Furthermore, the solution satisfies that there exists  $C > 0$  depending only on  $\|\rho^0\|_p$  and  $\alpha$  such that*

$$\|\rho(t)\|_p \leq C \quad \text{for all } t \in [0, T]. \quad (4.40)$$

The velocity field generated by  $\rho$ , given by  $u = -\nabla W * \rho$ , is bounded and Lipschitz continuous in space uniformly on  $[0, T]$ , and  $\rho$  is determined as the push-forward of the initial density through the flow map generated by  $u$ .

Moreover, if  $\rho_i$ ,  $i = 1, 2$ , are two such solutions to (3.15) with initial conditions  $\rho_i^0 \in \mathcal{P}_1(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$ ,  $1 < p \leq \infty$ , we have the following stability estimate:

$$\frac{d}{dt} d_1(t) \leq C \max(\|\rho_1\|, \|\rho_2\|) d_1(t),$$

where  $d_1(t) := d_1(\rho_1(t), \rho_2(t))$ .

*Proof.* Let us start by proving the uniqueness. Given two weak solutions  $\rho_i \in L^\infty(0, T; L^1 \cap L^p(\mathbb{R}^d)) \cap \mathcal{C}([0, T], \mathcal{P}_1(\mathbb{R}^d))$ ,  $i = 1, 2$ , to the continuous aggregation equations (3.15), consider the two flow maps  $\Psi_i : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ ,  $i = 1, 2$ , generated by the two velocity fields, i.e.,

$$\begin{cases} \frac{d}{dt} (\Psi_i(t; s, x)) = u_i(t; s, \Psi_i(t; s, x)), \\ \Psi_i(s; s, x) = x, \end{cases}$$

where  $u_i := -\nabla W * \rho_i$ ,  $t, s \in [0, T]$  and  $x \in \mathbb{R}^d$ . We know that the solutions are constructed by transporting the initial measures through the velocity fields  $\rho_i = \Psi_i \# \rho_i^0$ ,  $i = 1, 2$ .

Let  $\mathcal{T}^0$  be the optimal transportation between  $\rho_1(0)$  and  $\rho_2(0)$  for the  $d_1$ -distance. Then we define a transport (not necessarily optimal) between  $\rho_1(t)$  and  $\rho_2(t)$  by

$$\mathcal{T}^t(x) = \Psi_2(t; 0, x) \circ \mathcal{T}^0(x) \circ \Psi_1(0; t, x), \quad \mathcal{T}^t \# \rho_1(t) = \rho_2(t),$$

and  $\frac{d}{dt}d_1(t) \leq Q(t)$ , where  $d_1(t) := d_1(\rho_1(t), \rho_2(t))$  and

$$Q(t) := \int_{\mathbb{R}^d \times \mathbb{R}^d} |\nabla W(\mathcal{T}^t(x) - \mathcal{T}^t(y)) - \nabla W(x - y)| \rho_1(t, x) \rho_1(t, y) dx dy,$$

where we have used a similar argument as in Step A of the proof of Theorem 3.1. To simplify the notation, let us not make explicit the dependence on time. Note by symmetry that

$$\begin{aligned} Q(t) &\leq 4 \int_{\mathbb{R}^d \times \mathbb{R}^d} \left( \frac{|\mathcal{T}(x) - x|}{|\mathcal{T}(x) - \mathcal{T}(y)|^{1+\alpha}} + \frac{|\mathcal{T}(x) - x|}{|x - y|^{1+\alpha}} \right) \rho_1(x) \rho_1(y) dx dy \\ &:= \mathcal{J}_1 + \mathcal{J}_2. \end{aligned}$$

Straightforward computation using the near- and far-field decomposition as in (3.26) shows that

$$\begin{aligned} \mathcal{J}_1 &= 4 \int_{\mathbb{R}^d} |\mathcal{T}(x) - x| \rho_1(x) \left( \int_{\mathbb{R}^d} \frac{\rho_2(y)}{|\mathcal{T}(x) - y|^{1+\alpha}} dy \right) dx \\ &\leq C \|\rho_2\| \int_{\mathbb{R}^d} |\mathcal{T}(x) - x| \rho_1(x) dx = C \|\rho_2\| d_1(t). \end{aligned}$$

Similarly using again (3.26), we have  $\mathcal{J}_2 \leq C \|\rho_1\| d_1(t)$ . It yields that

$$\frac{d}{dt}d_1(t) \leq C \max(\|\rho_1\|, \|\rho_2\|) d_1(t),$$

from which we conclude the uniqueness part of the statement.

Let us now show the existence of weak solution. Let  $\varepsilon > 0$  and  $\theta$  be a standard mollifier:

$$\theta \geq 0, \quad \theta \in \mathcal{C}_0^\infty(\mathbb{R}^d), \quad \text{supp } \theta \subset B(0, 1), \quad \int_{\mathbb{R}^d} \theta(x) dx = 1,$$

and we set a sequence of smooth mollifiers:

$$\theta_\varepsilon(x) := \frac{1}{\varepsilon^d} \theta\left(\frac{x}{\varepsilon}\right).$$

We first regularize  $\nabla W$  such as  $\nabla W_\varepsilon := (\nabla W) * \theta_\varepsilon$ . Then since  $\nabla W_\varepsilon$  is a globally Lipschitz, we can apply the theory of Braun and Hepp (1977); Dobrushin (1979); Laurent (2007) which says that there exists a unique global solution  $\rho_\varepsilon$  to the following system

$$\begin{cases} \partial_t \rho_\varepsilon + \nabla \cdot (\rho_\varepsilon u_\varepsilon) = 0, & t > 0, \quad x \in \mathbb{R}^d, \\ u_\varepsilon(t, x) := -\nabla W_\varepsilon * \rho_\varepsilon, & t > 0, \quad x \in \mathbb{R}^d, \\ \rho_\varepsilon(0, x) := \rho^0(x), & x \in \mathbb{R}^d, \end{cases} \quad (4.41)$$

A standard calculation, see Bertozzi, Laurent, and Rosado (2010), implies that

$$\frac{d}{dt} \|\rho_\varepsilon\|_{L^1 \cap L^p} \leq C \|\rho_\varepsilon\|_{L^1 \cap L^p}^2, \quad (4.42)$$

where  $C$  is a uniform constant in  $\varepsilon$ . Note that the inequality (4.42) holds only formally for the non regularized problem, but it is fully rigorous for the regularized one with  $W_\varepsilon$ . This yields that the time of blow-up depends only on the initial data, more precisely  $\|\rho^0\|$ , and not on  $\varepsilon$ . Thus, there exists a  $T > 0$  such that

$$\sup_{\varepsilon > 0} \|\rho_\varepsilon\| < \infty. \quad (4.43)$$

It follows from (4.43) and the evolution in time of the first momentum of  $\rho$ , that this first moment is also uniformly bounded:

$$\sup_{\varepsilon > 0} \|x \rho_\varepsilon\|_{L^\infty(0, T; L^1(\mathbb{R}^d))} \leq C,$$

where  $C$  depends only on  $T$ ,  $\|x \rho^0\|_1$ , and  $\|\rho^0\|$ . We leave the details to the reader. Next, we show an estimate on the growth of the  $d_1$  distance  $\eta_{\varepsilon, \varepsilon'}(t) := d_1(\rho_\varepsilon(t), \rho_{\varepsilon'}(t))$  between  $\rho_\varepsilon$  and  $\rho_{\varepsilon'}$ , for  $\varepsilon, \varepsilon' > 0$ :

$$\frac{d}{dt} \eta_{\varepsilon, \varepsilon'}(t) \leq C \max(\|\rho_\varepsilon\|, \|\rho_{\varepsilon'}\|) (\eta_{\varepsilon, \varepsilon'}(t) + \varepsilon + \varepsilon'), \quad (4.44)$$

where  $C$  is a uniform constant in  $\varepsilon$  and  $\varepsilon'$ . We remark that the above estimate (4.44) implies that  $\{\rho_\varepsilon\}_{\varepsilon > 0}$  is a Cauchy sequence in  $\mathcal{C}([0, T], \mathcal{P}_1(\mathbb{R}^d))$ .

Let us remark that the weak solutions to the regularized problems (4.41) can be written in terms of characteristics. This is a consequence of the fact that the associated velocity field  $u_\varepsilon$  is bounded and Lipschitz in space, uniformly in time and some standard duality arguments. This strategy is explained in detail at the end of the proof of the present Theorem applied to the solution of the original problem, and we refer the reader there for details. Since solutions are constructed by characteristics, for the proof of

(4.44) we can proceed as in the part of uniqueness. Therefore, not making explicit the time dependency, we get

$$\begin{aligned}
\frac{d}{dt}\eta_{\varepsilon,\varepsilon'}(t) &\leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |\nabla W_\varepsilon(\mathcal{T}(x) - \mathcal{T}(y)) - \nabla W_{\varepsilon'}(x - y)| \rho_{\varepsilon'}(x) \rho_{\varepsilon'}(y) dx dy \\
&\leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |\nabla W_\varepsilon(\mathcal{T}(x) - \mathcal{T}(y)) - \nabla W_\varepsilon(x - y)| \rho_{\varepsilon'}(x) \rho_{\varepsilon'}(y) dx dy \\
&\quad + \int_{\mathbb{R}^d \times \mathbb{R}^d} |\nabla W_\varepsilon(x - y) - \nabla W_{\varepsilon'}(x - y)| \rho_{\varepsilon'}(x) \rho_{\varepsilon'}(y) dx dy \\
&:= \mathcal{K}_1 + \mathcal{K}_2, \tag{4.45}
\end{aligned}$$

where  $\mathcal{T}$  is the optimal transportation between  $\rho_{\varepsilon'}(t)$  and  $\rho_\varepsilon(t)$  for the  $d_1$ -distance. To estimate  $\mathcal{K}_1$ , we notice that

$$\begin{aligned}
|\nabla W_\varepsilon(x)| &\leq \int_{\{y: |y| < \frac{|x|}{2}\}} \frac{\theta_\varepsilon(y)}{|x - y|^{1+\alpha}} dy + \int_{\{y: |y| \geq \frac{|x|}{2}\}} \frac{\theta_\varepsilon(y)}{|x - y|^{1+\alpha}} dy \\
&\leq \frac{2^{1+\alpha}}{|x|^{1+\alpha}} \int_{\mathbb{R}^d} \theta_\varepsilon(y) dy + \mathbf{1}_{\{|x| \leq 2\varepsilon\}} \int_{\{y: \varepsilon \geq |y|\}} \frac{\theta_\varepsilon(y)}{|x - y|^{1+\alpha}} dy \\
&\leq \frac{C}{|x|^{1+\alpha}} + \frac{C\varepsilon^{1+\alpha}}{|x|^{1+\alpha}} \int_{\{y: \varepsilon \geq |y|\}} \frac{\theta_\varepsilon(y)}{|x - y|^{1+\alpha}} dy \leq \frac{C}{|x|^{1+\alpha}}. \tag{4.46}
\end{aligned}$$

Then we now use again the decomposition (3.26) as in the part of uniqueness to find

$$\mathcal{K}_1 \leq C \max(\|\rho_\varepsilon\|, \|\rho_{\varepsilon'}\|) \eta_{\varepsilon,\varepsilon'}(t), \tag{4.47}$$

where  $C$ ,  $\|\rho_\varepsilon\|$ , and  $\|\rho_{\varepsilon'}\|$  are uniformly bounded in  $\varepsilon$  and  $\varepsilon'$  thanks to the estimate (4.43). For the estimate of  $\mathcal{K}_2$ , we claim that

$$|\nabla(W - W_\varepsilon)(x)| \leq \frac{C\varepsilon}{|x|^{1+\alpha}}, \tag{4.48}$$

where  $C$  is independent on  $\varepsilon$ .

*Proof of Claim:* It is a straightforward to obtain

$$\begin{aligned}
|\nabla W_\varepsilon(x) - \nabla W(x)| &\leq \int_{\mathbb{R}^d} |\nabla W(x - y) - \nabla W(x)| \theta_\varepsilon(y) dy \\
&\leq 2 \int_{\mathbb{R}^d} \left( \frac{1}{|x|^{1+\alpha}} + \frac{1}{|x - y|^{1+\alpha}} \right) |y| \theta_\varepsilon(y) dy \tag{4.49} \\
&:= \mathcal{L}_1 + \mathcal{L}_2.
\end{aligned}$$

Noticing that the mollifier properties allow to gain an  $\varepsilon$  factor in front of the integrals, we can estimate  $\mathcal{L}_i, i = 1, 2$  as follows

$$\begin{aligned}\mathcal{L}_1 &\leq \frac{C\varepsilon}{|x|^{1+\alpha}} \int_{\mathbb{R}^d} \theta_\varepsilon(y) dy = \frac{C\varepsilon}{|x|^{1+\alpha}}, \\ \mathcal{L}_2 &\leq 2\varepsilon \int_{\mathbb{R}^d} \frac{\theta_\varepsilon(y)}{|x-y|^{1+\alpha}} dy \leq \frac{C\varepsilon}{|x|^{1+\alpha}},\end{aligned}\tag{4.50}$$

where we used a similar argument to (4.46) for  $\mathcal{L}_2$ . We now combine (4.49) and (4.50) to have the inequality (4.48). Then we use (4.48) together with (3.26) to find the estimate of  $\mathcal{K}_2$

$$\mathcal{K}_2 \leq C(\varepsilon + \varepsilon') \int_{\mathbb{R}^d \times \mathbb{R}^d} \frac{\rho_{\varepsilon'}(t, x) \rho_{\varepsilon'}(t, y)}{|x-y|^{1+\alpha}} dx dy \leq C(\varepsilon + \varepsilon') \|\rho_{\varepsilon'}\|.\tag{4.51}$$

This completes the proof of the inequality (4.44) by combining (4.44), (4.45), (4.47), and (4.51).

Since  $\rho_\varepsilon$  is a Cauchy sequence in  $\mathcal{C}([0, T], \mathcal{P}_1(\mathbb{R}^d))$ , it converges toward a limit curve of measures  $\rho \in \mathcal{C}([0, T], \mathcal{P}_1(\mathbb{R}^d))$ , and we also have  $\rho \in L^\infty(0, T; L^1 \cap L^p(\mathbb{R}^d))$  from the uniform bounded estimate (4.43). It remains to show that  $\rho$  is a solution of the aggregation equations (3.15). Choose a test function  $\phi(t, x) \in C_c^\infty([0, T] \times \mathbb{R}^d)$ , then  $\rho_\varepsilon$  satisfies

$$\begin{aligned}\int_{\mathbb{R}^d} \rho^0(x) \phi^0(x) dx &= \int_{\mathbb{R}^d} \rho_\varepsilon(T, x) \phi(T, x) dx + \int_0^T \int_{\mathbb{R}^d} \rho_\varepsilon(t, x) \partial_t \phi(t, x) dx dt \\ &\quad - \int_0^T \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \rho_\varepsilon(t, x) \rho_\varepsilon(t, y) \nabla W_\varepsilon(x-y) \cdot \nabla \phi(t, x) dx dy dt.\end{aligned}\tag{4.52}$$

The first two terms in the rhs of (4.52) converges to

$$\int_{\mathbb{R}^d} \rho(T, x) \phi(T, x) dx + \int_0^T \int_{\mathbb{R}^d} \rho(t, x) \partial_t \phi(t, x) dx dt,$$

since  $\rho_\varepsilon \rightarrow \rho$  in  $\mathcal{C}([0, T], \mathcal{P}_1(\mathbb{R}^d))$ . For the third term in the rhs of (4.52), we use the estimates (4.48) and (4.43) to find

$$\left| \int_0^T \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \rho_\varepsilon(t, x) \rho_\varepsilon(t, y) (\nabla W_\varepsilon(x-y) - \nabla W(x-y)) \cdot \nabla \phi(t, x) dx dy dt \right| \rightarrow 0,$$

as  $\varepsilon \rightarrow 0$ . It remains to show that

$$\begin{aligned}\int_0^T \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \rho_\varepsilon(t, x) \rho_\varepsilon(t, y) \nabla W(x-y) \cdot \nabla \phi(t, x) dx dy dt \\ \rightarrow \int_0^T \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \rho(t, x) \rho(t, y) \nabla W(x-y) \cdot \nabla \phi(t, x) dx dy dt,\end{aligned}$$

as  $\varepsilon \rightarrow 0$ . For this, we introduce a cut-off function  $\chi_\delta \in \mathcal{C}_c^\infty(\mathbb{R})$  such that

$$\chi_\delta(x) = \begin{cases} 1 & \text{if } |x| \leq \delta \\ 0 & \text{if } |x| \geq 2\delta \end{cases} .$$

Then it follows from the weak convergence that

$$\begin{aligned} & \int_0^T \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \rho_\varepsilon(t, x) \rho_\varepsilon(t, y) (1 - \chi_\delta(x - y)) \nabla W(x - y) \cdot \nabla \phi(t, x) dx dy dt \\ & \rightarrow \int_0^T \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \rho(t, x) \rho(t, y) (1 - \chi_\delta(x - y)) \nabla W(x - y) \cdot \nabla \phi(t, x) dx dy dt, \end{aligned}$$

as  $\varepsilon \rightarrow 0$ , since  $(1 - \chi_\delta(x - y)) \nabla W(x - y) \cdot \nabla \phi(t, x)$  is a Lipschitz function. We estimate the remainder as follows:

$$\begin{aligned} & \left| \int_0^T \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \rho_\varepsilon(t, x) \rho_\varepsilon(t, y) \chi_\delta(x - y) \nabla W(x - y) \cdot \nabla \phi(t, x) dx dy dt \right| \\ & \leq C\delta \int_0^T \int_{\{(x, y) \in \mathbb{R}^d \times \mathbb{R}^d: |x - y| \leq 2\delta\}} \frac{1}{|x - y|^{1+\alpha}} \rho_\varepsilon(t, x) \rho_\varepsilon(t, y) dx dy dt \\ & \leq CT\delta \|\rho_\varepsilon\| \rightarrow 0 \quad \text{as } \delta \rightarrow 0. \end{aligned}$$

Similarly, we have

$$\lim_{\delta \rightarrow 0} \left| \int_0^T \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \rho(t, x) \rho(t, y) \chi_\delta(x - y) \nabla W(x - y) \cdot \nabla \phi(t, x) dx dy dt \right| = 0.$$

Hence, we conclude that  $\rho$  satisfies

$$\begin{aligned} \int_{\mathbb{R}^d} \rho^0(x) \phi^0(x) dx &= \int_{\mathbb{R}^d} \rho(T, x) \phi(T, x) dx + \int_0^T \int_{\mathbb{R}^d} \rho(t, x) \partial_t \phi(t, x) dx dt \quad (4.53) \\ &\quad - \int_0^T \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \rho(t, x) \rho(t, y) \nabla W(x - y) \cdot \nabla \phi(t, x) dx dy dt, \end{aligned}$$

for all  $\phi \in \mathcal{C}_c^\infty([0, T] \times \mathbb{R}^d)$ .

Now, We notice that a weak solution in  $\rho \in L^\infty(0, T; L^1 \cap L^p(\mathbb{R}^d))$  to (3.15) under the assumptions (3.19) has a well defined flow by using the same arguments as the ones at the beginning of Theorem 3.1. In fact, the velocity field is bounded and Lipschitz continuous in space with

$$|u(t, x) - u(t, y)| \leq C \|\rho\| |x - y|$$

for all  $x, y \in \mathbb{R}^d$  and  $t \in [0, T]$ . Thus, the flow map

$$\begin{cases} \frac{d}{dt}(\Psi(t; s, x)) = u(t; s, \Psi(t; s, x)), \\ \Psi(s; s, x) = x, \end{cases}$$

for all  $s, t \in [0, T]$  is well-defined. Choosing as test function in (4.53)  $\phi(t, x) = \varphi(\Psi(t; \bar{T}, x))$  for any  $\bar{T} \in (0, T]$  with  $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ , it is straightforward to check, due to the definition of the flow map, that

$$\int_{\mathbb{R}^d} \rho^0(x) \varphi(\Psi(0; \bar{T}, x)) dx = \int_{\mathbb{R}^d} \rho(\bar{T}, x) \varphi(x) dx,$$

for all  $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ , and thus by a density argument we conclude  $\rho(\bar{T}) = \Psi(\bar{T}; 0, \cdot) \# \rho^0$ . Since this argument can be done for all  $0 < \bar{T} \leq T$ , this completes the proof.  $\square$

## 5 Propagation of chaos

In most practical purposes to approximate the continuum model by particle systems, it is naturally expected that initial positions and velocities will randomly and independently be selected. We will show that the empirical measure at time 0 is then close to  $\rho^0$  with large probability in suitable weak norm.

In a seminal article Kac (1956), the propagation of chaos was introduced by Kac giving a proof for a simplified collision evolution process. He showed how the limit of many particles rigorously follows from the property of propagation of chaos. For a classical introduction to these topics, we refer to Sznitman (1991). Later, this property has been studied and developed in kinetic theory, McKean (1967, 1975); Graham and Méléard (1997); Hauray and Mischler (2012); Mischler and Mouhot (2013).

Let us introduce the notion of propagation of chaos. Let us consider  $\rho^N(t, x_1, \dots, x_N)$  being the image by the dynamics to the coupled system (3.16) with  $N$ -equal masses particles of the initial law  $(\rho^0)^{\otimes N}$ . We define the  $k$ -marginals as follows.

$$\rho_k^N(t, x_1, \dots, x_k) := \int_{\mathbb{R}^{d(N-k)}} \rho_N(t, x) dx_{k+1} \cdots dx_N.$$

Let us choose the initial positions  $X^{N,0} := \{X_i^0\}_{i=1}^N$  as independent identically distributed random variables (in short iid) with law  $\rho^0$ . We can



construct the associated empirical measure as in (3.17) by

$$\mu_N(t) = \frac{1}{N} \sum_{i=1}^N \delta_{X_i(t)},$$

but now understood as a random variable with values in the space of probability measures.

The propagation of chaos property is defined as follows: for any fixed  $k \in \mathbb{N}$ ,

$$\rho_k^N \rightharpoonup (\rho)^{\otimes k} \quad \text{weakly-* as measures as } N \rightarrow \infty.$$

It is classically known Sznitman (1991) that it is sufficient to check this property for  $k = 2$  to derive the propagation of chaos. In fact, this is based on the fact that propagation of chaos is equivalent to show that the empirical measures  $\mu_N(t)$  converge in law towards the constant random variable  $\rho(t)$ .

Theorem 5.1 gives a quantified version of the convergence in probability of  $\mu_N(t)$  towards  $\rho(t)$ . We refer to Hauray and Mischler (2012); Mischler and Mouhot (2013) for a detailed explanation of the quantified equivalence relations. The propagation of chaos for the Vlasov-Poisson equations with singular force has recently been investigated in Hauray and Jabin (2012). Here, we are only able to provide such a result in a more restrictive setting than in the previous section. Namely, we only show the propagation of chaos for  $d \geq 3$  and with a more restrictive condition on the allowed singularities  $\alpha \geq 0$  depending on the regularity of the initial data  $1 < p < \infty$ .

**Theorem 5.1.** *Given  $\rho(t) \in L^\infty(0, T; (L^1 \cap L^p)(\mathbb{R}^d)) \cap \mathcal{C}([0, T], \mathcal{P}_1(\mathbb{R}^d))$  the unique solution to (3.15) with initial data  $\rho^0 \in \mathcal{P}_1(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$ ,  $1 < p \leq \infty$ , up to time  $T > 0$ . Assume that  $\rho^0$  has compact support, that the initial positions  $X^{N,0} := \{X_i^0\}_{i=1}^N$  are iid with law  $\rho^0$ , and that*

$$(1 + \alpha)p' < \frac{p-1}{2p-1}d,$$

with  $\alpha \geq 0$ . Then the propagation of chaos holds in the sense that

$$\mathbb{P} \left( \sup_{t \in [0, T]} d_1(\mu_N(t), \rho(t)) \geq \frac{C}{N^{\gamma/d}} \right) \rightarrow 0, \quad \text{as } N \rightarrow +\infty,$$

where  $\gamma$  is a positive constant satisfying

$$\frac{p'(2p-1)(1+\alpha)}{d(p-1)} < \gamma < 1.$$

**Remark 5.2.** The condition on  $\alpha$  gets more and more restrictive as  $p$  gets smaller and smaller. In  $d = 2$ , even for  $p = \infty$  the condition is empty for  $\alpha \geq 0$ . In  $d = 3$ , you get the condition  $\alpha < 1/2$  for  $p = \infty$  and with  $p = \frac{5+\sqrt{13}}{2}$  the condition is already empty. We also notice that the existence and uniqueness of the solutions are guaranteed by Theorem 4.1 and taking expectations in the corresponding inequalities for the particle system. Finally, in case  $-1 \leq \alpha < 0$ , the propagation of chaos holds using the same strategy as in Corollary 3.4 by taking expectations in the inequalities for the evolution of the Wasserstein distance.

We will follow the strategy recently introduced in Hauray and Jabin (2012) for the Vlasov equation. We first find a deterministic version of the propagation of chaos. This means that we consider a regularized system of particles as a kind of middle ground between the solution of the mean-field equation (3.15) and the random particle evolution. More precisely, we define the “blob” initial data  $\rho_N^0$  as

$$\rho_N^0 := \mu_N^0 * \frac{\mathbf{1}_{B_\varepsilon(0)}}{|B_\varepsilon(0)|} = \frac{1}{c_d \varepsilon^d} (\mu_N^0 * \mathbf{1}_{B_\varepsilon(0)}), \quad (5.54)$$

where  $\varepsilon > 0$  to be chosen as a function of the number of particles  $N$  and  $c_d$  is the volume of the unit ball in dimension  $d$ . We also define the “blob” approximation  $\rho_N(t)$  to be the solution of the system (3.15) with the kernel  $W$  satisfying (3.19) given by Theorem 4.1 and “blob” initial data  $\rho_N^0$ .

In the rest,  $\varepsilon$  is chosen as a function of  $N$  as  $\varepsilon(N) = N^{-\gamma/d}$  with  $0 < \gamma < 1$ . It is easy to check that  $\|\rho_N^0\|_p \simeq N^{(\gamma-1)/p'}$  for  $N$  large enough, then we can wonder how far is the empirical measure to its blob approximation if we assume a bound on  $\|\rho_N^0\|_p$  independent of  $N$ .

**Proposition 5.3.** *Under the assumptions of Theorem 5.1 and assuming that there exists  $C_1 > 0$  independent of the number of particles  $N$  such that*

$$\|\rho_N^0\|_p \leq C_1, \quad \text{and} \quad \eta_m^0 \geq \frac{1}{C_1} \varepsilon^r,$$

*with  $1 \leq r < \frac{d}{p'(1+\alpha)}$ . Then, there exists  $T > 0$  such that the solutions  $\rho_N(t)$  and the empirical measure  $\mu_N(t)$  are well-defined for all  $t \in [0, T]$ , and*

$$d_\infty(\rho_N(t), \mu_N(t)) \leq d_\infty(\rho_N^0, \mu_N^0) e^{C_2 T} \leq \varepsilon(N) e^{C_2 T},$$

*where  $C_2 > 0$  is independent of  $N$ .*

*Proof.* We follow a similar argument to Theorem 3.1. We first notice from Theorem 4.1 that there exists a common time of existence  $T > 0$  of the

solutions  $\rho_N$  independent of  $N$  since it only depends on  $\|\rho_N^0\|_p$  and  $\alpha$ . The empirical measure also exists up to this time since it will be smaller than the possible first collision time of particles. Moreover, due to (4.40), we get that  $\|\rho_N(t)\|_p \leq C$ , for all  $t \in [0, T]$ , where  $C$  is independent of  $N$ . We next substitute  $\rho_N(t)$  for  $\rho(t)$  in the proof of Theorem 3.1, and thus all estimates in Step A and B hold to deduce

$$\frac{d\eta_N}{dt} \leq C\eta_N\|\rho_N\| \left(1 + \eta_N^{d/p'} \eta_m^{-(1+\alpha)}\right) \leq C\eta_N \left(1 + \eta_N^{d/p'} \eta_m^{-(1+\alpha)}\right),$$

and

$$\frac{d\eta_m}{dt} \geq -C\eta_m\|\rho_N\| \left(1 + \eta_N^{d/p'} \eta_m^{-(1+\alpha)}\right) \geq -C\eta_m \left(1 + \eta_N^{d/p'} \eta_m^{-(1+\alpha)}\right),$$

where  $\eta_N(t) := d_\infty(\rho_N(t), \mu_N(t))$ . Note that the condition  $r \geq 1$  makes sense since  $\varepsilon \approx \eta_N^0 \geq \eta_m^0 \geq C\varepsilon^r$  for  $\varepsilon$  small enough. We finally conclude the desired result using a similar argument as in Step C of the proof of Theorem 3.1 since

$$\frac{(\eta_N^0)^{d/p'}}{(\eta_m^0)^{1+\alpha}} \leq C\varepsilon^{d/p' - r(1+\alpha)} \rightarrow 0 \quad \text{as } N \rightarrow \infty,$$

by assumption. □

We now present two propositions showing that the assumptions on  $\rho_N^0$  and  $\eta_m^0$  in Proposition 5.3 are generic in a probability sense when the initial positions  $X^{N,0}$  are iid with law  $\rho^0$  in  $L^p$ . We first prove in Proposition 5.4 that  $\eta_m^0$  is roughly larger than  $N^{-\frac{2p-1}{d(p-1)}}$  if the  $X^{N,0}$  are iid with law  $\rho^0$ .

**Proposition 5.4.** *Let  $\rho^0 \in \mathcal{P}_1(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$ ,  $1 < p \leq \infty$ , and the initial positions  $X^{N,0}$  be iid with law  $\rho^0$ . Suppose there exists  $L > 0$  such that*

$$2c_d^{\frac{1}{p'}} \|\rho^0\|_p L^{\frac{d}{p'}} \leq N,$$

then  $\eta_m^0$  satisfies

$$\mathbb{P}\left(\eta_m^0 \geq LN^{-\frac{2p-1}{d(p-1)}}\right) \geq e^{-2c_d^{\frac{1}{p'}} \|\rho^0\|_p L^{\frac{d}{p'}}}.$$

*Proof.* Choose an  $r \in \mathbb{R}_+$ . Then  $\eta_m^0 \geq r$  holds if

$$X_k^0 \in \mathbb{R}^d \setminus A_k, \quad \text{with} \quad A_k = \bigcup_{1 \leq i \leq k-1} B(X_i^0, r),$$

for all  $1 \leq k \leq N$ . It implies from our assumption with  $r = LN^{-\frac{2p-1}{d(p-1)}}$  that

$$\begin{aligned} \mathbb{P}\left(\eta_m^0 \geq LN^{-\frac{2p-1}{d(p-1)}}\right) &\geq \prod_{k=1}^N \left[1 - \int_{A_k} \rho^0(x) dx\right] \\ &\geq \prod_{k=1}^{N-1} \left[1 - c_d^{\frac{1}{p'}} \|\rho^0\|_p L^{\frac{d}{p'}} N^{-2+\frac{1}{p}} k^{\frac{1}{p'}}\right], \end{aligned}$$

and thus using that  $\ln(1-x) \geq -2x$  if  $x \in [0, \frac{1}{2}]$ , we conclude

$$\ln \mathbb{P}(\eta_m^0 \geq r) \geq -2c_d^{\frac{1}{p'}} \|\rho^0\|_p L^{\frac{d}{p'}} N^{-2+\frac{1}{p}} \sum_{k=1}^{N-1} k^{\frac{1}{p'}} \geq -2c_d^{\frac{1}{p'}} \|\rho^0\|_p L^{\frac{d}{p'}}.$$

□

The next proposition gives some bound on the large deviation of  $\|\rho_N^0\|_p$ . It states roughly that  $\|\rho_N^0\|_p$  is of the same order that  $\|\rho^0\|_p$ , if the  $X^{N,0}$  are iid with law  $\rho^0$ .

**Proposition 5.5.** *Let  $\rho^0 \in \mathcal{P}_1(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$ ,  $1 < p \leq \infty$ , with compactly support included in  $[-R, R]^d$ . For any iid  $X^{N,0}$  with law  $\rho^0$ , the smoothed empirical measures  $\rho_N^0$  defined in (5.54) satisfy the explicit “large deviations” bound*

$$\mathbb{P}(L_d \|\rho^0\|_p \leq \|\rho_N^0\|_p) \leq [2(R+1)]^d N^\gamma e^{-c_R \|\rho\|_p N^{1-\gamma}},$$

where  $L_d$  and  $c_R$  are explicitly given by

$$c_R := \frac{2 \ln 2}{[2(R+1)]^{\frac{d}{p}}} \quad \text{and} \quad L_d := \frac{4(4[\sqrt{d}] + 1)^{d/p}}{c_d},$$

with  $[\cdot]$  denoting the integer part.

*Proof.* For any  $X_i \in \mathbb{R}^d$  and  $x \in \mathbb{R}^d$ , we have

$$\rho_N^0(x) = \frac{1}{N c_d \varepsilon^d} \sum_{i=1}^N \mathbf{1}_{B_\varepsilon}(x - X_i) = \frac{1}{N c_d \varepsilon^d} \#\{i \text{ s.t. } |x - X_i| \leq \varepsilon\},$$

where  $\#$  stands for the cardinal (of a finite set). Next, we cover  $[-R, R]^d$  by  $M$  disjoint cubes  $C_k$  of size  $\varepsilon^d$ , centered at the points  $(c_k)_{k \leq M}$ . The number  $M$  of square needed depends on  $N$  via  $\varepsilon$ , and is bounded by

$$M \leq \left\lceil \frac{2(R+1)}{\varepsilon} \right\rceil^d.$$

Assume that  $x \in C_k$  for some  $1 \leq k \leq M$ , i.e.,  $|x - c_k| \leq \frac{\sqrt{d}\varepsilon}{2}$ , then

$$\#\{i \text{ s.t. } |x - X_i| \leq \varepsilon\} \leq \#\{i \text{ s.t. } |c_k - X_i|_\infty \leq 2\sqrt{d}\varepsilon\},$$

and for any  $1 < p < \infty$  we obtain

$$\begin{aligned} \int_{C_k} (\rho_N^0(x))^p dx &\leq \frac{\varepsilon^{d(1-p)}}{(N c_d)^p} \#\{i \text{ s.t. } |c_k - X_i|_\infty \leq 2\sqrt{d}\varepsilon\}^p \\ &= \frac{\varepsilon^{d(1-p)}}{(N c_d)^p} \#\{i \text{ s.t. } x \in C_k^d\}^p, \end{aligned}$$

where  $C_k^d$  denotes the cube of center  $c_k$  and size  $(4\sqrt{d}\varepsilon)^d$ . Let us consider the set of cubes of the lattice that contains  $C_k^d$ , i.e.,

$$C_k^d \subset \bigcup_{j \in I_k} C_j$$

where  $I_k = \{j \text{ such that } C_k^d \cap C_j \neq \emptyset\}$ . It is direct to check that  $\#I_k \leq M_d$  with  $M_d = (4\lceil\sqrt{d}\rceil + 1)^d$ . Moreover, there are only  $M_d$  possible values of  $1 \leq k \leq M$  such that  $j \in I_k$  for a given  $1 \leq j \leq M$ . This yields

$$\begin{aligned} \int_{\mathbb{R}^d} (\rho_N^0(x))^p dx &\leq \frac{\varepsilon^{d(1-p)}}{(N c_d)^p} \sum_{k=1}^M \sum_{j \in I_k} \#\{i \text{ s.t. } x \in C_j\}^p \\ &\leq \frac{M_d \varepsilon^{d(1-p)}}{(N c_d)^p} \sum_{k=1}^M \#\{i \text{ s.t. } X_i \in C_k\}^p. \end{aligned} \quad (5.55)$$

Let us introduce the notation  $N_k := \#\{i \text{ s.t. } X_i \in C_k\}$ .  $N_k$  is a random variable which follows a binomial law  $B(N, s_k)$  with  $s_k := \int_{C_k} \rho^0(x) dx$ . If  $L\|\rho^0\|_p \leq \|\rho_N^0\|_p$ , then (5.55) together with Hölder's inequality imply that

$$\sum_{k=1}^M N_k^p \geq \frac{(c_d N)^p}{M_d} \varepsilon^{d\frac{p}{p'}} \|\rho_N^0\|_p^p \geq N^p \tilde{L}^p \varepsilon^{d\frac{p}{p'}} \|\rho^0\|_p^p \geq N^p \tilde{L}^p \sum_{k=1}^M s_k^p,$$

where  $\tilde{L} := c_d L / (M_d)^{1/p}$ . But, if this happens, it means that for at least one  $k \leq M$ ,

$$\begin{aligned} N_k &\geq \left( \frac{1}{2} M^{-1} (N \tilde{L})^p \varepsilon^{d\frac{p}{p'}} \|\rho^0\|_p^p + \frac{1}{2} N^p \tilde{L}^p s_k^p \right)^{\frac{1}{p}} \\ &\geq \frac{1}{2} M^{-\frac{1}{p}} N \varepsilon^{\frac{d}{p'}} \tilde{L} \|\rho^0\|_p + \frac{1}{2} N \tilde{L} s_k \geq \frac{N \tilde{L}}{2} (\check{c}_R \varepsilon^d \|\rho^0\|_p + s_k), \end{aligned}$$

with  $\tilde{c}_R := 1/[2(R+1)]^{\frac{d}{p}}$ , where the concavity of  $x^{1/p}$  was used. Then, we deduce that

$$\mathbb{P}(L\|\rho^0\|_p \leq \|\rho_N^0\|_p) \leq \sum_{k=1}^M \mathbb{P}\left(N_k \geq \frac{N\tilde{L}}{2}[\tilde{c}_R\varepsilon^d\|\rho^0\|_p + s_k]\right).$$

Since  $N_k$  is a random variable which follows a binomial law  $B(N, s_k)$ , then for any  $\lambda$ , the exponential moments of  $N_k$  are bounded by

$$\mathbb{E}(e^{\lambda N_k}) \leq [1 + (e^\lambda - 1)s_k]^N \leq e^{(e^\lambda - 1)Ns_k}.$$

This together with Chebyshev's inequality implies that

$$\begin{aligned} \mathbb{P}(L\|\rho^0\|_p \leq \|\rho_N^0\|_p) &\leq \sum_{k=1}^M \mathbb{E}(e^{\lambda N_k}) e^{-\lambda \frac{N\tilde{L}}{2}[\tilde{c}_R\varepsilon^d\|\rho^0\|_p + s_k]} \\ &\leq \sum_{k=1}^M e^{(e^\lambda - 1)Ns_k - \lambda \frac{N\tilde{L}}{2}[\tilde{c}_R\varepsilon^d\|\rho^0\|_p + s_k]}. \end{aligned}$$

Taking  $\lambda = \ln L'$  with the notation  $L' = \frac{\tilde{L}}{2}$ , we get

$$\begin{aligned} \mathbb{P}(L\|\rho^0\|_p \leq \|\rho_N^0\|_p) &\leq \sum_{k=1}^M e^{-(L' \ln L' + 1 - L')Ns_k - L' \ln L' \tilde{c}_R N \varepsilon^d \|\rho^0\|_p} \\ &\leq \sum_{k=1}^M e^{-L' \ln L' c_R N \varepsilon^d \|\rho^0\|_p} = M e^{-L' \ln L' \tilde{c}_R N \varepsilon^d \|\rho^0\|_p}, \end{aligned}$$

where we used  $x \ln x - x + 1 \geq 0$ , for  $x > 0$ . With the scaling  $\varepsilon(N) = N^{-\frac{\gamma}{d}}$ , we get

$$\mathbb{P}(L\|\rho^0\|_p \leq \|\rho_N^0\|_p) \leq [2(R+1)]^d N^\gamma e^{-\tilde{c}_R L' \ln L' \|\rho^0\|_p N^{1-\gamma}}.$$

In particular, choosing  $L = L_d = 4(M_d)^{\frac{1}{p}}/c_d$  so that  $L' = 2$ , we get the desired result

$$\mathbb{P}(L_d\|\rho^0\|_p \leq \|\rho_N^0\|_p) \leq [2(R+1)]^d N^\gamma e^{-c_R \|\rho^0\|_p N^{1-\gamma}},$$

for  $1 < p < \infty$ . In the case of  $p = \infty$ , we first notice that as in (5.55), we deduce

$$\|\rho_N^0\|_\infty \leq \frac{M_d}{N c_d \varepsilon^d} \sup_{1 \leq k \leq M} \#\{i \text{ s.t. } |c_k - X_i|_\infty \leq \varepsilon\} = \frac{M_d}{N c_d \varepsilon^d} \sup_{1 \leq k \leq M} N_k.$$

Since  $N_k$  follows a binomial law  $B(N, s_k)$  and  $s_k \leq \|\rho^0\|_\infty \varepsilon^d$ , above estimates allow us to conclude the desired inequality.  $\square$

We are now in a position to give the proof of propagation of chaos.

*Proof of Theorem 5.1.* We introduce several sets for the random initial data:

$$\omega_1 := \{X^{N,0} : \eta_m^0 \geq \varepsilon^r\}, \quad \omega_2 := \{X^{N,0} : L_d \|\rho^0\|_p \geq \|\rho_N^0\|_p\},$$

and

$$\omega_3 := \{X^{N,0} : d_1(\mu_N^0, \rho^0) \leq \varepsilon\},$$

where  $r$ ,  $\varepsilon$  and  $L_d$  are given in Propositions 5.3, 5.4, and 5.5. We first provide the estimate of  $\mathbb{P}(\omega_1^c)$ . Note that since the assumption on  $\gamma$ , we obtain

$$\frac{2p-1}{\gamma(p-1)} < \frac{d}{p'(1+\alpha)}.$$

This yields the existence of  $r$  verifying

$$1 < \frac{2}{\gamma} \leq \frac{2p-1}{\gamma(p-1)} < r < \frac{d}{p'(1+\alpha)}.$$

This again implies the existence of  $\beta > 0$  satisfying

$$\frac{d}{\gamma}\beta + \frac{2p-1}{\gamma(p-1)} < r.$$

From Proposition 5.4, if we choose  $L = N^{-\beta}$ ,  $\varepsilon = N^{-\gamma/d}$ , then

$$\begin{aligned} \mathbb{P}(\omega_1^c) &= \mathbb{P}(X^{N,0} : \eta_m^0 \leq \varepsilon^r) = \mathbb{P}\left(X^{N,0} : \eta_m^0 \leq N^{-\frac{\gamma r}{d}}\right) \\ &\leq \mathbb{P}\left(X^{N,0} : \eta_m^0 \leq LN^{-\frac{2p-1}{d(p-1)}}\right) \leq 1 - e^{-2c_d^{1/p'} \|\rho^0\|_p L^{d/p'}} \\ &\leq 2c_d^{1/p'} \|\rho^0\|_p L^{d/p'} \leq CN^{-s}, \end{aligned}$$

for a sufficiently large  $N$  such that  $N \geq (2c_d^{1/p'} \|\rho^0\|_p)^{\frac{p'}{p'+d\beta}}$ , where  $s = \frac{d\beta}{p'}$ . For the estimate of  $\mathbb{P}(\omega_2^c)$ , we use the result of Proposition 5.5 to obtain

$$\mathbb{P}(\omega_2^c) \leq CN^\gamma e^{-CN^{1-\gamma}}.$$

Finally the estimate of  $\mathbb{P}(\omega_3^c)$  follows from (Boissard, 2011b, Proposition 1.2 of Annexe A) (see also Boissard (2011a); Bolley, Guillin, and Villani (2007)) that

$$\mathbb{P}(X^{N,0} : d_1(\mu_N^0, \rho^0) \geq \varepsilon) \leq CN^{-s'},$$

where  $C$  and  $s'$  are positive constants. We now denote  $\omega := \omega_1 \cap \omega_2 \cap \omega_3$ . Then we have

$$\mathbb{P}(\omega^c) \leq CN^{-l},$$

for some positive constants  $C$  and  $l$ . If the initial data belongs to  $\omega$ , then we obtain from Proposition 5.3 that

$$d_1(\rho_N(t), \mu_N(t)) \leq d_\infty(\rho_N(t), \mu_N(t)) \leq \frac{Ce^{CT}}{N^{\gamma/d}}, \quad \text{for } t \in [0, T].$$

We also notice from Theorem 4.1 that

$$d_1(\rho(t), \rho_N(t)) \leq d_1(\rho^0, \rho_N^0) e^{CT} \leq (d_1(\rho^0, \mu_N^0) + d_\infty(\mu_N^0, \rho_N^0)) e^{CT},$$

for all  $t \in [0, T]$ . Since  $d_\infty(\mu_N^0, \rho_N^0) \leq \varepsilon$  and the initial data belongs to  $\omega$ , this yields

$$d_1(\rho(t), \rho_N(t)) \leq \frac{Ce^{CT}}{N^{\gamma/d}},$$

for all  $t \in [0, T]$  since

$$d_1(\rho^0, \rho_{N,\varepsilon}^0) \leq d_1(\rho^0, \mu_N^0) + d_\infty(\mu_N^0, \rho_{N,\varepsilon}^0) \leq \frac{Ce^{CT}}{N^{\gamma/d}}.$$

Hence, we have

$$\mathbb{P}(\omega) \leq \mathbb{P}\left(\sup_{t \in [0, T]} d_1(\rho(t), \rho_N(t)) \leq \frac{Ce^{CT}}{N^{\gamma/d}}\right),$$

and it implies the desired result

$$\mathbb{P}\left(\sup_{t \in [0, T]} d_1(\rho(t), \rho_N(t)) \geq \frac{Ce^{CT}}{N^{\gamma/d}}\right) \leq \mathbb{P}(\omega^c) \leq \frac{C}{N^l}.$$

□

### Acknowledgments

JAC was partially supported by the project MTM2011-27739-C04-02 DGI (Spain) and 2009-SGR-345 from AGAUR-Generalitat de Catalunya. JAC acknowledges support from the Royal Society by a Wolfson Research Merit Award. YPC was supported by Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (ref. 2012R1A6A3A03039496). JAC and YPC were supported by Engineering and Physical Sciences Research Council grants with references EP/K008404/1 (individual grant) and EP/I019111/1 (platform grant).



## Bibliography

- M. Agueh, R. Illner, and A. Richardson, *Analysis and simulations of a refined flocking and swarming model of Cucker-Smale type*, *Kinetic and Related Models* 4:1–16, 2011.
- L.A. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, *Lectures in Mathematics*, Birkhäuser, 2005.
- G. Albi, L. Pareschi, *Modelling self-organized systems interacting with few individuals: from microscopic to macroscopic dynamics*, *Applied Math. Letters*, 26:397–401, 2013.
- I. Aoki, *A Simulation Study on the Schooling Mechanism in Fish*, *Bull. Jap. Soc. Sci. Fisheries* 48:1081–1088, 1982.
- D. Balagué, and J. A. Carrillo, *Aggregation equation with growing at infinity attractive-repulsive potentials*, *Proceedings of the 13th International Conference on Hyperbolic Problems*, Series in Contemporary Applied Mathematics CAM 17, Higher Education Press, 1:136–147, 2012.
- D. Balagué, Carrillo, T. J. A., Laurent, and G. Raoul, *Nonlocal interactions by repulsive-attractive potentials: radial ins/stability*, to appear in *Physica D*, 2013.
- D. Balagué, Carrillo, T. J. A., Laurent, and G. Raoul, *Dimensionality of Local Minimizers of the Interaction Energy*, to appear in *Arch. Rat. Mech. Anal.*, 2013.
- A. Barbaro, K. Taylor, P. F. Trethewey, L. Youseff, and B. Birnir, *Discrete and continuous models of the dynamics of pelagic fish: application to the capelin*, *Math. and Computers in Simulation*, 79:3397–3414, 2009.
- A. L. Bertozzi and T. Laurent, *Finite-time blow-up of solutions of an aggregation equation in  $\mathbb{R}^n$* , *Comm. Math. Phys.*, 274:717–735, 2007.
- A. L. Bertozzi, J. A. Carrillo, and T. Laurent, *Blowup in multidimensional aggregation equations with mildly singular interaction kernels*, *Nonlinearity*, 22:683–710, 2009.
- A. L. Bertozzi, T. Laurent, and J. Rosado,  *$L^p$  theory for the multidimensional aggregation equation*, *Comm. Pure Appl. Math.*, 43:415–430, 2010.
- A. L. Bertozzi, T. Laurent, and F. Léger, *Aggregation and spreading via the newtonian potential: the dynamics of patch solutions*, *Mathematical Models and Methods in Applied Sciences*, 22(supp01):1140005, 2012.
- M. Bodnar, J.J.L. Velazquez, *Friction dominated dynamics of interacting particles locally close to a crystallographic lattice*, accepted in *Math. Methods Appl. Sci.*, 2012.
- E. Boissard, *Simple bounds for convergence of empirical and occupation measures in 1-Wasserstein distance*, *Electron. J. Probab.* 16:2296–2333, 2011.

- E. Boissard, *Problèmes d'interaction discret-continu et distances de Wasserstein*, Thesis, <http://thesesups.ups-tlse.fr/1389/>, 2011.
- F. Bolley, J. A. Cañizo, and J. A. Carrillo *Stochastic mean-field limit: non-Lipschitz forces & swarming*, *Math. Mod. Meth. Appl. Sci.*, 21:2179-2210, 2011.
- F. Bolley, A. Guillin, and C. Villani, *Quantitative concentration inequalities for empirical measures on non-compact spaces*, *Probab. Theory Related Fields*, 137:541–593, 2007.
- W. Braun and K. Hepp, *The Vlasov Dynamics and Its Fluctuations in the  $1/N$  Limit of Interacting Classical Particles*, *Commun. Math. Phys.*, 56:101–113, 1977.
- M. Burger, P. Markowich, and J. Pietschmann, *Continuous limit of a crowd motion and herding model: Analysis and numerical simulations*, *Kinetic and Related Methods*, 4:1025–1047, 2011.
- S. Camazine, J.-L. Deneubourg, N. R. Franks, J. Sneyd, G. Theraulaz, and E. Bonabeau, *Self-Organization in Biological Systems*, Princeton University Press, 2003.
- J.A. Cañizo, J.A. Carrillo, and J. Rosado, *Collective Behavior of Animals: Swarming and Complex Patterns*, *Arbor*, 186:1035–1049, 2010.
- J.A. Cañizo, J.A. Carrillo, and J. Rosado : *A well-posedness theory in measures for some kinetic models of collective motion*, *Math. Mod. Meth. Appl. Sci.*, 21:515–539, 2011.
- 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.*, 156:229–271, 2011.
- J. A. Carrillo, M. Di Francesco, A. Figalli, T. Laurent, and D. Slepčev, *Confinement in nonlocal interaction equations*, *Nonlinear Anal.*, 75(2):550–558, 2012.
- J. A. Carrillo, M. R. D’Orsogna, and V. Panferov, *Double milling in self-propelled swarms from kinetic theory*, *Kinetic and Related Models* 2:363-378, 2009.
- J.-A. Carrillo, M. Fornasier, J. Rosado, and G. Toscani, *Asymptotic Flocking Dynamics for the kinetic Cucker-Smale model*, *SIAM J. Math. Anal.*, 42:218–236, 2010.
- J.A. Carrillo, M. Fornasier, G. Toscani, and F. Vecil, *Particle, Kinetic, and Hydrodynamic Models of Swarming*, *Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences, Series: Modelling and Simulation in Science and Technology*, Birkhauser, pages 297-336, 2010.
- J.A. Carrillo, A. Klar, S. Martin, and S. Tiwari, *Self-propelled interacting particle systems with roosting force*, *Math. Mod. Meth. Appl. Sci.*, 20:1533–1552, 2010.

- J. A. Carrillo, V. Panferov, S. Martin, *A new interaction potential for swarming models, to appear in Physica D*, 2013.
- 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*, 19:1–48, 2003.
- 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.*, 179:217–263, 2006.
- T. Champion, L. D. Pascale, and P. Juutinen, *The  $\infty$ -Wasserstein distance: local solutions and existence of optimal transport maps*, *SIAM J. Math. Anal.*, 40:1–20, 2008.
- I. D. Couzin, J. Krause, N. R. Franks, and S. A. Levin, *Effective leadership and decision making in animal groups on the move*, *Nature* 433:513–516, 2005.
- F. Cucker and S. Smale, *Emergent behavior in flocks*, *IEEE Trans. Automat. Control* 52:852–862, 2007.
- C. De Lellis and L. Székelyhidi, *The Euler equations as a differential inclusion*, *Ann. of Math.* 170:1417–1436, 2009.
- R. Dobrushin, *Vlasov equations*, *Funct. Anal. Appl.* 13:115–123, 1979.
- M. R. D’Orsogna, Y. L. Chuang, A. L. Bertozzi, and L. Chayes, *Self-propelled particles with soft-core interactions: patterns, stability, and collapse*, *Phys. Rev. Lett.* 96, 2006.
- R. Eftimie, G. de Vries, and M.A. Lewis, *Complex spatial group patterns result from different animal communication mechanisms*, *Proceedings of the National Academy of Sciences*, 104:6974–6979, 2007.
- N. Fournier, M. Hauray, and S. Mischler, *Propagation of chaos for the 2D viscous vortex model*, preprint <http://arxiv.org/abs/1212.1437>
- I. Gallagher, L. St-Raymond, and B. Texier, *From Newton to Boltzmann: hard spheres and short-range potentials*, preprint <http://arxiv.org/abs/1208.5753>
- C. R. Givens and R. M. Shortt, *A class of Wasserstein metrics for probability distributions*, *Michigan Math. J.*, 31(2):231–240, 1984.
- F. Golse, *The Mean-Field Limit for the Dynamics of Large Particle Systems*, *Journées équations aux dérivées partielles*, 9:1–47, 2003.
- J. Goodman, T. Hou, and J. Lowengrub *Convergence of the point vortex method for the 2-D Euler equations*, *Comm. Pure Appl. Math.* 43:415–430, 1990.
- C. Graham, and S. Méléard, *Stochastic particle approximations for generalized Boltzmann models and convergence estimates*, *The Annals of Probability*, 25:115–132, 1997.

- S.-Y. Ha, J.-G. Liu, A simple proof of the Cucker-Smale flocking dynamics and mean-field limit, *Commun. Math. Sci.* 7 (2) (2009) 297–325.
- S.-Y. Ha and E. Tadmor, *From particle to kinetic and hydrodynamic descriptions of flocking*, *Kinetic and Related Models* 1:415–435, 2008.
- J. Haskovec, *Flocking dynamics and mean-field limit in the Cucker-Smale-type model with topological interactions*, preprint, 2013.
- M. Hauray, *Wasserstein distances for vortices approximation of Euler-type equations*, *Math. Mod. Meth. Appl. Sci.*, 19:1357–1384, 2009.
- M. Hauray and P.-E. Jabin, *Particles approximations of Vlasov equations with singular forces : Propagation of chaos*, preprint.
- M. Hauray and S. Mischler. *On Kac’s chaos and related problems*, preprint.
- C. K. Hemelrijk and H. Hildenbrandt, *Self- Organized Shape and Frontal Density of Fish Schools*, *Ethology* 114, 2008.
- A. Huth and C. Wissel, *The Simulation of the Movement of Fish Schools*, *J. Theo. Bio.*, 1992.
- M. Kac, *Foundations of kinetic theory*, In Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III (Berkeley and Los Angeles, 1956), University of California Press, pp. 171–197.
- H. P. McKean, *An exponential formula for solving Boltzmann’s equation for a Maxwellian gas*, *J. Combinatorial Theory*, 2:358–382, 1967.
- H. P. McKean, *The central limit theorem for Carleman’s equation*, *Israel J. Math.*, 21:54–92, 1975.
- T. Kolokonikov, H. Sun, D. Uminsky, and A. Bertozzi. *Stability of ring patterns arising from 2d particle interactions*, *Physical Review E*, 84:015203, 2011.
- O.E. III. Lanford *Time evolution of large classical systems*. Lecture Notes in Phys. 38, Springer Verlag 1975, p. 1–111.
- T. Laurent, *Local and global existence for an aggregation equation*, *Communications in Partial Differential Equations*, 32:1941–1964, 2007.
- H. Levine, W.-J. Rappel and I. Cohen, *Self-organization in systems of self-propelled particles*, *Phys. Rev. E*, 63:017101-1/4, 2000.
- Y. X. Li, R. Lukeman, and L. Edelstein-Keshet, *Minimal mechanisms for school formation in self-propelled particles*, *Physica D*, 237:699–720, 2008.
- R. J. McCann, *Stable rotating binary stars and fluid in a tube*, *Houston J. Math.*, 32(2):603–631, 2006.
- A. Majda and A.L. Bertozzi, *Vorticity and Incompressible Flow*, Cambridge Texts in Applied Mathematics, Cambridge University Press, United Kingdom, 2002.

- C. Marchioro and M. Pulvirenti, *Mathematical theory of incompressible non-viscous fluids*, Applied Mathematical Sciences 96, Springer-Verlag, New York, 1994.
- S. Méléard, *Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models*, In Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995), Lecture Notes in Math. 1627. Springer, Berlin, 1996.
- S. Mischler and C. Mouhot, *Kac's program in kinetic theory*, to appear in *Inventiones mathematicae*, 2013.
- A. Mogilner and L. Edelstein-Keshet, *A non-local model for a swarm*, *J. Math. Bio.*, 38:534–570, 1999.
- A. Mogilner, L. Edelstein-Keshet, L. Bent, and A. Spiros, *Mutual interactions, potentials, and individual distance in a social aggregation*, *J. Math. Biol.*, 47:353–389, 2003.
- H. Neunzert, *An introduction to the nonlinear Boltzmann-Vlasov equation*, In *Kinetic theories and the Boltzmann equation* (Montecatini Terme, 1981), Lecture Notes in Math. 1048. Springer, Berlin, 1984.
- H. Osada, *Propagation of chaos for the two-dimensional Navier-Stokes equation*, In *Probabilistic methods in mathematical physics* (Katata/Kyoto, 1985), Academic Press, Boston, 1987, 303–334.
- F. Otto, *The geometry of dissipative evolution equations: the porous medium equation*, *Comm. Partial Differential Equations* 26:101–174, 2001.
- J. Parrish, and L. Edelstein-Keshet, *Complexity, pattern, and evolutionary trade-offs in animal aggregation*, *Science*, 294: 99–101, 1999.
- L. Perea, G. Gómez, P. Elosegui, *Extension of the cuckoo-smale control law to space flight formations*, *AIAA Journal of Guidance, Control, and Dynamics*, 32:527–537, 2009.
- S. Schochet, *The point-vortex method for periodic weak solutions of the 2-D Euler equations*, *Comm. Pure Appl. Math.* 49:911–965, 1996.
- H. Spohn, *Large scale dynamics of interacting particles*, *Texts and Monographs in Physics*, Springer, 1991.
- A.-S. Sznitman, *Topics in propagation of chaos*, In *Ecole d'Été de Probabilités de Saint-Flour XIX 1989*, Lecture Notes in Math. 1464. Springer, Berlin, 1991.
- C.M. Topaz and A.L. Bertozzi, *Swarming patterns in a two-dimensional kinematic model for biological groups*, *SIAM J. Appl. Math.*, 65:152–174, 2004.
- C.M. Topaz, A.L. Bertozzi, and M.A. Lewis, *A nonlocal continuum model for biological aggregation*, *Bulletin of Mathematical Biology*, 68:1601–1623, 2006.

- T. Vicsek, A. Czirok, E. Ben-Jacob, I. Cohen, and O. Shochet, *Novel type of phase transition in a system of self-driven particles*, *Phys. Rev. Lett.*, 75:1226–1229, 1995.
- C. Villani, *Limite de champ moyen*, *Lecture notes*, 2002.
- C. Villani, *Topics in optimal transportation*, volume 58 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2003.