# First-order aggregation models and zero inertia limits

R. C. Fetecau *   W. Sun †

May 12, 2015

**Abstract**

We consider a first-order aggregation model in both discrete and continuum formulations and show rigorously how it can be obtained as zero inertia limits of second-order models. In the continuum case the procedure consists in a macroscopic limit, enabling the passage from a kinetic model for aggregation to an evolution equation for the macroscopic density. We work within the general space of measure solutions and use mass transportation ideas and the characteristic method as essential tools in the analysis.

**Keywords** : aggregation models; kinetic equations; macroscopic limit; measure solutions; mass transportation; particle methods

## 1 Introduction

The focus of the present paper is a certain mathematical model for emerging self-collective behaviour in biological (and other) aggregations. There has been a surge of activity in this area of research during the past decade, and in fact the goals have extended well beyond biology. For biological applications, the primary motivation has been to understand and model the mechanisms behind the formation of the various spectacular groups observed in nature (fish schools, bird flocks, insect swarms) [11]. In terms of expansion of this research into collateral areas, we mention studies on robotics and space missions [27], opinion formation [33], traffic and pedestrian flow [23] and social networks [26].

Aggregation models can be classified in two main classes: i) individual/ particle-based, where the movements of all individuals in the group are being tracked, and ii) partial differential equations (PDE) models, formulated as evolution equations for the population density field. We refer to [14] for a recent review of models for aggregation behaviour, where the various microscopic/ macroscopic descriptions of collective motion are discussed and connected. In the present work we deal with a model that has both a discrete/ODE and a continuum/PDE formulation.

The continuum aggregation model considered in this article is given by the following evolution equation for the population density $\rho(t, x)$ in $\mathbb{R}^d$:

$$\rho_t + \nabla \cdot (\rho u) = 0, \tag{1.1a}$$

$$u = -\nabla K * \rho, \tag{1.1b}$$

*Department of Mathematics, Simon Fraser University, 8888 University Dr., Burnaby, BC V5A 1S6, Canada. Email: van@math.sfu.ca

†Department of Mathematics, Simon Fraser University, 8888 University Dr., Burnaby, BC V5A 1S6, Canada. Email: weirans@math.sfu.ca

where $K$ represents an interaction potential and $*$ denotes convolution. The potential $K$ typically incorporates social interactions such as short-range repulsion and long-range attraction. We consider $K$ to be radial, meaning that the inter-individual interactions are assumed to be isotropic.

Equation (1.1) appears in various contexts related to mathematical models for biological aggregations; we refer to [32, 36] and references therein for an extensive background and review of the literature on this topic. It also arises in a number of other applications such as material science and granular media [37], self-assembly of nanoparticles [24] and molecular dynamics simulations of matter [22]. The model has become widely popular and there has been intensive research on it during recent years.

The particular appeal of model (1.1) has lain in part in its simple form, which allowed rapid progress in terms of both numerics and analysis. Numerical simulations demonstrated a wide variety of self-collective or "swarm" behaviours captured by model (1.1), resulting in aggregations on disks, annuli, rings, soccer balls, etc [28, 39, 40]. Analysis-oriented studies addressed the well-posedness of the initial-value problem for (1.1) [9, 10, 30, 5, 13, 6], as well as the long time behaviour of its solutions [10, 31, 17, 4, 20, 19]. Also, there has been increasing interest lately on the analysis of (1.1) by variational methods [3, 2, 15].

Equation (1.1) is frequently regarded as the continuum approximation, when the number of particles increases to infinity, of the following individual-based model. Consider $N$ particles in $\mathbb{R}^d$ whose positions $x_i$ $(i = 1, \ldots, N)$ evolve according to the ODE system

$$\frac{dx_i}{dt} = v_i, \tag{1.2a}$$

$$v_i = -\frac{1}{N}\sum_{j\neq i} \nabla_{x_i} K(x_i - x_j), \tag{1.2b}$$

where $K$ denotes the same interaction potential as in (1.1).

Model (1.2) was justified and formally derived in [8], starting from the following second-order model in Newton's law form $(i = 1, \ldots, N)$:

$$\epsilon\frac{d^2x_i}{dt^2} + \frac{dx_i}{dt} = F_i, \qquad \text{with} \qquad F_i = -\frac{1}{N}\sum_{j\neq i} \nabla_{x_i} K(x_i - x_j), \tag{1.3}$$

and $\epsilon > 0$ small. From a biological point of view, (1.3) considers some small inertia/response time of individuals. By neglecting the $\epsilon$-term in (1.3), one can formally derive model (1.2). However, as noted in [8], making $\epsilon = 0$ translates to instantaneous changes in velocities, assumption which, quote, "is probably too restrictive in many cases".

In view of (1.2), one can write (1.3) more conveniently as

$$\frac{dx_i}{dt} = v_i, \tag{1.4a}$$

$$\epsilon\frac{dv_i}{dt} = -v_i - \frac{1}{N}\sum_{j\neq i} \nabla_{x_i} K(x_i - x_j). \tag{1.4b}$$

We point out that despite being at the origin of the extensively-studied models (1.2) and (1.1), the second-order model (1.3) (or (1.4)) and in particular, the $\epsilon \to 0$ limit of its solutions, have been overlooked completely. As briefly demonstrated in Section 2, a rigorous passage from model from (1.4) to (1.2) can be obtained in the $\epsilon \to 0$ limit by using a classical theorem due to Tikhonov [35].

The main focus of the present work is the analogous $\epsilon \to 0$ limit at the PDE level. Specifically, we investigate a zero inertia limit that yields the continuum model (1.1). Using techniques reviewed in [14], one can formally take the limit $N \to \infty$ and associate to (1.4) the following kinetic equation for the density $f(t, x, v)$ of individuals at position $x \in \mathbb{R}^d$ with velocity $v \in \mathbb{R}^d$:

$$f_t + v \cdot \nabla_x f = \frac{1}{\epsilon}\nabla_v \cdot (vf) + \frac{1}{\epsilon}\nabla_v \cdot ((\nabla_x K * \rho)f), \tag{1.5}$$

where

$$\rho(t, x) = \int_{\mathbb{R}^d} f(t, x, v)dv. \tag{1.6}$$

We consider measure-valued solutions of the kinetic model (1.5), in the framework of the well-posedness theory developed in [12], and study their macroscopic limit $\epsilon \to 0$. Passage from kinetic to macroscopic equations has been extensively studied in the hydrodynamic limits of the nonlinear Botlzmann equations for both classical and renormalized solutions. It is beyond the scope of this introduction to give a detailed account of the work that has been done in this vast and well-established research area, we simply refer here to a recent review paper [34] and the references therein.

The recurring theme and the key approach in the paper is the method of characteristics and measure-valued solutions to (1.1) and (1.5) defined in the mass transportation sense. Viewed from this perspective, the subtlety lies in the fact that the $\epsilon \to 0$ limit is *degenerate*, as characteristics of a second-order model[1] collapse to first-order characteristics. Theorem 5.2 proves this singular limit. Within the mass transportation framework, we then arrive in Theorem 5.3 at the result that solutions $f_\epsilon(t, x, v)$ to (1.5) converge weak-* as measures to $f(t, x, v) = \rho(t, x)\delta(v - u(t, x))$ as $\epsilon \to 0$, where $u$ is given in terms of $\rho$ by (1.1b) and $\rho$ satisfies the continuum equation (1.1a).

We note that a related study was done in [25] in the context of the macroscopic limit of a Vlasov type equation with friction. Our results, derived independently of [25], consider a more general space setup for solutions of the kinetic model (4.1); our solution space includes measures, while in [25] solutions of the Vlasov equation are taken to be bounded and integrable. Allowing for measure-valued solutions in aggregation models is important for connecting the discrete and the kinetic models, e.g., in associating an empirical distribution density to a discrete configuration. Also, the analysis in [25] follows exclusively a PDE approach, while ours is a mixture of PDE analysis and the method of characteristics. Once more, our approach is more suitable and in a certain sense essential in unifying the continuum and the discrete formulations of the aggregation model under study. In terms of PDE estimates, similar second-order fluctuation terms (that are shown to vanish with $\epsilon$) are considered in both [25] and the present work. Controlling such terms is essential in passing to the zero $\epsilon$ limit.

A key motivation for the zero inertia limits investigated in this article is the following. The recent work [16] of one of the authors showed that the second-order model (1.3) is absolutely essential provided one wants to include anisotropy in model (1.2). Specifically, [16] considers anisotropic inter-individual interactions in model (1.2) by replacing the explicit

[1]Strictly speaking, (1.5) is a first-order PDE, but we refer to it as a second-order model as it is essentially based on Newton's law (1.3). Furthermore, for $\epsilon > 0$ fixed, the monokinetic closure of (1.5) yields a momentum equation for the velocity, also in the form of Newton's second law [14].

representation (1.2b) of the velocities by a *weighted* sum, with weights that depend on a restricted visual perception of the individuals. Hence, these weights depend on the velocity vectors themselves, and the anisotropic analogue of (1.2b) becomes an *implicit* equation to be solved for $v_i$. It is shown in [16] that solutions of such implicit equations are generally non-unique and additionally, encounter discontinuities through the time evolution. The relaxation system (1.4), along with its small inertia/response time, is proposed in [16] as a biologically meaningful mechanism to select unique solutions and physically *correct* velocity jumps.

As for the ODE case, the present study sets the stage for generalizations of the continuum model (1.1) to include anisotropic interactions. In such an extension, (1.1b) would become an implicit equation for $u$ and issues such as non-uniqueness and loss of smoothness are again expected to arise. We argue that understanding how to approximate first-order models such as (1.1) and (1.2) (and subsequently, their generalizations) in the $\epsilon \to 0$ limit of second-order models is entirely essential for making further progress in this area of research.

Finally, we point out that we work in this paper with smooth potentials $K$ that satisfy $\nabla K \in W^{1,\infty}(\mathbb{R}^d)$. First, the assumption $\nabla K \in L^\infty(\mathbb{R}^d)$ is needed to show a uniform (in $\epsilon$) bound on the support of solutions to the kinetic equation (4.1) (Proposition 4.1). A further requirement ($\nabla^2 K \in L^\infty_{loc}(\mathbb{R}^d)$) is then needed to estimate the second-order fluctuation term in Proposition 4.2. The results in this paper do not apply for instance to quadratic potentials (due to growth at infinity) or pointy potentials, such as Morse potentials (due to the smoothness requirement). However, from the point of view of applications, some of these restrictions should not play a big role. As noted in [12] for instance, as far as numerical simulations are concerned, it makes little difference to distinguish between a pointy potential and its smooth regularization, as the two would give qualitatively similar aggregation behaviour.

The summary of the paper is as follows. Section 2 presents the $\epsilon \to 0$ limit of the ODE model (1.4). Section 3 contains a brief formal derivation of the kinetic model (1.5) from (1.4) and summarizes the results from [12] regarding the well-posedness of measure-valued solutions of (1.5). In Section 4 we derive uniform in $\epsilon$ estimates for solutions to (1.5) needed for passing the limit $\epsilon \to 0$. Section 5 contains the major results of this paper, which is the convergence of solutions to (1.5) as $\epsilon \to 0$ and how solutions of (1.1) are recovered in this limit.

## 2 Convergence as $\epsilon \to 0$ of the ODE model (1.4)

The limit $\epsilon \to 0$ of solutions to (1.4) can be carried out by a straightforward application of the general theory originally developed by Tikhonov [35]. An excellent account of this theory can be found in [38]. Since the application of Tikhonov's theorem does not appear in any of the works on aggregation models, we find it worthwhile to summarize the concepts, as well as to state the convergence result in the context of models (1.4) and (1.2). In fact, this framework will be used again in Section 5, when we study the $\epsilon \to 0$ limit of the PDE model (1.5).

Following the setup in [38], consider the general system

$$\begin{cases} \dfrac{d\mathbf{x}}{dt} = \mathbf{v}, \\ \epsilon \dfrac{d\mathbf{v}}{dt} = \mathcal{F}(\mathbf{x}, \mathbf{v}), \end{cases} \tag{2.1}$$

where $\mathbf{x}, \mathbf{v} \in \mathbb{R}^{Nd}$ and $\epsilon > 0$.

Note that indeed, system (1.4) can be written compactly as (2.1) provided $\mathbf{x}, \mathbf{v}$ denote the concatenation of the space and velocity vectors:

$$\mathbf{x} = (x_1, \dots, x_N), \qquad \mathbf{v} = (v_1, \dots, v_N), \tag{2.2}$$

and

$$\mathcal{F}(\mathbf{x}, \mathbf{v}) = \left(\mathcal{F}_1(\mathbf{x}, v_1), \dots, \mathcal{F}_N(\mathbf{x}, v_N)\right), \tag{2.3}$$

with

$$\mathcal{F}_i(\mathbf{x}, v_i) = -v_i - \frac{1}{N}\sum_{j\neq i} \nabla_{x_i} K(x_i - x_j), \qquad i = 1, \dots, N. \tag{2.4}$$

System (2.1) is a typical two-scale equation, where $t$ and $\tau = t/\epsilon$ represent the slow and fast time scales, respectively. Using classical terminology in singular perturbation theory for ODEs [18], the slow dynamics of (2.1) is given by

$$\begin{cases} \dfrac{d\mathbf{x}}{dt} = \mathbf{v}, \\ \\ \mathbf{v} = \Gamma(\mathbf{x}), \end{cases} \tag{2.5}$$

where $\mathbf{v} = \Gamma(\mathbf{x})$ is a root of the equation

$$\mathcal{F}(\mathbf{x}, \mathbf{v}) = 0. \tag{2.6}$$

Equation (2.6) defines the centre manifold for (2.1). Note that roots $\mathbf{v} = \Gamma(\mathbf{x})$ of (2.6) are in general, non-unique. In addition, for a fixed configuration $\mathbf{x}^*$, the fast dynamics is governed by

$$\frac{d\mathbf{v}}{d\tau} = \mathcal{F}(\mathbf{x}^*, \mathbf{v}). \tag{2.7}$$

Tikhonov's theorem establishes conditions on a root $\mathbf{v} = \Gamma(\mathbf{x})$ which guarantee that solutions of the two-scale system (2.1) converge, via a fast initial layer governed by (2.7), to solutions of the degenerate/slow system (2.5).

In our problem (see (2.2)-(2.4)), model (1.2) is the slow dynamics system for (1.4), associated to the *unique* root $\mathbf{v} = \Gamma(\mathbf{x})$, where $\Gamma(\mathbf{x}) = (\gamma_1(\mathbf{x}), \dots, \gamma_N(\mathbf{x}))$ is given *explicitly* by

$$v_i = \gamma_i(\mathbf{x}) := -\frac{1}{N}\sum_{j\neq i} \nabla_{x_i} K(x_i - x_j). \tag{2.8}$$

In general, take a closed and bounded set $D \subset \mathbb{R}^{Nd}$, and consider a root $\mathbf{v} = \Gamma(\mathbf{x})$, $\Gamma : D \to \mathbb{R}^{Nd}$. The root $\Gamma$ is called *isolated* if there is a $\delta > 0$ such that for all $\mathbf{x} \in D$, the only element in $B(\Gamma(\mathbf{x}), \delta)$ that satisfies $\mathcal{F}(\mathbf{x}, \mathbf{v}) = 0$ is $\mathbf{v} = \Gamma(\mathbf{x})$. An isolated root $\Gamma$ is called *positively stable* in $D$, if $\mathbf{v}^* = \Gamma(\mathbf{x}^*)$ is an asymptotically stable stationary point of (2.7) as $\tau \to \infty$, for each $\mathbf{x}^* \in D$. The *domain of influence* of an isolated positively stable root $\Gamma$ is the set of points $(\mathbf{x}^*, \tilde{\mathbf{v}})$ such that the solution of (2.7) satisfying $\mathbf{v}|_{\tau=0} = \tilde{\mathbf{v}}$ tends to $\mathbf{v}^* = \Gamma(\mathbf{x}^*)$ as $\tau \to \infty$.

Tikhonov's theorem [35] states the following:

**Theorem 2.1** (Tikhonov [35, 38]). *Assume that a root* $\mathbf{v} = \Gamma(\mathbf{x})$ *of* (2.6) *is isolated positively stable in some bounded closed domain* $D$. *Consider a point* $(\mathbf{x}_0, \mathbf{v}_0)$ *in the domain of influence of this point, and assume that the slow equation* (2.5) *has a solution* $\mathbf{x}(t)$ *initialized at* $\mathbf{x}(t_0) = \mathbf{x}_0$*, that lies in* $D$ *for all* $t \in [t_0, T]$. *Then, as* $\epsilon \to 0$*, the solution* $(\mathbf{x}_\epsilon(t), \mathbf{v}_\epsilon(t))$ *of* (2.1) *initialized at* $(\mathbf{x}_0, \mathbf{v}_0)$*, converges to* $(\mathbf{x}(t), \mathbf{v}(t)) := (\mathbf{x}(t), \Gamma(\mathbf{x}(t)))$ *in the following sense:*

*i)* $\lim_{\epsilon \to 0} \mathbf{v}_\epsilon(t) = \mathbf{v}(t)$ *for all* $t \in (t_0, T^*]$*, and*

*ii)* $\lim_{\epsilon \to 0} \mathbf{x}_\epsilon(t) = \mathbf{x}(t)$ *for all* $t \in [t_0, T^*]$,

*for some* $T^* < T$.

*Remark* 2.1. Note that the convergence of $\mathbf{v}_\epsilon(t)$ to $\mathbf{v}(t)$ occurs via a fast initial layer and normally does not occur at the initial time $t_0$, unless the initial data is on the centre manifold, i.e., $\mathbf{v}_0 = \Gamma(\mathbf{x}_0)$.

Applying Theorem 2.1 to models (1.4) and (1.2) is immediate. Given any spatial configuration $\mathbf{x}$, the root $\mathbf{v} = \Gamma(\mathbf{x})$ given by (2.8) is unique, hence isolated. Fix now an arbitrary spatial configuration $\mathbf{x}^* = (x_1^*, x_2^*, \ldots, x_N^*)$ and inspect the fast equation (2.7) with $\mathcal{F}$ given by (2.3)-(2.4). It is clear that each component of the fast system

$$\frac{dv_i}{d\tau} = -v_i - \frac{1}{N}\sum_{j\neq i} \nabla_{x_i} K(x_i^* - x_j^*)$$

has a globally attracting equilibrium $v_i^* = \gamma_i(\mathbf{x}^*)$ (one can simply interpret the interaction term, with $\mathbf{x}^*$ fixed, as an externally given force). Consequently, $\mathbf{v}^* = (v_1^*, \ldots, v_N^*)$ is positively stable and its domain of influence is $\{\mathbf{x}^*\} \times \mathbb{R}^{Nd}$.

**Theorem 2.2** (Convergence of the ODE model). *Consider a point* $(\mathbf{x}_0, \mathbf{v}_0) \in \mathbb{R}^{2Nd}$*, and suppose the first-order model* (1.2) *has a solution* $\mathbf{x}(t) = (x_1(t), \ldots, x_N(t)) \in \mathbb{R}^{Nd}$ *initialized at* $\mathbf{x}_0$ *for* $t \in [t_0, T]$. *Then, as* $\epsilon \to 0$*, the solution* $(\mathbf{x}_\epsilon(t), \mathbf{v}_\epsilon(t)) \in \mathbb{R}^{2Nd}$ *of the second-order model* (1.4) *initialized at* $(\mathbf{x}_0, \mathbf{v}_0)$*, converges to* $(\mathbf{x}(t), \mathbf{v}(t))$*, with* $\mathbf{v}(t) = (v_1(t), \ldots, v_N(t))$ *defined in terms of* $\mathbf{x}(t)$ *by* (2.8).

*Specifically, we have the convergence i) and ii) listed in Theorem 2.1, with the caveat that the convergence of* $\mathbf{v}_\epsilon(t)$ *does not hold initially, unless* $\mathbf{v}_0$ *and* $\mathbf{x}_0$ *are related by* (2.8).

# 3 Kinetic model and its well-posedness

## 3.1 Formal derivation of the kinetic model

The kinetic equation associated to the particle model (1.4) can be derived using the techniques reviewed in [14]. We present here the derivation via the mean-field limit [14].

Consider the empirical distribution density $f_N$ associated to the solution $(x_i(t), v_i(t))$ $(i = 1, \ldots, N)$ of (1.4), that is,

$$f_N(t, x, v) = \frac{1}{N}\sum_{i=1}^{N} \delta(x - x_i(t))\delta(v - v_i(t)).$$

Take a test function $\varphi \in C_0^1(\mathbb{R}^{2d})$ and compute, using (1.4):

$$\begin{aligned}\frac{d}{dt}\langle f_N(t),\varphi\rangle &= \frac{1}{N}\sum_{i=1}^{N}\frac{d}{dt}\varphi(x_i(t),v_i(t))\\ &= \frac{1}{N}\sum_{i=1}^{N}\nabla_x\varphi(x_i(t),v_i(t))\cdot v_i(t)\\ &\quad+\frac{1}{N}\sum_{i=1}^{N}\nabla_v\varphi(x_i(t),v_i(t))\cdot\frac{1}{\epsilon}\Big(-v_i-\frac{1}{N}\sum_{j\neq i}\nabla_{x_i}K(|x_i-x_j|)\Big).\end{aligned}$$

Denote by $\rho_N(t,x)$ the macroscopic density of $f_N$:

$$\rho_N(t,x) := \int f_N(t,x,v)dv = \frac{1}{N}\sum_{i=1}^{N}\delta(x-x_i(t)).$$

Since

$$\begin{aligned}\nabla K*\rho_N(x) &= \int \nabla K(x-y)\rho_N(y)dy\\ &= \frac{1}{N}\sum_{j=1}^{N}\nabla K(x-x_j),\end{aligned}$$

we get

$$\frac{d}{dt}\langle f_N(t),\varphi\rangle = \langle f_N(t),\nabla_x\varphi\cdot v\rangle + \left\langle f_N(t),\nabla_v\varphi\cdot\frac{1}{\epsilon}\left(-v-\nabla K*\rho_N(x)\right)\right\rangle.$$

Hence, after integration by parts in $x$ and $v$ we get

$$\left\langle \frac{\partial f_N}{\partial t}+v\cdot\nabla_x f_N-\frac{1}{\epsilon}\nabla_v\cdot\left(vf_N+(\nabla K*\rho_N)f_N\right),\varphi\right\rangle = 0.$$

In strong form, $f_N$ satisfies

$$\frac{\partial f_N}{\partial t}+v\cdot\nabla_x f_N = \frac{1}{\epsilon}\nabla_v\cdot(vf_N)+\frac{1}{\epsilon}\nabla_v\cdot\left((\nabla K*\rho_N)f_N\right).$$

Assuming that $f_N$ converges (on a subsequence) to a density $f$, taking the limit $N\to\infty$, formally, in the equation above, yields the kinetic equation (1.5).

### 3.2 Well-posedness for (1.5) with $\epsilon > 0$ fixed.

We discuss first the well-posedness theory of measure-valued solutions of (1.5), as developped in [12]. Since for later purposes (to send $\epsilon\to 0$) we need to work with smooth solutions, we present briefly the existence theory for classical solutions as well. A key ingredient is the method of characteristics, which is eventually used in Section 5 to connect the PDE analysis with the ODE theory via Tikhonov's theorem.

**Measure solutions.** In [12] the authors consider various kinetic models for aggregation and study the well-posedness of measure-valued solutions. The results there use the following measure space setup. Denote by $\mathcal{P}_1(\mathbb{R}^k)$ the space of probability measures on $\mathbb{R}^k$ that have finite first moment, i.e.,

$$\mathcal{P}_1(\mathbb{R}^k) = \left\{ f \in \mathcal{P}(\mathbb{R}^k) : \int_{\mathbb{R}^k} |x| f(x) dx < \infty \right\}.$$

We note that the convention adopted in [12], which is also used throughout the present paper, is to write $\int \varphi(x)\mu(x)dx$ as the integral of $\varphi$ with respect to the measure $\mu$, regardless of whether $\mu$ is absolutely continuous with respect to the Lebesgue measure.

*Remark* 3.1. Endowed with the 1-Wasserstein distance $W_1$, the space $\mathcal{P}_1(\mathbb{R}^k)$ is a complete metric space, and convergence in the $W_1$ metric relates to the usual weak-* convergence of measures. Specifically, for $\{f_n\}_{n\geq 1}$ and $f$ in $\mathcal{P}_1(\mathbb{R}^k)$, the following are equivalent:

i) $f_n \xrightarrow{W_1} f$, as $n \to \infty$

ii) $f_n \xrightarrow{w^*} f$ as measures as $n \to \infty$ and $\sup_{n\geq 1} \int_{|x|>R} |x| f_n(x) dx \to 0, \qquad$ as $R \to \infty$.

Results in [12] use $\mathcal{P}_1(\mathbb{R}^d \times \mathbb{R}^d)$ endowed with the $W_1$ distance as the measure space where solutions of the various kinetic models are sought for. Model (1.5), with $\epsilon > 0$ fixed, is in fact a particular case of the general class of models studied considered there and the results from [12] can be applied directly. We summarize briefly the results from [12].

Consider the characteristic equations associated to model (1.5):

$$\begin{aligned} \frac{\mathrm{d}x}{\mathrm{d}t} &= v\,, \\ \frac{\mathrm{d}v}{\mathrm{d}t} &= -\frac{1}{\epsilon}v - \frac{1}{\epsilon}\nabla_x K * \rho, \end{aligned} \tag{3.1}$$

initialized at $(x, v)_{|t=0} = (x_0, v_0)$. The main idea in [12] is to define a measure solution to (1.5) in a mass transportation sense, using the flow map defined by (3.1).

Suppose $E(t, x)$ is a given continuous vector field on $[0, T] \times \mathbb{R}^d$ which is locally Lipschitz with respect to $x$ . Take the characteristic system associated to $E$:

$$\begin{aligned} \frac{\mathrm{d}x}{\mathrm{d}t} &= v\,, \\ \frac{\mathrm{d}v}{\mathrm{d}t} &= -\frac{1}{\epsilon}v - \frac{1}{\epsilon}E(t, x), \end{aligned} \tag{3.2}$$

with initial data $(x, v)_{|t=0} = (x_0, v_0)$. Then standard ODE theory guarantees existence and uniqueness of smooth trajectories $(x_\epsilon, v_\epsilon) \in C^1([0, T], \mathbb{R}^d \times \mathbb{R}^d)$ originating from $(x_0, v_0)$. Furthermore, one can define the flow map $\mathcal{T}_E^{t,\epsilon}$ of (3.2) by

$$(x_0, v_0) \xrightarrow{\mathcal{T}_E^{t,\epsilon}} (x, v), \qquad (x, v) = (x_\epsilon(t), v_\epsilon(t)),$$

where $(x_\epsilon(t), v_\epsilon(t))$ is the unique solution of (3.2) that starts at $(x_0, v_0)$.

Now take a measure $f_0 \in \mathcal{P}_1(\mathbb{R}^d \times \mathbb{R}^d)$ and $T > 0$, and consider the mass-transport (or push-forward) of $f_0$ by $\mathcal{T}_E^{t,\epsilon}$. By definition, the push-forward $f_t = \mathcal{T}_E^{t,\epsilon} \# f_0$ is a measure-valued function $f : [0,T] \to \mathcal{P}_1(\mathbb{R}^d \times \mathbb{R}^d)$, that satisfies

$$\int_{\mathbb{R}^{2d}} \zeta(x,v) f(t,x,v) dxdv = \int_{\mathbb{R}^{2d}} \zeta(\mathcal{T}_E^{t,\epsilon}(X,V)) f_0(X,V) dXdV, \tag{3.3}$$

for all $\zeta \in C_b(\mathbb{R}^{2d})$.

Return to (3.1) and define the vector field $E[f]$ associated to a measure $f$ as

$$E[f] = -\nabla K * \rho. \tag{3.4}$$

Here, $\rho$ denotes the first marginal of $f$ defined by

$$\int_{\mathbb{R}^d} \tilde{\psi}(x)\rho(t,x)\,\mathrm{d}x = \int_{\mathbb{R}^{2d}} \tilde{\psi}(x) f(t,x,v)\,\mathrm{d}x\,\mathrm{d}v, \tag{3.5}$$

for all $\tilde{\psi} \in C_b(\mathbb{R}^d)$. Note that throughout the paper, by an abuse of notation, we also write $\rho$ as in (1.6).

The following definition of a measure solution of (1.5) is adopted from [12].

**Definition 3.1.** Take an initial measure $f_0 \in \mathcal{P}_1(\mathbb{R}^d \times \mathbb{R}^d)$ and $T > 0$. A function $f : [0,T] \to \mathcal{P}_1(\mathbb{R}^d \times \mathbb{R}^d)$ is a solution of the kinetic equation (1.5) with initial condition $f_0$ if:

1. The field $E[f]$ defined by (3.4) is locally Lipschitz with respect to $x$ and $E[f](t,x) < C(1+|x|)$, for all $t, x \in [0,T] \times \mathbb{R}^d$, for some $C > 0$, and
2. $f_t = \mathcal{T}_{E[f]}^{t,\epsilon} \# f_0$.

The main result in [12] establishes the existence and uniqueness for measure solutions via a fixed point argument. We state and discuss the result below.

**Theorem 3.1** (Measure-valued solutions [12])**.** *Assume the following properties on the potential $K$:*

$$\nabla K \text{ is locally Lipschitz, and } |\nabla K(x)| \le C(1+|x|) \text{ for all } x \in \mathbb{R}^d,$$

*for some $C > 0$. Consider $f_0 \in \mathcal{P}_1(\mathbb{R}^d \times \mathbb{R}^d)$ with compact support.*

*Then there exists a unique solution $f_\epsilon \in C([0,\infty); \mathcal{P}_1(\mathbb{R}^d \times \mathbb{R}^d))$ of (1.5), in the sense of Definition 3.1, whose support grows at a controlled rate. Specifically, there exists an increasing function $R_\epsilon(T)$ such that for all $T > 0$,*

$$\operatorname{supp} f_\epsilon(t) \subset B_{R_\epsilon(T)} \quad \text{for all } t \in [0,T].$$

The proof of this result in [12] relies on a fixed point argument. Briefly, the setup in [12, Theorem 3.10] is the following. Fix $T > 0$ and consider the metric space $\mathcal{F}$ made of all $f \in C([0,T]; \mathcal{P}_1(\mathbb{R}^d \times \mathbb{R}^d))$ such that the support of $f_t$ is contained in a fixed ball $B_R$ for all $t \in [0,T]$. The distance in $\mathcal{F}$ is taken to be

$$\mathcal{W}_1(f,g) = \sup_{t \in [0,T]} W_1(f_t, g_t),$$

where $W_1$ denotes the 1-Wasserstein distance. For $f \in \mathcal{F}$ fixed, define the map (see (3.4)):

$$\mathcal{G}[f](t) := \mathcal{T}^{t,\epsilon}_{E[f]} \# f_0.$$

It is shown in [12] that this map is contractive and hence, is has a unique fixed point in $\mathcal{F}$. This fixed point is the desired solution to (1.5).

We also note that by results in [29, 13], $f_\epsilon$ given by Theorem 3.1 is also a weak solution of (1.5), i.e., it satisfies

$$\begin{aligned} &\int_0^T \iint_{\mathbb{R}^d\times\mathbb{R}^d} \partial_t \phi(t,x,v) f_\epsilon(t,x,v)\,\mathrm{d}x\,\mathrm{d}v\,\mathrm{d}t + \int_0^T \iint_{\mathbb{R}^d\times\mathbb{R}^d} \nabla_x\phi\cdot v f_\epsilon\,\mathrm{d}x\,\mathrm{d}v\,\mathrm{d}t \\ &\qquad - \frac{1}{\epsilon}\int_0^T \iint_{\mathbb{R}^d\times\mathbb{R}^d} \nabla_v\phi\cdot(v+\nabla_x K * \rho_\epsilon) f_\epsilon\,\mathrm{d}x\,\mathrm{d}v\,\mathrm{d}t + \iint_{\mathbb{R}^d\times\mathbb{R}^d} \phi(0,x,v) f_0(x,v)\,\mathrm{d}x\,\mathrm{d}v = 0, \end{aligned} \tag{3.6}$$

for any $\phi \in C^1_c([0,T); C^1_b(\mathbb{R}^d\times\mathbb{R}^d))$. Here, we use the standard notation where the subindex $c$ refers to compact support with respect to the time variable $t$. For instance, $\xi \in C^1_c([0,T))$ may have $\xi(0) \neq 0$, while $\xi \in C^1_c((0,T))$ necessarily has $\xi(0) = 0$. In both cases, $\xi(T) = 0$.

*Remark* 3.2. Take a measure solution $f_\epsilon$ of (1.5) for $\epsilon > 0$ fixed. By definition, $f_\epsilon$ is the mass transport of $f_0$ along trajectories $(x_\epsilon(t), v_\epsilon(t))$ that satisfy the characteristic equations (see (3.1))

$$\begin{aligned} \frac{\mathrm{d}x}{\mathrm{d}t} &= v\,, \\ \epsilon\frac{\mathrm{d}v}{\mathrm{d}t} &= -v - \nabla_x K * \rho_\epsilon, \end{aligned} \tag{3.7}$$

$(x_\epsilon(0), v_\epsilon(0)) = (x_0, v_0)$. We are interested in this paper in the $\epsilon \to 0$ limit of $f_\epsilon$, $\rho_\epsilon$, as well as in the limiting behaviour of the characteristic trajectories $x_\epsilon(t)$, $v_\epsilon(t)$. This will require the use of the ODE framework from Section 2 combined with PDE estimates on (1.5) itself. To this end, we need to work first with classical solutions of (1.5) and derive uniform in $\epsilon$ estimates (see Theorem 4.2 for instance). Below is a brief account on existence theory for classical solutions of (1.5).

**Smooth solutions.** The existence of smooth solutions to (1.5) for $\epsilon > 0$ can be inferred using the classical framework for Vlasov type equations [21]. We state the theorem below and explain the steps of its proof. The full details of the proof can be found in [21, Chapter 4] for Vlasov-Maxwell equation.

**Theorem 3.2** (Existence of smooth solutions)**.** *Suppose $f_0 \in C^2(\mathbb{R}^d\times\mathbb{R}^d)\cap L^1(\mathbb{R}^d\times\mathbb{R}^d)$ and $\nabla_x K \in W^{1,\infty}(\mathbb{R}^d)$. Let $T > 0$ be arbitrary. Then equation (1.5) has a solution $f_\epsilon \in C([0,T]; C^1(\mathbb{R}^d\times\mathbb{R}^d))$ with initial data $f_\epsilon\big|_{t=0} = f_0$.*

*Sketch of Proof.* The proof is divided in three steps: construct an approximating sequence $f^{(n)}_\epsilon$ in $C([0,T]; C^2(\mathbb{R}^d\times\mathbb{R}^d))$ by iterations, prove a uniform bound of $f^{(n)}_\epsilon$ in $C([0,T]; C^1(\mathbb{R}^d\times\mathbb{R}^d))$, and show that $f^{(n)}_\epsilon$ is a Cauchy sequence in $C([0,T]; C^1(\mathbb{R}^d\times\mathbb{R}^d))$ which converges to the desired solution of (1.5). □

We end this section by pointing out that, similar to the analysis in [12] for measure-valued solutions, the classical results invoked to prove Theorem 3.2 also use the characteristic equations (3.7). For smooth solutions however the mass transportation formula (3.3) is equivalent to solving the equation by the method of characteristics.

## 4 Uniform in $\epsilon$ estimates

We present in this section all the (uniform in $\epsilon$) estimates needed to prove the convergence as $\epsilon \to 0$ of solutions to (1.5). Throughout the rest of the paper we assume $\nabla K \in W^{1,\infty}(\mathbb{R}^d)$ and that the initial density $f_0$ has compact support. Note that compared to [12] we require a slightly stronger condition on $K$, as for our analysis we need a global Lipschitz bound to obtain the uniform bound in $\epsilon$ for the support of the solution (Proposition 4.1).

For further reference, let us write the initial-value problem for (1.5), with explicit $\epsilon$-dependence indicated for its solution $f_\epsilon$:

$$\begin{aligned} \partial_t f_\epsilon + v \cdot \nabla_x f_\epsilon &= \frac{1}{\epsilon} \nabla_v \cdot \left( \left( v + \nabla_x K * \rho_\epsilon \right) f_\epsilon \right), \\ f_\epsilon \big|_{t=0} &= f_0(x, v), \end{aligned} \tag{4.1}$$

where

$$\rho_\epsilon(t, x) = \int_{\mathbb{R}^d} f_\epsilon(t, x, v) dv.$$

**Solutions of compact support.** We first make the observation that $f_0$ being compactly supported implies that solutions $f_\epsilon$ (either smooth or measure-valued) remain compactly supported for all times (see Theorem 3.1). We show below that the support of $f_\epsilon$ is in fact *independent* of $\epsilon$.

**Proposition 4.1** (Uniform estimate for the support). *Suppose $\nabla_x K \in L^\infty(\mathbb{R}^d)$. Consider a solution $f_\epsilon$ to (4.1) as provided by Theorem 3.1. Then, there exists an increasing function $R(T)$ (independent of $\epsilon$) such that for all $T > 0$,*

$$\operatorname{supp} f_\epsilon(t) \subset B_{R(T)} \qquad \text{for all } t \in [0, T] \text{ and } \epsilon > 0. \tag{4.2}$$

*The function $R(T)$ depends only on the support of $f_0$ and $\|\nabla K\|_{L^\infty}$.*

*Proof.* The support of $f_\epsilon$ evolves with the flow governed by the characteristic equations (3.7), initialized at points $(x_0, v_0)$ in the support of $f_0$. Since

$$|\nabla K * \rho_\epsilon| \leq \|\nabla K\|_{L^\infty},$$

from (3.7) we infer that the Euclidean norm $|v_\epsilon(t)|$ of the $v$-trajectories satisfies

$$\frac{d|v_\epsilon|}{dt} \leq -\frac{1}{\epsilon}|v_\epsilon| + \frac{1}{\epsilon}\|\nabla K\|_{L^\infty}, \qquad v_\epsilon(0) = v_0.$$

Consequently, there exists a constant $C$ that depends only on the support of $f_0$ and $\|\nabla K\|_{L^\infty}$, such that all characteristic paths that start from within supp $f_0$ satisfy

$$|v_\epsilon(t)| \leq C, \qquad \text{for all } t > 0 \text{ and } \epsilon > 0.$$

Since $\frac{d|x_\epsilon|}{dt} \leq |v_\epsilon|$, the $x$-trajectories grow at most linearly in time. Hence there exists a function $R(T)$ that depends only on the support of $f_0$, $\|\nabla K\|_{L^\infty}$ and $T$ such that (4.2) holds. □

### 4.1 Estimates for smooth solutions

Consider the smooth case and take solutions $f_\epsilon \in C([0,T];C^1(\mathbb{R}^d \times \mathbb{R}^d))$, as provided by Theorem 3.2. The key estimate needed for the convergence is provided by the following result. A similar estimate is shown in [25].

**Proposition 4.2** (Main estimate for smooth solutions)**.** *Suppose $\nabla_x K \in W^{1,\infty}(\mathbb{R}^d)$. Let $f_\epsilon$ be the classical solution to (4.1), as provided by Theorem 3.2. Assume additionally that the initial data $f_0$ has a finite first moment in $v$, that is, $|v|f_0 \in L^1(\mathbb{R}^d \times \mathbb{R}^d)$. Then there exist constants $C_0$, $\epsilon_0$ such that for any $\epsilon \le \epsilon_0$,*

$$\iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} |v + \nabla_x K * \rho_\epsilon| \, f_\epsilon \,\mathrm{d}x\,\mathrm{d}v \le C_0\epsilon, \qquad \text{for all } t \in [0,T], \tag{4.3}$$

*where $C_0$ and $\epsilon_0$ only depend on $\|\nabla_x K\|_{W^{1,\infty}}$ and $\|(1+|v|)f_0\|_{L^1(\mathbb{R}^d\times\mathbb{R}^d)}$.*

*Proof.* Denote the quantity on the left-hand side of (4.3) as

$$I(t) = \iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} |v + \nabla_x K * \rho_\epsilon| \, f_\epsilon \,\mathrm{d}x\,\mathrm{d}v.$$

Hence our goal is to show that there exists $C_0$ such that

$$\sup_{t\in[0,T]} I(t) \le C_0\epsilon, \tag{4.4}$$

for small enough $\epsilon$.

Multiply equation (4.1) by $|v + \nabla_x K * \rho_\epsilon|$ and integrate in $x, v$ to get

$$\frac{\mathrm{d}}{\mathrm{d}t} I(t) = -\frac{1}{\epsilon} I(t) + \underbrace{\iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} (\partial_t |v + \nabla_x K * \rho_\epsilon|) \, f_\epsilon \,\mathrm{d}x\,\mathrm{d}v}_{=I_1(t)} + \underbrace{\iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} (v\cdot\nabla_x |v + \nabla_x K * \rho_\epsilon|) \, f_\epsilon \,\mathrm{d}x\,\mathrm{d}v}_{=I_2(t)}. \tag{4.5}$$

Denote the two remainder terms in the right-side as

$$\begin{aligned} I_1(t) &= \iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} (\partial_t |v + \nabla_x K * \rho_\epsilon|) \, f_\epsilon \,\mathrm{d}x\,\mathrm{d}v, \\ I_2(t) &= \iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} (v\cdot\nabla_x |v + \nabla_x K * \rho_\epsilon|) \, f_\epsilon \,\mathrm{d}x\,\mathrm{d}v. \end{aligned}$$

The strategy is to show that $I_1(t)$ and $I_2(t)$ are bounded linearly by $I(t)$ and then derive a differential inequality from (4.5) to bound $I(t)$.

By integrating (4.1) in $v$ one finds that the macroscopic density $\rho_\epsilon \in C([0,T];C^1(\mathbb{R}^d))$ satisfies

$$\begin{aligned} \partial_t\rho_\epsilon + \nabla_x\cdot\langle v f_\epsilon\rangle &= 0, \\ \rho_\epsilon\big|_{t=0} &= \langle f_0\rangle. \end{aligned} \tag{4.6}$$

Here, angle brackets denote integration with respect to $v$. Equation (4.6) conserves mass:

$$\|\rho_\epsilon(t)\|_{L^1(\mathbb{R}^d)} = \|f_0\|_{L^1(\mathbb{R}^d\times\mathbb{R}^d)}, \qquad \text{for all } t > 0.$$

Hence,

$$
\begin{aligned}
\left|\partial_t \left|v + \nabla_x K * \rho_\epsilon\right|\right| &\le |\nabla_x K * \partial_t \rho_\epsilon| = \left|\nabla_x^2 K * \langle v f_\epsilon\rangle\right| \\
&\le \left|\nabla_x^2 K * \langle (v + \nabla_x K * \rho_\epsilon) f_\epsilon\rangle\right| + \left\|\nabla_x^2 K * \rho_\epsilon\right\|_{L^\infty} \|\nabla_x K * \rho_\epsilon\|_{L^\infty} \\
&\le \left\|\nabla_x^2 K\right\|_{L^\infty} I(t) + \left\|\nabla_x^2 K\right\|_{L^\infty} \|\nabla_x K\|_{L^\infty} \|\rho_\epsilon\|_{L^1(\mathbb{R}^d)}^2 .
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
|I_1(t)| &\le \left\|\nabla_x^2 K\right\|_{L^\infty} \|\rho_\epsilon\|_{L^1(\mathbb{R}^d)} I(t) + \left\|\nabla_x^2 K\right\|_{L^\infty} \|\nabla_x K\|_{L^\infty} \|\rho_\epsilon\|_{L^1(\mathbb{R}^d)}^3 \\
&\le C_1 I(t) + C_2,
\end{aligned} \tag{4.7}
$$

where $C_1, C_2$ only depend on $\|\nabla_x K\|_{W^{1,\infty}}$ and $\|f_0\|_{L^1(\mathbb{R}^d\times\mathbb{R}^d)}$.

A similar estimate can be derived for $I_2(t)$. Indeed,

$$
\begin{aligned}
|I_2(t)| &\le \iint_{\mathbb{R}^d\times\mathbb{R}^d} |v| \left|\nabla_x^2 K * \rho_\epsilon\right| f_\epsilon \, \mathrm{d}x \, \mathrm{d}v \\
&\le \left\|\nabla_x^2 K\right\|_{L^\infty} \|\rho_\epsilon\|_{L^1(\mathbb{R}^d)} \iint_{\mathbb{R}^d\times\mathbb{R}^d} |v| f_\epsilon \, \mathrm{d}x \, \mathrm{d}v \\
&\le C_3 \left( \iint_{\mathbb{R}^d\times\mathbb{R}^d} |v + \nabla_x K * \rho_\epsilon| f_\epsilon \, \mathrm{d}x \, \mathrm{d}v + \|\nabla_x K\|_{L^\infty} \|\rho_\epsilon\|_{L^1(\mathbb{R}^d)}^2 \right) \\
&\le C_3 I(t) + C_4,
\end{aligned} \tag{4.8}
$$

where $C_3, C_4$ only depend on $\|\nabla_x K\|_{W^{1,\infty}}$ and $\|f_0\|_{L^1(\mathbb{R}^d\times\mathbb{R}^d)}$.

Combining (4.5), (4.7) and (4.8) we obtain the following differential inequality for $I(t)$:

$$
\frac{\mathrm{d}}{\mathrm{d}t} I(t) \le -\frac{1}{\epsilon} I(t) + C_5 I(t) + C_6, \tag{4.9}
$$

where $C_5 = C_1 + C_3$ and $C_6 = C_2 + C_4$, both of which depending only on $\|\nabla_x K\|_{W^{1,\infty}}$ and $\|f_0\|_{L^1(\mathbb{R}^d\times\mathbb{R}^d)}$.

Note that initially,

$$
I(0) \le \||v| f_0\|_{L^1(\mathbb{R}^d\times\mathbb{R}^d)} + C_7, \tag{4.10}
$$

where $C_7$ only depends on $\|\nabla_x K\|_{L^\infty}$ and $\|f_0\|_{L^1(\mathbb{R}^d\times\mathbb{R}^d)}$.

Finally, for small enough $\epsilon$, one can derive (4.4) from (4.9) and (4.10), where the constant $C_0$ depends only on $\|\nabla_x K\|_{W^{1,\infty}}$ and $\|(1+|v|) f_0\|_{L^1(\mathbb{R}^d\times\mathbb{R}^d)}$. □

## 4.2 Estimates for measure-valued solutions

Next, we show that a similar bound as in (4.3) holds for a measure-valued solution $f_\epsilon$ as well. The strategy we employ here is to use a smooth approximating sequence for which the results in Section 4.1 are valid, and then pass to the limit to infer results for measure solutions.

Take an initial measure $f_0 \in \mathcal{P}_1(\mathbb{R}^d \times \mathbb{R}^d)$ with compact support and fix $\epsilon > 0$. Let $f_0^{(n)}$ be a sequence of mollifications of $f_0$ such that

$$
f_0^{(n)} = f_0 * \eta^{(n)} \in C^2(\mathbb{R}^d \times \mathbb{R}^d). \tag{4.11}
$$

Here the mollifier can be chosen such that $\eta^{(n)}(x,v) = n^{2d}\eta^{(1)}(nx,nv) \in C_c^\infty(\mathbb{R}^d\times\mathbb{R}^d)$, where $\eta^{(1)}$ is compactly supported over the unit ball in $\mathbb{R}^{2d}$ and satisfies

$$\eta^{(1)} \geq 0, \qquad \iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} \eta^{(1)}(x,v)\,\mathrm{d}x\,\mathrm{d}v = 1, \qquad \iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} |v|\eta^{(1)}(x,v)\,\mathrm{d}x\,\mathrm{d}v \leq 1.$$

The following mollification lemma is classical (see for example [1]).

**Lemma 4.3.** *Suppose $f_0 \in \mathcal{P}_1(\mathbb{R}^d\times\mathbb{R}^d)$ with supp $f_0 \subseteq B(R_0)$. Then the approximating sequence $f_0^{(n)}$ satisfies*

- *(a) supp $f_0^{(n)} \subseteq B(R_0+1)$ for all $n \geq 1$.*
- *(b) $f_0^{(n)} \in \mathcal{P}_1(\mathbb{R}^d\times\mathbb{R}^d)$ and the first moments $\iint |v| f_0^{(n)}(x,v)dxdv$ are uniformly bounded.*
- *(c) $\{f_0^{(n)}\}_{n\geq1}$ is a Cauchy sequence in $\mathcal{P}_1(\mathbb{R}^d\times\mathbb{R}^d)$ endowed with the $W_1$ distance, and $\|f_0^{(n)} - f_0\|_{W_1} \to 0$ as $n\to\infty$.*

Now we construct the approximating sequence $f_\epsilon^{(n)}$ such that

$$\begin{aligned} \partial_t f_\epsilon^{(n)} + v\cdot\nabla_x f_\epsilon^{(n)} &= \frac{1}{\epsilon}\nabla_v\cdot\left(\left(v+\nabla_x K * \rho_\epsilon^{(n)}\right) f_\epsilon^{(n)}\right), \\ f_\epsilon^{(n)}\big|_{t=0} &= f_0^{(n)}(x,v), \end{aligned} \tag{4.12}$$

where $\rho_\epsilon^{(n)} = \int_{\mathbb{R}^d} f_\epsilon^{(n)}\,\mathrm{d}v$.

**Lemma 4.4.** *Fix $\epsilon > 0$. Suppose $\nabla_x K \in W^{1,\infty}(\mathbb{R}^d)$ and $f_0 \in \mathcal{P}_1(\mathbb{R}^d\times\mathbb{R}^d)$ with compact support. Let $f_0^{(n)}$ be the sequence of mollifications of $f_0$ given by (4.11). Then for each $T>0$, there exists a sequence of solutions $f_\epsilon^{(n)} \in C([0,T];C^1(\mathbb{R}^d\times\mathbb{R}^d))$ to (4.12) whose supports only depend on $T$ and $\|\nabla_x K\|_{W^{1,\infty}}$. In particular, the supports are uniformly bounded in both $n$ and $\epsilon$.*

*Moreover, if we let $f_\epsilon \in C([0,T),\mathcal{P}_1(\mathbb{R}^d\times\mathbb{R}^d))$ be the unique measure solution to (4.1) in the sense of Definition 3.1, then*

$$f_\epsilon^{(n)}(t,\cdot,\cdot) \xrightarrow{W_1} f_\epsilon(t,\cdot,\cdot) \quad \text{in } \mathcal{P}_1(\mathbb{R}^d\times\mathbb{R}^d), \quad \text{uniformly in } t \text{ as } n\to\infty. \tag{4.13}$$

*Proof.* Since $f_0^{(n)} \in C^2(\mathbb{R}^d\times\mathbb{R}^d)$ has compact support, we can apply the existence theory in Theorem 3.2 and deduce that there exists a smooth solution $f_\epsilon^{(n)} \in C([0,T];C^1(\mathbb{R}^d\times\mathbb{R}^d))$ of (4.12) for every $n\geq 1$ and every $\epsilon>0$. Each $f_\epsilon^{(n)}$ is compactly supported and integrates to 1. Proposition 4.1 yields that the support of $f_\epsilon^{(n)}$ is independent of $\epsilon$ and depends only on $T$, $\|\nabla_x K\|_{W^{1,\infty}}$, and the support of $f_0^{(n)}$. By part (a) in Lemma 4.3, we further conclude that the supports of $f_\epsilon^{(n)}(t,\cdot,\cdot)$ are uniformly bounded in *both* $n$ and $\epsilon$ for all $t\in[0,T]$.

Let $f_\epsilon \in C([0,T];\mathcal{P}_1(\mathbb{R}^d\times\mathbb{R}^d))$ be the unique measure-valued solution to (4.1) in the sense of Definition 3.1. By the stability result in [12, Theorem 3.16], we have that for all times $t\geq 0$,

$$\left\|f_\epsilon^{(n)}(t,\cdot,\cdot) - f_\epsilon(t,\cdot,\cdot)\right\|_{W_1} \leq r(T)\left\|f_0^{(n)} - f_0\right\|_{W_1}, \tag{4.14}$$

where $r(T)$ only depends on $T$ and the support of $f_0$. From (4.14) and Lemma 4.3 we then conclude (4.13). □

It follows from (4.13) that

$$\rho_\epsilon^{(n)}(t,\cdot,\cdot) \longrightarrow \rho_\epsilon(t,\cdot,\cdot) \qquad \text{as measures,} \qquad \text{for each } t \in [0,T) \text{ as } n \to \infty, \tag{4.15}$$

where $\rho_\epsilon$ is the first marginal of $f_\epsilon$. Next we show the following $L^\infty$-convergence of $\nabla_x K * \rho_\epsilon^{(n)}$.

**Lemma 4.5.** *Suppose $\nabla_x K \in W^{1,\infty}(\mathbb{R}^d)$. Consider the measure solution $f_\epsilon$ of (4.1) obtained as a limit of mollifications $f_\epsilon^{(n)}$ as in Lemma 4.4. Then for each $t \geq 0$, we have $\nabla_x K * \rho_\epsilon \in C(\mathbb{R}^d)$ and*

$$\nabla_x K * \rho_\epsilon^{(n)}(t,\cdot) \longrightarrow \nabla_x K * \rho_\epsilon(t,\cdot) \qquad \textit{strongly in } L^\infty_{loc}(\mathbb{R}^d), \textit{ as } n \to \infty.$$

*Proof.* Given the regularity of $f_\epsilon^{(n)}$ we have that $\nabla_x K * \rho_\epsilon^{(n)}(t,\cdot) \in C(\mathbb{R}^d)$ for each time $t \geq 0$. Also, by mass conservation of $f_\epsilon^{(n)}$ and the properties of the mollifiers in Lemma 4.3, $\nabla_x K * \rho_\epsilon^{(n)}$ satisfies

$$|\nabla_x K * \rho_\epsilon^{(n)}(x)| \leq \|\nabla_x K\|_{L^\infty},$$
$$|\nabla_x K * \rho_\epsilon^{(n)}(x_1) - \nabla_x K * \rho_\epsilon^{(n)}(x_2)| \leq \left\|\nabla_x^2 K\right\|_{L^\infty} |x_1 - x_2|,$$

for any $x, x_1, x_2 \in \mathbb{R}^d$. Hence the sequence $\{\nabla_x K * \rho_\epsilon^{(n)}\}_{n\geq 1}$ is uniformly bounded and equicontinuous. By Ascoli-Arzelá theorem, we have that $\{\nabla_x K * \rho_\epsilon^{(n)}\}_{n\geq 1}$ converges on a subsequence in the strong topology of $C(\Omega)$, for any compact set $\Omega \subset \mathbb{R}^d$. Meanwhile, by (4.15), the limit function is $\nabla_x K * \rho_\epsilon$ (to see that $\nabla_x K * \rho_\epsilon \in C(\mathbb{R}^d)$ one can use the second of the two inequalities displayed above, with $\rho_\epsilon$ instead of $\rho_\epsilon^{(n)}$, and infer that $\nabla K * \rho_\epsilon$ is in fact Lipschitz continuous). It then follows that the entire sequence $\{\nabla_x K * \rho_\epsilon^{(n)}\}_{n\geq 1}$ converges to $\nabla_x K * \rho_\epsilon$, as desired. □

Now we establish the analogue for measure solutions of the main estimate (4.3). The approach is to fix $\epsilon > 0$ and use Proposition 4.2 for the smooth approximating sequence $f_\epsilon^{(n)}$. Then, with $n \geq 1$ fixed, there exist constants $\epsilon_0$, $C_0$ which depend on $\|\nabla_x K\|_{W^{1,\infty}}$ and $\iint (1+|v|) f_0^{(n)}\, dxdv$, such that for $\epsilon < \epsilon_0$,

$$\left| \iint_{\mathbb{R}^d\times\mathbb{R}^d} \left(v + \nabla_x K * \rho_\epsilon^{(n)}\right) f_\epsilon^{(n)}\, \mathrm{d}x\, \mathrm{d}v \right| \leq C_0 \epsilon, \qquad \text{for all } t \in [0,T].$$

However, by part (b) in Lemma 4.3, the constants $C_0$ and $\epsilon_0$ can be chosen to be independent of $n$, and hence

$$\left| \iint_{\mathbb{R}^d\times\mathbb{R}^d} \left(v + \nabla_x K * \rho_\epsilon^{(n)}\right) f_\epsilon^{(n)}\, \mathrm{d}x\, \mathrm{d}v \right| \leq C_0 \epsilon, \qquad \text{for all } n \geq 1, \text{ and } t \in [0,T], \tag{4.16}$$

where $\epsilon < \epsilon_0$, and $C_0$, $\epsilon_0$ depend only on $\|\nabla_x K\|_{W^{1,\infty}}$ and $\iint |v| f_0\, dxdv$.

**Proposition 4.6** (Main estimate for measure-valued solutions). *Fix $\epsilon > 0$ small, such that (4.16) is satisfied, and assume the hypotheses in Lemma 4.4. Then for any $\tilde{\phi} \in C_b(\mathbb{R}^d \times \mathbb{R}^d)$, there exists a constant $\tilde{C}_0$ such that,*

$$\left| \iint_{\mathbb{R}^d\times\mathbb{R}^d} \tilde{\phi}(x,v)\,(v + \nabla_x K * \rho_\epsilon)\, f_\epsilon\, \mathrm{d}x\, \mathrm{d}v \right| \leq \tilde{C}_0 \epsilon, \qquad \textit{for all } t \in [0,T]. \tag{4.17}$$

*Specifically, $\tilde{C}_0 = C_0 \|\tilde{\phi}\|_{L^\infty(\mathbb{R}^d\times\mathbb{R}^d)}$, where $C_0$ is a constant which depends only on $\|\nabla_x K\|_{W^{1,\infty}}$ and $\iint |v| f_0 \, dx dv$. In particular, $\tilde{C}_0$ is independent of $\epsilon$ and $t$.*

*Proof.* By (4.16) we have

$$\Big| \iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} \tilde{\phi}(x,v) \left( v + \nabla_x K * \rho_\epsilon^{(n)} \right) f_\epsilon^{(n)} \, \mathrm{d}x \, \mathrm{d}v \Big| \le C_0 \|\tilde{\phi}\|_{L^\infty} \epsilon, \tag{4.18}$$

where $C_0$ is a constant that depends only on $\|\nabla_x K\|_{W^{1,\infty}}$ and $\iint |v| f_0 \, dx dv$.

Denote by $\Omega(T) \subset \mathbb{R}^{2d}$ the common support of $f_\epsilon^{(n)}(t)$ for all $\epsilon > 0$, $n \ge 1$ at any $t \in [0,T]$. Then, by (4.13) (also recall here Remark 3.1) and Lemma 4.5, we have that for each $t \in [0,T]$,

$$\begin{aligned} &\iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} \tilde{\phi}(x,v) \left( v + \nabla_x K * \rho_\epsilon^{(n)} \right) f_\epsilon^{(n)} \, \mathrm{d}x \, \mathrm{d}v \\ &= \iint\limits_{\Omega(T)} \tilde{\phi}(x,v) \left( v + \nabla_x K * \rho_\epsilon^{(n)} \right) f_\epsilon^{(n)} \, \mathrm{d}x \, \mathrm{d}v \to \iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} \tilde{\phi}(x,v) \left( v + \nabla_x K * \rho_\epsilon \right) f_\epsilon \, \mathrm{d}x \, \mathrm{d}v, \end{aligned}$$

as $n \to \infty$. Hence by taking the limit $n \to \infty$ in (4.18), we obtain the desired bound in (4.17). □

## 5 Convergence as $\epsilon \to 0$ of solutions to (1.5)

By Proposition 4.6, measure-valued solutions $f_\epsilon$ of the transport equation (4.1), in the sense of Definition 3.1, satisfy the uniform (in $\epsilon$) estimate (4.17). In this section we use this key estimate to pass the limit $\epsilon \to 0$ in (4.1), that is, in the initial-value problem for (1.5).

First we explain the setting for well-posedness of the macroscopic equation (1.1). Consider the initial value problem for (1.1):

$$\begin{aligned} \rho_t - \nabla_x \cdot ((\nabla_x K * \rho)\rho) = 0 \,, \\ \rho\big|_{t=0} = \rho_0(x) \,. \end{aligned} \tag{5.1}$$

Similar to the kinetic equation, there exist several concepts of solutions to (5.1) over an arbitrary time interval $[0,T)$.

One concept is the measure-valued solution as defined in [12]. More specifically, assuming that $\nabla_x K \in W^{1,\infty}(\mathbb{R}^d)$, one can apply the framework in [12] and obtain a unique measure-valued solution $\rho \in C([0,T);\mathcal{P}_1(\mathbb{R}^d))$ in the mass transportation sense (similar to how a measure solution for the kinetic equation (1.5) has been introduced in Definition 3.1).

Another notion is the weak solution in $C([0,T);\mathcal{P}(\mathbb{R}^d))$ where the continuity in time is in the narrow sense. In particular, a weak solution $\rho \in C([0,T);\mathcal{P}(\mathbb{R}^d))$ to (5.1) satisfies

$$\int_0^T \int\limits_{\mathbb{R}^d} \partial_t \psi(t,x) \rho(t,x) \, \mathrm{d}x \, \mathrm{d}t - \int_0^T \int\limits_{\mathbb{R}^d} \nabla_x \psi \cdot (\nabla_x K * \rho) \, \rho \, \mathrm{d}x \, \mathrm{d}t + \int\limits_{\mathbb{R}^d} \psi(0,x) \rho_0(x) \, \mathrm{d}x = 0, \tag{5.2}$$

for any $\psi \in C_c^1([0,T);C_b^1(\mathbb{R}^d))$. A global-in-time well-posedness theory of weak measure solutions to (5.1) was established in [13] for a very general class of (nonsmooth) potentials. In their setting, as well as ours, the two concepts of solutions are in fact equivalent (see Step 3 in the proof of Theorem 5.1 for a more detailed account of this fact).

## 5.1 Convergence to the macroscopic equation

The following theorem is one of the main results of this paper.

**Theorem 5.1.** *Let $T > 0$ be arbitrary, $\nabla_x K \in W^{1,\infty}(\mathbb{R}^d)$ and $f_0 \in \mathcal{P}_1(\mathbb{R}^d \times \mathbb{R}^d)$ with compact support. Suppose $f_\epsilon \in C([0,T); \mathcal{P}_1(\mathbb{R} \times \mathbb{R}^d))$ is the measure-valued solution to (4.1) obtained in Theorem 3.1. Let $\rho_\epsilon$ be the first marginal of $f_\epsilon$ as defined in (3.5).*

*Then, there exists $\rho \in C([0,T); \mathcal{P}_1(\mathbb{R}^d))$ such that for each $t \in [0,T)$,*

$$\rho_\epsilon(t,\cdot) \xrightarrow{W_1} \rho(t,\cdot) \qquad \text{in } \mathcal{P}_1(\mathbb{R}^d) \text{ as } \epsilon \to 0. \tag{5.3}$$

*Moreover, $\rho$ is the unique solution to the initial value problem (5.1) in the weak sense, i.e., it satisfies (5.2), where $\rho_0$ is the first marginal of $f_0$.*

*Proof.* We break the proof into several steps.

*Step 1.* First we show a uniform in time convergence of $\rho_\epsilon$. Recall that the measure solution $f_\epsilon$ of (4.1) is also a weak solution, cf. (3.6). For any fixed $\psi_1 \in C_c^1(0,T)$ and $\psi_2 \in C_b^1(\mathbb{R}^d)$, let $\phi(t,x,v) = \psi_1(t)\psi_2(x)$ and use it as a test function in (3.6), to get:

$$\int_0^T \psi_1'(t) \int_{\mathbb{R}^d} \psi_2(x)\rho_\epsilon(t,x)\,\mathrm{d}x\,\mathrm{d}t = -\int_0^T \psi_1(t) \iint_{\mathbb{R}^d\times\mathbb{R}^d} \nabla_x\psi_2 \cdot v f_\epsilon \,\mathrm{d}x\,\mathrm{d}v\,\mathrm{d}t. \tag{5.4}$$

Denote

$$\xi_\epsilon(t) = \int_{\mathbb{R}^d} \psi_2(x)\rho_\epsilon(t,x)\,\mathrm{d}x. \tag{5.5}$$

By (5.4), the weak derivative of $\xi_\epsilon$ is given by

$$\xi_\epsilon'(t) = \iint_{\mathbb{R}^d\times\mathbb{R}^d} \nabla_x\psi_2 \cdot v f_\epsilon \,\mathrm{d}x\,\mathrm{d}v \in L^\infty(0,T).$$

Since $f_\epsilon$ is uniformly supported on $\Omega(T)$, we have

$$\|\xi_\epsilon\|_{W^{1,\infty}(0,T)} \le C(T)\,\|\psi_2\|_{C_b^1(\mathbb{R}^d)}, \tag{5.6}$$

where $C(T)$ only depends on $T$ (in particular, $C(T)$ is independent of $\epsilon$).

Since $\xi_\epsilon(t)$ is uniformly bounded in $W^{1,\infty}(0,T)$, we infer from Ascoli-Arzelá theorem that given any $\psi_2 \in C_b^1(\mathbb{R}^d)$, there exists a subsequence $\epsilon_k$ and $\xi(t) \in C([0,T))$ such that

$$\int_{\mathbb{R}^d} \psi_2(x)\rho_{\epsilon_k}(t,x)\,\mathrm{d}x \to \xi(t) \quad \text{uniformly on } [0,T) \qquad \text{as } \epsilon_k \to 0. \tag{5.7}$$

On the other hand, note that Proposition 4.1 provides a uniform (in $\epsilon$) bound for the support of $\rho_\epsilon$, which implies that the sequence $\rho_\epsilon(\cdot,t)$ is tight. By Prokhorov's theorem (cf. [7, Theorem 4.1]), for each $t \in [0,T)$, $\rho_\epsilon(t,\cdot)$ converges weak-* on a subsequence to a probability measure $\rho(\cdot,t) \in \mathcal{P}(\mathbb{R}^d)$. By Remark 3.1, the convergence holds in fact in $\mathcal{P}_1(\mathbb{R}^d)$, with respect to the Wasserstein metric $W_1$.

Hence, at each $t > 0$, there exist a subsequence of $\rho_{\epsilon_k}$, denoted as $\rho_{\epsilon_{k_n}}$ ($k_n$ may depend on $t$), which satisfies

$$\rho_{\epsilon_{k_n}}(t,\cdot) \xrightarrow{W_1} \rho(t,\cdot) \qquad \text{in } \mathcal{P}_1(\mathbb{R}^d) \quad \text{as } \epsilon_{k_n} \to 0.$$

Consequently, for each $t \in [0,T)$ and any $\psi_2 \in C_b^1(\mathbb{R}^d)$,

$$\int_{\mathbb{R}^d} \psi_2(x)\rho_{\epsilon_{k_n}}(t,x)\,\mathrm{d}x \to \int_{\mathbb{R}^d} \psi_2(x)\rho(t,x)\,\mathrm{d}x \qquad \text{as } \epsilon_{k_n} \to 0. \tag{5.8}$$

The convergence in (5.7) provides a unique limit $\xi(t)$ (reached along the subsequence $\epsilon_k$) at each $t \in [0,T)$. Combining this fact with (5.8) we infer that the sequence $\rho_{\epsilon_k}(t,\cdot)$ (with $\epsilon_k$ independent of $t$) and $\rho(t,\cdot) \in \mathcal{P}_1(\mathbb{R}^d)$ satisfy

$$\int_{\mathbb{R}^d} \psi_2(x)\rho_{\epsilon_k}(t,x)\,\mathrm{d}x \to \int_{\mathbb{R}^d} \psi_2(x)\rho(t,x)\,\mathrm{d}x \qquad \text{uniformly on } [0,T) \quad \text{as } \epsilon_k \to 0, \tag{5.9}$$

for any $\psi_2 \in C_b^1(\mathbb{R}^d)$. In addition, we have that

$$\rho_{\epsilon_k}(t,\cdot) \xrightarrow{W_1} \rho(t,\cdot) \qquad \text{in } \mathcal{P}_1(\mathbb{R}^d) \quad \text{as } \epsilon_k \to 0. \tag{5.10}$$

Next we show that we can also allow $\psi_2$ to depend on $t$ in (5.9). Specifically, we claim that given any $\psi_3 \in C_c([0,T); C_b^1(\mathbb{R}^d))$,

$$\int_{\mathbb{R}^d} \psi_3(t,x)\rho_{\epsilon_k}(t,x)\,\mathrm{d}x \quad \text{is equicontinuous on } [0,T)\ . \tag{5.11}$$

Indeed, for any $t,s \in [0,T)$,

$$\begin{aligned}
&\left| \int_{\mathbb{R}^d} \psi_3(t,x)\rho_{\epsilon_k}(t,x)\,\mathrm{d}x - \int_{\mathbb{R}^d} \psi_3(s,x)\rho_{\epsilon_k}(s,x)\,\mathrm{d}x \right| \\
&\le \int_{\mathbb{R}^d} |\psi_3(t,x) - \psi_3(s,x)|\,\rho_{\epsilon_k}(t,x)\,\mathrm{d}x + \left| \int_{\mathbb{R}^d} \psi_3(s,x)\rho_{\epsilon_k}(t,x)\,\mathrm{d}x - \int_{\mathbb{R}^d} \psi_3(s,x)\rho_{\epsilon_k}(s,x)\,\mathrm{d}x \right|. \\
&\le \sup_x |\psi_3(t,x) - \psi_3(s,x)| + C(T) \sup_t \|\psi_3\|_{C_b^1(\mathbb{R}^d)} |t-s|,
\end{aligned}$$

where the last inequality follows from (5.6). Since $\psi_3$ is uniformly continuous on $[0,T) \times \mathbb{R}^d$, we have that

$$\sup_x |\psi_3(t,x) - \psi_3(s,x)| \to 0, \qquad \text{uniformly as } |t-s| \to 0.$$

This shows that (5.11) holds. Hence up to a subsequence, still denoted as $\rho_{\epsilon_k}$, we have

$$\int_{\mathbb{R}^d} \psi_3(t,x)\rho_{\epsilon_k}(t,x)\,\mathrm{d}x \to \int_{\mathbb{R}^d} \psi_3(t,x)\rho(t,x)\,\mathrm{d}x \qquad \text{uniformly on } [0,T) \quad \text{as } \epsilon_k \to 0, \tag{5.12}$$

for any $\psi_3 \in C_c([0,T); C_b^1(\mathbb{R}^d))$.

*Step 2.* In this step we pass the limit $\epsilon_k \to 0$ on the subsequence $\rho_{\epsilon_k}$ to find a limiting equation for $\rho$.

Define

$$\Omega_1(T) = \{x \in \mathbb{R}^d : (x, v) \in \Omega(T)\}, \tag{5.13}$$

where recall that $\Omega(T) \subset \mathbb{R}^{2d}$ represents the common support of $f_\epsilon(t)$ for all $\epsilon > 0$ and $t \in [0, T]$. We have that $\Omega_1(T)$ is bounded and that the supports of $\rho_{\epsilon_k}$ and $\rho$ are included in $\Omega_1(T)$ for all $t \in [0, T]$.

For any $\psi \in C_c^1([0, T); C_b^1(\mathbb{R}^d))$, let $\phi(t, x, v) = \psi(t, x)$ in (3.6). Then

$$\int_0^T \int_{\mathbb{R}^d} \partial_t \psi(t, x) \rho_{\epsilon_k}(t, x) \, dx \, dt + \int_0^T \iint_{\mathbb{R}^d \times \mathbb{R}^d} \nabla_x \psi \cdot v f_{\epsilon_k} \, dx \, dv \, dt + \int_{\mathbb{R}^d} \psi(0, x) \rho_0(x) \, dx = 0. \tag{5.14}$$

We want to pass $\epsilon_k \to 0$ in (5.14). By (5.12),

$$\int_0^T \int_{\mathbb{R}^d} \partial_t \psi(t, x) \rho_{\epsilon_k}(t, x) \, dx \, dt \to \int_0^T \int_{\mathbb{R}^d} \partial_t \psi(t, x) \rho(t, x) \, dx \, dt \qquad \text{as } \epsilon_k \to 0.$$

Next we rewrite the integrand of the second term in (5.14) as

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d} \nabla_x \psi \cdot v f_{\epsilon_k} \, dx \, dv = \iint_{\mathbb{R}^d \times \mathbb{R}^d} \nabla_x \psi \cdot (v + \nabla_x K * \rho_{\epsilon_k}) \, f_{\epsilon_k} \, dx \, dv - \int_{\mathbb{R}^d} \nabla_x \psi \cdot (\nabla_x K * \rho_{\epsilon_k}) \, \rho_{\epsilon_k} \, dx. \tag{5.15}$$

Use (4.17) to get

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d} \nabla_x \psi \cdot (v + \nabla_x K * \rho_{\epsilon_k}) \, f_{\epsilon_k} \, dx \, dv \to 0 \qquad \text{as } \epsilon_k \to 0 \quad \text{uniformly in } t. \tag{5.16}$$

By the same argument as in Lemma 4.5, one can show that $\{\nabla_x K * \rho_{\epsilon_k}(t, \cdot)\}$ is a bounded family in $W^{1,\infty}(\mathbb{R}^d)$ for each $t \in [0, T)$. More precisely,

$$\|\nabla_x K * \rho_{\epsilon_k}(t, \cdot)\|_{W^{1,\infty}(\mathbb{R}^d)} \le \|\nabla_x K\|_{W^{1,\infty}} .$$

Now we want to show that $\nabla_x K * \rho_{\epsilon_k}$ is also equicontinuous in $t$. Note that $\nabla_x K$ does not have enough regularity for the bound in (5.6) to apply directly. To bypass this, we mollify $K$ by convolution and let $K_n = K * \eta^{(n)}$ where $\eta^{(n)}$ is the same mollifier defined in (4.11). Hence $\nabla_x K_n = \nabla_x K * \eta^{(n)}$ and $\nabla_x^2 K_n = \nabla_x^2 K * \eta^{(n)}$. This shows

$$\|\nabla_x K_n\|_{C_b^1} \le \|\nabla_x K\|_{W^{1,\infty}} \qquad \text{for all } n \ge 1.$$

Use $\psi_2 = \nabla_x K_n$ in (5.5). Then, bound (5.6) yields

$$\sup_x \|\nabla_x K_n * \rho_{\epsilon_k}\|_{W^{1,\infty}(0,T)} \le C(T) \, \|\nabla_x K_n\|_{C_b^1} \le C(T) \, \|\nabla_x K\|_{W^{1,\infty}} .$$

Together with

$$\|\nabla_x K_n * \rho_{\epsilon_k}(t, \cdot)\|_{W^{1,\infty}(\mathbb{R}^d)} \le \|\nabla_x K_n\|_{W^{1,\infty}} \le \|\nabla_x K\|_{W^{1,\infty}} ,$$

we have

$$\|\nabla_x K_n * \rho_{\epsilon_k}\|_{W^{1,\infty}([0,T)\times\mathbb{R}^d)} \le (C(T)+1)\,\|\nabla_x K\|_{W^{1,\infty}}\,.$$

Therefore, for any $t,s\in[0,T)$ and $x,y\in\mathbb{R}^d$,

$$|\nabla_x K_n * \rho_{\epsilon_k}(t,x) - \nabla_x K_n * \rho_{\epsilon_k}(s,y)| \le (C(T)+1)\,\|\nabla_x K\|_{W^{1,\infty}}\,(|t-s|+|x-y|)\,. \tag{5.17}$$

Since $\nabla_x K$ is continuous, we have that $\nabla_x K_n \to \nabla_x K$ uniformly on any compact set in $\mathbb{R}^d$. By

$$|\nabla_x K_n * \rho_{\epsilon_k}(t,x) - \nabla_x K * \rho_{\epsilon_k}(t,x)| \le \sup_x |\nabla_x K_n(x) - \nabla_x K(x)|\,,$$

we deduce that for any compact set $\widetilde{\Omega}\subset\mathbb{R}^d$

$$\nabla_x K_n * \rho_{\epsilon_k}(t,x) \overset{n\to\infty}{\longrightarrow} \nabla_x K * \rho_{\epsilon_k}(t,x)\,, \qquad \text{uniformly for } t\in[0,T), x\in\widetilde{\Omega}, \text{ and } k\in\mathbb{N}.$$

Hence, if we pass $n\to\infty$ in (5.17) over any compact set $\tilde{\Omega}\subset\mathbb{R}^d$, then

$$|\nabla_x K * \rho_{\epsilon_k}(t,x) - \nabla_x K * \rho_{\epsilon_k}(s,y)| \le (C(T)+1)\,\|\nabla_x K\|_{W^{1,\infty}}\,(|t-s|+|x-y|)\,, \tag{5.18}$$

for any $t,s\in[0,T)$ and $x,y\in\widetilde{\Omega}$.

By Ascoli-Arzelá theorem, there exists a further subsequence (also denoted as $\rho_{\epsilon_k}$) such that

$$\nabla_x K * \rho_{\epsilon_k} \to \nabla_x K * \rho \qquad \text{as } \epsilon_k\to 0 \qquad \text{strongly in } L^\infty([0,T)\times\widetilde{\Omega}), \tag{5.19}$$

for any compact set $\widetilde{\Omega}\subset\mathbb{R}^d$.

Now for every $t\in(0,T)$, we have

$$\begin{aligned}
&\left|\int_{\mathbb{R}^d} \nabla_x\psi\cdot\big((\nabla_x K*\rho_{\epsilon_k})\,\rho_{\epsilon_k} - (\nabla_x K*\rho)\,\rho\big)\,\mathrm{d}x\right| \\
&\le \int_{\Omega_1(T)} |\nabla_x\psi|\,|\nabla_x K*\rho_{\epsilon_k} - \nabla_x K*\rho|\,\rho_{\epsilon_k}\,\mathrm{d}x + \left|\int_{\mathbb{R}^d}\nabla_x\psi\cdot(\nabla_x K*\rho)\,(\rho_{\epsilon_k}-\rho)\,\mathrm{d}x\right| \\
&\le \|\nabla_x K*\rho_{\epsilon_k} - \nabla_x K*\rho\|_{L^\infty(\Omega_1(T))}\,\|\nabla_x\psi\|_{L^\infty} + \left|\int_{\Omega_1(T)}\nabla_x\psi\cdot(\nabla_x K*\rho)\,(\rho_{\epsilon_k}-\rho)\,\mathrm{d}x\right|,
\end{aligned}$$

where the first term converges to zero uniformly (in time) as $\epsilon_k\to 0$ by (5.19) and the second term converges to zero pointwise in $t$ by (5.10). This combined with (5.16) and (5.15) yields that for each $t\in(0,T)$,

$$\iint_{\mathbb{R}^d\times\mathbb{R}^d} \nabla_x\psi\cdot v f_{\epsilon_k}\,\mathrm{d}x\,\mathrm{d}v \to -\int_{\mathbb{R}^d}\nabla_x\psi\cdot(\nabla_x K*\rho)\,\rho\,\mathrm{d}x \qquad \text{as } \epsilon_k\to 0\,.$$

To conclude, let

$$\Omega_2(T) = \{v\in\mathbb{R}^d : (x,v)\in\Omega(T)\}\,.$$

Then $\Omega_2(T)$ is bounded for all $t \in (0,T)$ and we have the uniform (in $t$) bound

$$\left| \iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} \nabla_x\psi \cdot v f_{\epsilon_k}\,\mathrm{d}x\,\mathrm{d}v \right| \le C_2 \left\|\nabla_x\psi\right\|_{L^\infty(\mathbb{R}^d)},$$

where $C_2$ only depends on $\Omega_2(T)$. By the Lebesgue's dominated convergence theorem, we infer that

$$\int_0^T \iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} \nabla_x\psi \cdot v f_{\epsilon_k}\,\mathrm{d}x\,\mathrm{d}v\,\mathrm{d}t \to -\int_0^T \iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} \nabla_x\psi \cdot (\nabla_x K * \rho)\,\rho\,\mathrm{d}x\,\mathrm{d}t, \qquad \text{as } \epsilon_k \to 0.$$

Now (5.2) follows from (5.14) in the $\epsilon_k \to 0$ limit.

*Step 3.* Hence the limiting measure $\rho \in C([0,T);\mathcal{P}(\mathbb{R}^d))$ is a weak solution to (5.1). By [29] (Lemma 8.1.6 in Chapter 8), $\rho$ is the push-forward of the initial density $\rho_0$ by the characteristic flow, i.e., $\rho = \mathcal{T}^t_{E_1[\rho]}\#\rho_0$ with the vector field given by $E_1[\rho] = -\nabla_x K * \rho \in L^\infty([0,T)\times\mathbb{R}^d)$. Moreover, since $\rho(t,\cdot)$ is compactly supported and narrowly continuous in time, we have that $\rho(t,\cdot) \in C([0,T);\mathcal{P}_1(\mathbb{R}^d))$, where the continuity is in the $W_1$ metric cf. Remark 3.1.

We conclude that $\rho$ is the *unique* solution of (5.1) in the mass transportation sense [12]. Consequently, we infer that the *full* sequence $\rho_\epsilon(t,\cdot)$ converges to $\rho(t,\cdot)$ with respect to the $W_1$ distance, for each $t \in [0,T)$, as desired. □

## 5.2 Convergence of characteristic paths

Consider the solution $(x_\epsilon(t), v_\epsilon(t))$ of (3.7), that is, the characteristic paths defining the flow on $\mathbb{R}^d \times \mathbb{R}^d$ along which $f_\epsilon$ is being transported. We now investigate their limit as $\epsilon \to 0$.

**Theorem 5.2** (Convergence of characteristic paths)**.** *Suppose $\nabla_x K \in W^{1,\infty}(\mathbb{R}^d)$. Consider the measure-valued solution $f_\epsilon$ to (4.1) and a characteristic path $(x_\epsilon(t), v_\epsilon(t))$ that originates from $(x_0, v_0) \in \operatorname{supp} f_0$ at $t = 0$. Then,*

$$\lim_{\epsilon\to 0} x_\epsilon(t) = x(t), \qquad \textit{for all} \quad 0 \le t \le T, \tag{5.20}$$

*where $x(t)$ is the characteristic trajectory of the limiting macroscopic equation (5.1) that starts at $x_0$, i.e., $x(t)$ satisfies*

$$\frac{dx}{dt} = -\nabla_x K * \rho, \quad x(0) = x_0.$$

*Also,*

$$\lim_{\epsilon\to 0} v_\epsilon(t) = v(t), \qquad \textit{for all} \quad 0 < t \le T, \tag{5.21}$$

*where*

$$v(t) = -\nabla_x K * \rho(t, x(t)). \tag{5.22}$$

*Proof.* The task is to send $\epsilon \to 0$ in the characteristic system (3.7). Note that (3.7) does not fit immediately into the form (2.1) needed for a direct application of Tikhonov's theorem (Theorem 2.1), as the right-hand-side of the $v$-equation depends on $\epsilon$ as well.

Replace $\rho_\epsilon$ by $\rho$ in the right-hand-side of (3.7) to get the following system:

$$\begin{aligned} \frac{\mathrm{d}x}{\mathrm{d}t} &= v\,, \\ \epsilon\frac{\mathrm{d}v}{\mathrm{d}t} &= -v - \nabla_x K * \rho. \end{aligned} \tag{5.23}$$

Denote by $(\tilde{x}_\epsilon(t), \tilde{v}_\epsilon(t))$ the solution of (5.23) that starts from $(x_0, v_0)$. By Theorem 2.1, the convergence in (5.20)-(5.22), which needs to be shown for $x_\epsilon(t)$ and $v_\epsilon(t)$, holds for $\tilde{x}_\epsilon(t)$ and $\tilde{v}_\epsilon(t)$. Hence, it would be enough to show that for a fixed $t > 0$,

$$\lim_{\epsilon\to 0} |x_\epsilon(t) - \tilde{x}_\epsilon(t)| = 0 \quad \text{and} \quad \lim_{\epsilon\to 0} |v_\epsilon(t) - \tilde{v}_\epsilon(t)| = 0. \tag{5.24}$$

Indeed, from (3.7) and (5.23) we get

$$\epsilon\frac{\mathrm{d}}{\mathrm{d}t}(v_\epsilon(t) - \tilde{v}_\epsilon(t)) = -(v_\epsilon(t) - \tilde{v}_\epsilon(t)) - \nabla_x K * \rho_\epsilon(t, x_\epsilon(t)) + \nabla_x K * \rho(t, \tilde{x}_\epsilon(t)).$$

By integrating the above equation we find

$$v_\epsilon(t) - \tilde{v}_\epsilon(t) = \frac{1}{\epsilon}\int_0^t e^{\frac{1}{\epsilon}(s-t)}\left(-\nabla_x K * \rho_\epsilon(s, x_\epsilon(s)) + \nabla_x K * \rho(s, \tilde{x}_\epsilon(s))\right) ds, \tag{5.25}$$

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

We now estimate the term in round brackets in the integrand of the right-hand-side of (5.25). Add and subtract $\nabla_x K * \rho_\epsilon(s, x_\epsilon(s))$ to this term, and bound

$$\begin{aligned} |\nabla_x K * \rho_\epsilon(s, x_\epsilon(s)) - \nabla_x K * \rho(s, \tilde{x}_\epsilon(s))| &\le \sup_{t\in[0,T]} \|\nabla_x K * (\rho_\epsilon - \rho)\|_{L^\infty(\Omega_1(T))} \\ &\quad + \sup_{t\in[0,T]} \|\nabla_x^2 K * \rho\|_{L^\infty(\Omega_1(T))} |x_\epsilon(s) - \tilde{x}_\epsilon(s)| \\ &\le C_\epsilon + C_1 |x_\epsilon(s) - \tilde{x}_\epsilon(s)| \end{aligned}$$

where $\Omega_1(T)$ is defined in (5.13), and $C_\epsilon$ and $C_1$ denote constants that are dependent, and respectively independent, of $\epsilon$. By the uniform convergence shown in (5.19) (which holds for the full sequence $\rho_\epsilon$), $C_\epsilon \to 0$ as $\epsilon \to 0$.

By using the estimate above in (5.25) we find

$$|v_\epsilon(t) - \tilde{v}_\epsilon(t) \le C_\epsilon + \frac{C_1}{\epsilon}\int_0^t e^{\frac{1}{\epsilon}(s-t)} |x_\epsilon(s) - \tilde{x}_\epsilon(s)| ds,$$

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

The next step is to use the equations (3.7) and (5.23) for the evolution of the $x$-characteristics $x_\epsilon$ and $\tilde{x}_\epsilon$ and estimate further:

$$|v_\epsilon(t) - \tilde{v}_\epsilon(t) \le C_\epsilon + \frac{C_1}{\epsilon}\int_0^t e^{\frac{1}{\epsilon}(s-t)} \int_0^s |v_\epsilon(\tau) - \tilde{v}_\epsilon(\tau)|\, d\tau ds.$$

By changing the order of integration in the double integral above we get

$$|v_\epsilon(t) - \tilde{v}_\epsilon(t) \leq C_\epsilon + \frac{C_1}{\epsilon} \int_0^t |v_\epsilon(\tau) - \tilde{v}_\epsilon(\tau)| \int_\tau^t e^{\frac{1}{\epsilon}(s-t)} ds\, d\tau$$

$$\leq C_\epsilon + C_1 \int_0^t |v_\epsilon(\tau) - \tilde{v}_\epsilon(\tau)| d\tau.$$

The convergence of the $v$-charateristics in (5.24) follows now from Gronwall's inequality, given that $C_\epsilon$ vanishes with $\epsilon$. The convergence of trajectories follows immediately as well. $\square$

The convergence of characteristic paths yields the limiting flow map $\mathcal{T}^t$ given by $x_0 \xrightarrow{\mathcal{T}^t} x(t)$. It is convenient in the calculations below to use the notation $x(t; x_0)$ to denote the limiting characteristic path $x(t)$ that starts at $x_0$.

The next result characterizes the limiting densities.

**Theorem 5.3** (Characterization of the limiting density)**.** *Suppose $\rho_0$ and $\nabla_x K$ satisfy the assumptions in Theorem 5.1. Then the limiting macroscopic density $\rho$ identified in Theorem 5.1 is the push-forward of the initial density $\rho_0$ by the limiting flow map $\mathcal{T}^t$,*

$$\rho = \mathcal{T}^t \# \rho_0. \tag{5.26}$$

*In addition, for each $t \in [0, T)$, $f_\epsilon$ converges in the $W_1$ metric to a probability density $f(t, \cdot, \cdot) \in \mathcal{P}_1(\mathbb{R}^d \times \mathbb{R}^d)$:*

$$f_\epsilon \xrightarrow{W_1} f \qquad \text{in } \mathcal{P}_1(\mathbb{R}^d \times \mathbb{R}^d) \qquad \text{as } \epsilon \to 0. \tag{5.27}$$

*The limiting density $f$, with first marginal $\rho$, is given explicitly by:*

$$f(t, x, v) = \rho(t, x)\delta(v + \nabla K * \rho(t, x)). \tag{5.28}$$

*Proof.* The first part, expressed by equation (5.26), follows from considerations made in Theorem 5.1 (see in particular Step 3 in the proof of Theorem 5.1). However, we show below how it can be derived directly in the $\epsilon \to 0$ limit of the kinetic equation.

The limiting behaviour of $f_\epsilon$ was not explicitly stated or needed in Theorem 5.1, but follows by arguments similar to those used for $\rho_\epsilon$ in Step 1 of the proof of Theorem 5.1. Let us sketch this argument briefly.

Fix $\phi_1 \in C_c^1(0, T)$ and $\phi_2 \in C_b^1(\mathbb{R}^d \times \mathbb{R}^d)$, and let $\phi(t, x, v) = \phi_1(t)\phi_2(x, v)$ in (3.6). Find

$$\int_0^T \phi_1'(t) \iint_{\mathbb{R}^d \times \mathbb{R}^d} \phi_2(x, v) f_\epsilon(t, x, v)\, \mathrm{d}x\, \mathrm{d}v\, \mathrm{d}t =$$

$$- \int_0^T \phi_1(t) \iint_{\mathbb{R}^d \times \mathbb{R}^d} \left[ \nabla_x \phi_2 \cdot v - \frac{1}{\epsilon} \nabla_v \phi_2 \cdot (v + \nabla_x K * \rho_\epsilon) \right] f_\epsilon\, \mathrm{d}x\, \mathrm{d}v\, \mathrm{d}t. \tag{5.29}$$

Denoting

$$\tilde{\xi}_\epsilon(t) = \iint_{\mathbb{R}^d \times \mathbb{R}^d} \phi_2(x, v) f_\epsilon(t, x, v)\, \mathrm{d}x\, \mathrm{d}v,$$

then, by (5.29), the weak derivative of $\tilde{\xi}_\epsilon$ is given by

$$\tilde{\xi}_\epsilon'(t) = \iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} \left[\nabla_x\phi_2 \cdot v - \frac{1}{\epsilon}\nabla_v\phi_2 \cdot (v + \nabla_x K * \rho_\epsilon)\right] f_\epsilon \,\mathrm{d}x\,\mathrm{d}v \in L^\infty(0,T).$$

In the above, the boundedness of the term that contains $\epsilon$ follows from Proposition 4.6.

Since $\tilde{\xi}_\epsilon(t)$ is uniformly bounded in $W^{1,\infty}(0,T)$, it converges uniformly on a subsequence. Hence, given any $\phi_2 \in C_b^1(\mathbb{R}^d \times \mathbb{R}^d)$, there exists a subsequence $\epsilon_k$ and $\tilde{\xi}(t) \in C([0,T))$ such that

$$\iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} \phi_2(x,v) f_{\epsilon_k}(t,x,v)\,\mathrm{d}x\,\mathrm{d}v \to \tilde{\xi}(t) \qquad \text{uniformly on } [0,T) \quad \text{as } \epsilon_k \to 0. \tag{5.30}$$

Similar to arguments used for $\rho_\epsilon$, we note that the sequence $f_\epsilon(t,\cdot,\cdot) \in \mathcal{P}_1(\mathbb{R}^d \times \mathbb{R}^d)$ is tight, and hence, for each $t \in [0,T)$, $f_\epsilon(t,\cdot,\cdot)$ converges in the $W_1$ metric on a subsequence to a probability measure $f(t,\cdot,\cdot) \in \mathcal{P}_1(\mathbb{R}^d \times \mathbb{R}^d)$.

By using the uniqueness of the limit $\tilde{\xi}(t)$ reached along the subsequence $\epsilon_k$ in (5.30), one can then argue as in Step 1 of the proof of Theorem 5.1, and infer that

$$\iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} \phi_2(x,v) f_{\epsilon_k}(t,x,v)\,\mathrm{d}x\,\mathrm{d}v \to \iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} \phi_2(x,v) f(t,x,v)\,\mathrm{d}x\,\mathrm{d}v \quad \text{uniformly on } [0,T) \quad \text{as } \epsilon_k \to 0,$$

for any $\phi_2 \in C_b^1(\mathbb{R}^d \times \mathbb{R}^d)$. In addition, at each $t \in [0,T)$,

$$f_{\epsilon_k}(t,\cdot,\cdot) \xrightarrow{W_1} f(t,\cdot,\cdot) \qquad \text{in } \mathcal{P}_1(\mathbb{R}^d \times \mathbb{R}^d) \quad \text{as } \epsilon_k \to 0. \tag{5.31}$$

This proves the convergence (5.27) on a subsequence. To upgrade it to convergence on the full sequence $\epsilon \to 0$ we use the uniqueness of $f$, as derived from the arguments below.

Since $f_\epsilon(t) = \mathcal{T}^{t,\epsilon}_{E[f_\epsilon]} \# f_0$, by (3.3) it holds that

$$\iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} \zeta(x,v) f_{\epsilon_k}(t,x,v) dx dv = \iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} \zeta(\mathcal{T}^{t,\epsilon_k}_{E[f_{\epsilon_k}]}(X,V)) f_0(X,V) dX dV, \tag{5.32}$$

for all $\zeta \in C_b(\mathbb{R}^d \times \mathbb{R}^d)$.

By the weak convergence of $f_{\epsilon_k}$, the left-hand-side of (5.32) converges as $\epsilon_k \to 0$:

$$\iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} \zeta(x,v) f_{\epsilon_k}(t,x,v) dx dv \to \iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} \zeta(x,v) f(t,x,v) dx dv.$$

Due to convergence of trajectories (5.20)-(5.22), the right-hand-side of (5.32) converges by Lebesgue's dominated convergence theorem,

$$\iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} \zeta(\mathcal{T}^{t,\epsilon_k}_{E[f_{\epsilon_k}]}(X,V)) f_0(X,V) dX dV \to \iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} \zeta(x(t;X), -\nabla K * \rho(t, x(t;X))) f_0(X,V) dX dV,$$

as $\epsilon_k \to 0$. Combining the two, we find

$$\iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} \zeta(x,v) f(t,x,v) dx dv = \int\limits_{\mathbb{R}^d} \zeta(x(t;X), -\nabla K * \rho(t, x(t;X))) \rho_0(X) dX, \tag{5.33}$$

for all $\zeta \in C_b(\mathbb{R}^d \times \mathbb{R}^d)$.

First note that (5.26) can be derived from (5.33). Indeed, choose $\zeta(x, v) = \varphi(x)$ in (5.33) to find

$$\int_{\mathbb{R}^d} \varphi(x)\rho(t, x)dx = \int_{\mathbb{R}^d} \varphi(x(t; X))\rho_0(X)dX, \tag{5.34}$$

for all $\varphi \in C_b(\mathbb{R}^d)$. The equation above represents exactly the mass transport given by (5.26).

Now, observe that (5.28) is equivalent to

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d} \zeta(x, v) f(t, x, v)dxdv = \int_{\mathbb{R}^d} \zeta(x, -\nabla K * \rho(t, x))\rho(t, x)dx,$$

for all test functions $\zeta \in C_b(\mathbb{R}^d \times \mathbb{R}^d)$, which can be inferred immediately from (5.33) and (5.34).

Finally, the *unique* explicit representation of the limiting density $f$ implies that the convergence in (5.31) holds on the full sequence $f_\epsilon$, as desired in (5.27). □

# References


[1] L. Ambrosio. Lecture notes on optimal transport problem. In J. Rodrigues P. Colli, editor, *Mathematical aspects of evolving interfaces*, volume 1812 of *CIME summer school in Madeira (Pt)*, pages 1–52. Springer, 2003.

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

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

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

[5] Andrea L. Bertozzi and Thomas Laurent. Finite-time blow-up of solutions of an aggregation equation in $\mathbf{R}^n$. *Comm. Math. Phys.*, 274(3):717–735, 2007.

[6] Andrea L. Bertozzi, Thomas Laurent, and Jesús Rosado. $L^p$ theory for the multidimensional aggregation equation. *Comm. Pure Appl. Math.*, 64(1):45–83, 2011.

[7] P. Billingsley. *Weak convergence of measures: Applications in probability*. Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1971.

[8] M. Bodnar and J. J. L. Velasquez. Derivation of macroscopic equations for individual cell-based models: a formal approach. *Math. Meth. Appl. Sci. (M3AS)*, 28:1757–1779, 2005.

[9] M. Bodnar and J. J. L. Velazquez. An integro-differential equation arising as a limit of individual cell-based models. *J. Differential Equations*, 222(2):341–380, 2006.

[10] Martin Burger and Marco Di Francesco. Large time behavior of nonlocal aggregation models with nonlinear diffusion. *Netw. Heterog. Media*, 3(4):749–785, 2008.

[11] Scott Camazine, Jean-Louis Deneubourg, Nigel R. Franks, James Sneyd, Guy Theraulaz, and Eric Bonabeau. *Self-organization in biological systems*. Princeton Studies in Complexity. Princeton University Press, Princeton, NJ, 2003. Reprint of the 2001 original.

[12] J. A. Cañizo, J. A. Carrillo, and J. Rosado. A well-posedness theory in measures for some kinetic models of collective motion. *Math. Models Methods Appl. Sci.*, 21(3):515–539, 2011.

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

[14] José A. Carrillo, Massimo Fornasier, Giuseppe Toscani, and Francesco Vecil. Particle, kinetic, and hydrodynamic models of swarming. In *Mathematical modeling of collective behavior in socio-economic and life sciences*, Model. Simul. Sci. Eng. Technol., pages 297–336. Birkhäuser Boston, Inc., Boston, MA, 2010.

[15] R. Choksi, R. C. Fetecau, and I. Topaloglu. On minimizers of interaction functionals with competing attractive and repulsive potentials. *Ann. I. H. Poincaré - AN*, 2014. DOI: 10.1016/j.anihpc.2014.09.004 (in press).

[16] J. Evers, R. C. Fetecau, and L. Ryzhik. Anisotropic interactions in a first-order aggregation model: a proof of concept,. 2014. preprint.

[17] Klemens Fellner and Gaël Raoul. Stable stationary states of non-local interaction equations. *Math. Models Methods Appl. Sci.*, 20(12):2267–2291, 2010.

[18] Neil Fenichel. Geometric singular perturbation theory for ordinary differential equations. *J. Differential Equations*, 31(1):53–98, 1979.

[19] R. C. Fetecau and Y. Huang. Equilibria of biological aggregations with nonlocal repulsive-attractive interactions. *Phys. D*, 260:49–64, 2013.

[20] R. C. Fetecau, Y. Huang, and T. Kolokolnikov. Swarm dynamics and equilibria for a nonlocal aggregation model. *Nonlinearity*, 24(10):2681–2716, 2011.

[21] R. Glassey. *The Cauchy Problem in Kinetic Theory*. SIAM, Philadelphia, PA, 1996.

[22] J.M. Haile. *Molecular Dynamics Simulation: Elementary Methods*. John Wiley and Sons, Inc., New York, 1992.

[23] D. Helbing. Traffic and related self-driven many particle systems. *Rev. Mod. Phys.*, 73:1067–1141, 2001.

[24] Darryl D. Holm and Vakhtang Putkaradze. Aggregation of finite-size particles with variable mobility. *Phys Rev Lett.*, 95:226106, 2005.

[25] Pierre-Emmanuel Jabin. Macroscopic limit of Vlasov type equations with friction. *Annales de l'Institut Henri Poincare (C) Non Linear Analysis*, 17(5):651–672, 2000.

[26] M. O. Jackson. *Social and Economic Networks*. Princeton University Press, 2010.

[27] M. Ji and M. Egerstedt. Distributed coordination control of multi-agent systems while preserving connectedness. *IEEE Trans. Robot.*, 23(4):693–703, 2007.

[28] Theodore Kolokolnikov, Hui Sun, David Uminsky, and Andrea L. Bertozzi. A theory of complex patterns arising from 2D particle interactions. *Phys. Rev. E, Rapid Communications*, 84:015203(R), 2011.

[29] G. Savaré L. Ambrosio, N. Gigli. *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Birkhäuserauser, 2005.

[30] Thomas Laurent. Local and global existence for an aggregation equation. *Comm. Partial Differential Equations*, 32(10-12):1941–1964, 2007.

[31] Andrew J. Leverentz, Chad M. Topaz, and Andrew J. Bernoff. Asymptotic dynamics of attractive-repulsive swarms. *SIAM J. Appl. Dyn. Syst.*, 8(3):880–908, 2009.

[32] A. Mogilner and L. Edelstein-Keshet. A non-local model for a swarm. *J. Math. Biol.*, 38:534–570, 1999.

[33] Sebastien Motsch and Eitan Tadmor. Heterophilious dynamics enhances consensus. *SIAM Review*, 56:577–621, 2014.

[34] L. Saint-Raymond. A mathematical PDE perspective on the Chapman-Enskog expansion. *Bull. Amer. Math. Soc.*, 51:247–275, 2014.

[35] A. N. Tikhonov. Systems of differential equations containing small parameters in the derivatives. *Mat. Sb. (N.S.)*, 31(73):575–586, 1952.

[36] C. M. Topaz, A. L. Bertozzi, and M. A. Lewis. A nonlocal continuum model for biological aggregation. *Bull. Math. Bio.*, 68:1601–1623, 2006.

[37] Giuseppe Toscani. One-dimensional kinetic models of granular flows. *M2AN Math. Model. Numer. Anal.*, 34(6):1277–1291, 2000.

[38] A. B. Vasil′eva. Asymptotic behaviour of solutions of certain problems for ordinary nonlinear differential equations with a small parameter multiplying the highest derivatives. *Uspehi Mat. Nauk*, 18(3 (111)):15–86, 1963.

[39] James von Brecht, David Uminsky, Theodore Kolokolnikov, and Andrea Bertozzi. Predicting pattern formation in particle interactions. *Math. Models Methods Appl. Sci.*, 22(Supp. 1):1140002, 2012.

[40] James H. von Brecht and David Uminsky. On soccer balls and linearized inverse statistical mechanics. *J. Nonlinear Sci.*, 22(6):935–959, 2012.