Mathematical Models and Methods in Applied Sciences
Vol. 27, No. 4 (2017) 745–770

DOI: 10.1142/S0218202517500154

A quest toward a mathematical theory of the dynamics of swarms

Nicola Bellomo

*Department of Mathematics,
Faculty Sciences, King Abdulaziz University,
Jeddah, Saudi Arabia*

*Politecnico di Torino, Torino, Italy
nicola.bellomo@polito.it*

Seung-Yeal Ha*

*Department of Mathematical Sciences
and Research Institute of Mathematics,
Seoul National University, Seoul 151-747, Korea*

*Korea Institute for Advanced Study,
Hoegiro 87, Seoul, 130-722, Korea
syha@snu.ac.kr*

Received 23 September 2016
Revised 31 December 2016
Accepted 30 January 2017
Published 10 March 2017
Communicated by F. Brezzi

This paper addresses some preliminary steps toward the modeling and qualitative analysis of swarms viewed as living complex systems. The approach is based on the methods of kinetic theory and statistical mechanics, where interactions at the microscopic scale are nonlocal, nonlinearly additive and modeled by theoretical tools of stochastic game theory. Collective learning theory can play an important role in the modeling approach. We present a kinetic equation incorporating the Cucker–Smale flocking force and stochastic game theoretic interactions in collision operators. We also present a sufficient framework leading to the asymptotic velocity alignment and global existence of smooth solutions for the proposed kinetic model with a special kernel. Analytic results on the global existence and flocking dynamics are presented, while the last part of the paper looks ahead to research perspectives.

Keywords: Collective dynamics; Cucker–Smale flocking; learning; living complex systems; self-organization; swarming; collective behavior; nonlinear interactions.

AMS Subject Classification: 82D99, 91D10

*Corresponding author

1. Introduction

Modeling of swarm dynamics is a challenging research objective, which requires mathematical tools suitable to capture the complexity features of large systems of self-propelled particles. The approach should describe them as living systems due to their individual ability to develop specific strategies, individual and collective, based on their interactions with other entities. A number of celebrated phenomenological models have been proposed and studied in literature focusing also on various applications, e.g. Refs. 2, 5, 6, 13, 21–26, 35, 37, 40–42, and many others. The modeling approach has, in general, taken advantage of some analogy with crowd modeling,[8,10,11] while a system theory approach to social systems[1] can contribute to the understanding of social dynamics appearing in swarms. Despite remarkable efforts by various mathematicians who have enlightened the properties of these models, a general mathematical theory is still far from a complete theory.

The first step in the search of a theory should look for a mathematical structure suitable to include the main features of a swarm, based on a phenomenological interpretation of their collective dynamics, while the second step consists in deriving specific models consistent with such structure. A theory, once developed, should show how specific models, known in the literature, can be viewed as special cases of the ultimate universal theory. This quest should march together with computational methods leading to quantitative description of the dynamics. Traditional deterministic methods might present technical difficulties to be operative. Therefore, stochastic methods need to be developed, such as Monte Carlo particle methods.[3,7,36]

The purpose of this paper is to incorporate the flocking mechanism together with social interactions between agents which are registered by some collision mechanism. The terminology "*flocking*" in this paper is used in a broad sense, namely it denotes the phenomenon in which self-propelled agents using only limited environmental information and simple rules, organize into an ordered motion.[40] The challenging objective of our paper consists in proposing a contribution to the development of a mathematical theory of swarms focused in various fields of life sciences including crowd dynamics, financial markets, etc.

To incorporate two interaction mechanisms such as flocking mechanism and social agent–agent interactions, we propose a kinetic model which can be viewed as a "*collisional counterpart*" of the kinetic Cucker–Smale (CS) equation introduced in Refs. 30 and 31, while for issues related to the CS model, we refer to Refs. 28, 29 and 33. The term "*collision*" is here used as a "jargon" to indicate that both short and long-range interactions can be included in the kinetic equation.

In our paper, we are interested in the dynamics of the ensemble of many self-propelled particles (agents) exhibiting a collective flocking dynamics. In this situation, the dynamics of the large ensemble can be effectively described by a mesoscopic equation. To be more specific, we introduce a kinetic function $f = f(t, \mathbf{x}, \mathbf{v})$ which represents a one-particle probability density function at the phase space

$(\mathbf{x}, \mathbf{v}) \in \Omega \times D_{\mathbf{v}}$ at time t whose temporal–spatial dynamics can be described by the following kinetic equation:

$$\begin{cases} \partial_t f + \mathbf{v} \cdot \nabla_{\mathbf{x}} f + \kappa \nabla_{\mathbf{v}} \cdot (\mathcal{F}_a(f) f) = \varepsilon Q(f, f), \quad \mathbf{x} \in \Omega, \ \mathbf{v} \in D_{\mathbf{v}}, \ t > 0, \\ \mathcal{F}_a(f)(t, \mathbf{x}, \mathbf{v}) := -\displaystyle\int_{\Omega \times D_{\mathbf{v}}} \psi(|\mathbf{x} - \mathbf{x}^*|)(\mathbf{v} - \mathbf{v}^*) f(t, \mathbf{x}^*, \mathbf{v}^*) d\mathbf{v}^* \, d\mathbf{x}^*, \end{cases} \tag{1.1}$$

where $\mathcal{F}_a(f)$ and $Q(f, f)$ are CS flocking force and interaction operator for agent–agent interactions, respectively. The non-negative constants κ and ε denote a coupling strength and a constant inversely proportional to the mean free path between agents, respectively. The detailed flocking interactions between agents are registered by the non-negative communication weight function ψ.

In the absence of agent–agent interactions, i.e. $\varepsilon = 0$, Eq. (1.1) reduces to the kinetic CS equation[17,19,30,31]:

$$\begin{cases} \partial_t f + \mathbf{v} \cdot \nabla_{\mathbf{x}} f + \kappa \nabla_{\mathbf{v}} \cdot (\mathcal{F}_a(f) f) = 0, \quad \mathbf{x} \in \Omega, \ \mathbf{v} \in D_{\mathbf{v}}, \ t > 0, \\ \mathcal{F}_a(f)(t, \mathbf{x}, \mathbf{v}) := -\displaystyle\int_{\Omega \times D_{\mathbf{v}}} \psi(|\mathbf{x} - \mathbf{x}^*|)(\mathbf{v} - \mathbf{v}^*) f(t, \mathbf{x}^*, \mathbf{v}^*) d\mathbf{v}^* \, d\mathbf{x}^*. \end{cases} \tag{1.2}$$

The rest of this paper is organized as follows. In Sec. 2, we provide a description of the features of swarms that we aim at inserting in the mathematical description. This section will also propose a well-defined strategy toward a general mathematical theory of swarm modeling which can be specialized for each type of swarms. In Sec. 3, we tackle the first step of the aforementioned mathematical theory which specifically consists in designing a differential structure suitable to capture the phenomenological features presented in Sec. 2. In Sec. 4, we introduce a kinetic CS equation with agent–agent collisions and study some basic structures. In Sec. 5, we study the propagation of velocity moments and *a priori* flocking and $W^{k,\infty}$ estimates for (1.1) in a very restricted framework. Finally, Sec. 6 is devoted to the summary of our main results and a critical analysis which looks ahead to research perspectives.

2. On a Strategy Toward a Mathematical Approach

In this section, we identify some specific features of swarms, viewed as large living systems, to be properly accounted for in the mathematical approach. The contents are not yet formalized, but aim at providing the basis for a possible mathematical structure suitable to capture these features. The rationale of the approach proposed in the following is that although swarms may be composed of different types of agents (entities), some common features can be identified. Some of these features can be subsequently specialized according to the entities that compose the swarm. These will be called, hereinafter, "*active particles*" or, for short, *particles* (*agents*). These have the ability to develop specific strategies and hence a "behavioral" dynamics according to rules that differ from those of the classical particles, see Ref. 12. In detail, we consider the following specific features:

(1) The swarm consists of a large number of active particles interacting with other particles within a sensitivity domain to be properly defined in Sec. 3.2.
(2) Interactions are nonlocal and nonlinearly additive.
(3) The swarm occupies an initial domain which, due to the overall dynamics, evolves in time.
(4) The approach might include the possible presence of different groups, for instance related to the presence of leaders, distinguished by a different strategy and interaction rules with other groups they can develop.
(5) The rules of the dynamics in interactions depend on the specific type of entities composing of the swarm, but it is uniformly distributed within each group.

This paper pursues the derivation of a general mathematical structure suitable for the framework to be used toward the derivation of specific models corresponding to different types of swarms. A general strategy can be defined to derive such structure in view of the mathematical formalization to be dealt with in the next section.

- The approach to be used is that of the so-called *kinetic theory of active particles*, which has been applied to the modeling of crowd dynamics with somewhat different features.[8,10,11] However, due to important technical differences that swarms show with respect to crowds, a straightforward extension to our setting is not sufficient, and additional work should be done.
- The interacting entities are viewed as *active particles* due to their ability to develop specific strategies. Their state is defined at the *microscopic scale* by the variables appropriate to depict their physical state. This state can include both mechanical and activity (or internal) variables, where the *activity* models the possible heterogeneous behaviors of the particles.
- The overall system can be subdivided into a number of groups, called *functional subsystems*, characterized by different ways of expressing their strategy. As an example, this subdivision can include predators and prey, or a possible hierarchy within the swarm. The state of each functional subsystem is defined by a *probability distribution function over the microscopic state* of active particles.
- Each active particle has a *sensitivity domain* within which the presence of other particles can be sensed. This zone is generally symmetric with respect to the velocity, however this is not a general rule. The development of a strategy needs a *sufficient amount of informations*, namely a sufficient number of interacting particles.
- Particles interact within an *effective interaction domain* somewhat related to the aforementioned sensitivity domain and amount of information. Two types of interactions are developed, *adjustment of the velocity to that of the interaction domain* and *attraction or repulsion* with respect to the other particles depending on their reciprocal distances.

3. Search for a Mathematical Structure

In this section, we deal with the derivation, based on the suitable developments of the methods of the kinetic theory, of a mathematical structure deemed to provide the conceptual background for the derivation of a broad variety of swarming models. The contents are presented in a sequel of subsections focused on the following topics: representation, modeling the individual sensitivity domain, modeling of individual-based and mean-field interactions, and derivation of the mathematical structure for a one-component swarm, derivation of the structure for a swarm with several components.

3.1. *Representation*

We consider a swarm occupied on initial spatial domain $\Sigma_0 \subseteq \mathbb{R}^d$, and let ℓ be the diameter of the sphere with center zero that includes Σ_0. Such a domain evolves in time with a shape denoted by Σ_t. Dimensionless quantities can be used by referring the real space coordinates to ℓ, the speed to the maximal velocity v_L that can be reached by a fast isolated agent, the time to a critical time T_c obtained by dividing ℓ by v_L. In addition, the dimensional distribution function multiplied by v_L^3 is normalized with respect to the maximal density n_L corresponding to packing conditions due to the finite size of the particles.

Consider a swarm where all individual entities develop the same strategy. Then, according to classical methods of the kinetic theory, the representation of the overall system is delivered by the one particle distribution function $f = f(t, \mathbf{x}, \mathbf{v})$ of so-called *test particle*, where $\mathbf{x} \in \Sigma_t \subset \mathbb{R}^d$ is the spatial coordinate and $\mathbf{v} \in D_{\mathbf{v}} \subset \mathbb{R}^d$ is the velocity. The *microscopic state* is defined by $(\mathbf{x}, \mathbf{v})$, which is a point of the *phase space* in $\mathbb{R}^d \times \mathbb{R}^d$. The test particle is assumed to be the representative of the whole system. Once the kinetic density function $f = f(t, \mathbf{x}, \mathbf{v})$ is determined, then macroscopic quantities can be computed, under suitable integrability relations, by the velocity moments of f. For example, the local mass density ρ and local bulk velocity $\boldsymbol{\xi}$ are given by the following relations, respectively:

$$\rho(t, x) := \int_{D_v} f(t, \mathbf{x}, \mathbf{v}) d\mathbf{v} \quad \text{and} \quad \boldsymbol{\xi}(t, x) = \frac{1}{\rho(t, \mathbf{x})} \int_{D_v} \mathbf{v} f(t, \mathbf{x}, \mathbf{v}) d\mathbf{v},$$

while the local mass flux $\mathbf{q}$ is computed by $\rho \boldsymbol{\xi}$.

Coordinates are referred to a fixed frame of orthogonal axes. In some cases, it can be useful using spherical coordinates for the velocity

$$\mathbf{v} = v\,\boldsymbol{\omega}, \quad \boldsymbol{\omega} = \{\theta, \varphi\}, \quad \theta \in [-\pi, \pi],\ \varphi \in [0, 2\pi),$$

where $v \in [0, 1]$ is the real *speed* divided by v_L, while $\boldsymbol{\omega}$ is a unit vector defining the direction of the velocity. In the planar case $\theta = 0$, while $\varphi \in [0, 2\pi)$.

3.2. *Sensitivity and interaction domains*

Consider a point $\mathbf{x} \in \Sigma_t$ and let $\mathbf{v}$ be the local microscopic velocity in $\mathbf{x}$. The test particle has a *visibility domain* Γ, a *sensitivity domain* Ω_s, a *critical sensitivity domain* Ω_c, and *effective interaction domain* Ω. In detail:

- Ω_s is the domain within which a particle has the potential ability of sensing the presence of other particles, where the size of this domain also depends on the visibility situation;
- Ω_c is the domain within which a particle can acquire the information necessary to develop the strategy from other particles;
- Ω is the domain within which interactions effectively occur depending on the intersection between Ω_c and Ω_s.

All domains can evolve in time due to the dynamics of local flow conditions. For instance, Ω_s can be reduced by the presence of obstacles, while Ω_c depends on the local density conditions. We are interested in computing the interaction domain Ω.

A general assumption is that Ω-type domains have a well-defined *qualitative shape*, and they are symmetric with respect to the velocity direction $\boldsymbol{\omega}$. Then, detailed calculations can be developed by exploiting the conjecture proposed in paper Ref. 6 which states that Ω_c is related to a critical amount of information, namely amount of interactions, to be achieved to fully develop the decision process that supports the interaction rules between particles. This critical amount corresponds to a critical density m_c which is a constant value for each type of swarm. Namely, an active particle takes into account of only a fixed number of neighbors. This quantity needs to be related to m_s corresponding to the amount of information that can be acquired within the sensitivity domain. Here densities m_c and m_s are determined, for $t \geq 0$, by the following relations:

$$m_c(t) = \int_{D_{\mathbf{v}}} \int_{\Omega_c[f]} f(t, \mathbf{x}, \mathbf{v}) d\mathbf{x}\, d\mathbf{v}, \tag{3.1}$$

$$m_s(t) = \int_{D_{\mathbf{v}}} \int_{\Omega_s[f]} f(t, \mathbf{x}, \mathbf{v}) d\mathbf{x}\, d\mathbf{v}. \tag{3.2}$$

Equation (3.1) defines the map $f \mapsto \Omega_c[f]$, which if m_c is given, has a unique solution only under appropriate assumptions (and symmetry) of the shape of the Ω-type domains.

Interactions are effective only within Ω_s depending on the amount of information related to Ω_c. In more detail, let us refer to Fig. 1 representing the projection on a plane through the velocity $\mathbf{v}$, and shows that the overlapping region of domains is specialized, when their shape is a cone. The following result appears:

(1) If $\Omega_c \subseteq \Omega_s$, see Fig. 1 (Particle B), the information acquired by a particle is sufficient to organize its dynamics with interactions in $\Omega = \Omega_c$, and $m_c \leq m_s$ and $m_c > m_s$.

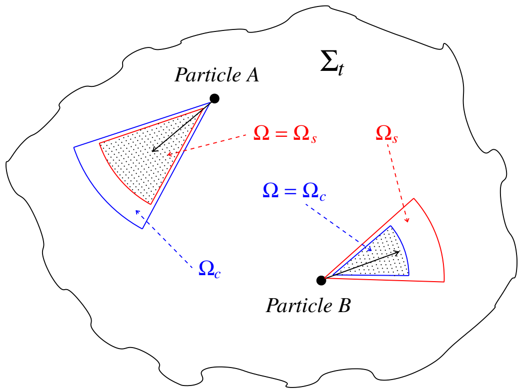

Fig. 1. Schematic diagram of sensitivity and interaction domains.

(2) If $\Omega_s \subset \Omega_c$ the information is not complete, see Fig. 1 (Particle A), the information acquired by a particle is not sufficient to organize its dynamics with interactions in $\Omega = \Omega_s$.

Remark 3.1. A parameter $\alpha \in [0, 1]$ can be introduced to quantify the *amount of information* received in the interactions, where $\alpha = 0$ and $\alpha = 1$ correspond to the absence of information and complete information, respectively. In general, α can depend on f, namely $\alpha = \alpha[f]$.

We have mentioned that detailed calculations can be developed only under suitable assumptions on the shape of the domains. Bearing this in mind, let us suppose that Ω, Ω_s, and Ω_c are included in a cone with vertex in $\mathbf{x}$ and axis $\boldsymbol{\omega}$. In addition, we suppose that the opening ψ of the cone is the same for all domains that, consequently, are defined by the three radii R_s, R_c, and R. The polar coordinates that have been defined above are used also for both variables $\mathbf{x}$ and $\mathbf{v}$ inside the cone, the relation (3.1) corresponds to

$$m_c = \int_{D_{\mathbf{v}}} \int_{-\pi}^{\pi} \int_0^{2\pi} \int_0^{R_c[f]} f(t, \mathbf{x}, \mathbf{v}) d\mathbf{x}\, d\mathbf{v}, \tag{3.3}$$

while analogous calculations in the sensitivity cone provide m_s.

The amount of information that each active particle can receive might be related to the aforementioned parameter α as follows:

(1) The domain Ω_c which is necessary to obtain a complete information is included in the sensitivity domain. Therefore, active particles receive sufficient information to organize their dynamics:

$$m_c \leq m_s \Rightarrow R = R_c[f] : \alpha[f] = 1.$$

(2) The domain necessary to obtain a complete information includes the sensitivity domain. Therefore, the active particles can receive only a partial information

to organize the dynamics:

$$m_c > m_s \Rightarrow R = R_s : \alpha[f] = \frac{m_s}{m_c}[f].$$

In particular, if $m_s = 0$, then $\alpha[f] = 0$.

3.3. *From the modeling of interactions to the derivation of mathematical structures*

According to the phenomenological description of Sec. 2, two types of interactions should be included into the modeling approach corresponding to consensus/dissent toward a common velocity, and attraction/repulsion, respectively.

The *former type* involves with the following three types of particles at each time t: *test* particles with microscopic state $(\mathbf{x}, \mathbf{v})$, *field* particles with microscopic state $(\mathbf{x}^*, \mathbf{v}^*)$, and *candidate* particles with microscopic state $(\mathbf{x}, \mathbf{v}_*)$. The candidate particle can acquire the state of the test particle with some probability as a consequence of interactions with field particles, while the test particle can also lose its state due to the interactions with field particles. Interactions can be modeled by the following quantities:

- $\eta[f]$ is the rate of encounters supposed to be dependent on f, as the number of particles involved in the interactions depends on the visibility and sensitivity zone. A simple model is as follows: $\eta[f] = \eta_0\, \alpha[f]$, which is a constant $\eta = \eta_0$ when the particle receives a complete information $\alpha = 1$, while if the information is not complete, it decays to $\eta = 0$, where $\alpha = 0$, corresponding to complete lack of information. A practical alternative consists in assuming $\eta = \eta_0$. In fact, the lack of information is already taken into account by $\Omega[f]$, which can also be equal to zero due to the excessive distance, presence of obstacles, etc. Equation (3.5) written in the following is based on this assumption.
- $\mathcal{A}[f](\mathbf{v}_* \to \mathbf{v}* \,|\, \mathbf{v}_*, \mathbf{v}^*)$ is the turning probability density that a candidate particle with velocity $\mathbf{v}_*$ changes into a particle with velocity $\mathbf{v}$ due to interactions with a field particle. In general, $\mathcal{A}$ might depend not only on the velocity of the interacting particles $\mathbf{v}_*, \mathbf{v}^*$, but also on f itself and satisfies the probability density condition:

$$\int_{D_{\mathbf{v}}} \mathcal{A}[f](\mathbf{v}_* \to \mathbf{v} \,|\, v_*, v^*) d\mathbf{v} = 1, \tag{3.4}$$

for all conditioning inputs.

The *latter type* involves, at each time t, the test and the field particles in the domain Ω. A general expression of the acceleration applied to the test particle is as follows:

$$\mathcal{F}[f](t, \mathbf{x}) = \mu_0 \int_{\Omega[f]} \int_{D_{\mathbf{v}}} \varphi[f](\mathbf{x}, \mathbf{x}^*) f(t, \mathbf{x}^*, \mathbf{v}^*) d\mathbf{x}^* d\mathbf{v}^*, \tag{3.5}$$

where φ is the individual action of the field particles over the test particle. This acceleration term can also depend on the distribution function inducing an flocking action, but also a repulsion depending on the distance. φ can be related to a pair interaction potential depending on the distance between the test and field particles:

$$\varphi[f](\mathbf{x}, \mathbf{x}^*) = \partial_z \mathcal{U}(\mathbf{x}, \mathbf{x}^*), \quad z = |\mathbf{x} - \mathbf{x}^*|.$$

Remark 3.2. Some general constraints toward the modeling of $\mathcal{U}$ can be given as follows:

$$\mathbf{x} \in \Omega[f] \Rightarrow \mathcal{U} \neq 0; \quad \mathbf{x} \notin \Omega[f] \Rightarrow \mathcal{U} = 0;$$

and

$$|\mathbf{x} - \mathbf{x}^*| \geq a \Rightarrow \partial_z \mathcal{U} < 0; \quad |\mathbf{x} - \mathbf{x}^*| > a \Rightarrow \partial_z \mathcal{U} > 0.$$

Therefore, the field force is repulsive when the distance is below a threshold, while it becomes weakly attractive when the distance is greater than such a threshold.

Suppose that all aforementioned quantities are known, the mathematical structure can be derived by the conservation of particles in the elementary volume of the phase space. Technical calculations yield

$$\partial_t f + \mathbf{v} \cdot \nabla_{\mathbf{x}} f + \kappa \nabla_{\mathbf{v}} \cdot (\mathcal{F}_a(f) f) = \varepsilon(Q^+(f,f) - Q^-(f,f)), \tag{3.6}$$

where Q^+ and Q^- correspond, respectively, to the *gain* and *loss* terms in the elementary space of the microscopic states due to short-range interactions. Their expressions are as follows:

$$\begin{aligned} Q^+(f,f)(t,\mathbf{x},\mathbf{v}) := \int_{\Omega[f]\times(D_{\mathbf{v}})^2} & \mathcal{A}[f](\mathbf{v}_* \to \mathbf{v} \,|\, \mathbf{v}_*, \mathbf{v}^*) \\ & \times f(t,\mathbf{x},\mathbf{v}_*) f(t,\mathbf{x}^*,\mathbf{v}^*) d\mathbf{v}_* \, d\mathbf{x}^* d\mathbf{v}^*, \end{aligned} \tag{3.7}$$

and

$$Q^-(f,f)(t,\mathbf{x},\mathbf{v}) := f(t,\mathbf{x},\mathbf{v}) \underbrace{\int_{\Omega[f]\times D_{\mathbf{v}}} f(t,\mathbf{x}^*,\mathbf{v}^*) d\mathbf{x}^* d\mathbf{v}^*}_{=:L(f)}. \tag{3.8}$$

Further, κ and ε are characteristic dimensionless parameters that are also related to η_0 and μ_0 as well as to the scaling to use dimensionless variables, while the two terms $\mathcal{A}$ and $\mathcal{F}_a(f)$ correspond to two different defined dynamics, respectively, flocking and consensus toward a common velocity. The detailed discussion of the mathematical structure (3.9) will be presented in next section referring to specific models that can be derived within such general framework.

3.4. *Mathematical structures for a heterogeneous system*

An interesting literature indicates the need of including in the swarms different types of active particles. In particular, the presence of leaders and predators.[13,27] Therefore, it is useful in view of possible modeling developments, generalizing the structure (3.9) to model large systems that include several interacting populations. There have been called in Ref. 12 *functional subsystems* to stress that each subsystem follows interaction rules which differs across subsystems.

Bearing all this in mind, let us consider as warm of n interacting subsystems labeled by the subscript $i = 1, \dots, n$, the general structure can be written as follows:

$$\partial_t f_i + \mathbf{v} \cdot \nabla_{\mathbf{x}} f_i + \kappa \sum_{j=1}^{n} \nabla_{\mathbf{v}} \cdot (\mathcal{F}_{ij}[f] f) = \varepsilon \sum_{j=1}^{n} (Q_{ij}^{+}(f,f) - Q_{ij}^{-}(f,f)), \tag{3.9}$$

where

$$Q_{ij}^{+}(f,f)(t,\mathbf{x},\mathbf{v}) := \int_{\Omega[f]\times(D_{\mathbf{v}})^2} \mathcal{A}_{ij}[f](\mathbf{v}_* \to \mathbf{v} \,|\, \mathbf{v}_*, \mathbf{v}^*) \times f_i(t,\mathbf{x},\mathbf{v}_*) f_j(t,\mathbf{x}^*,\mathbf{v}^*) d\mathbf{v}_*\, d\mathbf{x}^* d\mathbf{v}^* \tag{3.10}$$

and

$$Q_{ij}^{-}(f,f)(t,\mathbf{x},\mathbf{v}) := f_i(t,\mathbf{x},\mathbf{v}) \int_{\Omega[f]\times D_{\mathbf{v}}} f_j(t,\mathbf{x}^*,\mathbf{v}^*) d\mathbf{x}^* d\mathbf{v}^*. \tag{3.11}$$

4. A Kinetic Cucker–Smale Equation

In this section, we present a kinetic CS equation with short-range interactions (collisions) (3.9), (3.10), (3.8), (3.11), where agent–agent interactions are incorporated in the original kinetic CS model.[30,31] To motivate the corresponding kinetic equation, we first briefly recall the CS model and present its kinetic counterpart added with short-range interaction effects.

4.1. *From particle CS model to kinetic CS equation*

Cucker and Smale introduced a Newton-type agent-based model[22] for a system of interacting particles exhibiting flocking phenomenon, and provided several sufficient conditions for the asymptotic flockings in terms of initial configuration, coupling strength and communication weights. Let $\mathbf{x}_i$ and $\mathbf{v}_i$ be the position and velocity of the ith CS particle, respectively. Then, the particle CS model with metric dependent communication weights is given by the following ODE system:

$$\begin{cases} \dfrac{d\mathbf{x}_i}{dt} = \mathbf{v}_i, \quad t > 0, \quad i = 1, \dots, N, \\[2ex] \dfrac{d\mathbf{v}_i}{dt} = \dfrac{\kappa}{N} \displaystyle\sum_{j=1}^{N} \psi(|\mathbf{x}_j - \mathbf{x}_i|)(\mathbf{v}_j - \mathbf{v}_i), \end{cases} \tag{4.1}$$

where κ is a non-negative coupling strength and ψ is a communication weight function measuring the degree of communications (interactions) between agents. For the large CS system (4.1) with $N \gg 1$, it is not reasonable to integrate the particle model directly due to the computational cost. Thus, it is natural to introduce its kinetic counterpart for (4.1). Let $f = f(t, \mathbf{x}, \mathbf{v})$ be a one-particle probability density function at phase space position $(\mathbf{x}, \mathbf{v})$ and at time t. Then the spatial–temporal evolution of f is governed by the following Vlasov-type equation (1.2):

$$\begin{cases} \partial_t f + \mathbf{v} \cdot \nabla_{\mathbf{x}} f + \kappa \nabla_{\mathbf{v}} \cdot (\mathcal{F}_a(f) f) = 0, \quad (\mathbf{x}, \mathbf{v}) \in \Omega[f] \times D_{\mathbf{v}}, \quad t > 0, \\ \mathcal{F}_a[f](t, \mathbf{x}, \mathbf{v}) := -\displaystyle\int_{\Omega[f] \times D_{\mathbf{v}}} \psi(|\mathbf{x} - \mathbf{x}^*|)(\mathbf{v} - \mathbf{v}^*) f(t, \mathbf{x}^*, \mathbf{v}^*) d\mathbf{v}^* d\mathbf{x}^*. \end{cases} \tag{4.2}$$

Note that the direct particle–particle interactions are not taken into account in the above equation (4.1), and Eq. (4.2) admits a global smooth solution, as long as initial datum is compactly supported in $\mathbf{x}$ and $\mathbf{v}$ and sufficiently regular (see Ref. 31).

Lemma 4.1. (Conservation laws and dissipation) *Let $f = f(t, \mathbf{x}, \mathbf{v})$ be a $\mathcal{C}^1$-solution to* (4.2) *decaying at infinity in the phase space sufficiently fast. Then, the velocity moments of f satisfy the following estimates*: *for $t > 0$*,

(i) $\frac{d}{dt}\int_{\Omega[f] \times D_{\mathbf{v}}} f(t, \mathbf{x}, \mathbf{v}) d\mathbf{v}\, d\mathbf{x} = 0$,

(ii) $\frac{d}{dt}\int_{\Omega[f] \times D_{\mathbf{v}}} \mathbf{v} f(t, \mathbf{x}, \mathbf{v}) d\mathbf{v}\, d\mathbf{x} = 0$,

(iii) $\frac{d}{dt}\int_{\Omega[f] \times D_{\mathbf{v}}} |\mathbf{v}|^2 f(t, \mathbf{x}, \mathbf{v}) d\mathbf{v}\, d\mathbf{x} = -\kappa \int_{\Omega[f]^2 \times D_{\mathbf{v}}^2} \psi(|\mathbf{x} - \mathbf{x}^*|)|\mathbf{v} - \mathbf{v}^*|^2 f(t, \mathbf{x}, \mathbf{v})$ $f(t, \mathbf{x}^*, \mathbf{v}^*) d\mathbf{v}^* d\mathbf{v}\, d\mathbf{x}^* d\mathbf{x}$.

Proof. For a proof, we refer to Refs. 30 and 31. □

We are now ready to return to the case with short-range interactions, where the contents in Sec. 3.3 play a key role.

4.2. *Adding short-range interactions to the kinetic CS equation*

In this subsection, we augment the collisionless Vlasov equation (4.2) by adding a collision operator responsible for short-range agent–agent interactions. Our collision operator will be made to take into account of candidate and field particles, and is assumed to satisfy a minimal conservation law (conservation of mass). Since our collision terms are modeled to incorporate social and stochastic interactions between particles, in general, it does not satisfy the conservation of momentum and energy unlike to the elastic interactions between mechanical particles. More precisely, we define

$$Q(f, f) := Q^+(f, f) - Q^-(f, f) = Q^+(f, f) - fL(f),$$

$$
\begin{aligned}
Q^+(f,f)(t,\mathbf{x},\mathbf{v}) &:= \int_{\Omega[f]\times D_{\mathbf{v}}} \left(\int_{D_{\mathbf{v}}} \mathcal{A}[f](\mathbf{v}_* \to \mathbf{v} \,|\, \mathbf{v}_*, \mathbf{v}^*) f(t,\mathbf{x},\mathbf{v}_*) d\mathbf{v}_* \right) \\
&\qquad \times f(t,\mathbf{x}^*,\mathbf{v}^*) d\mathbf{v}^* d\mathbf{x}^*, \\
(fL(f))(t,\mathbf{x},\mathbf{v}) &:= f(t,\mathbf{x},\mathbf{v}) \int_{\Omega[f]\times D_{\mathbf{v}}} f(t,\mathbf{x}^*,\mathbf{v}^*) d\mathbf{v}^* d\mathbf{x}^*.
\end{aligned} \tag{4.3}
$$

Note that the terms inside the parenthesis denote the local mass of field particles which will interact with field particles, and the collision operator (4.3) is reminiscent of Povzner's collision operator in the kinetic theory of gases.[4,38] Then, our "collisional" kinetic CS model is the one combining of (4.2) and (4.3):

$$
\partial_t f + \mathbf{v}\cdot\nabla_{\mathbf{x}} f + \kappa \nabla_{\mathbf{v}} \cdot (\mathcal{F}_a(f) f) = \varepsilon Q(f,f). \tag{4.4}
$$

Next, we study the conservation law and dissipation related to Eq. (4.4) in the following lemma. In general, due to the nature of non-symmetry and non-mechanical interactions of the collision operator $Q(f,f)$, we only have a conservation of total mass. The total momentum and energy will not be conserved along the flow (4.4). For $t \geq 0$, we set the first three velocity moments of f as follows:

$$
\begin{aligned}
M_0(t) &:= \int_{\Omega[f]\times D_{\mathbf{v}}} f(t,\mathbf{x},\mathbf{v}) d\mathbf{v}\, d\mathbf{x}, \quad \text{total mass}, \\
M_1(t) &:= \int_{\Omega[f]\times D_{\mathbf{v}}} \mathbf{v} f(t,\mathbf{x},\mathbf{v}) d\mathbf{v}\, d\mathbf{x}, \quad \text{total momentum}, \\
M_2(t) &:= \int_{\Omega[f]\times D_{\mathbf{v}}} |\mathbf{v}|^2 f(t,\mathbf{x},\mathbf{v}) d\mathbf{v}\, d\mathbf{x}, \quad \text{total energy}.
\end{aligned} \tag{4.5}
$$

Lemma 4.2. (Conservation of mass) *Let $f = f(t,\mathbf{x},\mathbf{v})$ be a $\mathcal{C}^1$-solution to* (4.4) *decaying to zero sufficiently fast in phase space. Then, the total mass is conserved along the flow*:

$$
\frac{dM_0(t)}{dt} = 0, \quad t > 0.
$$

Proof. First, we note that

$$
\begin{aligned}
&\int_{\Omega[f]\times D_{\mathbf{v}}} Q(f,f)(t,\mathbf{x},\mathbf{v}) d\mathbf{v}\, d\mathbf{x} \\
&\quad = \int_{\Omega[f]^2\times (D_{\mathbf{v}})^3} \mathcal{A}[f](\mathbf{v}_* \to \mathbf{v} \,|\, \mathbf{v}_* \mathbf{v}^*) f(t,\mathbf{x},\mathbf{v}_*) f(t,\mathbf{x}^*,\mathbf{v}^*) d\mathbf{x}^* d\mathbf{v}_*\, d\mathbf{v}^*\, d\mathbf{v}\, d\mathbf{x} \\
&\qquad - \int_{\Omega[f]^2\times (D_{\mathbf{v}})^2} f(t,\mathbf{x},\mathbf{v}) f(t,\mathbf{x}^*,\mathbf{v}^*) d\mathbf{x}^* d\mathbf{v}^*\, d\mathbf{v}\, d\mathbf{x}
\end{aligned}
$$

$$= \int_{\Omega[f]^2\times(D_{\mathbf{v}})^2} \left(\int_{D_{\mathbf{v}}} \mathcal{A}[f](\mathbf{v}_* \to \mathbf{v} \,|\, \mathbf{v}_*, \mathbf{v}^*) d\mathbf{v} \right)$$

$$\times f(t,\mathbf{x},\mathbf{v}_*) f(t,\mathbf{x}^*,\mathbf{v}^*) d\mathbf{x}^* d\mathbf{v}_*\, d\mathbf{v}^*\, d\mathbf{x}$$

$$- \int_{\Omega[f]^2\times(D_{\mathbf{v}})^2} f(t,\mathbf{x},\mathbf{v}) f(t,\mathbf{x}^*,\mathbf{v}^*) d\mathbf{x}^* d\mathbf{v}^*\, d\mathbf{v}\, d\mathbf{x} = 0. \tag{4.6}$$

Then, the conservation of mass follows from the direct integration of Eq. (4.4) using the relation (4.6). □

Throughout the paper, we will assume that the total mass is unity, i.e.

$$\int_{\Omega[f]\times D_{\mathbf{v}}} f(t,\mathbf{x},\mathbf{v}) d\mathbf{v}\, d\mathbf{x} = 1. \tag{4.7}$$

In the subsequent two sections, we will study *a priori* flocking estimates, propagation of velocity moments $M_i(t),\ i = 1,2$ under the restricted structural condition of the turning probability density $\mathcal{A}[f](\mathbf{v}_* \to \mathbf{v})$.

5. *A Priori* Estimates for Flocking and Regularity Propagation

In this section, we present several *a priori* estimates for (4.4) under very restricted conditions on the turning probability $\mathcal{A}$ appearing in the gain operator $Q^+(f,f)$. More precisely, we are interested in the regime where collisional-type effects registered by $Q(f,f)$ are negligible compared to the flocking term $\nabla_{\mathbf{v}} \cdot (\mathcal{F}_a(f)f)$, i.e. weakly collisional regime. From now on, as long as there is no confusion, we suppress t-dependence in f and other quantities involving with f, i.e.

$$f(\mathbf{x},\mathbf{v}) := f(t,\mathbf{x},\mathbf{v}).$$

Below, we will state several structural assumptions on the communication weight ψ and $\mathcal{A}$ and then, we will derive a coupled differential system for velocity moments M_1, M_2 and the flocking functional $\mathcal{L}$ to be defined later. As noted in (4.4), we need to impose some conditions on the kernel functions ψ and $\mathcal{A}[f]$ to obtain a flocking estimate and existence theory for (4.4).

5.1. *Assumptions on the interaction domain and kernel functions*

In this subsection, for the analytical treatment, we list simplifying assumptions on domains $\Omega[f], D_{\mathbf{v}}$, the communication weight function ψ and $\mathcal{A}[f](\mathbf{v}_* \to \mathbf{v} \,|\, \mathbf{v}_*, \mathbf{v}^*)$ appearing in $\mathcal{F}_a(f)$ and $Q^+(f,f)$, respectively as follows:

($\mathcal{H}_1$) The interaction domain $\Omega[f]$ and admissible velocity set $D_{\mathbf{v}}$ are assumed to be the whole space

$$\Omega[f] = \mathbb{R}^d, \quad D_{\mathbf{v}} = \mathbb{R}^d.$$

($\mathcal{H}_2$) The communication weight function ψ is assumed to be constant, which denotes the all-to-all coupling:

$$\psi(s) = 1, \quad s \geq 0. \tag{5.1}$$

($\mathcal{H}_3$) The turning probability $\mathcal{A}[f](\mathbf{v}_* \to \mathbf{v}\,|\,\mathbf{v}_*, \mathbf{v}^*)$ is independent of f, and is localized to the neighborhood of the set $\{\mathbf{v} = \mathbf{v}_*\}$ in the sense that there exists a small positive constant $\eta \ll 1$ satisfying:

$$\begin{aligned}
&\text{(i)}\ \ \mathcal{A}[f](\mathbf{v}_* \to \mathbf{v}) = \mathcal{A}(\mathbf{v}_* \to \mathbf{v}), \quad \int_{D_{\mathbf{v}}} \mathcal{A}(\mathbf{v}_* \to \mathbf{v}\,|\,\mathbf{v}_*, \mathbf{v}^*) d\mathbf{v} = 1,\\
&\text{(ii)}\ \ \left| \int_{D_{\mathbf{v}}} \mathbf{v}\mathcal{A}(\mathbf{v}_* \to \mathbf{v}\,|\,\mathbf{v}_*, \mathbf{v}^*) d\mathbf{v} - \mathbf{v}_* \right| \leq \eta |\mathbf{v}_*|,\\
&\text{(iii)}\ \ \left| \int_{D_{\mathbf{v}}} |\mathbf{v}|^2 \mathcal{A}(\mathbf{v}_* \to \mathbf{v}\,|\,\mathbf{v}_*, \mathbf{v}^*) d\mathbf{v} - |\mathbf{v}_*|^2 \right| \leq \eta |\mathbf{v}_*|^2,\\
&\text{(iv)}\ \ \|\mathcal{A}(\mathbf{v}_* \to \cdot)\|_{C^k(\mathbb{R}^d)} < \infty, \quad k \geq 1,\\
&\text{(v)}\ \ \left| \int_{D_{\mathbf{v}_*}} \partial_{\mathbf{v}}^{\beta} \mathcal{A}(\mathbf{v}_* \to \mathbf{v}) g(\mathbf{v}_*) d\mathbf{v}_* - g(\mathbf{v}) \right| \leq \eta g(\mathbf{v}),\\
&\quad\ \forall\, \mathbf{v} \in D_{\mathbf{v}}, \quad g \in \mathcal{C}_c^1(D_{\mathbf{v}}), \quad |\beta| \leq k.
\end{aligned} \tag{5.2}$$

Remark 5.1. (1) For the flocking estimate for the CS model, serval communication weights were employed in literature, e.g. algebraically decaying communication weights ψ_{cs}:

$$\psi_{\mathrm{cs}}(s) = \frac{1}{(1 + s^2)^{\frac{\beta}{2}}}, \quad \beta \geq 0,$$

has been used in Refs. 4, 17, 20, 22, 31, 39, and more general Lipshitz continuous and nonincreasing weight function was also used in Ref. 30. Thus, constant communication weight (5.1) corresponds to $\beta = 0$ in the CS communication weight.

(2) As a possible candidate for the kernel $\mathcal{A}$, we might take

$$\mathcal{A}[f](\mathbf{v}_* \to \mathbf{v}\,|\,\mathbf{v}_*, \mathbf{v}^*) = \frac{1}{(\pi\theta)^{\frac{d}{2}}} e^{-\frac{|\mathbf{v}-\mathbf{v}_*|^2}{\theta}}, \quad 0 < \theta \ll 1.$$

Under the assumption (5.2), the gain term $Q^+(f,f)$ satisfies

$$\begin{aligned}
Q^+(f,f)(\mathbf{x},\mathbf{v}) &= \int_{\Omega[f]} \int_{(D_{\mathbf{v}})^2} \mathcal{A}[f](\mathbf{v}_* \to \mathbf{v}\,|\,\mathbf{v}_*, \mathbf{v}^*) f(\mathbf{x}, \mathbf{v}_*) f(\mathbf{x}^*, \mathbf{v}^*) d\mathbf{x}^* d\mathbf{v}_*\, d\mathbf{v}^*\\
&\approx f(\mathbf{x},\mathbf{v}) \int_{\Omega[f] \times D_{\mathbf{v}}} f(\mathbf{x}^*, \mathbf{v}^*) d\mathbf{x}^* d\mathbf{v}^*\\
&= (fL(f))(\mathbf{x},\mathbf{v}).
\end{aligned} \tag{5.3}$$

Thus, the assumption (5.2) makes the effect of $Q(f,f)$ negligible pointwise so that (1.1) can be effectively approximated by the collisionless flow (4.2).

(3) Technical alternatives can be delivered by a variety of consensus models widely used in literature of active particles such as those delivered by modeling of human crowds.[8,10]

5.2. A priori *flocking estimate*

In this subsection, we will provide asymptotic flocking estimate by deriving a coupled system of differential inequalities under the assumptions $(\mathcal{H}_1)$–$(\mathcal{H}_3)$ and unit mass condition (4.7). Note that under the assumptions $(\mathcal{H}_1)$ and $(\mathcal{H}_2)$, flocking force $\mathcal{F}_a(f)$ becomes

$$\mathcal{F}_a(f)(t, \mathbf{x}, \mathbf{v}) = -\mathbf{v} + M_1(t), \quad L(f) = \int_{\mathbb{R}^{2d}} f(t, \mathbf{x}^*, \mathbf{v}^*) d\mathbf{v}^* d\mathbf{x}^* = 1,$$

while Eq. (1.1) becomes

$$\begin{aligned} &\partial_t f + \nabla_{\mathbf{x}} \cdot (\mathbf{v} f) + \kappa \nabla_{\mathbf{v}} \cdot ((-\mathbf{v} + M_1) f) = \varepsilon\, Q^+(f, f) - \varepsilon f, \\ &\quad \text{or equivalently} \\ &\partial_t f + \mathbf{v} \cdot \nabla_{\mathbf{x}} f + \kappa(-\mathbf{v} + M_1(t)) \cdot \nabla_{\mathbf{v}} f = (\kappa d - \varepsilon) f + \varepsilon Q^+(f, f). \end{aligned} \tag{5.4}$$

For any $\mathcal{C}^1$-solution f with a compact support, we introduce a Lyapunov functional $\mathcal{L}$ measuring the degree of velocity flocking:

$$\mathcal{L}(t) := \int_{\mathbb{R}^{4d}} |\mathbf{v}_* - \mathbf{v}^*|^2 f(\mathbf{x}, \mathbf{v}_*) f(\mathbf{x}^*, \mathbf{v}^*) d\mathbf{v}^* d\mathbf{v}\, d\mathbf{x}^* d\mathbf{x}. \tag{5.5}$$

Then, it is easy to see that the Lyapunov functional (5.5) can be rewritten in terms of velocity moments:

$$\mathcal{L}(t) = 2(M_2(t)\|f_0\|_{L^1} - |M_1(t)|^2) = 2(M_2(t) - |M_1(t)|^2). \tag{5.6}$$

When the total momentum M_1 is zero, Lyapunov functional $\mathcal{L}$ is a constant multiple of the total energy M_2 and is the same as the functional employed in Ref. 31. Unfortunately, under the assumption $(\mathcal{H}_3)$, we cannot guarantee the constancy of M_1 and dissipation of M_2. In the sequel, we will derive a coupled system of differential inequalities:

$$\begin{cases} \left|\dfrac{dM_1}{dt}\right| \le \varepsilon\eta\sqrt{M_2}, \quad \left|\dfrac{dM_2}{dt} + \kappa\mathcal{L}\right| \le \varepsilon\eta M_2, \quad t > 0, \\ -2(\kappa + \varepsilon\eta)\mathcal{L} \le \dfrac{d\mathcal{L}}{dt} \le -2(\kappa - \varepsilon\eta)\mathcal{L}. \end{cases} \tag{5.7}$$

Below, we will present the derivation of (5.7) one-by-one.

5.2.1. *Derivation of* $(5.7)_1$

For the estimate of time-rate change of M_1, we multiply $\mathbf{v}$ to (5.4) to find

$$\begin{aligned} &\partial_t(\mathbf{v} f) + \nabla_{\mathbf{x}} \cdot (\mathbf{v} \otimes \mathbf{v} f) + \kappa \nabla_{\mathbf{v}} \cdot (\mathbf{v} \otimes (-\mathbf{v} + M_1) f) \\ &\quad = \kappa d(-\mathbf{v} + M_1) f + \varepsilon \mathbf{v} Q^+(f, f) - \varepsilon \mathbf{v} f. \end{aligned}$$

We integrate the above relation in phase space $\mathbb{R}^{2d}$ to obtain

$$\begin{aligned}\frac{d}{dt}\int_{\mathbb{R}^{2d}} \mathbf{v}f\,d\mathbf{v}\,d\mathbf{x} &= \kappa d\int_{\mathbb{R}^{2d}}(-\mathbf{v}+M_1)f\,d\mathbf{v}\,d\mathbf{x} + \varepsilon\int_{\mathbb{R}^{2d}} \mathbf{v}Q^+(f,f)d\mathbf{v}\,d\mathbf{x}\\ &\quad -\varepsilon\int_{\mathbb{R}^{2d}} \mathbf{v}f\,d\mathbf{v}\,d\mathbf{x} =: \mathcal{I}_{11}+\mathcal{I}_{12}-\varepsilon M_1.\end{aligned}$$

Below, we estimate the terms $\mathcal{I}_{1i}$ separately.

Case A.1. (Estimate of $\mathcal{I}_{11}$) We use the unit mass condition (4.7) to obtain

$$\mathcal{I}_{11} = \kappa d\int_{\mathbb{R}^{2d}}(-\mathbf{v}+M_1)f\,d\mathbf{v}\,d\mathbf{x} = 0.$$

Case A.2. (Estimate of $\mathcal{I}_{12}$) Again, the unit mass condition (4.7) implies

$$\begin{aligned}\mathcal{I}_{12}-\varepsilon M_1 &= \varepsilon\int_{\mathbb{R}^{2d}} \mathbf{v}Q^+(f,f)d\mathbf{v}\,d\mathbf{x} - \varepsilon M_1\\ &= \varepsilon\int_{\mathbb{R}^{5d}} \mathbf{v}\mathcal{A}(\mathbf{v}_*\to\mathbf{v}\,|\,\mathbf{v}_*,\mathbf{v}^*)f(\mathbf{x},\mathbf{v}_*)f(\mathbf{x}^*,\mathbf{v}^*)d\mathbf{x}^*d\mathbf{v}_*\,d\mathbf{v}^*d\mathbf{v}\,d\mathbf{x} - M_1\\ &= \varepsilon\int_{\mathbb{R}^{4d}}\left(\int_{\mathbb{R}^d}\mathbf{v}\mathcal{A}(\mathbf{v}_*\to\mathbf{v}\,|\,\mathbf{v}_*,\mathbf{v}^*)d\mathbf{v}-\mathbf{v}_*\right)\\ &\qquad\times f(\mathbf{x},\mathbf{v}_*)f(\mathbf{x}^*,\mathbf{v}^*)d\mathbf{x}^*d\mathbf{v}_*d\mathbf{v}^*d\mathbf{x}.\end{aligned}\tag{5.8}$$

The relations (4.7), (5.11) and Cauchy–Schwarz inequality yield

$$\begin{aligned}|\mathcal{I}_{12}-\varepsilon M_1| &\le \varepsilon\int_{\mathbb{R}^{4d}}\left|\int_{\mathbb{R}^d}\mathbf{v}\mathcal{A}(\mathbf{v}_*\to\mathbf{v}\,|\,\mathbf{v}_*,\mathbf{v}^*)d\mathbf{v}-\mathbf{v}_*\right|\\ &\qquad\times f(\mathbf{x},\mathbf{v}_*)f(\mathbf{x}^*,\mathbf{v}^*)d\mathbf{x}^*d\mathbf{v}_*\,d\mathbf{v}^*d\mathbf{x}\\ &\le \varepsilon\eta\int_{\mathbb{R}^{4d}}|\mathbf{v}_*|f(\mathbf{x},\mathbf{v}_*)f(\mathbf{x}^*,\mathbf{v}^*)d\mathbf{x}^*d\mathbf{v}_*\,d\mathbf{v}^*d\mathbf{v}\,d\mathbf{x}\\ &\le \varepsilon\eta\sqrt{M_2}.\end{aligned}$$

Finally, as we shall see in (5.10), we can combine all estimates in Cases A.1 and A.2 to find a desired Gronwall's inequality:

$$\left|\frac{dM_1(t)}{dt}\right| \le \varepsilon\eta\sqrt{M_2(t)}.$$

5.2.2. *Derivation of* $(5.7)_2$

For the time-rate of change of M_2, we multiply $|\mathbf{v}|^2$ to (5.4) to obtain

$$\begin{aligned}&\partial_t(|\mathbf{v}|^2f)+\nabla_{\mathbf{x}}(|\mathbf{v}|^2\mathbf{v}f)+\kappa\nabla_{\mathbf{v}}\cdot(|\mathbf{v}|^2(-\mathbf{v}+M_1)f)\\ &\quad = 2\kappa\mathbf{v}\cdot(-\mathbf{v}+M_1)f+\varepsilon|\mathbf{v}|^2Q^+(f,f)-\varepsilon|\mathbf{v}|^2f.\end{aligned}$$

We integrate the above relation with respect to $(\mathbf{x}, \mathbf{v})$ to obtain

$$\begin{aligned}\frac{dM_2}{dt} &= 2\kappa \int_{\mathbb{R}^{2d}} \mathbf{v} \cdot (-\mathbf{v} + M_1) f + \varepsilon \int_{\mathbb{R}^{2d}} |\mathbf{v}|^2 Q^+(f, f) - \varepsilon \int_{\mathbb{R}^{2d}} |\mathbf{v}|^2 f \\ &=: \mathcal{I}_{21} + \mathcal{I}_{22} - \varepsilon M_2. \end{aligned} \tag{5.9}$$

Case B.1. (Estimate of $\mathcal{I}_{21}$) We use the relation (5.6) to obtain

$$\mathcal{I}_{21} = 2\kappa(-M_2 + |M_1|^2) = -\kappa\mathcal{L}.$$

Case B.2. (Estimate of $\mathcal{I}_{22}$) Similar to the estimate in Case A.2, we have

$$\begin{aligned}\mathcal{I}_{22} - \varepsilon M_2 = \varepsilon \int_{\mathbb{R}^{4d}} &\left(\int_{\mathbb{R}^d} |\mathbf{v}|^2 \mathcal{A}(\mathbf{v}_* \to \mathbf{v} \,|\, \mathbf{v}_*, \mathbf{v}^*) d\mathbf{v} - |\mathbf{v}_*|^2 \right) f(\mathbf{x}, \mathbf{v}_*) \\ &\times f(\mathbf{x}^*, \mathbf{v}^*) d\mathbf{x}^* d\mathbf{v}_* \, d\mathbf{v}^* d\mathbf{x}.\end{aligned}$$

This yields

$$|\mathcal{I}_{22} - \varepsilon M_2| \le \varepsilon \eta M_2.$$

In (5.9), we combine all estimates in Cases B.1 and B.2 to obtain the desired estimate.

5.2.3. *Derivation of* $(5.7)_3$

In this subsection, we provide a decay estimate of $\mathcal{L}$. Before we study the variation of $\mathcal{L}$, we study the dynamics of the two-particle probability density function f_2:

$$f_2(\mathbf{x}, \mathbf{x}^*, \mathbf{v}, \mathbf{v}^*) := f(\mathbf{x}, \mathbf{v}) f(\mathbf{x}^*, \mathbf{v}^*).$$

Lemma 5.1. *Let $f = f(\mathbf{x}, \mathbf{v})$ be a $\mathcal{C}^1$-solution to* (5.4). *Then, f_2 satisfies*

$$\begin{aligned}&\partial_t f_2 + \nabla_{(\mathbf{x},\mathbf{x}^*)} \cdot ((\mathbf{v}, \mathbf{v}^*) f_2) + \kappa \nabla_{(\mathbf{v},\mathbf{v}^*)} \cdot [(-\mathbf{v} + M_1, -\mathbf{v}^* + M_1) f_2] \\ &\quad = \varepsilon [Q^+(f, f)(\mathbf{x}, \mathbf{v}) f(\mathbf{x}^*, \mathbf{v}^*) + Q^+(f, f)(\mathbf{x}^*, \mathbf{v}^*) f(\mathbf{x}, \mathbf{v})] - 2\varepsilon f_2.\end{aligned}$$

Proof. By direct calculation, we have

$$\begin{aligned}\partial_t f_2(\mathbf{x}, \mathbf{x}^*, \mathbf{v}, \mathbf{v}^*) &= (\partial_t f(\mathbf{x}, \mathbf{v})) f(\mathbf{x}^*, \mathbf{v}^*) + f(\mathbf{x}, \mathbf{v})(\partial_t f(\mathbf{x}^*, \mathbf{v}^*)) \\ &= [-\mathbf{v} \cdot \nabla_{\mathbf{x}} f(\mathbf{x}, \mathbf{v}) - \kappa \nabla_{\mathbf{v}} \cdot ((-\mathbf{v} + M_1) f(\mathbf{x}, \mathbf{v})) \\ &\quad + \varepsilon Q^+(f, f)(\mathbf{x}, \mathbf{v}) - \varepsilon f(\mathbf{x}, \mathbf{v})] f(\mathbf{x}^*, \mathbf{v}^*) \\ &\quad + [-\mathbf{v}^* \cdot \nabla_{\mathbf{x}^*} f(\mathbf{x}^*, \mathbf{v}^*) - \kappa \nabla_{\mathbf{v}^*} \cdot ((-\mathbf{v}^* + M_1) f(\mathbf{x}^*, \mathbf{v}^*)) \\ &\quad + \varepsilon Q^+(f, f)(\mathbf{x}^*, \mathbf{v}^*) - \varepsilon f(\mathbf{x}^*, \mathbf{v}^*)] f(\mathbf{x}, \mathbf{v})\end{aligned}$$

$$
\begin{aligned}
&= -\nabla_{(\mathbf{x},\mathbf{x}^*)} \cdot ((\mathbf{v},\mathbf{v}^*) f_2(\mathbf{x},\mathbf{x}^*,\mathbf{v},\mathbf{v}^*)) \\
&\quad - \kappa \nabla_{(\mathbf{v},\mathbf{v}^*)} \cdot [((-\mathbf{v}+M_1),(-\mathbf{v}^*+M_1)) f_2(\mathbf{x},\mathbf{x}^*,\mathbf{v},\mathbf{v}^*)] \\
&\quad + \varepsilon[Q^+(f,f)(t,\mathbf{x},\mathbf{v}) f(\mathbf{x}^*,\mathbf{v}^*) + Q^+(f,f)(\mathbf{x}^*,\mathbf{v}^*) f(\mathbf{x},\mathbf{v})] \\
&\quad - 2\varepsilon f_2(\mathbf{x},\mathbf{x}^*,\mathbf{v},\mathbf{v}^*)].
\end{aligned}
$$

This yields the desired estimate. □

We now set

$$
L(\mathbf{x},\mathbf{x}^*,\mathbf{v},\mathbf{v}^*) := |\mathbf{v}-\mathbf{v}^*|^2 f_2(\mathbf{x},\mathbf{x}^*,\mathbf{v},\mathbf{v}^*).
$$

Then, we use Lemma 5.1 to find

$$
\begin{aligned}
\partial_t L &= -\nabla_{(\mathbf{x},\mathbf{y})} \cdot ((\mathbf{v}_*,\mathbf{v}^*)|\mathbf{v}_*-\mathbf{v}^*|^2 f_2) \\
&\quad - \kappa \nabla_{(\mathbf{v}_*,\mathbf{v}^*)} \cdot [|\mathbf{v}_*-\mathbf{v}^*|^2((-\mathbf{v}_*+M_1),(-\mathbf{v}^*+M_1)) f_2] \\
&\quad - 2\kappa|\mathbf{v}_*-\mathbf{v}^*|^2 f_2 + \varepsilon[|\mathbf{v}_*-\mathbf{v}^*|^2 Q^+(f,f)(\mathbf{x},\mathbf{v}_*) f(\mathbf{x}^*,\mathbf{v}^*) \\
&\quad + |\mathbf{v}_*-\mathbf{v}^*|^2 Q^+(f,f)(\mathbf{x}^*,\mathbf{v}^*) f(\mathbf{x},\mathbf{v}_*)] \\
&\quad - 2\varepsilon[|\mathbf{v}_*-\mathbf{v}^*|^2 f_2].
\end{aligned}
$$

We integrate the above relation to find the following estimate:

$$
\begin{aligned}
\frac{d\mathcal{L}}{dt} &= \int_{\mathbb{R}^{4d}} \partial_t L \, d\mathbf{v}_* \, d\mathbf{v}^* d\mathbf{x}^* d\mathbf{x} \\
&= -2\kappa\mathcal{L} + \varepsilon \int_{\mathbb{R}^{4d}} [|\mathbf{v}_*-\mathbf{v}^*|^2 Q^+(f,f)(\mathbf{x},\mathbf{v}_*) f(\mathbf{x}^*,\mathbf{v}^*) \\
&\quad + |\mathbf{v}_*-\mathbf{v}^*|^2 Q^+(f,f)(\mathbf{x}^*,\mathbf{v}^*) f(\mathbf{x},\mathbf{v}_*)] d\mathbf{v}_* \, d\mathbf{v}^* d\mathbf{x}^* d\mathbf{x} \\
&\quad - 2\varepsilon \int_{\mathbb{R}^{4d}} |\mathbf{v}_*-\mathbf{v}^*|^2 f_2 \, d\mathbf{v}_* \, d\mathbf{v}^* d\mathbf{x}^* d\mathbf{x} \\
&=: -2(\kappa+\varepsilon)\mathcal{L} + \mathcal{I}_3.
\end{aligned} \tag{5.10}
$$

Below, we estimate the terms $\mathcal{I}_3$ as follows. By definition of $Q^+(f,f)$ and the property of $\mathcal{A}$ in (v) of (5.2), we have

$$
\begin{aligned}
\mathcal{I}_3 &\le 2\varepsilon(1+\eta) \int_{\mathbb{R}^{4d}} |\mathbf{v}_*-\mathbf{v}^*|^2 f(\mathbf{x},\mathbf{v}_*) f(\mathbf{x}^*,\mathbf{v}^*) d\mathbf{v}_* \, d\mathbf{v}^* d\mathbf{x}^* d\mathbf{x} \\
&= 2\varepsilon(1+\eta)\mathcal{L}.
\end{aligned} \tag{5.11}
$$

Finally, we combine (5.10) and (5.11) to obtain desired estimate. We are ready to provide a decay estimate for $\mathcal{L}$ and growth estimates for M_1 and M_2.

Proposition 5.1. (Flocking estimate) *Suppose that the coupling strength κ is sufficiently large so that $\kappa - \varepsilon\eta > 0$. Then, for any compactly supported $\mathcal{C}^1$-solution f to* (5.4), *we have*:

(i) $\mathcal{L}(0)e^{-2(\kappa+\varepsilon\eta)t} \le \mathcal{L}(t) \le \mathcal{L}(0)e^{-2(\kappa-\varepsilon\eta)t}, \quad t \ge 0.$

(ii) $(M_2(0) - \frac{\kappa\mathcal{L}(0)}{2\kappa-3\varepsilon\eta})e^{-\varepsilon\eta t} + \frac{\kappa\mathcal{L}(0)}{2\kappa-3\varepsilon\eta}e^{-2(K-\varepsilon\eta)t} \le M_2(t) \le (M_2(0) - \frac{\kappa\mathcal{L}(0)}{2\kappa+3\varepsilon\eta})e^{\varepsilon\eta t} + \frac{\kappa\mathcal{L}(0)}{2\kappa+3\varepsilon\eta}e^{-2(\kappa+\varepsilon\eta)t}$.

(iii) $|M_1(t)| \le |M_1(0)| + \frac{\varepsilon\eta}{\kappa+\varepsilon\eta}\sqrt{\frac{\kappa\mathcal{L}(0)}{2\kappa+3\varepsilon\eta}} + 2\sqrt{|M_2(0) - \frac{\kappa\mathcal{L}(0)}{2\kappa+3\varepsilon\eta}|}e^{\frac{\varepsilon\eta t}{2}}$.

Proof. (i) The first estimate directly follows from Gronwall's inequality for $\mathcal{L}$ in (5.7).

(ii) Note that Gronwall's inequality for M_2 in (5.7) and estimate (i) imply

$$-\kappa\mathcal{L}(0)e^{-2(\kappa-\varepsilon\eta)t} - \varepsilon\eta M_2(t) \le \frac{dM_2}{dt} \le -\kappa\mathcal{L}(0)e^{-2(\kappa+\varepsilon\eta)t} + \varepsilon\eta M_2.$$

This yields the desired estimate.

(iii) Note that M_1 satisfies

$$\begin{aligned}\left|\frac{dM_1(t)}{dt}\right| &\le \varepsilon\eta\sqrt{M_2(t)}\\ &\le \varepsilon\eta\left(\sqrt{\left|M_2(0) - \frac{\kappa\mathcal{L}(0)}{2\kappa+3\varepsilon\eta}\right|}e^{\frac{\varepsilon\eta t}{2}} + \sqrt{\frac{\kappa\mathcal{L}(0)}{2\kappa+3\varepsilon\eta}}e^{-(\kappa+\varepsilon\eta)t}\right), \end{aligned}\tag{5.12}$$

where we used $\sqrt{a+b} \le \sqrt{a} + \sqrt{b}, a, b \ge 0$. Then, we use (5.12) to obtain

$$\begin{aligned}&|M_1(t) - M_1(0)|\\ &\quad\le \int_0^t \left|\frac{dM_1(s)}{ds}\right| ds\\ &\quad\le 2\sqrt{\left|M_2(0) - \frac{\kappa\mathcal{L}(0)}{2\kappa+3\varepsilon\eta}\right|}(e^{\frac{\varepsilon\eta t}{2}} - 1) + \frac{\varepsilon\eta}{\kappa+\varepsilon\eta}\sqrt{\frac{\kappa\mathcal{L}(0)}{2\kappa+3\varepsilon\eta}}(1 - e^{-(\kappa+\varepsilon\eta)t})\\ &\quad\le 2\sqrt{\left|M_2(0) - \frac{\kappa\mathcal{L}(0)}{2\kappa+3\varepsilon\eta}\right|}e^{\frac{\varepsilon\eta t}{2}} + \frac{\varepsilon\eta}{\kappa+\varepsilon\eta}\sqrt{\frac{\kappa\mathcal{L}(0)}{2\kappa+3\varepsilon\eta}}.\end{aligned}$$

This yields the desired growth estimate for M_1. □

Remark 5.2. For a large coupling strength $\kappa \gg 1$,

$$M_2(0) - \frac{\kappa\mathcal{L}(0)}{2\kappa+3\varepsilon\eta} \approx M_2(0) - \frac{\mathcal{L}(0)}{2} > 0,$$

where the last inequality follows from the relation:

$$\mathcal{L}(0) = 2(M_2(0) - |M_1(0)|^2) \le 2M_2(0), \quad \text{i.e. } M_2(0) > \frac{\mathcal{L}(0)}{2}.$$

Hence, M_2 grows at most exponentially fast as $t \to \infty$:

$$M_2(t) \lesssim e^{\varepsilon\eta t}.$$

5.3. *Propagation of $W^{k,\infty}$-bound*

In this subsection, we discuss the propagation of regularity along the flow (5.4), which is sufficient for a global existence theory of smooth solutions under the suitable regularity assumptions $\mathcal{H}$ in Sec. 5.1.

For a given $(\mathbf{x}_0, \mathbf{v}_0) \in \mathbb{R}^d \times \mathbb{R}^d$, we set the forward characteristics (or particle trajectory) $[\mathbf{x}(t), \mathbf{v}(t)] \equiv [\mathbf{x}(t; 0, \mathbf{x}_0, \mathbf{v}_0), \mathbf{v}(t; 0, \mathbf{x}_0, \mathbf{v}_0)]$ by the unique solution of the characteristic system:

$$\begin{aligned} &\frac{d\mathbf{x}(t)}{dt} = \mathbf{v}(t), \quad \frac{d\mathbf{v}(t)}{dt} = \kappa(-\mathbf{v}(t) + M_1(t)), \quad t > 0, \\ &(\mathbf{x}(0), \mathbf{v}(0)) = (\mathbf{x}_0, \mathbf{v}_0). \end{aligned} \tag{5.13}$$

We next study the variation of the velocity trajectory $\mathbf{v}(t)$ in the following lemma.

Lemma 5.2. *Let* $(\mathbf{x}(t), \mathbf{v}(t))$ *be a forward characteristics* (5.13) *issued from* $(\mathbf{x}_0, \mathbf{v}_0) \in \operatorname{Supp} f_0$ *at time* $t = 0$. *Then, the velocity trajectory* $\mathbf{v} = (v^1, \ldots, v^d)$ *satisfies*

$$|v^i(t)| \le C_1 e^{\frac{\varepsilon\eta t}{2}}, \quad |x^i(t)| \le \tilde{C}_1 e^{\frac{\varepsilon\eta t}{2}}, \quad t \ge 0,$$

where $C_1 = C_1(\kappa, \max_i |\mathbf{v}^i(0)|, \mathcal{L}(0), \varepsilon\eta, M_1(0), M_2(0))$ *and* $\tilde{C}_1$ *are positive constants independent of* t.

Proof. (i) Note that the ith component of $\mathbf{v}(t)$ satisfies

$$\begin{aligned} \frac{dv^i(t)}{dt} &= -\kappa v^i(t) + \kappa M_1^i(t) \\ &\le -\kappa v^i(t) + \kappa |M_1(t)| \\ &\le -\kappa v^i(t) + \kappa C_0 \left(1 + e^{\frac{\varepsilon\eta t}{2}}\right), \end{aligned} \tag{5.14}$$

where $C_0 = C_0(\kappa, \mathcal{L}(0), \varepsilon\eta, M_1(0), M_2(0))$ is a positive constant independent of t and we used (iii) of Proposition 5.1. Then, the relation (5.14) yields

$$\max_{1 \le i \le d} |v^i(t)| \le C_1 \left(1 + e^{\frac{\varepsilon\eta t}{2}}\right),$$

where $C_1 = C_1(\kappa, \max_i |v^i(0)|, \mathcal{L}(0), \varepsilon\eta, M_1(0), M_2(0))$ is a positive constant independent of t.

(ii) Since $\frac{dx^i(t)}{dt} = v^i(t)$, we have

$$|x^i(t) - x^i(0)| \le \int_0^t |v^i(s)| ds \le C_1 \left[t + \frac{2}{\varepsilon\eta}\left(e^{\frac{\varepsilon\eta t}{2}} - 1\right)\right] \le \tilde{C}_1 \left(t + e^{\frac{\varepsilon\eta t}{2}}\right). \qquad \square$$

Proposition 5.2. *Suppose that the assumptions* $(\mathcal{H}_1)$–$(\mathcal{H}_3)$ *hold, and let* $f = f(\mathbf{x}, \mathbf{v}, t)$ *be a solution to* (5.4). *Then, along the characteristics* $(\mathbf{x}(t), \mathbf{v}(t))$ *issued* *from* $(\mathbf{x}_0, \mathbf{v}_0)$, *we have*

$$f_0(\mathbf{x}_0, \mathbf{v}_0)e^{(\kappa d-\eta)t} \le f(t, \mathbf{x}(t), \mathbf{v}(t)) \le f_0(\mathbf{x}_0, \mathbf{v}_0)e^{(\kappa d+\eta)t}, \quad t \ge 0.$$

Proof. In (5.4), it is easy to see that along the characteristics $(\mathbf{x}(t), \mathbf{v}(t))$, the density function $f(t) = f(t, \mathbf{x}(t), \mathbf{v}(t))$ satisfies

$$\frac{df(t)}{dt} = (\kappa - \varepsilon)f + \varepsilon Q^+(f, f), \quad t > 0. \tag{5.15}$$

On the other hand, since $Q^+(f, f)$ satisfies

$$(1 - \eta)f \le Q^+(f, f) \le (1 + \eta)f. \tag{5.16}$$

Then, we combine (5.15) and (5.16) to obtain

$$(\kappa d - \eta)f \le \frac{df(t)}{dt} \le (\kappa d + \eta)f, \quad t > 0.$$

This yields the desired estimates. $\square$

Remark 5.3. Note that the result in Proposition 5.2 implies

$$\|f(t)\|_{L^\infty} \le \|f_0\|_{L^\infty} e^{(Kd+\eta)t}, \quad t \ge 0.$$

Next, we discuss the propagation of higher-order $W^{k,\infty}$-estimates. For this, we introduce a transport operator $\mathcal{T}$:

$$\mathcal{T} := \partial_t + \mathbf{v} \cdot \nabla_{\mathbf{x}} + K\underbrace{(-\mathbf{v} + M_1(t))}_{=\mathcal{F}_a(f)} \cdot \nabla_{\mathbf{v}}, \quad M_1(t) = \int_{\mathbb{R}^{2d}} \mathbf{v} f \, d\mathbf{v} \, d\mathbf{x}.$$

Then, the relation (5.4) implies

$$\mathcal{T}f = (\kappa d - \varepsilon)f + \varepsilon Q^+(f, f). \tag{5.17}$$

Note that $\mathcal{F}_a(f), L(f)$ and $Q^+(f, f)$ satisfy:

$$\begin{aligned}
&\partial^\alpha_{\mathbf{x}} \mathcal{F}_a(f) = 0, \quad |\alpha| \ge 1, \quad \partial^\beta_{\mathbf{v}} \mathcal{F}_a(f) = 0, \quad |\beta| \ge 2,\\
&\partial^\alpha_{\mathbf{x}} L(f) = 0, \quad |\alpha| \ge 1, \quad \partial^\beta_{\mathbf{v}} L(f) = 0, \quad |\beta| \ge 1,\\
&\partial^\alpha_{\mathbf{x}} Q^+(f, f) = \int_{\mathbb{R}^{3d}} \mathcal{A}(\mathbf{v}_* \to \mathbf{v})(\partial^\alpha_{\mathbf{x}} f(\mathbf{x}, \mathbf{v}_*))f(\mathbf{x}^*, \mathbf{v}^*)d\mathbf{x}^* d\mathbf{v}_* \, d\mathbf{v}^*,\\
&\partial^\alpha_{\mathbf{v}} Q^+(f, f) = \int_{\mathbb{R}^{3d}} (\partial^\alpha_{\mathbf{v}} \mathcal{A}(\mathbf{v}_* \to \mathbf{v}))f(\mathbf{x}, \mathbf{v}_*)f(\mathbf{x}^*, \mathbf{v}^*)d\mathbf{x}^* d\mathbf{v}_* \, d\mathbf{v}^*.
\end{aligned} \tag{5.18}$$

Case C.1. (Estimate on $\partial^\alpha_{\mathbf{x}} f$, $|\alpha| \ge 1$) We apply $\partial^\alpha_{\mathbf{x}}$ to (5.4) and use (5.18) to obtain

$$\begin{aligned}
\mathcal{T}\partial^\alpha_{\mathbf{x}} f &= \partial_t(\partial^\alpha_{\mathbf{x}} f) + \mathbf{v} \cdot \nabla_{\mathbf{x}}(\partial^\alpha_{\mathbf{x}} f) + \kappa \mathcal{F}_a(f) \cdot \nabla_{\mathbf{v}}(\partial^\alpha_{\mathbf{x}} f)\\
&= (\kappa d - \varepsilon)\partial^\alpha_{\mathbf{x}} f + \varepsilon \partial^\alpha_{\mathbf{x}} Q^+(f, f).
\end{aligned} \tag{5.19}$$

We use (5.18) and (5.19) to obtain

$$\mathcal{T}|\partial_{\mathbf{x}}^{\alpha} f| \leq (\kappa d + \varepsilon\eta)|\partial_{\mathbf{x}}^{\alpha} f|. \tag{5.20}$$

Case C.2. (Estimate on $\partial_{\mathbf{v}}^{\beta} f,\ |\beta| \geq 1$) We apply $\partial_{\mathbf{v}}^{\beta}$ derivative to (5.4) and use (5.18) to obtain

$$\begin{aligned}
&\partial_t(\partial_{\mathbf{v}}^{\beta} f) + \mathbf{v}\cdot\nabla_{\mathbf{x}}(\partial_{\mathbf{v}}^{\beta} f) + \kappa\mathcal{F}_a(f)\cdot\nabla_{\mathbf{v}}(\partial_{\mathbf{v}}^{\beta} f) \\
&\quad = -\sum_{|\gamma|=1}\binom{\beta}{\gamma}(\partial_{\mathbf{v}}^{\gamma}\mathbf{v})\cdot\nabla_{\mathbf{x}}(\partial_{\mathbf{v}}^{\beta-\gamma} f) - \kappa\sum_{|\gamma|=1}\binom{\beta}{\gamma}\partial_{\mathbf{v}}^{\gamma}(\mathcal{F}_a(f))\cdot\nabla_{\mathbf{v}}(\partial_{\mathbf{v}}^{\beta-\gamma} f) \\
&\qquad + (\kappa d - \varepsilon)\partial_{\mathbf{v}}^{\beta} f + \varepsilon\partial_{\mathbf{v}}^{\beta} Q^{+}(f,f).
\end{aligned} \tag{5.21}$$

The relation (5.21) implies

$$\mathcal{T}|\partial_{\mathbf{v}}^{\beta} f| \lesssim \sum_{|\gamma|=1}|\partial_{\mathbf{x}}(\partial_{\mathbf{v}}^{\beta-\gamma} f)| + (\kappa d - \varepsilon)\sum_{|\beta'|=|\beta|}|\partial_{\mathbf{v}}^{\beta'} f| + \varepsilon(1+\eta)f. \tag{5.22}$$

Case C.3. (*Estimate on* $\partial_{\mathbf{x}}^{\alpha}\partial_{\mathbf{v}}^{\beta} f, |\alpha|, |\beta| \geq 1$) We apply $\partial_{\mathbf{x}}^{\alpha}$ to (5.21) and use (5.18) to obtain

$$\begin{aligned}
&\partial_t(\partial_{\mathbf{x}}^{\alpha}\partial_{\mathbf{v}}^{\beta} f) + \mathbf{v}\cdot\nabla_{\mathbf{x}}(\partial_{\mathbf{x}}^{\alpha}\partial_{\mathbf{v}}^{\beta} f) + \kappa\mathcal{F}_a(f)\cdot\nabla_{\mathbf{v}}(\partial_{\mathbf{x}}^{\alpha}\partial_{\mathbf{v}}^{\beta} f) \\
&\quad = -\sum_{|\gamma|=1}\binom{\beta}{\gamma}(\partial_{\mathbf{v}}^{\gamma}\mathbf{v})\cdot\nabla_{\mathbf{x}}(\partial_{\mathbf{x}}^{\alpha}\partial_{\mathbf{v}}^{\beta-\gamma} f) \\
&\qquad - \kappa\sum_{|\gamma|=1}\binom{\beta}{\gamma}\partial_{\mathbf{v}}^{\gamma}(\mathcal{F}_a(f))\cdot\nabla_{\mathbf{v}}(\partial_{\mathbf{x}}^{\alpha}\partial_{\mathbf{v}}^{\beta-\gamma} f) \\
&\qquad + (\kappa d - \varepsilon)\partial_{\mathbf{x}}^{\alpha}\partial_{\mathbf{v}}^{\beta} f + \varepsilon\partial_{\mathbf{x}}^{\alpha}\partial_{\mathbf{v}}^{\beta} Q^{+}(f,f).
\end{aligned} \tag{5.23}$$

Again, we use (5.23) to obtain

$$\mathcal{T}|\partial_{\mathbf{x}}^{\alpha}\partial_{\mathbf{v}}^{\beta} f| \lesssim (1+\kappa)\left(\sum_{|\gamma_1|+|\gamma_2|=|\alpha|+|\beta|}|\partial_{\mathbf{x}}^{\gamma_1}\partial_{\mathbf{v}}^{\gamma_2} f|\right) + |\kappa d - \varepsilon||\partial_{\mathbf{x}}^{\alpha}\partial_{\mathbf{v}}^{\beta} f| + \varepsilon|\partial_{\mathbf{x}}^{\alpha} f|. \tag{5.24}$$

We set a functional $\mathcal{E}$ which is equivalent to $W^{k,\infty}$-norm of $f(t)$:

$$\mathcal{E}(t) := \sum_{0\leq|\alpha|+|\beta|\leq k}\|\nabla_{\mathbf{x}}^{\alpha}\nabla_{\mathbf{v}}^{\beta} f(t)\|_{L^{\infty}_{\mathbf{x},\mathbf{v}}}.$$

Then, it follows from (5.20), (5.22) and (5.24) that we have

$$\frac{d}{dt}\mathcal{E}(t) \lesssim C\mathcal{E}(t), \quad t > 0.$$

This yields

$$\mathcal{E}(t) \lesssim \mathcal{E}(0)e^{Ct}, \quad t \geq 0. \tag{5.25}$$

Thus, for any $T \in (0, \infty)$, we have

$$\sup_{0 \le t < T} \|f(t)\|_{W^{k,\infty}} < \infty.$$

This *a priori* estimate is sufficient to construct a global existence of a smooth C^1-solution as long as initial data is sufficiently smooth and compactly supported in phase space by the standard continuation argument of local smooth solution.

6. Critical Analysis and Perspectives

In this paper, we presented a "collisional" kinetic model for the dynamic interplay between flocking force and collisions between particles. When the effect of particle–particle collision is neglected, the emergence of mono-cluster (global) flocking has been intensively studied in flocking literature. In fact, several sufficient conditions have been proposed for the particle, kinetic and macroscopic CS models. However, in real modeling of swarming behavior of active particles, the collision effect cannot be ignored any more. To take into account of these interaction effects between active particles, we employed the Boltzmann-like collision operator in the original kinetic CS model so that the resulting kinetic model can describe the dynamic interplay between flocking force and collision mechanism. In a mean-field setting, when the communication weight and turning probability are independent of f, we showed that the asymptotic mono-cluster flocking can emerge in a large coupling strength regime.

Of course, more interesting situations will be the case where the turning probability depends on the density function f itself, internal variables of the active particles and some decision mechanism of intelligent active particles. Thus, our proposed modeling and results provide a first footstep toward the flocking model of active particles such as a human being, animals, etc. This interesting issue for the marriage of flocking theory and game theory appears to be an interesting and challenging problem to be investigated in future.

Modeling issues can look at swarm mixtures with presence of leaders or predators. This presence requires further on the interaction rules that should also account for theoretical tools of learning dynamics.[15,16] Since the mathematical structure proposed in this paper introduces substantial differences from that of the original CS model, the asymptotic limit to derive hydrodynamic equations from the underlying description at the microscopic scale[32] introduces new conceptual difficulties. Mathematical tools used for the asymptotic limits for swarms of cells[9] can possibly be investigated to tackle the aforementioned challenging problem.

In addition, applied mathematics should also look at control problems of swarms. This topic is recently receiving the interest of applied mathematicians as witnessed by various papers.[14,18,43] An additional topic which deserves attention is the development of computational methods. These might take advantage of the knowledge from simulation tools for human and cell crowds.[11,34] The additional difficulty to be accounted for is the nonlinearity of interaction domain depending on the distribution function.

Acknowledgments

The work of N.B. has been partially supported by MIUR, Italian Minister for University and Research, PRIN Project coordinated by M. Pulvirenti, and the work of S.-Y.H. is supported by the National Research Foundation of Korea (NRF2014R1A2A205002096).

References

1. G. Ajmone Marsan, N. Bellomo and L. Gibelli, Towards a systems approach to behavioral social dynamics, *Math. Models Methods Appl. Sci.* **26** (2016) 1051–1093.
2. G. Albi and L. Pareschi, Modeling of self-organized systems interacting with a few individuals: From microscopic to macroscopic dynamics, *Appl. Math. Lett.* **26** (2013) 397–401.
3. V. V. Aristov, *Direct Methods for Solving the Boltzmann Equation and Study of Nonequilibrium Flows* (Springer, 2001).
4. L. Arkeryd and A. Nouri, On the stationary Povzner equation in three space variables, *J. Math. Kyoto Univ.* **39** (1999) 115–153.
5. H.-O. Bae, S.-Y. Ha, Y. Kim, S.-H. Lee, H. Lim and J. Yoo, A mathematical model for volatility flocking with a regime switching mechanism in a stock market, *Math. Models Methods Appl. Sci.* **25** (2015) 1299–1335.
6. M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, M. Viale and V. Zdravkovic, Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study, *Proc. Natl. Acad. Sci. USA* **105** (2008) 1232–1237.
7. P. Barbante, A. Frezzotti and L. Gibelli, A kinetic theory description of liquid menisci at the microscale, *Kinet. Relat. Models* **8** (2015) 235–254.
8. N. Bellomo and A. Bellouquid, On multiscale models of pedestrian crowds from mesoscopic to macroscopic, *Commun. Math. Sci.* **13** (2015) 1649–1664.
9. N. Bellomo, A. Bellouquid and N. Chouhad, From a multiscale derivation of nonlinear cross-diffusion models to Keller–Segel models in a Navier–Stokes fluid, *Math. Models Methods Appl. Sci.* **26** (2016) 2041–2069.
10. N. Bellomo, A. Bellouquid and D. Knopoff, From the micro-scale to collective crowd dynamics, *Multiscale Model. Simulat.* **11** (2013) 943–963.
11. N. Bellomo and L. Gibelli, Toward a behavioral-social dynamics of pedestrian crowds, *Math. Models Methods Appl. Sci.* **25** (2015) 2417–2437.
12. N. Bellomo, D. Knopoff and J. Soler, On the difficult interplay between life, "complexity", and mathematical sciences, *Math. Models Methods Appl. Sci.* **23** (2013) 1861–1913.
13. N. Bellomo and J. Soler, On the mathematical theory of the dynamics of swarms viewed as complex systems, *Math. Models Methods Appl. Sci.* **22** (2012) 1140006.
14. A. Borzi and S. Wongkaew, Modeling and control through leadership of a refined flocking system, *Math. Models Methods Appl. Sci.* **25** (2015) 1193–1215.
15. D. Burini, S. De Lillo and L. Gibelli, Stochastic differential "nonlinear" games modeling collective learning dynamics, *Phys. Life Rev.* **26** (2016) 123–139.

16. D. Burini, S. De Lillo and L. Gibelli, Reply to "Stochastic differential 'nonlinear' games modeling collective learning dynamics", *Phys. Life Rev.* **26** (2016) 158–162.
17. J.-A. Canizo, 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** (2011) 515–539.
18. M. Caponigro, M. Fornasier, B. Piccoli and E. Trélat, Sparse stabilization and control of alignment models, *Math. Models Methods Appl. Sci.* **25** (2015) 521–564.
19. J.-A. Carrillo, M. Fornasier, J. Rosado and G. Toscani, Asymptotic flocking dynamics for the kinetic Cucker–Smale model, *SIAM J. Math. Anal.* **42** (2010) 218–236.
20. J. Cho, S.-Y. Ha, F. Huang, C. Jin and D. Ko, Emergence of bi-cluster flocking for the Cucker–Smale model, *Math. Models Methods Appl. Sci.* **26** (2016) 1191–1218.
21. F. Cucker and J.-G. Dong, On flocks influenced by closest neighbors, *Math. Models Methods Appl. Sci.* **26** (2016) 2685–2708.
22. F. Cucker and S. Smale, Emergent behavior in flocks, *IEEE Trans. Automat. Control* **52** (2007) 852–862.
23. P. Degond and S. Motsch, Macroscopic limit of self-driven particles with orientation interaction, *C. R. Math. Acad. Sci. Paris* **345** (2007) 555–560.
24. P. Degond and S. Motsch, Large-scale dynamics of the persistent Turing Walker model of fish behavior, *J. Statist. Phys.* **131** (2008) 989–1022.
25. P. Degond and S. Motsch, Continuum limit of self-driven particles with orientation interaction, *Math. Models Methods Appl. Sci.* **18** (2008) 1193–1215.
26. P. Degond and L. Navoret, A multi-layer model for self-propelled disks interacting through alignment and volume exclusion, *Math. Models Methods Appl. Sci.* **25** (2015) 2439–2475.
27. M. Di Francesco and S. Fagioli, A nonlocal swarm model for predators–prey interactions, *Math. Models Methods Appl. Sci.* **26** (2016) 319–355.
28. R. Duan, M. Fornasier and G. Toscani, A kinetic flocking model with diffusion, *Commun. Math. Phys.* **300** (2010) 95–145.
29. M. Fornasier, J. Haskovec and G. Toscani, Fluid dynamic description of flocking via Povzner–Boltzmann equation, *Physica D* **240** (2011) 21–31.
30. S.-Y. Ha and J.-G. Liu, A simple proof of Cucker–Smale flocking dynamics and mean field limit, *Commun. Math. Sci.* **7** (2009) 297–325.
31. S.-Y. Ha and E. Tadmor, From particle to kinetic and hydrodynamic description of flocking, *Kinet. Relat. Models* **1** (2008) 415–435.
32. T.-K. Karper, A. Mellet and K. Trivisa, Hydrodynamic limit of the kinetic Cucker–Smale flocking model, *Math. Models Methods Appl. Sci.* **25** (2015) 131–163.
33. S. Motsch and E. Tadmor, A new model for self-organized dynamics and its flocking behavior, *J. Statist. Phys.* **144** (2011) 923–947.
34. N. Outada, N. Vauchelet, T. Akrid and M. Khaladi, From kinetic theory of multicellular systems to hyperbolic tissue equations: Asymptotic limits and computing, *Math. Models Methods Appl. Sci.* **26** (2016) 2709–2734.
35. D.-A. Paley, N.-E. Leonard, R. Sepulchre, D. Grunbaum and J.-K. Parrish, Oscillator models and collective motion, *IEEE Control Syst.* **27** (2007) 89–105.
36. L. Pareschi and G. Toscani, *Interacting Multiagent Systems: Kinetic Equations and Monte Carlo Methods* (Oxford Univ. Press, 2014).
37. J. Park, H. Kim and S.-Y. Ha, Cucker–Smale flocking with inter-particle bonding forces, *IEEE Trans. Automat. Control* **55** (2010) 2617–2623.
38. A. Y. Povzner, The Boltzmann equation in kinetic theory of gases, *Amer. Math. Soc. Transl. Ser.* 2 **47** (1962) 193–216.

39. J. Shen, Cucker–Smale flocking under hierarchical leadership, *SIAM J. Appl. Math.* **68** (2007) 694–719.
40. J. Toner and Y. Tu, Flocks, herds and schools: A quantitative theory of flocking, *Phys. Rev. E* **58** (1998) 4828–4858.
41. C.-M. Topaz and A. L. Bertozzi, Swarming patterns in a two-dimensional kinematic model for biological groups, *SIAM J. Appl. Math.* **65** (2004) 152–174.
42. T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen and O. Schochet, Novel type of phase transition in a system of self-driven particles, *Phys. Rev. Lett.* **75** (1995) 1226–1229.
43. S. Wongkaew, M. Caponigro and A. Borzi, On the control through leadership of the Hegselmann–Krause opinion formation model, *Math. Models Methods Appl. Sci.* **25** (2015) 565–585.