MULTISCALE MODEL. SIMUL.
Vol. 19, No. 2, pp. 1115–1141

MULTIFLOCKS: EMERGENT DYNAMICS IN SYSTEMS WITH MULTISCALE COLLECTIVE BEHAVIOR*

ROMAN SHVYDKOY† AND EITAN TADMOR‡

To Björn Engquist with friendship and appreciation

Abstract. We study the multiscale description of large-time collective behavior of agents driven by alignment. The resulting *multiflock dynamics* arises naturally with realistic initial configurations consisting of multiple spatial scaling, which in turn peak at different time scales. We derive a "master-equation" which describes a complex multiflock congregations governed by two ingredients: (i) a fast inner-flock communication; and (ii) a slow(-er) interflock communication. The latter is driven by macroscopic observables which feature the *up-scaling* of the problem. We extend the current monoflock theory, proving a series of results which describe rates of multiflocking with natural dependencies on communication strengths. Both agent-based, kinetic, and hydrodynamic descriptions are considered, with particular emphasis placed on the discrete and macroscopic descriptions.

Key words. alignment, Cucker–Smale, large-time behavior, multiscale, multiflock, hydrodynamics

AMS subject classifications. 92D25, 35Q35, 76N10

DOI. 10.1137/20M1324454

1. Introduction.

We present (to our knowledge—a first) systematic study of multiscale analysis for the large-time behavior of collective dynamics. Different scales of the dynamics are captured by different descriptions. Our starting point is an agent-based description of *alignment dynamics* in which a crowd of N agents, each with unit mass, identified by (position, velocity) pairs $(\mathbf{x}_i(t), \mathbf{v}_i(t)) \in \mathbb{R}^d \times \mathbb{R}^d$, are governed by

$$\begin{aligned} &\dot{\mathbf{x}}_i(t) = \mathbf{v}_i(t), \\ &\dot{\mathbf{v}}_i(t) = \lambda \sum_{j\in\mathcal{C}} \phi(\mathbf{x}_i, \mathbf{x}_j)(\mathbf{v}_j(t) - \mathbf{v}_i(t)), \qquad i \in \mathcal{C} := \{1, 2, \dots, N\}. \end{aligned} \tag{1.1}$$

The alignment dynamics is dictated by the *symmetric* communication kernel $\phi(\cdot,\cdot) \geqslant 0$. It is tacitly assumed here that the initial configuration of the agents are equidistributed which justifies a scaling factor $\lambda = 1/N$, and thus (1.1) amounts to the celebrated Cucker–Smale (CS) model [6, 7]. The tendency to align velocities leads to the generic large-time formation of a flock.

In realistic scenarios, however, initial configurations are not equidistributed. Indeed, fluctuations in initial density may admit different scales of spatial concentrations. What is the collective behavior subject to such nonuniform initial densities? This is the main focus of our work.

The presence of different spatial scales leads to formation of separate flocks at different time scales, which are realized by mixing different formulations of alignment

*Received by the editors March 10, 2020; accepted for publication (in revised form) March 8, 2021; published electronically June 24, 2021.
https://doi.org/10.1137/20M1324454
Funding: The first author was supported in part by NSF grant DMS-1813351. The second author was supported in part by NSF grants DMS16-13911, RNMS11-07444 (KI-Net), and ONR grant N00014-1812465.
†Department of Mathematics, Statistics, and Computer Science, University of Illinois, Chicago, IL 60607-7045 USA (shvydkoy@uic.edu).
‡Department of Mathematics and Institute for Physical Sciences & Technology (IPST), University of Maryland, College Park, MD 207423289 USA (tadmor@umd.edu).

dynamics—from agent-based to hydrodynamic descriptions. In section 2 we make a systematic derivation, starting with the agent-based CS dynamics for a single flock (1.1) and ending with dynamics which involves several flocks $\mathcal{C}_\alpha, \alpha = 1, \dots A$: the α-flock consists of N_α agents, identified by (position, velocity) pairs $\{(\mathbf{x}_{\alpha i}, \mathbf{v}_{\alpha i})\}_{i \in \mathcal{C}_a}$, which is one part of a total crowd of size $N = \sum_{i=1}^A N_\alpha$. The resulting *multiflock* dynamics is governed by a "master equation"

$$
\begin{cases}
\dot{\mathbf{x}}_{\alpha i} = \mathbf{v}_{\alpha i}, \\
\dot{\mathbf{v}}_{\alpha i} = \lambda_\alpha \displaystyle\sum_{j=1}^{N_\alpha} m_{\alpha j} \phi_\alpha(\mathbf{x}_{\alpha i}, \mathbf{x}_{\alpha j})(\mathbf{v}_{\alpha j} - \mathbf{v}_{\alpha i}) + \mu \sum_{\substack{\beta=1 \\ \beta \neq \alpha}}^{A} M_\beta \psi(\mathbf{X}_\alpha, \mathbf{X}_\beta)(\mathbf{V}_\beta - \mathbf{v}_{\alpha i}).
\end{cases}
\tag{1.2}
$$

The system (1.2) arises naturally as an effective description for the alignment dynamics with multiple spatial scaling, which in turn, yields multiple *temporal* scalings. Such multiscaling appears when each α-flock undergoes evolution on a time scale much shorter than relative evolution between the flocks. Accordingly, the dynamics in (1.2) have two main parts. The first sum on the right encodes short-range alignment interactions among agents in flock α, dictated by *symmetric* communication kernel ϕ_α with amplitude λ_α. The new feature here is that spatial variations in initial density require us to trace the different masses $m_{\alpha j}$ attached to different agents located at $\mathbf{x}_{\alpha j}$. The second sum on the right encodes the interactions between agents in flock α and the "remote" flocks $\beta \neq \alpha$. The communication is dictated by symmetric kernel ψ with amplitude μ: since these are long-range interactions, they are scaled with relatively weak amplitude $\mu \ll 1$, and we therefore do not get into finer resolution of different kernels, $\psi_{\alpha\beta}$, to different flocks (interflocking interactions driven by different $\psi_{\alpha\beta}$ is the topic of a recent study on *multispecies* dynamics [11]). The new feature here is that the remote flocks in these long-range interactions, $\mathcal{C}_{\beta \neq \alpha}$, are encoded in terms of their macroscopic "observables"—their mass, $M_\beta = \sum_{i \in \mathcal{C}_\beta} m_{\beta i}$, and centers of mass and momentum

$$
\mathbf{X}_\beta := \frac{1}{M_\beta} \sum_{i \in \mathcal{C}_\alpha} m_{\beta i} \mathbf{x}_{\beta i}, \quad \mathbf{V}_\beta := \frac{1}{M_\beta} \sum_{i \in \mathcal{C}_\beta} m_{\beta i} \mathbf{v}_{\beta i}, \qquad M_\beta := \sum_{i \in \mathcal{C}_\beta} m_{\beta i}.
$$

These macroscopic quantities $\{(\mathbf{X}_\alpha, \mathbf{V}_\alpha)\}$ are determined by the slow interflocking dynamics: a weighted sum $\sum_i m_{\alpha i} (1.2)_i$ yields

$$
\begin{cases}
\dot{\mathbf{X}}_\alpha = \mathbf{V}_\alpha, \\
\dot{\mathbf{V}}_\alpha = \mu \displaystyle\sum_{\beta \neq \alpha} M_\beta \psi(\mathbf{X}_\alpha, \mathbf{X}_\beta)(\mathbf{V}_\beta - \mathbf{V}_\alpha).
\end{cases}
\tag{1.3}
$$

Thus, starting with agent dynamics (1.1) we end up with the same classical CS dynamics (1.3) for "super agents," weighted by their respective masses and representing macroscopic parameters of those flocks. This up-scaling—the process of bottom-up integration [12], naturally yields $\mathbf{X}_\beta, \mathbf{V}_\beta$, and M_β and their corresponding communication kernel, ψ, as the the effective parameters for our multiscale description. At the same time, our master equation (1.2) specifies how these up-scale parameters interact with the subscales parameters $\mathbf{x}_\alpha, \mathbf{v}_\alpha$, and ϕ_α, much like models for collective dynamics at the cellular scale that which include subcellular mechanisms [14]. The importance of multiscaling in collective dynamics was highlighted in a recent theme issue of the *Phil. Trans. Royal Soc. B* devoted to collective migration in biological systems; as the editors [8] indicate multiscale methods in collective migration uncover new unifying organization principles and in particular shed light on the transition from single to collective migration. These features are realized in our multiflocking approach.

Remark 1.1 (smooth and singular kernels). In the case when the interflock and internal communication kernels are smooth, the global existence of the system (1.2) follows by a trivial application of the Picard iteration and continuation. If the kernels ϕ_α are singular, however, collisions lead to finite time blowup, so this case needs to be addressed separately. In the appendix we show that multiflock dynamics governed by singular communication kernels with "fat head" so that $\int_0^1 \phi_\alpha(r)\,\mathrm{d}r = \infty$ experiences no internal collisions. Consequently, one can deduce global existence for systems with smooth ψ and a family of either smooth kernels or fat head kernels.

1.1. Statement of main results. Much of the theory available in the literature on monoscale flocking, e.g., [1, 2] and the references therein, admits proper extension to the framework of multiflocks. We chose to carry out proofs to three main aspects of (i) the large-time *alignment* behavior of (1.2); (ii) multiflocks in presence of additional *attractive forcing*; and (iii) large-crowd *hydrodynamics* of multiflocks. Below we highlight the main results.

We begin, in section 3, with the large-time alignment behavior of the multiflock dynamics (1.2). We assume that the short- and long-range communication kernels ϕ_α and ψ are *bounded* and *fat-tailed* in the sense that[1]

$$\phi_\alpha(\mathbf{x},\mathbf{y}) \gtrsim \langle|\mathbf{x}-\mathbf{y}|\rangle^{-\eta_\alpha}, \quad \psi(\mathbf{x},\mathbf{y}) \gtrsim \langle|\mathbf{x}-\mathbf{y}|\rangle^{-\zeta}, \qquad \eta_a, \zeta \leqslant 1. \tag{1.4}$$

They dictate the fast alignment rates insides flocks and slow cross-flocks rates, summarized in the following two theorems.

THEOREM 1.2 (fast local flocking). *Assume that the communication in an α-flock has a fat-tailed kernel $\phi_\alpha(\mathbf{x},\mathbf{y}) \gtrsim \langle|\mathbf{x}-\mathbf{y}|\rangle^{-\eta_\alpha}$, $\eta_\alpha \leqslant 1$. Then, the diameter of the α-flock is uniformly bounded in time, $\mathcal{D}_\alpha(t) := \max_{i,j}|\mathbf{x}_{\alpha i}(t)-\mathbf{x}_{\alpha j}(t)| \leqslant \overline{\mathcal{D}}_\alpha$, and the α-flock aligns exponentially fast toward its center of momentum*

$$\max_i |\mathbf{v}_{\alpha i}(t) - \mathbf{V}_\alpha(t)| \lesssim e^{-\delta_\alpha t}, \qquad \delta_\alpha = \lambda_\alpha M_\alpha(\overline{\mathcal{D}}_\alpha)^{-\eta_\alpha}. \tag{1.5}$$

The main message of this theorem is that the α-flock alignment towards $\mathbf{V}_\alpha$ depends only on the α-flock own parameters but not the global values. The global alignment has a slow(-er) rate reflecting weaker communication due to the smaller amplitude μ and the global diameter of the multiflock $\mathcal{D}$. Let $\mathbf{V}$ denote the center of momentum of the whole crowd, $\mathbf{V} := \frac{1}{M}\sum_\alpha M_\alpha \mathbf{V}_\alpha(t)$, and observe that it is time invariant.

THEOREM 1.3 (slow global flocking). *Suppose ψ has a fat tail, $\psi(\mathbf{x},\mathbf{y}) \gtrsim \langle|\mathbf{x}-\mathbf{y}|\rangle^{-\zeta}, \zeta \leqslant 1$. Then the diameter of the whole crowd is uniformly bounded in time, $\mathcal{D}(t) := \max_{\alpha,\beta}|\mathbf{X}_\alpha(t)-\mathbf{X}_\beta(t)| \leqslant \overline{\mathcal{D}}$, and solutions of* (1.2) *globally align with the global center of momentum* $\mathbf{V}$*,*

$$\max_{\alpha,i} |\mathbf{v}_{\alpha i}(t) - \mathbf{V}| \lesssim e^{-\delta t}, \qquad \delta = \mu M(\overline{\mathcal{D}})^{-\zeta}. \tag{1.6}$$

[1]Here and below we abbreviate $\langle X\rangle := (1+|X|^2)^{1/2}$, while $\langle \mathbf{v},\mathbf{u}\rangle$ or simply $\mathbf{v}\cdot\mathbf{u}$ stands for the usual scalar product of vectors. We also adopt the convention of approximate inequality signs to designate inequalities which hold up to a constant: $A \gtrsim B$ if $\exists c>0$ such that $A \geqslant cB$, etc.

As a consequence of the two theorems above we obtain what is called "strong flocking," that is, when all the displacements between agents stabilize, $\mathbf{x}_{\alpha i}(t) - \mathbf{x}_{\alpha j}(t) \to \bar{\mathbf{x}}_{\alpha ij}$ as $t \to \infty$.

Indeed, $\mathbf{x}_{\alpha i}(t) - \mathbf{x}_{\alpha j}(t) = \mathbf{x}_{\alpha i}(0) - \mathbf{x}_{\alpha j}(0) + \displaystyle\int_0^t [\mathbf{v}_{\alpha i}(s) - \mathbf{v}_{\alpha j}(s)]\,\mathrm{d}s$, hence

$$\bar{\mathbf{x}}_{\alpha ij} = \mathbf{x}_{\alpha i}(0) - \mathbf{x}_{\alpha j}(0) + \int_0^\infty [\mathbf{v}_{\alpha i}(s) - \mathbf{v}_{\alpha j}(s)]\,\mathrm{d}s,$$

and the rate of convergence is obviously the same as that claimed for the velocities.

In section 4 we study the multiflock dynamics (1.2) with additional *attractive forcing* (here we restrict attention to interactions determined by a radially symmetric kernels)

$$\begin{cases} \dot{\mathbf{x}}_{\alpha i} = \mathbf{v}_{\alpha i}, \\ \dot{\mathbf{v}}_{\alpha i} = \dfrac{1}{N_\alpha}\displaystyle\sum_{j=1}^{N_\alpha} m_{\alpha j}\phi_\alpha(|\mathbf{x}_{\alpha i} - \mathbf{x}_{\alpha j}|)(\mathbf{v}_{\alpha j} - \mathbf{v}_{\alpha i}) + \mu \sum_{\substack{\beta=1\\ \beta\neq\alpha}}^{A} M_\beta \psi(|\mathbf{X}_\alpha - \mathbf{X}_\beta|)(\mathbf{V}_\beta - \mathbf{v}_{\alpha i}) \\ \qquad + \mathbf{F}_{\alpha i}. \end{cases} \tag{1.7}$$

Here, $\mathbf{F}_{\alpha i}(t) = -\frac{1}{N_\alpha}\sum_{j=1}^{N_\alpha} \nabla U(|\mathbf{x}_{\alpha i} - \mathbf{x}_{\alpha j}|)$ is an external attractive forcing induced by a convex potential U which belongs to the class of potentials outlined in (4.6) below. Arguing along the lines of [16] we prove the following (the detailed result is outlined in Theorem 4.1 below).

THEOREM 1.4 (local flocking with attraction potential). *Consider the multiflock dynamics* (1.7) *with fat-tailed radial kernels,* $\phi_\alpha(r) \gtrsim \langle r\rangle^{-\eta}$ *and convex potential* $U(r) \gtrsim r^\beta$ *with tamed growth* $U^{(k)}(r) \lesssim r^{\beta-k}$, $k = 1, 2$, *for some* $\beta \geqslant 1$ *(further outlined in* (4.6) *below). There exists* η_β *specified in* (4.7)*, such that for* $\eta \leqslant \eta_\beta$*, the dynamics of each flock admits asymptotic aggregation,* $\limsup_{t\to\infty} \mathcal{D}_\alpha(t) \leqslant L$*, and alignment decay*

$$\frac{1}{2N_\alpha}\sum_{i=1}^{N_\alpha} |\mathbf{v}_{\alpha i} - \mathbf{V}_\alpha|^2 \lesssim \frac{C_\delta}{\langle t\rangle^{1-\delta}} \quad \forall \delta > 0, \qquad \alpha = 1, 2, \ldots, A.$$

It should be emphasized that the confining action of the attraction potential is assumed to act only on far-field, $r > L$, but otherwise is allowed to be "turned off" for $U(r) = 0,\ r \leqslant L$ as depicted in Figure 1.1. This offers an extension of the recent result [17] for the case $L = 0$. In fact, as noted in Theorem 4.2 below, if the potential U has a global support, then there is *exponential rate* alignment.

When $N_\alpha \gg 1$ one recovers the large-crowd dynamics in terms of the macroscopic density and velocity $(\rho_a, \mathbf{u}_\alpha)$, governed by the hydrodynamic multiflock system, which is the topic of section 5

$$\begin{cases} \partial_t \rho_\alpha + \nabla\cdot(\mathbf{u}_\alpha \rho_\alpha) = 0 \\ \partial_t \mathbf{u}_\alpha + \mathbf{u}_\alpha \cdot \nabla \mathbf{u}_\alpha = \lambda_\alpha \displaystyle\int_{\mathbb{R}^d} \phi_\alpha(\mathbf{x},\mathbf{y})(\mathbf{u}_\alpha(\mathbf{y}) - \mathbf{u}_\alpha(\mathbf{x}))\rho_\alpha(\mathbf{y})\,\mathrm{d}\mathbf{y} \\ \qquad\qquad + \mu \displaystyle\sum_{\beta\neq\alpha} M_\beta \psi(\mathbf{X}_\alpha, \mathbf{X}_\beta)(\mathbf{V}_\beta - \mathbf{u}_\alpha(\mathbf{x}, t)), \end{cases} \qquad \alpha = 1, \ldots, A.$$

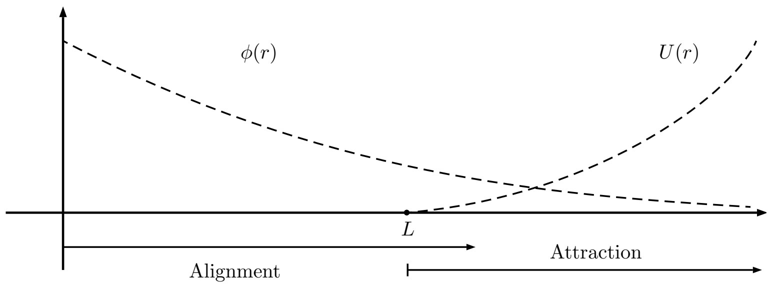

FIG. 1.1. *Two-zone attraction-alignment model.*

Here $\{(\mathbf{X}_\alpha, \mathbf{V}_\alpha)\}_\alpha$ are the macroscopic quantities which record the center of mass and momentum of α-flock governed by (1.3). The alignment dynamics reflects the discrete framework of Theorems 1.2 and 1.3, namely, if ϕ_α and ψ are fat-tailed then smooth solutions of the α-flock and, respectively, the whole crowd will align towards their respective averages. The details can be found in Theorem 5.1 below. In particular, we prove that the one-dimensional multiflock hydrodynamics with radial ϕ_α's, either smooth or singular, and subject to subcritical initial condition $u'_\alpha(x,0) + \lambda_\alpha \phi_\alpha * \rho_\alpha(x,0) \geqslant 0$, $\forall x \in \mathbb{R}$, admits global smooth solution and flocking insues.

2. From agents to multiflocks and back: Up-scaling.

2.1. Agent-based description. Our starting point is the alignment-based dynamics (1.1)

$$\begin{aligned} &\dot{\mathbf{x}}_i(t) = \mathbf{v}_i(t) \\ &\dot{\mathbf{v}}_i(t) = \sum_{j\in\mathcal{C}} \phi(\mathbf{x}_i, \mathbf{x}_j)(\mathbf{v}_j(t) - \mathbf{v}_i(t)), \qquad i \in \mathcal{C} := \{1, 2, \ldots, N\}. \end{aligned} \tag{2.1}$$

This expresses the tendency of agents to align their velocities with the rest of the crowd, dictated by the *symmetric* communication kernel $\phi(\cdot,\cdot) \geqslant 0$. Let us assume that each of the terms on the right is of the same order, $\mathcal{O}(1)$; then the total action on the right of order $\mathcal{O}(N)$ will peak at time $t = \mathcal{O}(1/N)$. Using the scaling parameter $\lambda = 1/N$, one arrives at the celebrated CS model [6, 7]

$$\dot{\mathbf{v}}_i = \lambda \sum_{j\in\mathcal{C}} \phi(\mathbf{x}_i, \mathbf{x}_j)(\mathbf{v}_j - \mathbf{v}_i), \qquad \lambda = \frac{1}{N},$$

where the dynamics are rescaled to peak at the desired $t \sim \mathcal{O}(1)$. But what happens when the terms on the right of (2.1) are of different order? Assume that the crowd consists of two mostly separated groups, $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2$, where $\mathcal{C}_1$ has a large crowd of N_1 agents whereas $\mathcal{C}_2$ has a much smaller crowd of $N_2 \ll N_1$ agents. By "mostly separated" we mean that the two groups have a very low level of communication so that $\{\phi(\mathbf{x}_i, \mathbf{x}_j) \ll 1 \mid (\mathbf{x}_i, \mathbf{x}_j) \in (\mathcal{C}_1, \mathcal{C}_2)\}$. We will quantify a precise statement of separation in section 2.3 below. Now the dynamics (2.1) will experience two time scales: the action of the larger crowd $\mathcal{C}_1$ will peak earlier at time $t_1 = \mathcal{O}(1/N_1)$, mostly ignoring the negligible effect of the "far way" crowd in $\mathcal{C}_2$. The crowd of $\mathcal{C}_2$ will peak *much later* at time $t_2 = \mathcal{O}(1/N_2) \gg t_1$. In [13] we suggested an adaptive scaling parameter

$$\dot{\mathbf{v}}_i = \lambda_i \sum_{j\in\mathcal{C}} \phi(\mathbf{x}_i, \mathbf{x}_j)(\mathbf{v}_j - \mathbf{v}_i), \qquad \lambda_i = \frac{1}{\sum_j \phi(\mathbf{x}_i, \mathbf{x}_j)};$$

here, λ_i adapts itself to the different clocks of both crowds: when in the larger crowd $i \in \mathcal{C}_1$, we have $\lambda_i \sim 1/N_1$, whereas for agents in the smaller crowd $i \in \mathcal{C}_2$ we have $\lambda_i \sim 1/N_2$

$$\dot{\mathbf{v}}_i = \begin{cases} \lambda_i \sum_{j\in\mathcal{C}_1} \phi(\mathbf{x}_i, \mathbf{x}_j)(\mathbf{v}_j - \mathbf{v}_i), & i \in \mathcal{C}_1: \ \lambda_i = \dfrac{1}{\sum_j \phi(\mathbf{x}_i, \mathbf{x}_j)} \sim \dfrac{1}{N_1} \\ \lambda_i \sum_{j\in\mathcal{C}_2} \phi(\mathbf{x}_i, \mathbf{x}_j)(\mathbf{v}_j - \mathbf{v}_i), & i \in \mathcal{C}_2: \ \lambda_i = \dfrac{1}{\sum_j \phi(\mathbf{x}_i, \mathbf{x}_j)} \sim \dfrac{1}{N_2}. \end{cases}$$

Thus, λ_i should be viewed as *time scaling* adapted for both crowds to peak at the desired $t = \mathcal{O}(1)$. While this scaling is satisfactory for $\mathcal{C}_1$, it neglects taking into account that the activity of the smaller $\mathcal{C}_2$ peaks much later after the peak of the larger crowd $\mathcal{C}_1$, which has an additional effect on the dynamics of $\mathcal{C}_2$.

2.2. Scale separation in time. We want to take both groups into account while being precise of using the same "clock." To this end, it will be convenient to observe the configurations of crowds $\mathcal{C}_1$ and $\mathcal{C}_2$ in terms of their empirical distribution

$$\mu_1(\mathbf{x}, \mathbf{v}, t) := \frac{1}{N_1}\sum_{k\in\mathcal{C}_1} \delta_{\mathbf{x}_k(t)}(\mathbf{x}) \otimes \delta_{\mathbf{v}_k(t)}(\mathbf{v}), \quad \mu_2(\mathbf{x}, \mathbf{v}, t) := \frac{1}{N_2}\sum_{k\in\mathcal{C}_2} \delta_{\mathbf{x}_k(t)}(\mathbf{x}) \otimes \delta_{\mathbf{v}_k(t)}(\mathbf{v}).$$

We distinguish between three time scales.

(i) Time $t \lesssim t_1$. The dynamics is captured by the agent-based description of the two separate groups which form the crowd $\mathcal{C}$ in (2.1).

(ii) Time $t_1 \ll t \lesssim t_2$. Since $t_2 \gg t_1$, crowd $\mathcal{C}_1$ is captured by its large-time dynamics, which is realized as a continuum with macroscopic density $\mu_1(\mathbf{x}, \mathbf{v}, t)\, \mathrm{d}\mathbf{v} \stackrel{N_1 \gg 1}{\longrightarrow} \rho_1(\mathbf{x}, t) : \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}_+$, and momentum $\mu_1(\mathbf{x}, \mathbf{v}, t)\mathbf{v}\, \mathrm{d}\mathbf{v} \stackrel{N_1 \gg 1}{\longrightarrow} (\rho_1\mathbf{u}_1)(\mathbf{x}, t) : \mathbb{R}^d \times \mathbb{R}_+ \mapsto \mathbb{R}^d$. Observe that the dynamics at this stage involves two groups with two different descriptions: crowd $\mathcal{C}_1$ is encoded in terms of its hydrodynamic observables, $(\rho_1, \rho_1\mathbf{u}_1)$, while crowd $\mathcal{C}_2$ is still encoded in terms of its agent-based description

$$\rho(\mathbf{y}, t) = \rho_1(\mathbf{y}, t) + \overbrace{\frac{1}{N_2}\sum_{k\in\mathcal{C}_2} \delta_{\mathbf{x}_k(t)}(\mathbf{y})}^{\rho_2(\mathbf{y},t)}, \quad \rho\mathbf{u}(\mathbf{y}, t) = \rho_1\mathbf{u}_1(\mathbf{y}, t) + \overbrace{\frac{1}{N_2}\sum_{k\in\mathcal{C}_2} \mathbf{v}_k(t)\delta_{\mathbf{x}_k(t)}(\mathbf{y})}^{\rho_2\mathbf{u}_2(\mathbf{y},t)}.$$

The large-time dynamics of $\mathcal{C}_1$ is governed by the hydrodynamic system [10, 3]

$$(2.2)_1 \qquad \begin{cases} (\rho_1)_t + \nabla_{\mathbf{x}} \cdot (\rho_1\mathbf{u}_1) = 0, \\ (\rho_1\mathbf{u}_1)_t + \nabla_{\mathbf{x}} \cdot (\rho_1\mathbf{u}_1 \otimes \mathbf{u}_1 + P_1) = \displaystyle\int_{\mathbb{R}^n} \phi(\mathbf{x}, \mathbf{y})\big\{(\rho\mathbf{u})(\mathbf{y}, t)\rho_1(\mathbf{x}, t) \\ \qquad\qquad - \rho(\mathbf{y}, t)(\rho_1\mathbf{u}_1)(\mathbf{x}, t))\big\}\, \mathrm{d}\mathbf{y}, \end{cases}$$

while crowd $\mathcal{C}_2$ is governed by the agent-based description (2.1) which takes the weak formulation

$$(2.2)_2 \qquad \begin{cases} (\rho_2)_t + \nabla_{\mathbf{x}} \cdot (\rho_2\mathbf{u}_2) = 0, \\ (\rho_2\mathbf{u}_2)_t + \nabla_{\mathbf{x}} \cdot (\rho_2\mathbf{u}_2 \otimes \mathbf{u}_2 + P_2) = \displaystyle\int_{\mathbb{R}^n} \phi(\mathbf{x}, \mathbf{y})\big\{(\rho\mathbf{u})(\mathbf{y}, t)\rho_2(\mathbf{x}, t) \\ \qquad\qquad - \rho(\mathbf{y}, t)(\rho_2\mathbf{u}_2)(\mathbf{x}, t))\big\}\, \mathrm{d}\mathbf{y}. \end{cases}$$

Here, $P_1 = P(\mathbf{v}-\mathbf{u}_1 \otimes \mathbf{v}-\mathbf{u}_1)$ is a second-order fluctuations pressure tensor which requires a closure relations between the microscopic and macroscopic variables. We shall not dwell on its specific form: the large time behavior of $\mathcal{C}_1$ in $(2.2)_1$ is *independent* of the specifics of this closure. It will suffice to observe the center of mass and average velocity of crowd $\mathcal{C}_1$:

$$\mathbf{X}_1(t) := \frac{1}{M_1}\int_{\mathcal{S}_1} \mathbf{x}\rho_1(\mathbf{x},t)\,\mathrm{d}\mathbf{x}, \qquad \mathbf{V}_1(t) := \frac{1}{M_1}\int_{\mathcal{S}_1} \rho_1(\mathbf{x},t)\mathbf{u}_1(\mathbf{x},t)\,\mathrm{d}\mathbf{x},$$
$$\mathcal{S}_1 := \mathrm{supp}\{\rho_1(t,\cdot)\}.$$

Integrating $(2.2)_1$ over the support of the first crowd $\mathcal{S}_1$: since the "self"-alignment of $\mathcal{C}_1$ with itself vanishes for $\mathbf{y}\in\mathcal{S}_1$, we are left with the contribution from the second crowd $\rho(\mathbf{y},t) \mapsto \rho_2 = \frac{1}{N_2}\sum_{k\in\mathcal{C}_2}\delta_{\mathbf{x}_k(t)}(\mathbf{y})$ and $(\rho\mathbf{u})(\mathbf{y},t) \mapsto \rho_2\mathbf{u}_2 = \frac{1}{N_2}\sum_{k\in\mathcal{C}_2}\mathbf{v}_k(t)\delta_{\mathbf{x}_k(t)}(\mathbf{y})$, which yields

$$\begin{aligned}
\dot{\mathbf{X}}_1 &= \mathbf{V}_1\\
M_1\dot{\mathbf{V}}_1 &= \int_{\mathbf{x}\in\mathcal{S}_1}\int_{\mathbf{y}\in\mathcal{S}_2} \phi(\mathbf{x},\mathbf{y})\big\{(\rho_2\mathbf{u}_2)(\mathbf{y},t)\rho_1(\mathbf{x},t) - \rho_2(\mathbf{y},t)(\rho_1\mathbf{u}_1)(\mathbf{x},t))\big\}\,\mathrm{d}\mathbf{y}\,\mathrm{d}\mathbf{x}\\
&= \frac{1}{N_2}\sum_{j\in\mathcal{C}_2}\mathbf{v}_{2j}(t)\int_{\mathbf{x}\in\mathcal{S}_1}\phi(\mathbf{x},\mathbf{x}_{2j})\rho_1(\mathbf{x},t)\,\mathrm{d}\mathbf{x}\\
&\quad - \frac{1}{N_2}\sum_{j\in\mathcal{C}_2}\int_{\mathbf{x}\in\mathcal{S}_1}\phi(\mathbf{x},\mathbf{x}_{2j})(\rho_1\mathbf{u}_1)(\mathbf{x},t)\,\mathrm{d}\mathbf{x}.
\end{aligned}$$

Due to assumed relatively large separation between the flocks, we can approximate the last two integrals by the values of the kernel integrands at the centers of mass:

$$\left\{\begin{aligned}
&\int_{\mathbf{x}\in\mathcal{S}_1}\phi(\mathbf{x},\mathbf{x}_{2j})\rho_1(\mathbf{x},t)\,\mathrm{d}\mathbf{x} =: \mu\psi(\mathbf{X}_1,\mathbf{X}_2)M_1,\\
&\qquad\qquad \mu\psi(\mathbf{X},\mathbf{Y}) \approx \phi(\mathbf{X},\mathbf{Y}), \quad \mu \ll 1\\
&\int_{\mathbf{x}\in\mathcal{S}_1}\phi(\mathbf{x},\mathbf{x}_{2j})(\rho_1\mathbf{u}_1)(\mathbf{x},t)\,\mathrm{d}\mathbf{x} =: \mu\psi(\mathbf{X}_1,\mathbf{X}_2)M_1\mathbf{V}_1,
\end{aligned}\right. \tag{2.3}$$

obtaining

$$\left\{\begin{aligned}
&\dot{\mathbf{X}}_1(t) = \mathbf{V}_1(t)\\
&\dot{\mathbf{V}}_1(t) = \mu\psi(\mathbf{X}_1,\mathbf{X}_2)\big(\mathbf{V}_2(t) - \mathbf{V}_1(t)\big), \quad \mathbf{V}_2(t) = \frac{1}{M_2}\int \rho_2\mathbf{u}_2(\mathbf{x},t)\,\mathrm{d}\mathbf{x}\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad = \frac{1}{N_2}\sum_{j\in\mathcal{S}_2}\mathbf{v}_{2j}(t).
\end{aligned}\right. \tag{2.4a}$$

For the dynamics of the second group $\mathcal{C}_2$ we may take $P_2 \equiv 0$ on the left of $(2.2)_2$. The cross-group interactions term $(\rho,\rho\mathbf{u}) \mapsto (\rho_1,\rho_1\mathbf{u}_1)$ on the right of $(2.2)_2$ yields

$$\int_{\mathbf{y}\in\mathcal{S}_1}\phi(\mathbf{x}_{2j},\mathbf{y})\rho_1(\mathbf{y},t)\,\mathrm{d}\mathbf{y} = \mu\psi(\mathbf{X}_2,\mathbf{X}_1)M_1,$$
$$\int_{\mathbf{y}\in\mathcal{S}_1}\phi(\mathbf{x}_{2j},\mathbf{y})(\rho_1\mathbf{u}_1)(\mathbf{y},t)\,\mathrm{d}\mathbf{y} = \mu\psi(\mathbf{X}_2,\mathbf{X}_1)M_1\mathbf{V}_1,$$

arriving at

$$\text{(2.4b)}\qquad \begin{cases} \dot{\mathbf{x}}_{2i} = \mathbf{v}_{2i}, \quad i \in \mathcal{C}_1, \\ \dot{\mathbf{v}}_{2i} = \displaystyle\sum_{j\in\mathcal{C}_2} \phi(\mathbf{x}_{2i}, \mathbf{x}_{2j})(\mathbf{v}_{2j} - \mathbf{v}_{2i}) + \mu\psi(\mathbf{X}_2, \mathbf{X}_1)M_1(\mathbf{V}_1 - \mathbf{v}_{2i}). \end{cases}$$

Thus, we end up with a new agent-based dynamics, (2.4b), in which the dynamics of group $\mathcal{C}_1$ is encoded as new agent governed by mean position $\mathbf{X}_1$ and a mean velocity $\mathbf{V}_1$. This is a precisely the system (1.2) written for the smaller flock $\mathcal{C}_2$.

(iii) Time $t \gg t_2$. Now the second crowd $\mathcal{C}_2$ is also captured by its large-time dynamics, realized in terms of macroscopic density $\mu_2(\mathbf{x}, \mathbf{v}, t)\,\mathrm{d}\mathbf{v} \to \rho_2(\mathbf{x}, t) : \mathbb{R}^d \times \mathbb{R}_+ \mapsto \mathbb{R}_+$, and momentum $\mu_2(\mathbf{x}, \mathbf{v}, t)\mathbf{v}\,\mathrm{d}\mathbf{v} \to (\rho_2\mathbf{u}_2)(\mathbf{x}, t) : \mathbb{R}^d \times \mathbb{R}_+ \mapsto \mathbb{R}^d$. Together, groups $\mathcal{C}_1$ and $\mathcal{C}_2$ form the crowd

$$\rho(\mathbf{y}, t) = \rho_1(\mathbf{y}, t) + \rho_2(\mathbf{y}, t) \qquad \rho\mathbf{u}(\mathbf{y}, t) = \rho_1\mathbf{u}_1(\mathbf{y}, t) + \rho_2\mathbf{u}_2(\mathbf{y}, t),$$
$$\mathcal{S}_i = \mathcal{S}_i(t) := \mathrm{supp}\{\rho_i(\cdot, t)\},$$

which is governed by $(2.2)_1$–$(2.2)_2$. Here, $P_2 = P(\mathbf{v} - \mathbf{u}_2 \otimes \mathbf{v} - \mathbf{u}_2)$ is a second-order fluctuations pressure tensor which requires a closure relations between the microscopic and macroscopic variables. But we do not dwell on its specific form, since the large time behavior of $\mathcal{C}_2$ in $(2.2)_2$ is captured by the center of mass and average velocity of crowd $\mathcal{C}_2$:

$$\mathbf{X}_2(t) := \frac{1}{M_2}\int_{\mathcal{S}_2} \mathbf{x}\rho_2(\mathbf{x}, t)\,\mathrm{d}\mathbf{x}, \qquad \mathbf{V}_2(t) := \frac{1}{M_2}\int_{\mathcal{S}_2} \mathbf{u}_2(\mathbf{x}, t)\rho_2(\mathbf{x}, t)\,\mathrm{d}\mathbf{x}.$$

Integrating $(2.2)_2$ over the support of the second crowd $\mathcal{S}_2$: since the "self"-alignment of $\mathcal{C}_2$ with itself vanishes for $\mathbf{y} \in \mathcal{S}_2$, and using (2.3) we are left with

$$\begin{aligned} \dot{\mathbf{X}}_2 &= \mathbf{V}_2, \\ M_2\dot{\mathbf{V}}_2 &= \int_{\mathbf{x}\in\mathcal{S}_2}\int_{\mathbf{y}\in\mathcal{S}_1} \phi(\mathbf{x}, \mathbf{y})\big\{(\rho\mathbf{u})(\mathbf{y}, t)\rho_2(\mathbf{x}, t) - \rho(\mathbf{y}, t)(\rho_2\mathbf{u}_2)(\mathbf{x}, t))\big\}\,\mathrm{d}\mathbf{y}\,\mathrm{d}\mathbf{x} \\ &= \int_{\mathbf{x}\in\mathcal{S}_2}\int_{\mathbf{y}\in\mathcal{S}_1} \phi(\mathbf{x}, \mathbf{y})\big\{(\rho_1\mathbf{u}_1)(\mathbf{y}, t)\rho_2(\mathbf{x}, t) - \rho_1(\mathbf{y}, t)(\rho_2\mathbf{u}_2)(\mathbf{x}, t))\big\}\,\mathrm{d}\mathbf{y}\,\mathrm{d}\mathbf{x} \\ &= \int_{\mathbf{x}\in\mathcal{S}_2} \phi(\mathbf{x}, \mathbf{X}_1)\big\{M_1\mathbf{V}_1\rho_2(\mathbf{x}, t) - M_1(\rho_2\mathbf{u}_2)(\mathbf{x}, t))\big\}\,\mathrm{d}\mathbf{x}. \end{aligned}$$

We approximate the last two integrals by the same principle as before

$$\text{(2.5)}\qquad \begin{aligned} \int_{\mathbf{x}\in\mathcal{S}_2} \phi(\mathbf{x}, \mathbf{X}_1)\rho_2(\mathbf{x}, t)\,\mathrm{d}\mathbf{x} &= \mu\psi(\mathbf{X}_2, \mathbf{X}_1)M_2, \\ \int_{\mathbf{x}\in\mathcal{S}_2} \phi(\mathbf{x}, \mathbf{X}_1)(\rho_2\mathbf{u}_2)(\mathbf{x}, t)\,\mathrm{d}\mathbf{x} &= \mu\psi(\mathbf{X}_2, \mathbf{X}_1)M_2\mathbf{V}_2, \end{aligned}$$

arriving at a 2-agent system described by the dynamics of their center of mass/momentum $(\mathbf{x}_\alpha, \mathbf{V}_\alpha)$,

$$\text{(2.6)}\qquad \begin{cases} \dot{\mathbf{X}}_\alpha(t) = \mathbf{V}_\alpha(t), \\ M_\alpha\dot{\mathbf{V}}_\alpha(t) = \mu\displaystyle\sum_{\beta\neq\alpha} \psi(\mathbf{X}_\alpha, \mathbf{X}_\beta)M_\alpha M_\beta(\mathbf{V}_\beta(t) - \mathbf{V}_\alpha(t)), \end{cases} \qquad \alpha, \beta \in \{1, 2\}.$$

In summary, we began with the agent based description for two crowds of $N_1 \gg N_2$ agents, (2.1) valid for $t \lesssim t_1$. It evolved into an agent-based description for crowd of N_2+1 agents (2.4) valid for $t_1 \ll t \lesssim t_2$ and ended with 2-agent description (2.6) valid for $t \gg t_2$. This is a process of *up-scaling* in which the notion of an "agent" is replaced with a "multiflock" — a larger blob made of agents, which is identified by its center of mass/momentum. The only difference is that the multiflock-based dynamics now takes into account only the up-scaled quantities of the multiflock. Let us recall that the more general system (1.2) permits up-scaling in the same way.

2.3. Scale separation in space. Following up on the idea of spatial separation between islands it is instructive to assess the scale on which approximation of mass/momentum given in (2.3), (2.5) is valid. To make analysis more precise we assume the large distance behavior of the communication kernel $\phi(\mathbf{x},\mathbf{y}) \sim |\mathbf{x}-\mathbf{y}|^{-\eta}$. We consider the prototypical integrals in (2.3) $\int_{\mathbf{x}\in\mathcal{S}_1} \phi(\mathbf{x},\mathbf{y})\rho_1(\mathbf{x},t)\,\mathrm{d}\mathbf{x}$ and $\int_{\mathbf{x}\in\mathcal{S}_1} \phi(\mathbf{x},\mathbf{y})\rho_1\mathbf{u}_1(\mathbf{x},t)\,\mathrm{d}\mathbf{x}$ for $\mathbf{y}\in\mathcal{S}_2$. We now fix $\mathbf{X}\in \operatorname{conv}\mathcal{S}_1$ and $\mathbf{Y}\in\operatorname{conv}\mathcal{S}_2$, and for any given pair of agents $\mathbf{x}\in\mathcal{S}_1$, $\mathbf{y}\in\mathcal{S}_2$ we decompose $\mathbf{x}-\mathbf{y} = (\mathbf{X}-\mathbf{Y}) + (\mathbf{Y}-\mathbf{y}) - (\mathbf{X}-\mathbf{x})$. Thus, $R := |\mathbf{X}-\mathbf{Y}|$ is the (fixed) long-range distance between the two groups, whereas $r := |(\mathbf{Y}-\mathbf{y})-(\mathbf{X}-\mathbf{x})|$ encapsulates the short-range distances within the crowds, $r \ll R$. Similar decomposition holds for the weighted integral of ρ_2 sought in (2.5). We have

$$\frac{1}{|\mathbf{x}-\mathbf{y}|} = \frac{1}{\sqrt{R^2+r^2-2r\cos\theta}} = \frac{1}{R}\sum_{k=0}^{\infty}\left(\frac{r}{R}\right)^k P_k(\cos\theta),$$

where $\cos\theta = \langle(\mathbf{X}-\mathbf{y})/R, (\mathbf{X}-\mathbf{x})/r\rangle$ and P_k are the k-degree Legendre polynomials, $P_0(x)=1, P_1(x)=x$, etc. We find

$$\begin{aligned}\frac{1}{|\mathbf{x}-\mathbf{y}|} &= \frac{1}{R} + \frac{r}{R^2}\left\langle \frac{\mathbf{X}-\mathbf{Y}}{R}, \frac{(\mathbf{Y}-\mathbf{y})-(\mathbf{X}-\mathbf{x})}{r}\right\rangle + \mathcal{O}\left(\frac{r^2}{R^3}\right)\\ &= \frac{1}{R}\left(1+\frac{1}{R^2}\Big\langle \mathbf{X}-\mathbf{Y}, (\mathbf{Y}-\mathbf{y})-(\mathbf{X}-\mathbf{x})\Big\rangle + \mathcal{O}\left(\frac{r^2}{R^2}\right)\right),\\ &\quad \mathbf{x}\in\mathcal{S}_1, \quad \mathbf{y}\in\mathcal{S}_2.\end{aligned}$$

Since the contribution of the second term on the right is of order $r/R \ll 1$ we can further approximate

$$\begin{aligned}\phi(\mathbf{x},\mathbf{y}) \sim \frac{1}{|\mathbf{x}-\mathbf{y}|^\eta} &= \frac{1}{R^\eta}\left(1+\frac{\eta}{R^2}\Big\langle \mathbf{X}-\mathbf{Y}, (\mathbf{Y}-\mathbf{y})-(\mathbf{X}-\mathbf{x})\Big\rangle + \mathcal{O}\left(\frac{r^2}{R^2}\right)\right)\\ &= \frac{1}{R^\eta} + \frac{\eta}{R^{2+\eta}}\langle \mathbf{X}-\mathbf{Y}, (\mathbf{Y}-\mathbf{y})-(\mathbf{X}-\mathbf{x})\rangle + \mathcal{O}\left(\frac{r^2}{R^{2+\eta}}\right).\end{aligned}$$

The first key point is that by choosing $\mathbf{X}=\mathbf{X}_1$ and $\mathbf{Y}=\mathbf{X}_2$ as the centers of mass of the flocks, so that $M_1\mathbf{X}_1 = \int_{\mathbf{x}\in\mathcal{S}_1}\mathbf{x}\rho_1(\mathbf{x},t)$; then the second term has a negligible contribution. Indeed,

$$\begin{aligned}\int_{\mathbf{x}\in\mathcal{S}_1}\phi(\mathbf{x},\mathbf{y})\rho_1(\mathbf{x},t)\,\mathrm{d}\mathbf{x} &\sim \int_{\mathbf{x}\in\mathcal{S}_1}\frac{1}{|\mathbf{x}-\mathbf{y}|^\eta}\rho_1(\mathbf{x},t)\,\mathrm{d}\mathbf{x}\\ &= \frac{1}{R^\eta}\int_{\mathbf{x}\in\mathcal{S}_1}\rho_1(\mathbf{x},t)\,\mathrm{d}\mathbf{x} + \frac{\eta}{R^{2+\eta}}\int_{\mathbf{x}\in\mathcal{S}_1}\langle \mathbf{X}_1-\mathbf{X}_2, (\mathbf{X}_2-\mathbf{y})\\ &\quad -(\mathbf{X}_1-\mathbf{x})\rangle\rho_1(\mathbf{x},t)\,\mathrm{d}\mathbf{x} + \mathcal{O}\left(\frac{r^2}{R^{2+\eta}}\right)\end{aligned}$$

$$= \frac{1}{R^\eta} M_1 + \frac{\eta}{R^{2+\eta}} \int_{\mathbf{x}\in\mathcal{S}_1} \langle \mathbf{X}_1 - \mathbf{X}_2, (\mathbf{X}_2 - \mathbf{y})\rangle \rho_1(\mathbf{x},t)\,\mathrm{d}\mathbf{x} + \mathcal{O}\left(\frac{r^2}{R^{2+\eta}}\right)$$
$$= \frac{1}{R^\eta} M_1 + \mathcal{O}\left(\frac{r}{R^{1+\eta}}\right) + \mathcal{O}\left(\frac{r^2}{R^{2+\eta}}\right).$$

Noting that $\phi(\mathbf{X}_1,\mathbf{X}_2) = R^{-\eta}$ we conclude with the first part of (2.3)

$$\int_{\mathbf{x}\in\mathcal{S}_1} \phi(\mathbf{x},\mathbf{y})\rho_1(\mathbf{x},t)\,\mathrm{d}\mathbf{x} = \phi(\mathbf{X}_1,\mathbf{y})M_1 + \mathcal{O}\left(\frac{r}{R^{1+\eta}}\right), \qquad \mathbf{y}\in\mathcal{S}_2. \tag{2.7a}$$

Similarly, we recover the asymptotic formula for momentum (2.3)

$$\int_{\mathbf{x}\in\mathcal{S}_1} \phi(\mathbf{x},\mathbf{y})(\rho_1\mathbf{u}_1)(\mathbf{x},t)\,\mathrm{d}\mathbf{x} = \phi(\mathbf{X}_1,\mathbf{X}_2)M_1\mathbf{V}_1(t) + \mathcal{O}\left(\frac{r}{R^{1+\eta}}\right). \tag{2.7b}$$

The same argument applies for crowd $\mathcal{C}_2$:

$$\int_{\mathbf{x}\in\mathcal{S}_2} \phi(\mathbf{x},\mathbf{y}) \left\{ \begin{array}{l} \rho_2(\mathbf{x},t) \\ (\rho_2\mathbf{u}_2)(\mathbf{x},t) \end{array} \right\} \mathrm{d}\mathbf{x} = \phi(\mathbf{X}_2,\mathbf{X}_1) \left\{ \begin{array}{l} M_2 \\ M_2\mathbf{V}_2(t) \end{array} \right\} + \mathcal{O}\left(\frac{r}{R^{1+\eta}}\right), \quad \mathbf{y}\in\mathcal{S}_1. \tag{2.8}$$

Remark 2.1. The bounds (2.7), (2.8) quantify first-order errors, $\mathcal{O}(\epsilon_{ij}) \ll 1$, provided the diameters of crowds $\mathcal{C}_i, \mathcal{C}_j$ are much smaller than their distance, $\epsilon_{ij} := \max\{r_i, r_j\}/R_{ij} \ll 1$.

3. Slow and fast alignment in multiflocks. In this section we focus on alignment dynamics for system (1.2) under conditions of Theorem 1.2 and 1.3. In fact, with a slight abuse of notation we will make a more general assumption that there exist radially symmetric subkernels

$$\phi_\alpha(\mathbf{x},\mathbf{y}) \geqslant \phi_\alpha(|\mathbf{x}-\mathbf{y}|), \qquad \psi(\mathbf{x},\mathbf{y}) \geqslant \psi(|\mathbf{x}-\mathbf{y}|), \tag{3.1}$$

which are positive, monotonely decreasing, and fat tail at infinity

$$\int_{r_0}^\infty \phi_\alpha(r)dr = \infty, \qquad \int_{r_0}^\infty \psi(r)dr = \infty. \tag{3.2}$$

We start by noting that any cluster system (1.2) satisfies the global maximum principle—maximum of each coordinate in the total family $\mathbf{v}_{\alpha i}$ is nonincreasing, and the minimum is nondecreasing. Therefore the system (1.2) is well prepared "as is" for establishing global flocking behavior. However, this is not the case for each individual flock. Each flock satisfies "internal maximum principle" relative to its own time-dependent momentum $\mathbf{V}_\alpha$. This dictates passage to the reference frame evolving with that momentum and center of mass:

$$\mathbf{w}_{\alpha i} = \mathbf{v}_{\alpha i} - \mathbf{V}_\alpha, \quad \mathbf{y}_{\alpha i} = \mathbf{x}_{\alpha i} - \mathbf{X}_\alpha. \tag{3.3}$$

Using (1.2) and (1.3) one readily obtains the system

$$\begin{cases} \dot{\mathbf{y}}_{\alpha i} = \mathbf{w}_{\alpha i}, \\ \dot{\mathbf{w}}_{\alpha i} = \lambda_\alpha \displaystyle\sum_{j=1}^{N_\alpha} m_{\alpha j}\phi_{\alpha ij}(\mathbf{w}_{\alpha i} - \mathbf{w}_{\alpha j}) - \mu R_\alpha(t)\mathbf{w}_{\alpha i}, \end{cases} \tag{3.4}$$

where

$$R_\alpha(t) := \sum_{\beta \neq \alpha} M_\beta \psi(|\mathbf{X}_\alpha - \mathbf{X}_\beta|),$$

and we abbreviate

$$\phi_{\alpha ij} = \phi_\alpha(\mathbf{y}_{\alpha i} + \mathbf{X}_\alpha, \mathbf{y}_{\alpha j} + \mathbf{X}_\alpha).$$

This system now does have a maximum principle and is well prepared for establishing flocking.

Let us denote individual flock parameters:

$$\mathcal{D}_\alpha(t) := \max_{i,j=1,\dots,N_\alpha} |\mathbf{x}_{\alpha i}(t) - \mathbf{x}_{\alpha j}(t)|, \quad \mathcal{A}_\alpha = \max_{i,j=1,\dots,N_\alpha} |\mathbf{w}_{\alpha i} - \mathbf{w}_{\alpha j}| = \max_{\substack{\boldsymbol{\ell} \in \mathbb{R}^n : |\boldsymbol{\ell}|=1 \\ i,j=1,\dots,N_\alpha}} \langle \boldsymbol{\ell}, \mathbf{w}_{\alpha i} - \mathbf{w}_{\alpha j} \rangle.$$

By Rademacher's lemma, we can evaluate the derivative of $\mathcal{A}_\alpha$ by considering $\boldsymbol{\ell}, i, j$ at which that maximum is achieved at any instance of time:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{A}_\alpha = \langle \boldsymbol{\ell}, \dot{\mathbf{w}}_{\alpha i} - \dot{\mathbf{w}}_{\alpha j} \rangle &= \lambda_\alpha \sum_{k=1}^{N_\alpha} m_{\alpha k} \phi_{\alpha i k} \langle \boldsymbol{\ell}, \mathbf{w}_{\alpha k} - \mathbf{w}_{\alpha i} \rangle \\ &\quad - \lambda_\alpha \sum_{k=1}^{N_\alpha} m_{\alpha k} \phi_{\alpha j k} \langle \boldsymbol{\ell}, \mathbf{w}_{\alpha k} - \mathbf{w}_{\alpha j} \rangle \\ &\quad - \mu R_\alpha(t) \langle \boldsymbol{\ell}, \mathbf{w}_{\alpha i} - \mathbf{w}_{\alpha j} \rangle \\ &= \lambda_\alpha \sum_{k=1}^{N_\alpha} m_{\alpha k} \phi_{\alpha i k} (\langle \boldsymbol{\ell}, \mathbf{w}_{\alpha k} - \mathbf{w}_{\alpha j} \rangle - \langle \boldsymbol{\ell}, \mathbf{w}_{\alpha i} - \mathbf{w}_{\alpha j} \rangle) \\ &\quad + \lambda_\alpha \sum_{k=1}^{N_\alpha} m_{\alpha k} \phi_{\alpha j k} (\langle \boldsymbol{\ell}, \mathbf{w}_{\alpha i} - \mathbf{w}_{\alpha k} \rangle - \langle \boldsymbol{\ell}, \mathbf{w}_{\alpha i} - \mathbf{w}_{\alpha j} \rangle) - \mu R_\alpha(t) \mathcal{A}_\alpha. \end{aligned}$$

Each difference of the action of $\boldsymbol{\ell}$ is negative due to maximality of $\boldsymbol{\ell}, i, j$. Hence, we replace values of ϕ_α's with the use of (3.1) and its minimal value at $\mathcal{D}_\alpha$:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{A}_\alpha \leqslant \lambda_\alpha \phi_\alpha(\mathcal{D}_\alpha) \sum_{k=1}^{N_\alpha} m_{\alpha k} (&\langle \boldsymbol{\ell}, \mathbf{w}_{\alpha k} - \mathbf{w}_{\alpha j} \rangle - \langle \boldsymbol{\ell}, \mathbf{w}_{\alpha i} - \mathbf{w}_{\alpha j} \rangle \\ &+ \langle \boldsymbol{\ell}, \mathbf{w}_{\alpha i} - \mathbf{w}_{\alpha k} \rangle - \langle \boldsymbol{\ell}, \mathbf{w}_{\alpha i} - \mathbf{w}_{\alpha j} \rangle) \\ &- \mu R_\alpha(t) \mathcal{A}_\alpha = -\lambda_\alpha M_\alpha \phi_\alpha(\mathcal{D}_\alpha) \mathcal{A}_\alpha - \mu R_\alpha(t) \mathcal{A}_\alpha. \end{aligned}$$

At the same time, $R_\alpha(t) \geqslant M\psi(\mathcal{D})$, where

$$\mathcal{D} := \max_{\alpha,\beta} |\mathbf{X}_\alpha - \mathbf{X}_\beta|, \quad \mathcal{A} := \max_{\alpha,\beta} |\mathbf{V}_\alpha - \mathbf{V}_\beta|.$$

Combining it with the system for $(\mathcal{D}, \mathcal{A})$ which follows a similar computation applied to macroscopic values (1.3), we arrive at the following system of ordinary differential inequalities (ODIs):

$$\begin{cases} \dot{\mathcal{A}}_\alpha \leqslant -\lambda_\alpha M_\alpha \phi_\alpha(\mathcal{D}_\alpha) \mathcal{A}_\alpha - \mu M \psi(\mathcal{D}) \mathcal{A}_\alpha \\ \dot{\mathcal{D}}_\alpha \leqslant \mathcal{A}_\alpha \\ \dot{\mathcal{A}} \leqslant -\mu M \psi(\mathcal{D}) \mathcal{A} \\ \dot{\mathcal{D}} \leqslant \mathcal{A}. \end{cases} \tag{3.5}$$

This system encompasses prototypical systems of the form

$$\begin{cases} \dot{A} \leqslant -\gamma\phi(D)A \\ \dot{D} \leqslant A. \end{cases} \tag{3.6}$$

Following Ha and Liu [9] we can define a Lyapunov function

$$L = A + \gamma \int_0^D \phi(r)\,\mathrm{d}r,$$

which is nonincreasing. Hence, there exists $\overline{D}$ and $\delta > 0$ such that

$$\int_{D_0}^{\infty} \phi(r)\,\mathrm{d}r > \frac{A_0}{\gamma} \quad \rightsquigarrow \quad D(t) \leqslant \overline{D},\ A(t) \leqslant A_0 e^{-\gamma\phi(\overline{D})t}. \tag{3.7}$$

Note that condition (3.7) is always satisfied for a fat tail ϕ.

Going back to (3.5) and ignoring the term $-\mu M\psi(\mathcal{D})\mathcal{A}_\alpha$ in the $\mathcal{A}_\alpha$ equation we observe that the α-flock completely decouples from the rest of the multiflock. We arrive at (3.6) for the pair $(\mathcal{D}_\alpha, \mathcal{A}_\alpha)$. One obtains the fast internal alignment result (1.5) asserted in Theorem 1.2

$$\max_i |\mathbf{v}_{\alpha i}(t) - \mathbf{V}_\alpha(t)| \lesssim e^{-\delta_\alpha t}, \qquad \delta_\alpha = \lambda_\alpha M_\alpha \phi_\alpha(\overline{\mathcal{D}}_\alpha).$$

As noted before, this indicates that the α-flock behavior depends solely only on its own parameters but not the global values. In particular, the α-flock alignment towards $\mathbf{V}_\alpha(t)$ occurs regardless whether these centers of momentum align or not. The latter will be guaranteed if the interflock communication ψ satisfies the fat tail condition (3.2). In fact, in this case the global alignment ensues even if internal communications are completely absent. This is evident from (3.5) where we ignore the $-\lambda_\alpha M_\alpha \phi_\alpha(\mathcal{D}_\alpha)\mathcal{A}_\alpha$ term and obtain boundedness of $\mathcal{D}$ from the last two equations, obtaining the slow alignment (1.6) asserted in Theorem 1.3

$$\max_{\alpha,i} |\mathbf{v}_{\alpha i}(t) - \mathbf{V}| \lesssim e^{-\delta t}, \qquad \delta = \mu M \psi(\overline{\mathcal{D}}).$$

Alignment rate in this case is slow since it depends on μ and the global diameter of the multiflock.

Remark 3.1 (asymptotic rate). Asymptotic dependence of the implied alignment rates for small μ and large λ_α for the CS kernel can be worked out from (3.7) (we omit the details). In the context of fast local alignment with $\phi_\alpha(r) \sim r^{-\eta_a}$ we obtain $\delta \sim \lambda_\alpha$ for all $\eta_\alpha \leqslant 1$, while in the context of slow alignment with $\psi(r) = r^{-\zeta}$ we obtain

$$\delta \sim \begin{cases} \mu^{\frac{1}{1-\zeta}}, & \zeta < 1, \\ \mu e^{-1/\mu}, & \zeta = 1. \end{cases}$$

4. Multiflocking driven by alignment and attraction. In this section we consider multiflock alignment model with additional attraction forces. Out goal is to show that each flock would aggregate towards its center of mass within the radius of influence of the potential. Our results present an extension of [16].

We assume that the interactions are determined by a radially symmetric smooth potential $U \in C^2(\mathbb{R}_+)$:

$$\begin{cases} \dot{\mathbf{x}}_{\alpha i} = \mathbf{v}_{\alpha i}, \\ \dot{\mathbf{v}}_{\alpha i} = \dfrac{1}{N_\alpha} \displaystyle\sum_{j=1}^{N_\alpha} \phi_\alpha(|\mathbf{x}_{\alpha i} - \mathbf{x}_{\alpha j}|)(\mathbf{v}_{\alpha j} - \mathbf{v}_{\alpha i}) + \mu \sum_{\substack{\beta=1 \\ \beta \neq \alpha}}^{A} \psi(|\mathbf{X}_\alpha - \mathbf{X}_\beta|)(\mathbf{V}_\beta - \mathbf{v}_{\alpha i}) + \mathbf{F}_\alpha(t), \end{cases} \tag{4.1}$$

where

$$\mathbf{F}_{\alpha i}(t) = -\frac{1}{N_\alpha} \sum_{j=1}^{N_\alpha} \nabla U(|\mathbf{x}_{\alpha i} - \mathbf{x}_{\alpha j}|).$$

Here we assumed for notational simplicity that all masses are $1/N_\alpha$, and potentials are the same. However, the statements below can easily be carried out for a general set of parameters.

Note that the system upscales to the same CS system (1.3) for the flock-level variables.

Using transformation (3.3), we rewrite the system in the new coordinate frame

$$\begin{cases} \dot{\mathbf{y}}_{\alpha i} = \mathbf{w}_{\alpha i}, \\ \dot{\mathbf{w}}_{\alpha i} = \dfrac{1}{N_\alpha} \displaystyle\sum_{j=1}^{N_\alpha} \phi_{\alpha ij}(\mathbf{w}_{\alpha i} - \mathbf{w}_{\alpha j}) - \mu R_\alpha(t) \mathbf{w}_{\alpha i} + \mathbf{F}_{\alpha i}(t). \end{cases} \tag{4.2}$$

The classical energy $\mathcal{E}_\alpha = \mathcal{K}_\alpha + \mathcal{P}_\alpha$ where[2]

$$\begin{aligned} \mathcal{K}_\alpha &:= \frac{1}{2N_\alpha} \sum_{i=1}^{N_\alpha} |\mathbf{w}_{\alpha i}|^2 = \frac{1}{4N_\alpha^2} \sum_{i=1}^{N_\alpha} |\mathbf{w}_{\alpha ij}|^2, \qquad \mathbf{w}_{\alpha ij} = \mathbf{w}_{\alpha i} - \mathbf{w}_{\alpha j}, \\ \mathcal{P}_\alpha &:= \frac{1}{2N_\alpha^2} \sum_{i,j=1}^{N_\alpha} U(|\mathbf{y}_{\alpha ij}|), \qquad \mathbf{y}_{\alpha ij} = \mathbf{y}_{\alpha i} - \mathbf{y}_{\alpha j}, \end{aligned} \tag{4.3}$$

satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{E}_\alpha = -\frac{1}{N_\alpha^2} \sum_{i,j=1}^{N_\alpha} \phi_{\alpha ij} |\mathbf{w}_{\alpha ij}|^2 - \mu R_\alpha(t) \mathcal{K}_\alpha.$$

Denoting the dissipation term by

$$\mathcal{I}_\alpha := \frac{1}{N_\alpha^2} \sum_{i,j=1}^{N_\alpha} \phi_{\alpha ij} |\mathbf{w}_{\alpha ij}|^2,$$

we obtain the energy law

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{E}_\alpha = -\mathcal{I}_\alpha - \mu R_\alpha(t) \mathcal{K}_\alpha. \tag{4.4}$$

At this stage already we can see that if $\mu > 0$ and ψ has a fat tail, then global slow exponential alignment will ensue regardless of internal flock communications.

[2]Here and in what follows we occasionally use a shortcut for $\mathbf{a}_{\alpha ij} = \mathbf{a}_{\alpha i} - \mathbf{a}_{\alpha j}$.

Indeed, the up-scaled dynamics (1.3) will stabilize the macroscopic values which implies boundedness of R_α. Hence, ignoring dissipation term $\mathcal{I}_\alpha$ in (4.4) we obtain exponential decay of all the energies:

$$\mathcal{E}_\alpha \lesssim e^{-c\mu t}.$$

In this section we show that flocking occurs also in each individual α-flock regardless of global communication, although it may be happening at a slower rate. To fix the notation we consider regular communication kernels with power-like decay:

$$\phi_\alpha'(r) \leqslant 0, \quad \phi_\alpha(r) \geqslant \frac{c_0}{\langle r\rangle^\gamma} \text{ for } r \geqslant 0. \tag{4.5}$$

For the potential we assume essentially a power law: for some $\beta > 1$ and $L' > L > 0$,

$$\begin{aligned} &\text{Support:} && U \in C^2(\mathbb{R}^+), \quad U(r) = 0 \quad \forall r \leqslant L, \\ &\text{Growth:} && U(r) \geqslant a_0 r^\beta, \quad |U'(r)| \leqslant a_1 r^{\beta-1}, \quad |U''(r)| \leqslant a_2 r^{\beta-2} \quad \forall r > L', \\ &\text{Convexity:} && U'(r), U''(r) \geqslant 0 \quad \forall r > 0. \end{aligned} \tag{4.6}$$

THEOREM 4.1 (local flocking with interaction potential). *Under the assumptions* (4.5) *and* (4.6) *on the kernel and potential in the range of parameters given by*

$$\gamma < \begin{cases} 1, & 1 < \beta < \dfrac{4}{3}, \\ \dfrac{3}{2}\beta - 1, & \dfrac{4}{3} \leqslant \beta < 2, \\ 2, & \beta \geqslant 2, \end{cases} \tag{4.7}$$

all solutions to the system (4.1) *flock with the bound independent of* N_α*:*

$$\mathcal{D}_\alpha(t) < \overline{\mathcal{D}}_\alpha \quad \forall t > 0,$$

asymptotically aggregate

$$\limsup_{t\to\infty} \mathcal{D}_\alpha(t) \leqslant L,$$

and align

$$\frac{1}{2N_\alpha}\sum_{i=1}^{N_\alpha} |\mathbf{v}_{\alpha i} - \mathbf{V}_\alpha|^2 + \frac{1}{2N_\alpha^2}\sum_{i,j=1}^{N_\alpha} U(|\mathbf{x}_{\alpha i} - \mathbf{x}_{\alpha j}|) \leqslant \frac{C_\delta}{\langle t\rangle^{1-\delta}} \quad \forall \delta > 0. \tag{4.8}$$

Note that the latter statement follows from local alignments (4.8) and the exponential alignment of the flock parameters governed by the upscaled system (1.3).

Proof. We will operate with the particle energy defined similarly to [16]

$$\mathcal{E}_{\alpha i} = \frac{1}{2}|\mathbf{w}_{\alpha i}|^2 + \frac{1}{N_\alpha}\sum_{k=1}^{N_\alpha} U(|\mathbf{y}_{\alpha i k}|), \quad \mathcal{E}_{\alpha\infty} = \max_i \mathcal{E}_{\alpha i}.$$

First, we observe that the particle energy controls the diameter of the flock. Indeed, by convexity and our assumptions on the growth of the potential, we have

$$\mathcal{E}_{\alpha i} \geqslant U(|\mathbf{y}_{\alpha i}|) \geqslant (|\mathbf{y}_{\alpha i}| - L')_+^\beta. \tag{4.9}$$

So,

$$\mathcal{D}_\alpha \leqslant \mathcal{E}_{\alpha\infty}^{1/\beta} + L'. \tag{4.10}$$

Let us now establish a bound on $\mathcal{E}_{\alpha\infty}$. For each i we test (4.2) with $\mathbf{w}_{\alpha i}$ and ignore that R_α-term:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_{\alpha i} \leqslant \frac{1}{N_\alpha}\sum_{k=1}^{N_\alpha} \phi_{\alpha ik}\mathbf{w}_{\alpha ki}\cdot \mathbf{w}_{\alpha i} - \frac{1}{N_\alpha}\sum_{k=1}^{N_\alpha} \nabla U(|\mathbf{y}_{\alpha ik}|)\cdot \mathbf{w}_{\alpha k}. \tag{4.11}$$

For the kinetic part we use the vector identity

$$\mathbf{a}_{ki}\cdot\mathbf{a}_i = -\frac{1}{2}|\mathbf{a}_{ki}|^2 - \frac{1}{2}|\mathbf{a}_i|^2 + \frac{1}{2}|\mathbf{a}_k|^2. \tag{4.12}$$

Discarding all the negative terms, we bound

$$\frac{1}{N_\alpha}\sum_{k=1}^{N_\alpha} \phi_{\alpha ik}\mathbf{w}_{\alpha ki}\cdot\mathbf{w}_{\alpha i} \leqslant |\phi_\alpha|_\infty \mathcal{K}_\alpha.$$

Due to the energy law $\mathcal{K}_\alpha$ will remain bounded, but we will keep it in the bound above for now. As for the potential term, there are several ways we can handle it.

For any $1 \leqslant \beta \leqslant \frac{4}{3}$ we apply a direct estimate from the first derivative:

$$\left|\frac{1}{N_\alpha}\sum_{k=1}^{N_\alpha}\nabla U(|\mathbf{y}_{\alpha ik}|)\cdot\mathbf{w}_{\alpha k}\right| \leqslant \sqrt{\mathcal{K}_\alpha}\left(\frac{1}{N_\alpha}\sum_{k=1}^{N_\alpha}|\nabla U(|\mathbf{y}_{\alpha ik}|)|^2\right)^{\frac{1}{2}} \leqslant \sqrt{\mathcal{K}_\alpha}\mathcal{D}_\alpha^{\beta-1}.$$

Consequently,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_{\alpha i} \leqslant c_1\mathcal{K}_\alpha + c_2\sqrt{\mathcal{K}_\alpha}\mathcal{D}_\alpha^{\beta-1} \lesssim \sqrt{\mathcal{K}_\alpha}(1+\mathcal{E}_{\alpha\infty}^{\frac{\beta-1}{\beta}}),$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_{\alpha\infty} \leqslant c_3\sqrt{\mathcal{K}_\alpha}(1+\mathcal{E}_{\alpha\infty}^{\frac{\beta-1}{\beta}}) \quad\Rightarrow\quad \mathcal{E}_{\alpha\infty} \lesssim \langle t\rangle^\beta \quad\Rightarrow\quad \mathcal{D}_\alpha \lesssim \langle t\rangle. \tag{4.13}$$

In the range $\frac{4}{3} \leqslant \beta \leqslant 2$ it is better to make use of the second derivative:

$$\begin{aligned}\left|\frac{1}{N_\alpha}\sum_{k=1}^{N_\alpha}\nabla U(|\mathbf{y}_{\alpha ik}|)\cdot\mathbf{w}_{\alpha k}\right| &= \frac{1}{N_\alpha}\sum_{k=1}^{N_\alpha}(\nabla U(|\mathbf{y}_{\alpha ik}|) - \nabla U(|\mathbf{y}_{\alpha i}|))\cdot\mathbf{v}_k \\ &\leqslant \|D^2U\|_\infty\sqrt{\mathcal{K}_\alpha}\left(\frac{1}{N_\alpha}\sum_{k=1}^{N_\alpha}|\mathbf{y}_{\alpha k}|^2\right)^{\frac{1}{2}} \\ &\leqslant c_4\sqrt{\mathcal{K}_\alpha}\left(\frac{1}{N_\alpha^2}\sum_{i,j=1}^{N_\alpha}|\mathbf{y}_{\alpha ij}|^2\right)^{\frac{1}{2}}.\end{aligned} \tag{4.14}$$

The following inequality will be used repeatedly

$$\frac{1}{N_\alpha^2}\sum_{i,j=1}^{N_\alpha}|\mathbf{y}_{\alpha ij}|^2 \leqslant (L')^2 + \frac{1}{N_\alpha^2}\sum_{i,j=1}^{N_\alpha}(|\mathbf{y}_{\alpha ij}| - L')_+^2 \leqslant C(1+\mathcal{D}_\alpha^{(2-\beta)_+}\mathcal{P}_\alpha). \tag{4.15}$$

Continuing the above,

$$\left|\frac{1}{N_\alpha}\sum_{k=1}^{N_\alpha}\nabla U(|\mathbf{y}_{\alpha ik}|)\cdot\mathbf{w}_{\alpha k}\right| \leqslant c_4\sqrt{\mathcal{K}_\alpha}(1+\mathcal{D}_\alpha^{2-\beta}\mathcal{P}_\alpha)^{1/2} \leqslant c_5\sqrt{\mathcal{K}_\alpha}(1+\mathcal{E}_{\alpha\infty})^{\frac{2-\beta}{2\beta}}.$$

In this case,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_{\alpha\infty} \leqslant c_6\sqrt{\mathcal{K}_\alpha}(1+\mathcal{E}_{\alpha\infty})^{\frac{2-\beta}{2\beta}} \quad\Rightarrow\quad \mathcal{E}_{\alpha\infty} \lesssim \langle t\rangle^{\frac{2\beta}{3\beta-2}} \quad\Rightarrow\quad \mathcal{D}_\alpha \leqslant \langle t\rangle^{\frac{2}{3\beta-2}}. \tag{4.16}$$

Finally, for $\beta > 2$, we argue similarly, using that $|D^2U(|\mathbf{y}_{\alpha ik}|)| \leqslant \mathcal{D}_\alpha^{\beta-2}$, and (4.15), to obtain

$$\left|\frac{1}{N_\alpha}\sum_{k=1}^{N_\alpha}\nabla U(|\mathbf{y}_{\alpha ik}|)\cdot\mathbf{w}_{\alpha k}\right| \leqslant \sqrt{\mathcal{K}_\alpha}\mathcal{D}_\alpha^{\beta-2},$$

and hence,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_{\alpha\infty} \leqslant c_7\sqrt{\mathcal{K}_\alpha}(1+\mathcal{E}_{\alpha\infty})^{\frac{\beta-2}{\beta}} \quad\Rightarrow\quad \mathcal{E}_{\alpha\infty} \lesssim \langle t\rangle^{\frac{\beta}{2}} \quad\Rightarrow\quad \mathcal{D}_\alpha \leqslant \langle t\rangle^{\frac{1}{2}}. \tag{4.17}$$

We have proved the following a priori estimate:

$$\mathcal{D}_\alpha(t) \lesssim \langle t\rangle^d, \qquad \text{where} \qquad d = \begin{cases} 1, & 1 \leqslant \beta < \frac{4}{3}, \\ \frac{2}{3\beta-2}, & \frac{4}{3} \leqslant \beta < 2, \\ \frac{1}{2}, & \beta \geqslant 2. \end{cases} \tag{4.18}$$

Denote $\zeta(t) = \langle t\rangle^{-\gamma d}$. Then according to the basic energy equation (4.4) we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_\alpha \leqslant -\frac{1}{2}\mathcal{I}_\alpha - c\zeta(t)\mathcal{K}_\alpha - \mu R_\alpha(t)\mathcal{K}_\alpha. \tag{4.19}$$

Considering this as a starting point, just like in the quadratic confinement case, we will build correctors to the energy to achieve full coercivity on the right-hand side of (4.19). We introduce one more auxiliary power function

$$\eta(t) = \langle t\rangle^{-a}, \quad \gamma d \leqslant a < 1.$$

First, we consider the same longitudinal momentum

$$\mathcal{X}_\alpha = \frac{1}{N_\alpha}\sum_{i=1}^{N_\alpha}\mathbf{y}_{\alpha i}\cdot\mathbf{w}_{\alpha i}.$$

It will come with a prefactor $\delta\eta(t)$, where δ is a small parameter. Let us estimate using (4.15):

$$\delta\eta(t)|\mathcal{X}_\alpha| \leqslant \delta\mathcal{K}_\alpha + \delta\eta^2(t)\frac{1}{N_\alpha^2}\sum_{i,j=1}^{N_\alpha}|\mathbf{y}_{\alpha ij}|^2 \leqslant \delta\mathcal{K}_\alpha + c\delta\eta^2(t) + \delta\eta^2(t)\mathcal{D}_\alpha^{(2-\beta)_+}\mathcal{P}_\alpha.$$

The potential term is bounded by $\delta\mathcal{P}_\alpha$ as long as $2a \geqslant d(2-\beta)_+$. Hence,

$$\delta\eta(t)|\mathcal{X}_\alpha| \leqslant \delta\mathcal{E}_\alpha + c\eta^2(t). \tag{4.20}$$

This shows that

$$\mathcal{E}_\alpha + \delta\eta(t)\mathcal{X}_\alpha + 2c\eta^2(t) \sim \mathcal{E}_\alpha + c\delta\eta^2(t).$$

Let us now consider the derivative

$$\mathcal{X}'_\alpha = \frac{1}{N_\alpha}\sum_{i=1}^{N_\alpha}|\mathbf{w}_{\alpha i}|^2 + \frac{1}{N_\alpha^2}\sum_{i,k=1}^{N_\alpha}\mathbf{y}_{\alpha ik}\cdot\mathbf{w}_{\alpha ki}\phi_{\alpha ki} - \frac{1}{N_\alpha^2}\sum_{i,k=1}^{N_\alpha}\mathbf{y}_{\alpha ik}\cdot\nabla U(|\mathbf{y}_{\alpha ik}|) - \mu R_\alpha(t)\mathcal{X}_\alpha$$
$$= \mathcal{K}_\alpha + A - B - \mu R_\alpha(t)\mathcal{X}_\alpha.$$

The gain term B, by convexity dominates the potential energy $B \geqslant \mathcal{P}_\alpha$. As to A:

$$|A| \leqslant \frac{|\phi|_\infty}{2\delta^{1/2}\eta(t)}\mathcal{I}_\alpha + \frac{\delta^{1/2}\eta(t)}{2}\frac{1}{N_\alpha^2}\sum_{i,j=1}^{N_\alpha}|\mathbf{y}_{\alpha ij}|^2 \lesssim \frac{1}{\delta^{1/2}\eta(t)}\mathcal{I} + \delta^{1/2}\eta(t) + \delta^{1/2}\eta(t)\mathcal{D}_\alpha^{(2-\beta)_+}\mathcal{P}_\alpha.$$

By requiring a more stringent assumption on parameters

$$\alpha \geqslant d(2-\beta)_+,$$

we can ensure that the potential term is bounded by $\sim \mu^{1/2}\mathcal{P}$, which can be absorbed by the gain term.

The interflock term in (4.19) helps abosrb the corresponding residual term $\mu R_\alpha(t)\mathcal{X}_\alpha$. Indeed,

$$\mu R_\alpha(t)\mathcal{X}_\alpha \leqslant \frac{1}{2\delta\eta(t)}\mu R_\alpha(t)\mathcal{K}_\alpha + \mu R_\alpha(t)\delta\eta(t)\sum_{i,j=1}^{N_\alpha}|\mathbf{y}_{\alpha ij}|^2$$
$$\leqslant \frac{1}{2\delta\eta(t)}\mu R_\alpha(t)\mathcal{K}_\alpha + C_1\delta\eta(t) + C_2\delta\eta(t)\mathcal{D}_\alpha^{(2-\beta)_+}\mathcal{P}_\alpha,$$

with the latter absorbed into the gain term as in the case of A.

So far, we have obtained

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{E}_\alpha + \delta\eta(t)\mathcal{X}_\alpha + 2c\eta^2(t)) \leqslant -c_1\delta\eta(t)\mathcal{E} + c_2\eta^2(t) + \delta\eta'(t)\mathcal{X}_\alpha. \tag{4.21}$$

In view of (4.20),

$$|\delta\eta'(t)\mathcal{X}_\alpha| \leqslant \delta\frac{1}{\langle t\rangle}\eta(t)|\mathcal{X}_\alpha| \leqslant \delta\frac{1}{\langle t\rangle}\mathcal{E}_\alpha + \delta\frac{\eta^2(t)}{\langle t\rangle}.$$

Since $a < 1$, the energy term will be absorbed, and the free term is even smaller then η^2. Denoting

$$E_\alpha = \mathcal{E}_\alpha + \delta\eta(t)\mathcal{X}_\alpha + 2c\eta^2(t),$$

we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}E_\alpha \leqslant -c_1\eta(t)E_\alpha + c_2\eta^2(t).$$

By Duhamel's formula,

$$E_\alpha(t) \lesssim \exp\{-\langle t\rangle^{1-a}\} + \exp\{-\langle t\rangle^{1-a}\}\int_0^t \frac{e^{\langle s\rangle^{1-a}}}{\langle s\rangle^{2a}}\,\mathrm{d}s.$$

By an elementary asymptotic analysis,

$$\int_0^t \frac{e^{\langle s\rangle^{1-a'}}}{\langle s\rangle^{a''}}\,\mathrm{d}s \sim \exp\{\langle t\rangle^{1-a'}\}\frac{1}{\langle t\rangle^{a''-a'}}.$$

Thus, we obtain an algebraic decay rate

$$E_\alpha(t) \lesssim \frac{1}{\langle t\rangle^a} \quad \forall a<1, \tag{4.22}$$

provided

$$d\gamma < 1 \quad \text{and} \quad d(2-\beta)_+ < 1. \tag{4.23}$$

This translates exactly into the conditions on γ given by (4.7), and (4.22) automatically implies (4.8)

Going back to the estimates (4.13) and (4.16) but keeping the kinetic energy with its established decay, we obtain a new decay rate for the diameter

$$\mathcal{D}_\alpha \leqslant C_\delta \langle t\rangle^{\frac{d}{2}+\delta} \quad \forall \delta>0.$$

At the next stage we prove flocking: $\mathcal{D}_\alpha(t) < \bar{D}_\alpha$. In order to achieve this we return again to the particle energy estimates. Let us denote

$$\mathcal{P}_{\alpha i} = \frac{1}{N_\alpha}\sum_{k=1}^{N_\alpha} U(|\mathbf{y}_{\alpha ik}|), \quad \mathcal{I}_{\alpha i} = \frac{1}{N_\alpha}\sum_{k=1}^{N_\alpha} \phi_{\alpha ik}|\mathbf{w}_{\alpha ki}|^2, \quad \mathcal{X}_{\alpha i} = \mathbf{y}_{\alpha i}\cdot \mathbf{w}_{\alpha i}.$$

Using (4.11), (4.12), (4.14), (4.15), and the fact that $\mathcal{D}_\alpha^{(2-\beta)_+}\mathcal{P}$ has a negative rate of decrease, we obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_{\alpha i} &\leqslant \mathcal{K}_\alpha - \frac12\phi_\alpha(\mathcal{D}_\alpha)|\mathbf{w}_{\alpha i}|^2 - \mathcal{I}_{\alpha i} + c\sqrt{\mathcal{K}_\alpha} - \mu R_\alpha(t)|\mathbf{w}_{\alpha i}|^2\\ &\lesssim -\frac12\phi_\alpha(\mathcal{D}_\alpha)|\mathbf{w}_{\alpha i}|^2 - \mathcal{I}_{\alpha i} + \frac{1}{\langle t\rangle^{\frac12-\delta}} - \mu R_\alpha(t)|\mathbf{w}_{\alpha i}|^2 \quad \forall\delta>0.\end{aligned}$$

In view of (4.23), we can pick a and small b such that

$$\begin{gathered}\frac{d\gamma}{2} + b\gamma < \frac12 - 2b < a < \frac12 - b\\ (2-\beta)_+ d + 2\delta(2-\beta)_+ < 2a.\end{gathered} \tag{4.24}$$

We use as before the auxiliary rate function $\eta(t) = \langle t\rangle^{-a}$. Let us estimate the corrector

$$\begin{aligned}|\delta\eta(t)\mathcal{X}_{\alpha i}| \leqslant \mu|\mathbf{w}_{\alpha i}|^2 + \delta\eta^2(t)|\mathbf{y}_{\alpha i}|^2 &\leqslant \delta|\mathbf{w}_{\alpha i}|^2 + \delta\eta^2(t)\mathcal{D}_\alpha^{2-\beta}\mathcal{P}_{\alpha i} + L^2\delta\eta^2(t)\\ &\leqslant \delta|\mathbf{w}_{\alpha i}|^2 + c\delta\mathcal{P}_{\alpha i} + L^2\delta\eta^2(t).\end{aligned}$$

So,

$$E_{\alpha i} := \mathcal{E}_{\alpha i} + \delta\eta(t)\mathcal{X}_{\alpha i} + 2L^2\delta\eta^2(t) \sim \mathcal{E}_{\alpha i} + L^2\delta\eta^2(t).$$

Differentiating,

$$\begin{aligned}\mathcal{X}'_{\alpha i} &= |\mathbf{w}_{\alpha i}|^2 + \frac{1}{N_\alpha}\sum_{k=1}^{N_\alpha}\mathbf{y}_{\alpha i}\cdot\mathbf{w}_{\alpha ki}\phi_{\alpha ki} - \frac{1}{N_\alpha}\sum_{k=1}^{N_\alpha}\mathbf{y}_{\alpha ik}\cdot\nabla U(|\mathbf{y}_{\alpha ik}|)\\ &\quad + \frac{1}{N_\alpha}\sum_{k=1}^{N_\alpha}\mathbf{y}_{\alpha k}\cdot(\nabla U(|\mathbf{y}_{\alpha ik}|) - \nabla U(|\mathbf{y}_{\alpha i}|)) - \mu R_\alpha(t)\mathcal{X}_{\alpha i}\\ &\leqslant |\mathbf{w}_{\alpha i}|^2 + \delta^{1/2}\eta(t)|\mathbf{y}_{\alpha i}|^2 + \frac{1}{\delta^{1/2}\eta(t)}\mathcal{I}_{\alpha i} - \mathcal{P}_{\alpha i} + \frac{1}{N_\alpha^2}\sum_{l,k=1}^{N_\alpha}|\mathbf{y}_{\alpha kl}|^2\\ &\quad + \frac{1}{2\delta\eta(t)}\mu R_\alpha(t)|\mathbf{w}_{\alpha i}|^2 + 2\delta\eta(t)\mu R_\alpha(t)|\mathbf{y}_{\alpha i}|^2,\end{aligned}$$

where the last term is already smaller than $\delta^{1/2}\eta(t)|\mathbf{y}_{\alpha i}|^2$ for small enough δ,

$$\leqslant |\mathbf{w}_{\alpha i}|^2 + \delta^{1/2}L^2\eta(t) + \delta^{1/2}\mathcal{D}_\alpha^{(2-\beta)_+}\eta(t)\mathcal{P}_{\alpha i} + \frac{1}{\delta^{1/2}\eta(t)}\mathcal{I}_{\alpha i} - \mathcal{P}_{\alpha i} + C + \frac{1}{2\delta\eta(t)}\mu R_\alpha(t)|\mathbf{w}_{\alpha i}|^2$$

in view of (4.24), $\mu^{1/2}\mathcal{D}^{(2-\beta)_+}\eta(t) \lesssim \mu^{1/2}$, so the potential term is absorbed by $-\mathcal{P}_i$,

$$\leqslant |\mathbf{w}_{\alpha i}|^2 + \frac{1}{\eta(t)}\mathcal{I}_{\alpha i} - \frac{1}{2}\mathcal{P}_{\alpha i} + C + \frac{1}{2\delta\eta(t)}\mu R_\alpha(t)|\mathbf{w}_{\alpha i}|^2.$$

Again in view of (4.24), $\eta(t)$ decays faster than $\phi_\alpha(\mathcal{D}_\alpha)$, so plugging into the energy equation we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}E_{\alpha i} \leqslant -\delta\eta(t)E_{\alpha i} + \eta(t) + \sqrt{\mathcal{K}_\alpha} + \delta\eta'(t)\mathcal{X}_{\alpha i},$$

and as before $\delta\eta'(t)\mathcal{X}_{\alpha i}$ is a lower-order term which is absorbed into the negative energy term and $+\eta^2$. So,

$$\frac{\mathrm{d}}{\mathrm{d}t}E_{\alpha i} \leqslant -\delta\eta(t)E_{\alpha i} + \eta(t) + \sqrt{\mathcal{K}_\alpha}.$$

By our choice of constants (4.24), $\sqrt{\mathcal{K}_\alpha}$ decays faster than $\eta(t)$; hence,

$$\frac{\mathrm{d}}{\mathrm{d}t}E_{\alpha i} \lesssim -\delta\eta(t)E_{\alpha i} + \eta(t).$$

This proves boundedness of $E_{\alpha i}$, and hence that of $\mathcal{E}_{\alpha i} + L^2\delta\eta^2(t)$, and hence that of $\mathcal{E}_{\alpha i}$. In view of (4.10), this implies the flocking bound $\mathcal{D}_\alpha(t) < \overline{\mathcal{D}}_\alpha$ for all $t > 0$. ∎

It is interesting to note that when the support of the potential spans the entire line, $L = 0$, and U lands at the origin with at least a quadratic touch:

$$U(r) \geqslant a_0 r^2, r < L', \tag{4.25}$$

then we can establish exponential alignment in terms of the energy $\mathcal{E}_\alpha$. Indeed, since we already know that the diameter is bounded, the basic energy equation reads

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_\alpha \leqslant -c_0\mathcal{K}_\alpha - \frac{1}{2}\mathcal{I}_\alpha.$$

The momentum corrector needs only an δ-prefactor to satisfy the bound

$$|\delta\mathcal{X}_\alpha| \leqslant \delta\mathcal{K}_\alpha + \delta c\mathcal{P}_\alpha.$$

This is due to the assumed quadratic order of the potential near the origin and, again, boundedness of the diameter. Hence, $\mathcal{E}_\alpha + \delta\mathcal{X}_\alpha \sim \mathcal{E}_\alpha$. The rest of the argument is similar to the general case. We obtain

$$\mathcal{X}_\alpha \lesssim \mathcal{K}_\alpha + \delta^{1/2}\mathcal{P}_\alpha + \frac{1}{\delta^{1/2}}\mathcal{I}_\alpha - \mathcal{P}_\alpha \leqslant \mathcal{K}_\alpha - \frac{1}{2}\mathcal{P}_\alpha\frac{1}{\delta^{1/2}}\mathcal{I}_\alpha.$$

Thus,

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{E}_\alpha + \delta\mathcal{X}_\alpha) \leqslant -c_1\mathcal{E}_\alpha \sim -c_1(\mathcal{E}_\alpha + \delta\mathcal{X}_\alpha).$$

This proves exponential decay of the energy $\mathcal{E}_\alpha$. Going further to consider the individual particle energies, we discover similar decays. Indeed, denoting by $\mathrm{Exp}(t)$ any quantity that decays exponentially fast, we follow the same scheme:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_{\alpha i} \leqslant -c_1|\mathbf{w}_{\alpha i}|^2 - \frac{1}{2}\mathcal{I}_{\alpha i} + \mathrm{Exp}(t), \qquad \mathrm{Exp}(t) \lesssim e^{-Ct}.$$

In view of $|\mathbf{y}_{\alpha i}|^2 \lesssim \mathcal{P}_{\alpha i}$,

$$\delta|\mathcal{X}_{\alpha i}| \leqslant \delta|\mathbf{w}_{\alpha i}|^2 + \delta\mathcal{P}_{\alpha i},$$

so $\mathcal{E}_{\alpha i} + \delta\mathcal{X}_{\alpha i} \sim \mathcal{E}_{\alpha i}$. Further following the estimates as in the proof,

$$\mathcal{X}'_{\alpha i} \lesssim |\mathbf{w}_{\alpha i}|^2 + \frac{1}{\delta^{1/2}}\mathcal{I}_{\alpha i} - \frac{1}{2}\mathcal{P}_{\alpha i}.$$

Thus,

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{E}_{\alpha i} + \delta\mathcal{X}_{\alpha i}) \leqslant -c_1(\mathcal{E}_{\alpha i} + \delta\mathcal{X}_{\alpha i}) + \mathrm{Exp}(t).$$

This establishes exponential decay for $\mathcal{E}_{\alpha\infty}$, and hence for the individual velocities. This also proves that $\mathcal{D}_\alpha(t) = \mathrm{Exp}$. So, the long time behavior here is characterized by exponential aggregation to a point.

THEOREM 4.2. *Let us assume that the support of the potential spans the entire space and* (4.25). *Then the solutions aggregate exponentially fast:*

$$\mathcal{D}_\alpha(t) + \max_i |\mathbf{v}_{\alpha i}(t) - \mathbf{V}_\alpha(t)|_\infty \leqslant Ce^{-\delta t}$$

for some $C, \delta > 0$.

5. Hydrodynamics of multiflocks. In the case of smooth communication kernels, one can formally derive the corresponding kinetic model from (1.2) via the classical BBGKY (Bogoliubov–Born–Green–Kirkwood–Yvon) hierarchy. Let $f_\alpha(x, v, t)$ denote a density distribution of the α-flock, and define the corresponding flock parameters:

$$\begin{aligned} &M_\alpha = \int_{\mathbb{R}^{2d}} f_\alpha(\mathbf{x},\mathbf{v},t)\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{v}, \quad \mathbf{X}_\alpha = \frac{1}{M_\alpha}\int_{\mathbb{R}^{2d}} \mathbf{x} f_\alpha(\mathbf{x},\mathbf{v},t)\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{v}, \\ &\mathbf{V}_\alpha = \frac{1}{M_\alpha}\int_{\mathbb{R}^{2d}} \mathbf{v} f_\alpha(\mathbf{x},\mathbf{v},t)\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{v}. \end{aligned} \tag{5.1}$$

The kinetic model reads as follows:

$$\partial_t f_\alpha + \mathbf{v}\cdot\nabla_{\mathbf{x}} f_\alpha + \lambda\nabla_{\mathbf{v}}\cdot\mathbf{Q}_\alpha(f_\alpha, f_\alpha) + \mu\nabla_{\mathbf{v}}\cdot\left[\sum_{\beta\neq\alpha} M_\beta\psi(\mathbf{X}_\alpha,\mathbf{X}_\beta)(\mathbf{V}_\beta - \mathbf{v})f_\alpha\right] = 0, \tag{5.2}$$

where

$$\mathbf{Q}_\alpha(f,f)(\mathbf{x},\mathbf{v},t) = f(\mathbf{x},\mathbf{v},t)\int_{\mathbb{R}^{2d}} \phi_\alpha(\mathbf{x},\mathbf{x}')(\mathbf{v}'-\mathbf{v})f(\mathbf{x}',\mathbf{v},t')\,\mathrm{d}\mathbf{x}'\,\mathrm{d}\mathbf{v}'. \tag{5.3}$$

The macroscopic system can be obtained, again formally, from (5.2) by considering monokinetic closure $f_\alpha = \delta_0(\mathbf{v} - \mathbf{u}_\alpha(\mathbf{x},t))\rho_\alpha(\mathbf{x},t)$. The resulting system presents as hybrid of hydrodynamic and discrete parts, where the hydrodynamic part corresponds to the classical CS dynamics within flocks, while the discrete part governs communication of a given flock with other flocks' averaged quantities. To write down the equations, we denote macroscopic variables by $(\rho_\alpha, \mathbf{u}_\alpha)_{\alpha=1}^A$,

$$\rho_\alpha(\mathbf{x},t) = \int_{\mathbb{R}^d} f_\alpha(\mathbf{x},\mathbf{v},t)\,\mathrm{d}\mathbf{v}, \quad \rho_\alpha \mathbf{u}_\alpha = \int_{\mathbb{R}^d} \mathbf{v} f_\alpha(\mathbf{x},\mathbf{v},t)\,\mathrm{d}\mathbf{v},$$

while (5.1) represent upscale parameters of the flocks. The full hydrodynamic system reads

$$\begin{cases} \partial_t \rho_\alpha + \nabla\cdot(\mathbf{u}_\alpha\rho_\alpha) = 0 \\ \partial_t \mathbf{u}_\alpha + \mathbf{u}_\alpha\cdot\nabla\mathbf{u}_\alpha = \lambda_\alpha \int_{\mathbb{R}^d} \phi_\alpha(\mathbf{x},\mathbf{y})(\mathbf{u}_\alpha(\mathbf{y})-\mathbf{u}_\alpha(\mathbf{x}))\rho_\alpha(\mathbf{y})\,\mathrm{d}\mathbf{y} \\ \qquad\qquad + \mu \sum_{\beta\neq\alpha} M_\beta \psi(\mathbf{X}_\alpha,\mathbf{X}_\beta)[\mathbf{V}_\beta - \mathbf{u}_\alpha(\mathbf{x},t)]. \end{cases} \quad \alpha = 1,\dots,A. \tag{5.4}$$

Writing the momentum equation in conservative form we obtain

$$\begin{aligned} \partial_t(\rho_\alpha\mathbf{u}_\alpha) + \nabla_x(\rho_\alpha\mathbf{u}_\alpha\otimes\mathbf{u}_\alpha) = \lambda_\alpha \int_{\mathbb{R}^d} \phi_\alpha(\mathbf{x},\mathbf{y})(\mathbf{u}_\alpha(\mathbf{y})-\mathbf{u}_\alpha(\mathbf{x}))\rho_\alpha(\mathbf{x})\rho_\alpha(\mathbf{y})\,\mathrm{d}\mathbf{y} \\ + \mu \sum_{\beta\neq\alpha} M_\beta\psi(\mathbf{X}_\alpha,\mathbf{X}_\beta)[\mathbf{V}_\beta - \mathbf{u}_\alpha(\mathbf{x},t)]\rho_\alpha(\mathbf{x}). \end{aligned} \tag{5.5}$$

Integrating (5.5) over $\mathbb{R}^d$, (5.4) up-scales to the same discrete CS system (1.3) for macroscopic parameters $\{\mathbf{X}_\alpha, \mathbf{V}_\alpha\}_\alpha$.

5.1. Slow and fast alignment of hydrodynamic multiflocks. As in the discrete case, we will deal with kernels that admit fat tail subkernels (3.1). Alignment dynamics for hydrodynamic description mimics that of the discrete one once we pass to Lagrangian coordinates. Denote by $\mathbf{x}_\alpha(\mathbf{x},t)$ the characteristic flow map of the $\mathbf{u}_\alpha$. From the continuity equation we conclude that the mass measure $\rho_\alpha(\mathbf{y},t)\,\mathrm{d}\mathbf{y}$ is the push-forward of the initial measure $\rho_\alpha(\mathbf{y},0)\,\mathrm{d}\mathbf{y}$ by the flow. So, passing to the Lagrangian coordinates $\mathbf{v}_\alpha(\mathbf{x},t) = \mathbf{u}_\alpha(\mathbf{x}_\alpha(\mathbf{x},t),t)$ we obtain

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{v}_\alpha = \lambda_\alpha \int_{\mathbb{R}^d} \phi_\alpha(\mathbf{x}_\alpha(\mathbf{x},t),\mathbf{x}_\alpha(\mathbf{y},t))(\mathbf{v}_\alpha(\mathbf{y})-\mathbf{v}_\alpha(\mathbf{x}))\rho_\alpha(\mathbf{y},0)\,\mathrm{d}\mathbf{y} \\ + \mu\sum_{\beta\neq\alpha} M_\beta\psi(\mathbf{X}_\alpha,\mathbf{X}_\beta)[\mathbf{V}_\beta - \mathbf{v}_\alpha(\mathbf{x},t)]. \end{aligned}$$

Passing to the reference frame moving with the average velocity in each flock

$$\mathbf{w}_\alpha(\mathbf{x},t) := \mathbf{v}_\alpha(\mathbf{x},t) - \mathbf{V}_\alpha(t), \tag{5.6}$$

we obtain the momentum system quite similar to its discrete counterpart (3.4)

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{w}_\alpha(\mathbf{x},t) = \lambda_\alpha \int_{\mathbb{R}^d} \phi_\alpha(\mathbf{x}_\alpha(\mathbf{x},t),\mathbf{x}_\alpha(\mathbf{y},t))(\mathbf{w}_\alpha(\mathbf{y},t) - \mathbf{w}_\alpha(\mathbf{x},t))\rho_\alpha(\mathbf{y},0)\,\mathrm{d}\mathbf{y} - \mu R_\alpha(t)\mathbf{w}_\alpha.$$

Thus, all the alignment statements of Theorem 1.2 and Theorem 1.3 carry over directly to these settings. In the original variables these translate into the following.

THEOREM 5.1. *Assuming that the initial diameter of the α-flock is finite, and ϕ_α has fat tail, the α-flock aligns at a rate dependent on λ_α:*

$$\operatorname{diam}(\operatorname{supp}\rho_\alpha(\cdot,t)) < \overline{\mathcal{D}}_\alpha, \qquad \max_{\mathbf{x}\in\operatorname{supp}\rho_\alpha(\cdot,t)} |\mathbf{u}_\alpha(\mathbf{x},t) - \mathbf{V}_\alpha(t)| \lesssim e^{-\delta_\alpha t},$$

where $\delta_\alpha = \lambda_\alpha M_\alpha \phi_\alpha(\overline{\mathcal{D}}_\alpha)$. Furthermore, if ψ has a fat tail, the kernels $\phi_\alpha \geqslant 0$ are arbitrary, and the multiflock has a finite diameter initially; then global alignment occurs at a rate dependent on μ:

$$\operatorname{diam}(\cup_\alpha \operatorname{supp}\rho_\alpha(\cdot,t)) < \overline{\mathcal{D}}, \qquad \max_{\mathbf{x}\in\operatorname{supp}\rho_\alpha(\cdot,t),\alpha=1,\dots,A} |\mathbf{u}_\alpha(\mathbf{x},t) - \mathbf{V}| \lesssim e^{-\delta t},$$

where $\delta = \mu M \psi(\overline{\mathcal{D}})$.

5.2. External forcing. Theorems 4.1 and 4.2 have similar analogues for the system with additional external interaction forces [17]

$$\mathbf{F}_\alpha = -\nabla_{\mathbf{x}} U * \rho_\alpha.$$

This is due to the fact that our arguments establish rates independent of the number of agents. The hydrodynamic proofs repeat the discrete case ad verbatim; we therefore leave them out entirely.

5.3. Global existence and one-dimensional multiflocking: Smooth kernel case. We restrict attention to radial communication kernels $\phi_\alpha, \psi \in W^{2,\infty}$. The most convenient form of (5.4) to study regularity is in the shifted reference frame attached to the flock:

$$v_\alpha(x,t) := u_\alpha(x - X_\alpha(t), t) - V_\alpha(t), \qquad r_\alpha := \rho_\alpha(x - X_\alpha(t), t).$$

The new pair satisfies

$$\begin{cases} \partial_t r_\alpha + (v_\alpha r_\alpha)' = 0 \\ \quad \partial_t v_\alpha + v_\alpha v_\alpha' = \lambda_\alpha \displaystyle\int_{\mathbb{R}^d} \phi_\alpha(|x-y|)(v_\alpha(y) - v_\alpha(x)) r_\alpha(y)\,\mathrm{d}y - \mu R_\alpha(t) v_\alpha, \\ \qquad\qquad R_\alpha(t) = \displaystyle\sum_{\beta\neq\alpha} M_\beta \psi(|X_\alpha - X_\beta|). \end{cases} \tag{5.7}$$

In the case of the classical hydrodynamic alignment system the global well-posedness in one dimension relies on a threshold condition for the auxiliary quantity $e = v' + \phi * \rho$, which satisfies the same continuity as the density; see [23]. For multiflocks we define, accordingly, the family of such quantities

$$e_\alpha(x,t) = v_\alpha' + \lambda_\alpha \phi_\alpha * r_\alpha.$$

By virtue of (5.7), e_α satisfies

$$\partial_t e_\alpha + (v_\alpha e_\alpha)' = -\mu R_\alpha(t) v_\alpha',$$

which can be written as a nonautonomous logistic equation along characteristics of v_α:

$$\frac{\mathrm{d}}{\mathrm{d}t} e_\alpha = (\mu R_\alpha + e_\alpha)(\phi_\alpha * r_\alpha - e_\alpha), \qquad \frac{\mathrm{d}}{\mathrm{d}t} := \partial_t + v_\alpha \partial_x. \tag{5.8}$$

It is therefore natural to a expect threshold condition to guarantee global existence. We elaborate on that in the next result.

THEOREM 5.2 (global existence). *Let $\psi, \phi_\alpha \in W^{2,\infty}(\mathbb{R})$. For any initial conditions $(u_\alpha, \rho_\alpha) \in W^{2,\infty} \times (W^{1,\infty} \cap L^1)$ satisfying*

$$u'_\alpha(x,0) + \lambda_\alpha \phi_\alpha * \rho_\alpha(x,0) \geqslant 0 \quad \forall x \in \mathbb{R},\ \alpha = 1,\dots,A \tag{5.9}$$

there exists a unique global solution $(u_\alpha, \rho_\alpha) \in L^\infty_{\rm loc}([0,\infty); W^{2,\infty} \times (W^{1,\infty} \cap L^1))$. On the other hand, if for some $x_0 \in \mathbb{R}$ and $\alpha \in \{1,\dots,A\}$

$$u'_\alpha(x_0,0) + \lambda_\alpha \phi_\alpha * \rho_\alpha(x_0,0) < -\mu M \psi(0), \tag{5.10}$$

then the solution develops a finite time blowup.

The gap between the threshold levels is due to the fact that it is hard to predict the dumping coefficient $\mu R_\alpha(t)$, which may fluctuate in time. In particular, if ψ has a fat tail, then the argument below shows that the threshold for global existence is improvable to

$$e_\alpha(x,0) \geqslant -\mu M \psi(\overline{\mathcal{D}}) \quad \forall x \in \mathbb{R},\ \alpha = 1,\dots,A, \tag{5.11}$$

where $\bar{D}$ is determined from the initial conditions by (5.12):

$$\mu \int_{\mathcal{D}_0}^{\overline{\mathcal{D}}} \psi(r)\,\mathrm{d}r = \mathcal{A}_0. \tag{5.12}$$

Proof. Let us start with the negative result. Noting that $\mu M \psi(0)$ is the global upper bound for μR_α, from (5.8) we conclude that $\mathrm{d}/\mathrm{d}t e_\alpha \leqslant 0$. So, e_α will remain below $-\mu(1+\delta)M\psi(0)$ for some $\delta > 0$ along the characteristics starting at x_0. Hence,

$$\frac{\mathrm{d}}{\mathrm{d}t} e_\alpha \leqslant \frac{\delta}{1+\delta} e_\alpha(\phi_\alpha * r_\alpha - e_\alpha) \lesssim -e_\alpha^2.$$

Hence, e_α blows up in finite time.

On the other hand, if (5.9) holds initially, then since

$$e_\alpha(\phi_\alpha * r_\alpha - e_\alpha) \leqslant \dot{e}_\alpha \leqslant (\mu R_\alpha + e_\alpha)(|\phi_\alpha|_\infty M_\alpha - e_\alpha),$$

e_α will remain nonnegative and asymptotically bounded from above by $|\phi_\alpha|_\infty M_\alpha$. Hence, $\|v'_\alpha\|_\infty$ is uniformly bounded. Next, solving the continuity equation along characteristics

$$r_\alpha(x_\alpha(x_0;t),t) = r_\alpha(x_0,0) \exp\left\{ -\int_0^t v'_\alpha(x_\alpha(x_0;s),s)\,\mathrm{d}s \right\},$$

we conclude that r_α remains a priori bounded on any finite time interval.

Next, differentiating the e-equation,

$$\frac{\mathrm{d}}{\mathrm{d}t} e'_\alpha + v''_\alpha e_\alpha + 2 v'_\alpha e'_\alpha + v_\alpha e''_\alpha = -\mu R_\alpha v''_\alpha,$$

passing to Lagrangian coordinates and replacing $v''_\alpha = e'_\alpha - \lambda_\alpha \phi'_\alpha * \rho_\alpha$ we obtain, in view of already known information,

$$\frac{\mathrm{d}}{\mathrm{d}t} |e'_\alpha|^2 \leqslant f(t)|e'_\alpha|^2 + g(t),$$

where f and g are bounded functions. Hence, e'_α remain bounded as well, and consequently so does v''_α. Finally, $r'_\alpha \in L^\infty$ follows from differentiating and integrating the continuity equation.

The obtained a priori estimates lead to a construction of global solutions by the standard approximation and continuation argument (see [18] for systematic exposition). ∎

We proceed with two strong flocking results that demonstrate alignment in cluster with interflock slow and inner-flock fast rates as expected.

THEOREM 5.3 (strong flocking). *Suppose the threshold condition* (5.9) *holds so the solution exists globally. If for some* $\alpha \in \{1, \dots, A\}$ *the* α*-flock has compact support and the kernel* ϕ_α *has a fat tail, then there exists* $\delta_\alpha = \delta_\alpha(\phi_\alpha, \lambda_\alpha, u_\alpha(0), \rho_\alpha(0))$ *such that*

$$\sup_{x \in \operatorname{supp} \rho_\alpha(\cdot,t)} |u_\alpha(x,t) - U_\alpha(t)| + |u'_\alpha(x,t)| + |u''_\alpha(x,t)| \lesssim e^{-\delta_\alpha t},$$

and the density ρ_α *converges to a traveling wave with profile* $\bar\rho_\alpha$ *in the metric of* C^γ *for any* $0 < \gamma < 1$*:*

$$\|\rho_\alpha(\cdot,t) - \bar\rho_\alpha(\cdot - X_\alpha(t))\|_{C^\gamma} \lesssim e^{-\delta_\alpha t}.$$

Furthermore, if ψ *has a fat tail, the kernels* $\phi_\alpha \geqslant 0$ *are arbitrary, and the multiflock has a finite diameter initially, then global alignment occurs at a rate* $\delta = \delta(\psi, \mu, u(0), \rho(0))$*:*

$$\sup_{x \in \operatorname{supp} \rho_\alpha(\cdot,t), \alpha=1,\dots,A} |u_\alpha(x,t) - U| + |u'_\alpha(x,t)| + |u''_\alpha(x,t)| \lesssim e^{-\delta t},$$
$$\|\rho_\alpha(\cdot,t) - \bar\rho_\alpha(\cdot - Ut)\|_{C^\gamma} \lesssim e^{-\delta t}.$$

Proof. Let us prove the local statement first. Note that the alignment itself is a consequence of Theorem 5.1. Plus we have a global bound $\overline{\mathcal{D}}_\alpha$ on the diameter of the α-flock. Next, let us make the following observation: since

$$\phi_\alpha * \rho_\alpha(x) \geqslant M_\alpha \phi_\alpha(\overline{\mathcal{D}}_\alpha) = c_0 \quad \forall x \in \operatorname{supp} r_\alpha,$$

then from (5.8) we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} e_\alpha \geqslant e_\alpha(c_0 - e_\alpha).$$

Consequently, there exists a time t_0 starting from which $e_\alpha(x) \geqslant c_0/2$ for all $x \in \operatorname{supp} r_\alpha$. This follows by direct solution of the ODI.

Let us now write the equation for v'_α

$$\frac{\mathrm{d}}{\mathrm{d}t} v'_\alpha + v_\alpha v''_\alpha = \int_{\mathbb{R}} \phi'_\alpha(x-y)(v_\alpha(y) - v_\alpha(x)) r_\alpha(y)\,\mathrm{d}y - (\mu R_\alpha(t) + e_\alpha) v'_\alpha. \tag{5.13}$$

We already know from Theorem 5.1 that the velocity variations are exponentially decaying with the desired rate. Let us denote, as before, by $\operatorname{Exp}(t)$ a generic function with such exponential decay. Then, in Lagrangian coordinates,

$$\frac{\mathrm{d}}{\mathrm{d}t} |v'_\alpha|^2 \leqslant \operatorname{Exp}(t) v'_\alpha - \frac{c_0}{2} |v'_\alpha|^2 \leqslant \operatorname{Exp}(t) - \frac{c_0}{4} |v'_\alpha|^2.$$

This establishes the decay for v'_α on the support of r_α. Next,

$$\frac{\mathrm{d}}{\mathrm{d}t}v''_\alpha + 2v'_\alpha v''_\alpha = \int_{\mathbb{R}} \phi''_\alpha(x-y)(v_\alpha(y) - v_\alpha(x))r_\alpha(y)\,\mathrm{d}y \\ - 2v'_\alpha \phi'_\alpha * r_\alpha - (\mu R_\alpha(t) + e_\alpha)v''_\alpha. \tag{5.14}$$

So, similar to the previous

$$\frac{\mathrm{d}}{\mathrm{d}t}|v''_\alpha|^2 \leqslant \mathrm{Exp}(t) - \frac{c_0}{4}|v''_\alpha|^2.$$

Thus, $|v''_\alpha| \sim \mathrm{Exp}(t)$. As to the density,

$$\frac{\mathrm{d}}{\mathrm{d}t}r'_\alpha = -2v'_\alpha r'_\alpha - v''_\alpha r_\alpha = \mathrm{Exp}(t) r'_\alpha + E(t), \tag{5.15}$$

and we obtain uniform in time control over $\|r'_\alpha\|_\infty$.

To conclude strong flocking we write

$$\frac{\mathrm{d}}{\mathrm{d}t}r_\alpha = -v_\alpha r'_\alpha - v'_\alpha r_\alpha = \mathrm{Exp}(t). \tag{5.16}$$

This shows that $r_\alpha(t)$ is Cauchy in t in the metric of L^∞. Hence, there exists $\bar{r}_\alpha \in L^\infty$ such that $\|r_\alpha(t) - \bar{r}_\alpha\|_\infty = \mathrm{Exp}(t)$. Since r'_α is uniformly bounded, this also shows that $\bar{r}_\alpha$ is Lipschitz. Convergence in C^γ, $\gamma < 1$, follows by interpolation. Finally, passing to the original coordinate frame gives the desired result.

As to the global statement, the result follows from exact same argument above by noting that $\mu R_\alpha(t) \geqslant \mu M \psi(\bar{D}) = c_0$, and all the macroscopic momenta U_α align by Theorem 5.1. ∎

Remark 5.4. We note that the strong flocking result is new even in the classical monoflock context. The work [20] treats the more restrictive case of a kernel with positive infimum, while [23] only claims bounded diameter.

5.4. Global existence and one-dimensional multiflocking: Singular kernel case. In the case when ψ is smooth and inner communication kernels are singular

$$\phi_\alpha(r) = \frac{1}{r^{1+s}}, \quad 0 < s < 2, \tag{5.17}$$

the system (5.7) becomes of fractional parabolic type with bounded drift (due to the maximum principle) and bounded dumbing term. Considered under periodic settings $\mathbb{T}$ with no vaccum initial condition $\rho_\alpha > 0$ for all $\alpha = 1, \ldots, A$, we encounter no additional issues in the application of the regularity results obtained in [19, 20, 21]. Indeed, the dumping term $\mu R_\mu \mathbf{v}_\alpha$ has no effect on the continuity equation written in parabolic form

$$\partial_t r_\alpha + v_\alpha r'_\alpha + e_\alpha r_\alpha = r_\alpha \Lambda_s r_\alpha,$$

where $\Lambda_s = -(-\Delta)^{s/2}$ is the fractional s-Laplacian. As to the momentum equation it can be viewed as a bounded force for the initial Hölder regularization applied from [22, 15] in the way identical to our previous works. Further adaptation of the nonlocal maximum principal estimates of Constantin–Vicol [5] and continuation criteria for higher-order Sobolev spaces is straightforward.

THEOREM 5.5. *Let ψ be a smooth kernel and ϕ_α be the kernel of Λ_s on $\mathbb{T}^1$. Then system* (5.4) *admits a global solition for any initial data in $u_\alpha \in H^4(\mathbb{T}^1)$, $\rho_\alpha \in H^{3+s}(\mathbb{T}^1)$ with no vacuum:*

$$\min_{\alpha, x \in \mathbb{T}^1} \rho_\alpha(x, 0) > 0.$$

The solution belongs locally to

$$u_\alpha \in C([0,\infty), H^4) \cap L^2([0,\infty), H^{4+\frac{s}{2}}), \quad \rho_\alpha \in C([0,\infty), H^{3+s}) \cap L^2([0,\infty), H^{3+\frac{3s}{2}}).$$

6. Appendix. Global existence for singular kernels. Although collisions between the agents are possible with smooth kernels, this does not cause issues from the point of view of proving global existence of (1.2), using Picard iteration and continuation. If the kernels ϕ_α are singular, however, collisions lead to finite time blowup, so this case needs to be addressed separately. As was shown in [4], if the kernel is sufficiently singular collisions are prevented by strong close range alignment. We revisit this result in the context of multiflocks.

THEOREM 6.1 (singular communication kernels). *Suppose the α-flock is governed by a singular communication so that*

$$\int_0^1 \phi_\alpha(r)\, dr = \infty. \tag{6.1}$$

Then the flock experiences no internal collisions between agents.

Proof. The proof given below is a simplified version of the argument given in [4]. First, we assume for notational simplicity that all the masses are unity. Let us assume that for a given noncollisional initial condition $(\mathbf{x}_{\alpha i}, \mathbf{v}_{\alpha i})_{i,\alpha}$ a collision occurs at time T^* for the first time. Let $\Omega^*_\alpha \subset \Omega_\alpha = \{1, \dots, N_\alpha\}$ be the indexes of the agents that collided at one point. Hence, there exists a $\delta > 0$ such that $|\mathbf{x}_{\alpha i}(t) - \mathbf{x}_{\alpha k}(t)| \geqslant \delta$ for all $i \in \Omega^*_\alpha$ and $k \in \Omega_\alpha \backslash \Omega^*_\alpha$. Denote

$$\mathcal{D}^*_\alpha(t) = \max_{i,j\in\Omega^*_\alpha} |\mathbf{x}_{\alpha i}(t) - \mathbf{x}_{\alpha j}(t)|, \quad \mathcal{A}^*_\alpha(t) = \max_{i,j\in\Omega^*_\alpha} |\mathbf{v}_{\alpha i}(t) - \mathbf{v}_{\alpha j}(t)| = \max_{\substack{\boldsymbol{\ell}\in\mathbb{R}^n:|\boldsymbol{\ell}|=1\\ i,j\in\Omega^*_\alpha}} \langle \boldsymbol{\ell}, \mathbf{v}_{\alpha i} - \mathbf{v}_{\alpha j}\rangle.$$

Directly from the characteristic equation we obtain $|\dot{\mathcal{D}}^*_\alpha| \leqslant \mathcal{A}^*_\alpha$, and hence

$$-\dot{\mathcal{D}}^*_\alpha \leqslant \mathcal{A}^*_\alpha. \tag{6.2}$$

Let us fix a maximizing triple (ℓ, i, j) for $\mathcal{A}^*_\alpha(t)$ and compute using the momentum equation

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{A}^*_\alpha &= \sum_{k=1}^N m_{\alpha k}[\phi_\alpha(|\mathbf{x}_{ik}|)\ell(\mathbf{v}_{\alpha k i}) - \phi_\alpha(|\mathbf{x}_{jk}|)\ell(\mathbf{v}_{\alpha k j})] - \mathcal{A}^*_\alpha R_\alpha(t)\\ &= \sum_{k\in\Omega^*_\alpha} m_{\alpha k}[\phi_\alpha(|\mathbf{x}_{ik}|)\ell(\mathbf{v}_{\alpha k j} - \mathbf{v}_{\alpha i j}) + \phi_\alpha(|\mathbf{x}_{jk}|)\ell(-\mathbf{v}_{\alpha k i} - \mathbf{v}_{\alpha i j})]\\ &\quad + \sum_{k\notin\Omega^*_\alpha} m_{\alpha k}[\phi_\alpha(|\mathbf{x}_{ik}|)\ell(\mathbf{v}_{\alpha k i}) - \phi_\alpha(|\mathbf{x}_{jk}|)\ell(\mathbf{v}_{\alpha k j})] - \mathcal{A}^*_\alpha R_\alpha(t).\end{aligned}$$

The term $-\mathcal{A}^*_\alpha R_\alpha(t)$ is negative and will be dropped. In the first sum all terms are negative, so we can pull out the minimal value of the kernel which is $\phi_\alpha(\mathcal{D}^*_\alpha)$. In the second sum, all the distances $|\mathbf{x}_{ik}|$, $|\mathbf{x}_{jk}|$ are separated by δ up to the critical time T^*. So, the kernel will remain bounded. Putting together these remarks we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{A}^*_\alpha \leqslant C_1 - C_2\phi_\alpha(\mathcal{D}^*_\alpha)\mathcal{A}^*_\alpha.$$

Let us consider the energy functional

$$E_\alpha(t) = \mathcal{A}^*_\alpha(t) + C_2\int_{\mathcal{D}^*_\alpha(t)}^1 \phi_\alpha(r)\,\mathrm{d}r.$$

From the above we obtain $\mathrm{d}/\mathrm{d}t E_\alpha(t) \leqslant C_1$. So, E_α remains bounded up to the critical time, which implies that $\mathcal{D}^*_\alpha(t)$ stays away from zero. ∎

COROLLARY 6.2. *Suppose ψ is a smooth kernel, and each kernel ϕ_α is either smooth or condition* (6.1) *holds. Then the system* (1.2) *admits a unique global solution from any initial datum.*

We conclude by noting that this does not preclude collisions between agents from different flocks.

Acknowledgment. RS thanks CSCAMM and University of Maryland for hospitality during the preparation of this paper.

REFERENCES

[1] N. Bellomo, P. Degond, and E. Tadmor, *Active Particles, Volume* 1: *Advances in Theory, Models, and Applications*, Birkhäuser, Basel, 2017.

[2] N. Bellomo, P. Degond, and E. Tadmor, *Active Particles, Volume* 2: *Advances in Theory, Models, and Applications*, Springer, New York, 2019.

[3] J. A. Carrillo, M. Fornasier, G. Toscani, and F. Vecil, *Particle, kinetic, and hydrodynamic models of swarming*, in Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences, Naldi G, Pareschi L, Toscani G, eds., Modeling and Simulation in Science, Engineering and Technology, Birkhäuser, Boston, 2010.

[4] J. A. Carrillo, Y.-P. Choi, P. B. Mucha, and J. Peszek, *Sharp conditions to avoid collisions in singular Cucker-Smale interactions*, Nonlinear Anal. Real World Appl., 37 (2017), pp. 317–328.

[5] P. Constantin and V. Vicol, *Nonlinear maximum principles for dissipative linear nonlocal operators and applications*, Geom. Funct. Anal., 22 (2012), pp. 1289–1321.

[6] F. Cucker and S. Smale, *Emergent behavior in flocks*, IEEE Trans. Automat. Control, 52 (2007), pp. 852–862.

[7] F. Cucker and S. Smale, *On the mathematics of emergence*, Jpn. J. Math., 2 (2007), pp. 197–227.

[8] A. Deutsch, P. Friedl, L. Preziosi, and G. Theraulaz, *Multi-scale analysis and modelling of collective migration in biological systems*, Phil. Trans. R. Soc. B, 375 (2020), 20190377, https://doi.org/10.1098/rstb.2019.0377.

[9] S.-Y. Ha and J.-G. Liu, *A simple proof of the Cucker-Smale flocking dynamics and mean-field limit*, Commun. Math. Sci., 7 (2009), pp. 297–325.

[10] S.-Y. Ha and E. Tadmor, *From particle to kinetic and hydrodynamic descriptions of flocking*, Kinet. Relat. Models, 1 (2008), pp. 415–435.

[11] S. He and E. Tadmor, *A game of alignment: Collective behavior of multi-species*, Ann. I. H. Poincare-AN, 38 (2021), pp. 1031–1053.

[12] A. Lesne, *Multiscale analysis of biological systems*, Acta Biotheor., 61 (2013), pp. 3–19, https://doi.org/10.1007/s10441-013-9170-z.

[13] S. Motsch and E. Tadmor, *A new model for self-organized dynamics and its flocking behavior*, J. Stat. Phys., 144 (2011), pp. 923–947.

[14] L. Preziosi and A. Tosin, *Multiphase and multiscale trends in cancer modelling*, Math. Model. Nat. Phenom., 4 (2009), pp. 1–11, https://doi.org/10.1051/mmnp/20094301.

[15] R. W. Schwab and L. Silvestre, *Regularity for parabolic integro-differential equations with very irregular kernels*, Anal. PDE, 9 (2016), pp. 727–772.

[16] R. Shu and E. Tadmor, *Anticipation breeds alignment*, Arch. Rat. Mech. Anal., 240 (2021), pp. 203–241, https://doi.org/10.1007/s00205-021-01609-8.

[17] R. Shu and E. Tadmor, *Flocking hydrodynamics with external potentials*, Arch. Rat. Mech. Anal., 238 (2020), pp. 347–381.

[18] R. Shvydkoy, *Dynamics and Analysis of Alignment Models of Collective Behavior*, Birkhäuser, Cham, 2021.

[19] R. Shvydkoy and E. Tadmor, *Eulerian dynamics with a commutator forcing*, Trans. Math. Appl., 1 (2017), tnx001.

[20] R. Shvydkoy and E. Tadmor, *Eulerian dynamics with a commutator forcing* II: *Flocking*, Discrete Contin. Dyn. Syst., 37 (2017), pp. 5503–5520.

[21] R. Shvydkoy and E. Tadmor, *Eulerian dynamics with a commutator forcing* III. *Fractional diffusion of order* $0 < \alpha < 1$, Phys. D, 376/377 (2018), pp. 131–137.

[22] L. Silvestre, *Hölder estimates for advection fractional-diffusion equations*, Ann. Sc. Norm. Super. Pisa Cl. Sci., 11 (2012), pp. 843–855.

[23] E. Tadmor and C. Tan, *Critical thresholds in flocking hydrodynamics with non-local alignment*, Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 372 (2014), 20130401.