MATHEMATICS OF EMERGENCE: DEVELOPING TRENDS IN ALIGNMENT DYNAMICS

ROMAN SHVYDKOY

Abstract. These notes are based on a series of lectures given at ICMAT, Madrid in November 2018, the Gran Sasso Science Institute, L’Aquila (Italy) in June 2018, and the joint working group at Ecole Normale Superieure and Laboratoire Jacques-Louis Lions at Universite Pierre et Marie Curie (Paris 6), Paris in May 2018, and Charles University in Prague in May 2019.

Contents

1. Background, Basic concepts, Levels of description 2
   1.1. Emergent phenomena 2
   1.2. Types of communication 3

2. Discrete Systems 5
   2.1. Momentum, Energy, and Maximum Principle 5
   2.2. How much communication is needed? Connectivity and spectral method 6
   2.3. Alignment on $\mathbb{R}^n$. Non-degenerate case 8
   2.4. Stability 11
   2.5. Singular kernels and the issue of collisions 12
   2.6. Degenerate communication. Corrector Method 14
   2.7. Dynamics under external potential forces: confinement 18
   2.8. Attraction-Alignment models 21
   2.9. Multi-flocks 27

3. Kinetic models 29
   3.1. BBGKY hierarchy: formal derivation 30
   3.2. Weak formulation and basic principles of kinetic dynamics 31
   3.3. Kinetic maximum principle and flocking 33
   3.5. Mean-field limit 37
   3.6. Macroscopic description. Hydrodynamic limit. 38

4. Euler Alignment System 39
   4.1. Hydrodynamic flocking and stability 40
   4.2. Spectral method. Hydrodynamic connectivity 42
   4.3. Topological models. Adaptive diffusion 44

5. Regularity I: local well-posedness and continuation criteria 50
   5.1. Smooth models 50
   5.2. Singular models 54

6. One-dimensional theory 61
   6.1. Smooth kernels: threshold for GWP and stability. 61
   6.2. Entropy. Csiszár-Kullback inequality. Distribution of the limiting Flock 63

Date: December 4, 2019.

1991 Mathematics Subject Classification. 92D25, 35Q35, 76N10.
Key words and phrases. Flocking, alignment, fractional dissipation, Cucker-Smale.
Acknowledgment. Research is supported in part by NSF grant DMS-1813351 and Simons Foundation.
1.1. Emergent phenomena. Agent-based models of collective behavior describe dynamics of a number of objects:
\[ x_i \in \Omega \subset \mathbb{R}^n, \quad i = 1, \ldots, N \]
\[ v_i = \dot{x}_i \]
governed by mutual communication - adjustment of velocity or position to that of nearby neighbors.

Emergence is a phenomenon of self-organization of a system of agents governed by local communication.

The long time dynamics of a self-organized system can be characterized by the following phenomena:

- **alignment**: \( \lim_{t \to \infty} \max_i |v_i - \bar{v}| = 0 \),
- **flocking**: \( \sup_{i,j} |x_i - x_j| \leq \mathcal{D} < \infty \),
- **strong flocking**: \( x_i - x_j \to \bar{x}_{ij} \), as \( t \to \infty \),
- **aggregation**: \( x_i - x_j \to 0 \), as \( t \to \infty \).

Note that alignment implies strong flocking provided it occurs at a sufficiently fast rate

\[ \int_0^\infty \max_{i,j} |v_i - v_j| \, dt < \infty. \]  

Indeed,
\[ x_i(t) - x_j(t) = x_i(0) - x_j(0) + \int_0^t [v_i(s) - v_j(s)] \, ds, \]

hence
\[ \dot{x}_{ij} = x_i(0) - x_j(0) + \int_0^\infty [v_i(s) - v_j(s)] \, ds. \]
Most common models that appear is various applications are

- Vicsek discrete model, 1995:

\[
\begin{align*}
 v_i(k+1) &= v_0 \frac{\sum_{j:|x_j-x_i|<r_0} v_j}{\sum_{j:|x_j-x_i|<r_0} v_j} + F_i, \\
 x_i(k+1) &= x_i(k) + v_i(k+1).
\end{align*}
\]

Here \(F_i\) indicate additional forces.

- Kuramoto synchronization model, \(\theta_i \in \mathbb{T}^1\):

\[
\dot{\theta}_i = \frac{\lambda}{N} \sum_{j \in N_i} \sin(\theta_j - \theta_i) + \omega_i.
\]

- 1st order environmental averaging:

\[
\dot{p}_i = \lambda \sum_{j \in N_i} a_{ij}(t)(p_j - p_i) + F_i, \quad \sum_{j} a_{ij}(t) = 1,
\]

where \(N_i\) is a set of 'active' agents in local proximity, \(p_i \in \mathbb{R}^n\), e.g. \(p = x\) or \(p = v\).

- 2nd order Cucker-Smale type systems:

\[
\begin{align*}
 \dot{x}_i &= v_i, \\
 \dot{v}_i &= \lambda \sum_{j=1}^{N} m_j \phi(x_i,x_j)(v_j - v_i) + F_i, \quad (x_i,v_i) \in \Omega \times \mathbb{R}^n
\end{align*}
\]

where, \(\phi(x,y) \geq 0\) is a communication kernel and \(\Omega = \mathbb{R}^n\) or \(\mathbb{T}^n\).

1.2. Types of communication. The focus of these lectures will be on systems of Cucker-Smale type (11). We distinguish between the following general types of communication:

- Absolute: \(\inf_{x,y \in \Omega} \phi(x,y) > 0\);
- Global: \(\phi(x,y) > 0\), for all \(x,y \in \Omega\);
- Local: there is a \(r_0 > 0\) so that \(\phi(x,y) = 0\) if \(|x - y| > r_0\);
- Symmetric: \(\phi(x,y) = \phi(y,x)\);
- Convolution type: \(\phi(x,y) = \phi(|x - y|)\).

Sometimes by a local kernel, we mean a lack of global assumptions, for example, when the only information available is

\[
\phi(x,y) \geq 1, \text{ for } |x - y| \leq r_0.
\]
A wealth of symmetric examples can be produced by convolution type kernels, where \( \phi \in C^\infty(\mathbb{R}^+) \) is a non-negative function. Here are some examples that we will be discussing:

(Cucker-Smale)

\[
\phi(r) = \frac{h(r)}{(1 + r^2)^{\beta/2}}, \quad \beta > 0
\]

(Singular)

\[
\phi(r) = \frac{h(r)}{r^\beta},
\]

(Extremely Singular)

\[
\phi(r) = \frac{h(r)}{|r - \delta|^\beta}, \quad r > \delta,
\]

where \( h \) is a possible cut-off function if we consider local kernels. A measure of strength of communication in long and short range can be expressed by the following more general conditions:

(Long range ("fat tail")

\[
\int_{r_0}^\infty \phi(r) \, dr = \infty
\]

(Short range)

\[
\int_{0}^{r_0} \phi(r) \, dr = \infty.
\]

One example of a non-symmetric communication kernel was introduced by Motsch and Tadmor [9]:

\[
\psi(x_i, x_j) = \frac{\phi(|x_i - x_j|)}{\sum_k m_k \phi(|x_i - x_k|)}.
\]

Note that MT-model defines an averaging communication protocol: \( \sum_j m_j \psi(x_i, x_j) = 1 \). Both singular and Motsch-Tadmor kernels are meant to emphasize local interactions over global ones whenever such communication is more realistic. These belong to a class of “metric” kernels, meaning that the dependencies on other agents \( x_k \)'s are stated in terms of Euclidean distances to the center \( x_i \).

A new class of “topological” kernels was introduced in [13]. Its construction is based on the following principle: communication between a pair of agents is determined by the density of crowd between them. Effectively, overcrowded areas impede communication, thus information spreads slower. The quantity

\[
d_{ij} = \left[ \sum_{k: x_k \in \Omega_{ij}} m_k \right]^{\frac{1}{n}}
\]

measures the distance between agents in terms of the mass of a communication domain \( \Omega_{ij} \) which is assumed to be symmetric \( \Omega_{ij} = \Omega_{ji} \). A kernel that incorporates both topological and a degree of
metric communications can be defined as
\begin{equation}
\phi_{ij}(x) = \frac{1}{\tau_{ij}} \psi(|x_i - x_j|),
\end{equation}
where $\psi$ is a non-negative decreasing function, and $\tau$ is a parameter that defines presence of topological component. A full description of this class of models will be presented in Section 4.3.

2. Discrete Systems

Much of the groundwork on the alignment will be laid out in the context of the most basic agent-based Cucker-Smale system given by
\begin{equation}
\begin{cases}
\dot{x}_i = v_i, \\
\dot{v}_i = \lambda \sum_{j=1}^{N} m_j \phi(x_i, x_j) (v_j - v_i),
\end{cases}
\quad (x_i, v_i) \in \Omega \times \mathbb{R}^n
\end{equation}

Let us first discuss fundamental properties of the system.

2.1. Momentum, Energy, and Maximum Principle. Cucker-Smale systems with symmetric kernels, as opposed to non-symmetric, preserve the average total momentum of the system:
\begin{equation}
\bar{v} = \frac{1}{M} \sum_i m_i v_i, \quad M = \sum_{i=1}^{N} m_i, \quad \frac{d}{dt} \bar{v} = 0.
\end{equation}

Due to conservation of momentum, the center of mass moves with a constant velocity
\begin{equation}
\bar{x} = \frac{1}{M} \sum_i m_i x_i, \quad \frac{d}{dt} \bar{x} = \bar{v}.
\end{equation}

For convolution type kernels, conservation of momentum can be used to shift the reference frame centered at $\bar{x}$, due to Galilean invariance of the system:
\begin{equation}
x_i \rightarrow x_i - t\bar{v}, \quad v_i \rightarrow v_i - \bar{v}.
\end{equation}

So, in this case we can assume without loss of generality that $\bar{v} = 0$. In general, such translational invariance is not available. Nonetheless, if alignment occurs, then necessarily all $v_i \rightarrow \bar{v}$. In other words, we can determine the limiting velocity from initial conditions.

Let us consider the following variation and dissipation functions:
\begin{align}
V_2 &= \frac{1}{2} \sum_{i,j} m_i m_j |v_i - v_j|^2 \\
I_2 &= \sum_{i,j} m_i m_j \phi(x_i, x_j) |v_i - v_j|^2.
\end{align}

The system has the classical kinetic energy as well defined by
\begin{equation}
E = \frac{1}{2} \sum_{i=1}^{N} m_i |v_i|^2.
\end{equation}

In case if $\bar{v} = 0$, then $V_2 = 2ME$, however it is not prudent to use energy as a measure of alignment $E$ in the non-symmetric case, simply because we do not know if 0 would remain to be the momentum of the system for all time.

The following energy law is easily verified:
\begin{equation}
\frac{d}{dt} V_2 = -\lambda MI_2.
\end{equation}
At this point one can obtain an $\ell^2$-base alignment result under absolute communication: if $\inf \phi = c_0 > 0$, then $\mathcal{I}_2 \geq c_0 \mathcal{V}_2$, and hence

$$\dot{\mathcal{V}}_2 \leq -2c_0 \lambda M \mathcal{V}_2.$$ 

Hence,

$$\mathcal{V}_2(t) \leq \mathcal{V}_2(0)e^{-2c_0 \lambda Mt}.$$ 

This result provides exponential alignment “on average”, specifically $L^2$-average, which does not translate well into individual information on agents. Indeed, one only obtains

$$(17) \quad |v_i - v_j| \leq \frac{1}{m_i m_j} \mathcal{V}_2(0)e^{-\delta t}.$$ 

This estimate clearly deteriorates in the large crowd limit $N \to \infty$ when all masses vanish $m_i \to 0$. In order to improve upon (17) we must resort to an $\ell^\infty$-based argument and use the maximum principle to be discussed later in Section 2.3.

2.2. How much communication is needed? Connectivity and spectral method. Let us assume for simplicity that all agents have the same mass $m_i = \frac{1}{N}$. It is clear that main ”enemy” of alignment is lack of communication, especially at long range. In fact, it is easy to produce an example of initial condition that would diverge and not align if the kernel is local, see the Figure 3. If global communication is not available, this situation can be remedied by either confined environmental condition, such as periodic boundary conditions to be discussed below in detail, or by an assumption of connectivity. The latter means that for any pair of agents $x_i$ and $x_j$ there exists a chain of agents $x_{k_1}, \ldots, x_{k_l}$ with end-points at $x_i$ and $x_j$ and such that all $|x_{k_p} - x_{k_{p+1}}| < r_0$. In this case one can recover alignment as follows. First, note that the shortest chain connecting any pair of agents has no repeated agents in the chain. Hence, every chain is limited to length $N$. Assuming that $\lambda \phi(r_0) = \varepsilon > 0$, we can estimate $\mathcal{V}_2$ as follows:

$$\mathcal{V}_2 = \frac{1}{2N^2} \sum_{i \neq j} |v_i - v_j|^2 \leq \frac{1}{2N^2} \sum_{i \neq j} \sum_{p=1}^{P_{ij}} |v_{k_p} - v_{k_{p+1}}|^2 \leq \frac{\lambda}{2\varepsilon N^2} \sum_{i \neq j} \sum_{p=1}^{P_{ij}} \phi(|x_{k_p} - x_{k_{p+1}}|)|v_{k_p} - v_{k_{p+1}}|^2 \leq \frac{N(N-1)}{2\varepsilon N^2} \lambda \sum_{k' \neq k''} \phi(|x_{k'} - x_{k''}|)|v_{k'} - v_{k''}|^2 \lesssim N^2 \mathcal{I}_2.$$ 

So, we obtain a desired differential inequality for $\mathcal{V}_2$: $\dot{\mathcal{V}}_2 \leq -\frac{1}{N^2} \mathcal{V}_2$.

The argument outlined above produces a bad dependence on $N$, and as such is not suitable in the limit $N \to \infty$. A continuous analogue of connectivity condition requires more elaboration and is discussed in [8].

The connectivity assumption at all times is hard to verify of course. However, it is guaranteed to hold provided the initial configuration is connected and the communication strength $\lambda$ is large enough. This is because one can ensure in this case that the system aligns almost instantaneously before any disconnection becomes possible. Indeed, suppose initially the system is $r_0/2$ connected, and $\lambda \phi(r_0)$ is large. Then for some large $\Lambda > 0$ we will have

$$\frac{d}{dt} \mathcal{V}_2 \leq -\Lambda \mathcal{V}_2,$$

for as long as the system is $r_0$-connected. So, if a pair $x_i$, $x_j$ is initially at most $r_0/2$ apart, then

$$|x_i(t) - x_j(t)| \leq |x_i(0) - x_j(0)| + \frac{C}{\Lambda}.$$
This shows that the same pair will never get \( r_0 \)-disconnected and hence the connectivity is preserved at all times.

Even if connectivity is intermittent, but reoccurring, one can still achieve the alignment effect if the flock reconnects “more often then not”. There is an elegant algebraic way to make this statement qualitatively precise and in fact it lies in the heart of the original spectral approach introduced by Cucker and Smale in [3], see also Motsch and Tadmor’s implementation to their model in [10], and to topological models in [13] and Section 4.2.

Let us consider the matrix \( A = \{a_{ij}(t)\}_{i,j=1}^N \otimes I_{n \times n} \), where \( a_{ij} = \lambda \phi(x_i(t), x_j(t)), \ i \neq j \), and \( a_{ij} = 0 \), if \( i = j \). Note that \( a_{ij} \neq 0 \) if and only if the corresponding agents and “connected” in the sense that they communication through the influence kernel \( \phi \). By analogy with the graph theory we call it the adjacency matrix of the flock. In fact, one can consider the actual adjacency matrix associated with the flock under the above connectivity definition: \( \hat{A} = \{\hat{a}_{ij}(t)\}_{i,j=1}^N \otimes I_{n \times n} \), where \( \hat{a}_{ij} = 1 \) if \( a_{ij} \neq 0 \), and 0 otherwise. Let \( D = \text{diag}\{b_1, \ldots, b_N\} \otimes I_{n \times n} \), where \( b_i = \sum_j a_{ij} \), and \( \hat{D} = \text{diag}\{\hat{b}_1, \ldots, \hat{b}_N\} \otimes I_{n \times n} \), where \( \hat{b}_i = \sum_j \hat{a}_{ij} \). Note that each \( \hat{b}_i \) is precisely the degree of the vertex \( x_i \), i.e. the number of other agents to which it is connected.

In this notation the system (11) can now be written in terms of the grand velocity vector \( \hat{V} = (v_1, \ldots, v_N) \), and the Laplacian associated to \( A, L = \hat{D} - \hat{A} \), namely,

\[
\frac{d}{dt} \hat{V} = -LV.
\]

The matrix \( L \) is non-negative definite, and hence the spectrum consists of a sequence \( 0 = \kappa_1 \leq \kappa_2 \leq \ldots \leq \kappa_N \). The average grand vector \( \bar{\hat{V}} = \langle \hat{V} \rangle \) is obviously a member of the kernel of \( L \), hence \( \kappa_1 = 0 \). The next eigenvalue \( \kappa_2 \) is called the Fiedler number, although the classical Fiedler number is one associated to the Laplacian \( \hat{L} = \hat{D} - \hat{A} \) is a similar way, we denote it \( \hat{\kappa}_2 \). By the min-max theorem \( \kappa_2 \) is given by

\[
\kappa_2 = \min_{\sum_i v_i = 0} \frac{\langle LV, V \rangle}{|V|^2}.
\]

A simple fact to verify is that \( \kappa_2 \neq 0 \) if and only if the graph is connected, and the relationship to \( \hat{\kappa}_2 \) is given by (see [3, Proposition 2]):

\[
\kappa_2 \geq \hat{\kappa}_2 \min_{i,j:a_{ij} \neq 0} a_{ij}.
\]

For this reason, we can call \( \kappa_2 \) a weighted Fiedler number, which captures not only algebraic connectivity of the flock as a graph, but also the collective strength of the connection weighted by the kernel \( \phi \). With the use of the weighted Fiedler number \( \kappa_2 = \kappa_2(t) \), which, let’s recall, depends on time, we can measure alignment in (18) by writing the energy law as

\[
\frac{d}{dt} |\hat{V} - \bar{\hat{V}}|^2 = -2\langle L(\hat{V} - \bar{\hat{V}}), (\hat{V} - \bar{\hat{V}}) \rangle \leq -2\kappa_2(t) |\hat{V} - \bar{\hat{V}}|^2.
\]

Consequently,

\[
|\hat{V}(t) - \bar{\hat{V}}|_2 \leq |\hat{V}_0 - \bar{\hat{V}}|_2 \exp \left\{ -\int_0^t \kappa_2(s) \, ds \right\}.
\]

We can see that the divergence of the integral inside leads to alignment, which can be seen as a quantitative measure of connectivity as a function of time. Let us state it precisely.

**Lemma 2.1.** If \( \int_0^\infty \kappa_2(s) \, ds = \infty \), then the flock aligns.

At the center of the original result of Cucker and Smale was a statement that for kernels of type (4) the integral of the weighted Fiedler number is indeed divergent. Here, clearly, the flock remains always algebraically connected, thus \( \hat{\kappa}_2 = N \), and one can appeal directly to (20) to restate the problem in terms of control on the decay of the adjacency matrix. The original theorem of Cucker and Smale [2, 3] is the following.
Figure 3. Connected and disconnected flocks

**Theorem 2.2.** Let $\phi(r) = \frac{1}{(1+r^2)^{\beta/2}}$. Then every solution aligns exponentially and flocks strongly for $\beta \leq 1$, and conditionally if $\beta > 1$.

In Section 2.3 this result will be proved using a more direct approach due to Ha and Liu, [6], which paves a way to extensions into meso- and macroscopic systems. Let us not, however, underestimate the spectral method as sometimes it is the only one available if no specific structural information is known about the kernel.

To see that the non-integrable decay rate of the kernel is necessary in the Cucker-Smale theorem, let us consider the following example.

**Example 2.3.** Let the kernel be $\phi(r) = \frac{1}{r^\beta}$ for $r > r_0$, for simplicity, and let $x = x_1 = -x_2 > r_0$ and $v = v_1 = -v_2 > 0$. This symmetry is preserved in time. Then the system (11) becomes

$$\frac{dx}{2^{\beta-1}x^\beta} + \frac{dv}{1} = 0.$$  

This equation admits a conservation law

$$J = v + \frac{1}{2^{\beta-1}(1-\beta)x^{\beta-1}}.$$  

If $\beta > 1$, and the initial velocity is large enough, then $J(t) = J_0 > 0$, and hence $v(t) \geq J_0$ holds true for all times. This creates permanent misalignment between the two velocities at hand: $v$ and $-v$.

2.3. **Alignment on $\mathbb{R}^n$. Non-degenerate case.** The main goal in this section will be to prove the Cucker-Smale Theorem 2.2 in more general settings of convolution type fat tail kernel using a method based on the maximal principle. We will see that the argument is also easily adaptable to the non-symmetric case of the Motsch-Tadmor kernel. So, we consider the classical Cucker-Smale system

$$\begin{cases} \dot{x}_i = v_i, \\ \dot{v}_i = \lambda \sum_{k=1}^{N} m_k \phi(x_i - x_k)(v_k - v_i), \quad (x_i, v_i) \in \mathbb{R}^n \times \mathbb{R}^n. \end{cases}$$  

We assume that $\phi$ is monotonely decreasing and everywhere positive. It will be useful to use sometimes the following shortcut notation:

$$x_{ij} = x_i - x_j, \quad v_{ij} = v_i - v_j, \quad \phi_{ij} = \phi(x_i - x_j), \quad \text{etc.}$$  

Let us also consider the amplitude and flock diameter:

$$D = \max_{i,j} |x_{ij}|, \quad A = \max_{i,j} |v_{ij}|.$$
**Theorem 2.4** (Alignment for Cucker-Smale model). Any solution to the system (21) with initial condition $D_0,A_0$ satisfying

$$\int_{D_0}^{\infty} \phi(r) \, dr > \frac{A_0}{\lambda M},$$

aligns and flock exponentially fast:

$$\sup_{t \geq 0} D(t) \leq \bar{D}, \quad A(t) \leq A_0 e^{-t \lambda M \phi(\bar{D})}.$$  \tag{25}

In particular, every solution flocks provided the kernel satisfies the fat tail condition (7).

To make the proof perfectly rigorous, let us recall the classical Redemacher Lemma, which we will use throughout.

Suppose $f(x,t) : X \times \mathbb{R}_+ \to \mathbb{R}$ is a Lipschitz in time function uniformly in $x$, where $X$ is an arbitrary index set. So, $\exists L > 0$ such that for all $t,s,x$ we have

$$|f(x,t) - f(x,s)| \leq L |t - s|.$$  \tag{24}

Suppose that at any time $t$ there is a point $x(t) \in X$ such that

$$f(x(t),t) = \sup_{x \in X} f(x,t) := M(t).$$

Note that $M(t)$ is a Lipschitz function with the same constant $L$. Indeed, let $t,s \in \mathbb{R}_+$ and $M(t) > M(s)$. Then

$$M(t) - M(s) = f(x(t),t) - f(x(t),s) + f(x(t),s) - f(x(s),s) \leq f(x(t),t) - f(x(t),s) \leq L |t - s|.$$  \tag{25}

Consequently, $M$ is absolutely continuous on any finite interval, i.e.

$$M(t) - M(s) = \int_s^t m(r) \, dr,$$

where $\|m\|_{\infty} \leq L_{\text{loc}}$, and hence $M' = m$ a.e.

**Lemma 2.5.** If $f(x,\cdot)$ is differentiable everywhere in $t$ for all $x \in X$, then $M'(t) = \partial_t f(x(t),t)$ holds at any point where $M'$ exists.

Indeed computing one-sided derivative from the right we have

$$M'(t) = \lim_{h \to 0^+} \frac{f(x(t+h),t+h) - f(x(t),t+h) + f(x(t),t+h) - f(x(t),t)}{h} \geq \lim_{h \to 0^+} \frac{f(x(t),t+h) - f(x(t),t)}{h} = \partial_t f(x(t),t).$$

Taking $h < 0$ proves the opposite inequality.

**Proof of Theorem 2.4.** Let us represent $A$ as

$$A = \max_{|\ell| = 1,i,j} \ell(v_{ij}).$$

Note that the maximum is takes over a fixed compact set not changing in time. So, using Rademacher’s lemma we can pick $\ell$ and $i,j$ at each time $t$ for which the maximum is achieved. Using the velocity equation we obtain

$$\frac{d}{dt} \ell(v_{ij}) = \lambda \sum_{k=1}^{N} m_k \phi_{ik} \ell(v_{ki}) + m_k \phi_{jk} \ell(v_{jk})$$  \tag{26}
Now notice that $\ell(v_{ki}) \leq 0$ and $\ell(v_{jk}) \leq 0$. One can see it by adding and subtracting $v_j$ in the first case and $v_i$ in the second and using maximality of $\ell(v_{ij})$. So, then we can pull out the minimal value of the kernel in both sums:

$$\frac{d}{dt} \ell(v_{ij}) \leq -\lambda \phi(D) \sum_{k=1}^{N} m_k \ell(v_{ij}) = -\lambda M \phi(D) \ell(v_{ij}).$$

So, we obtain

$$(27) \begin{cases} \frac{d}{dt} A \leq -\lambda M \phi(D) A \\ \frac{d}{dt} D \leq A. \end{cases}$$

This system of ordinary differential inequalities (ODIs) has a decreasing Lyapunov function given by $L = A + \lambda M \int_0^D \phi(r) \, dr$. This, in particular, implies that

$$\lambda M \int_0^D \phi(r) \, dr \leq A_0 + \lambda M \int_0^{D_0} \phi(r) \, dr, \quad \forall t > 0.$$  

Consequently, $D(t) \leq \bar{D}$, where $\bar{D}$ is obtained from the equation

$$(28) \quad \lambda M \int_{D_0}^{D} \phi(r) \, dr = A_0,$$

which is guaranteed to have a finite solution due to (126). Then, $\dot{A} \leq -\lambda M \phi(\bar{D}) A$ and the theorem follows.

Solving equation (28) allows one to provide explicit decay rates for solutions of (35) for some kernels. In particular, for the classical Cucker-Smale kernel

$$\phi(r) = \frac{1}{(1 + r^2)^{\frac{\beta}{2}}},$$

one obtains

$$\bar{D} \leq \left( \left[ \frac{1 - \beta}{\lambda M} A_0 + (1 + D_0^2)^{\frac{1-\beta}{2}} \right]^{\frac{2}{1-\beta}} - 1 \right)^{\frac{1}{2}}, \quad \beta < 1,$$

(29)

$$\text{rate} = \frac{\lambda M}{\left[ \frac{1 - \beta}{\lambda M} A_0 + (1 + D_0^2)^{\frac{1-\beta}{2}} \right]^{\frac{1}{1-\beta}}}.$$

(Here one replaces $\phi$ with a smaller but explicitly integrable kernel $\frac{r}{(1 + r^2)^{\frac{\beta+1}{2}}}$, and

$$(30) \quad \bar{D} \leq \left( e^{\frac{2}{\lambda M} A_0} (1 + D_0^2) - 1 \right)^{\frac{1}{2}}, \quad \text{rate} = \frac{\lambda M}{e^{\frac{A_0}{\lambda M}} (1 + D_0^2)^{\frac{1}{2}}}, \quad \beta = 1.$$

The proof of Theorem 2.4 is rather flexible and can be adapted to other models, including those with non-symmetric communication.

**Motsch-Tadmor model:**

$$(31) \quad \int_{D_0}^{\infty} \phi(r) \, dr > \frac{A_0 |\phi|_\infty}{\lambda} \quad \Rightarrow \quad A(t) \leq A_0 e^{-t\lambda |\phi|_\infty \phi(D)}.$$

**Topological model:** requires fat tail condition on the metric part,

$$(32) \quad \int_{D_0}^{\infty} \psi(r) \, dr > \frac{A_0}{\lambda} \quad \Rightarrow \quad A(t) \leq A_0 e^{-t\lambda \psi(D)}.$$
To see it for the MT model we compute
\[
\frac{d}{dt} \ell(v_{ij}) = \lambda \sum_{k=1}^{N} \frac{m_k \phi_{ik}}{\sum_{p=1}^{N} m_p \phi_{ip}} \ell(v_{ki}) + \frac{m_k \phi_{jk}}{\sum_{p=1}^{N} m_p \phi_{jp}} \ell(v_{jk})
\]
(33)
\[
\leq \lambda \left| \phi(D) \right| \sum_{k=1}^{N} \frac{m_k}{M} [\ell(v_{ki}) + \ell(v_{jk})] = -\lambda \left| \phi(D) \right| \ell(v_{ij}).
\]

Let us note that in the symmetric case we can predict the limiting velocity from the conserved momentum: \( \bar{v} \). For the MT model the limit is not determined by the initial condition, but rather becomes an emergent quantity of the dynamics. Since all the differences \( v_{ij} \) vanish exponentially fast, it implies that velocities do in fact converge to a time independent limit as seen from integrating the velocity equation:
\[
\lim_{t \to \infty} v_i(t) = v_i(0) + \int_{0}^{\infty} \lambda \sum_{k=1}^{N} \frac{m_k \phi_{ik}}{\sum_{p=1}^{N} m_p \phi_{ip}} v_{ki}(s) \, ds.
\]

A notable byproduct of the arguments presented above is that \( | \cdot | \) can be any norm defined on \( \mathbb{R}^n \), with unit functionals \( | \ell | = 1 \) chosen in the dual norm. Even though all norms on \( \mathbb{R}^n \) are equivalent, their equivalency constants can be large. However, the basic system of inequalities (35) would still hold true with exact same coefficients, independent of chosen norm.

2.4. Stability. Suppose now we have two close initial conditions for the same flock with the same momentum:
\[
V = (v_i)_{i=1}^{N}, \quad X = (x_i)_{i=1}^{N}, \quad \bar{V} = (\bar{v}_i)_{i=1}^{N}, \quad \bar{X} = (\bar{x}_i)_{i=1}^{N},
\]
and \( \bar{v} = \bar{\bar{v}} \). We show that if initially these parameters are close, then they will remain close uniformly for all time. To formulate this precisely let us introduce the measure of distance between flocks:
\[
\|X - \bar{X}\| = \max_i |x_i - \bar{x}_i|, \quad \|V - \bar{V}\| = \max_i |v_i - \bar{v}_i|.
\]
(34)

Here \( | \cdot | \) denotes any norm on \( \mathbb{R}^n \). Much in the spirit of the previous proofs let us derive a system of ODEs for these two quantities. First, we clearly have
\[
\frac{d}{dt} \|X - \bar{X}\| \leq \|V - \bar{V}\|.
\]

Second, let us represent
\[
\|V - \bar{V}\| = \max_{i, |\ell| = 1} \ell(v_i - \bar{v}_i),
\]
and by Rademacher’s lemma apply derivative in time with a maximizing pair \( i, \ell \):
\[
\frac{d}{dt} \ell(v_i - \bar{v}_i) = \lambda \sum_{k=1}^{N} m_k \phi(x_{ik}) \ell(v_{ki}) - m_k \phi(\bar{x}_{ik}) \ell(\bar{v}_{ki})
\]
\[
= \lambda \sum_{k=1}^{N} m_k (\phi(x_{ik}) - \phi(\bar{x}_{ik})) \ell(v_{ki}) + \lambda \sum_{k=1}^{N} m_k \phi(\bar{x}_{ik}) [\ell(v_k - \bar{v}_k) - \ell(v_i - \bar{v}_i)]
\]
\[
\leq 2M \lambda |\nabla \phi|_\infty A \|X - \bar{X}\| - \lambda \phi(\bar{D}) \sum_{k=1}^{N} m_k [\ell(v_k - \bar{v}_k) - \ell(v_i - \bar{v}_i)]
\]
\[
= 2M \lambda |\nabla \phi|_\infty A_0 e^{-tM \phi(\bar{D})} \|X - \bar{X}\| - M \lambda \phi(\bar{D}) \|V - \bar{V}\|
\]
Let us recall that the bounds on the diameters ultimately depend on the initial conditions via (28). If the initial flock $X, V$ is known and the perturbation $\tilde{X}, \tilde{V}$ is relatively small, then those diameters can be quantified by initial values of $D_0$ and $A_0$. So, let us simply denote

$$\gamma = \max\{\phi(D), \phi(\tilde{D})\}.$$  

We have obtained the system

$$\begin{align*}
\frac{d}{dt} &\|V - \tilde{V}\| \leq 2M\lambda \|\nabla \phi\|_{\infty} A_0 e^{-t\lambda \gamma M} \|X - \tilde{X}\| - \lambda \gamma M \|V - \tilde{V}\| \\
\frac{d}{dt} &\|X - \tilde{X}\| \leq \|V - \tilde{V}\|.
\end{align*}$$  

Let us simply rewrite it as

$$
x' \leq v, \quad v' \leq ae^{-bt}x - bv.
$$  

It is elementary to obtain a bound on solutions. Indeed, denoting $w = ve^{bt}$ we obtain

$$x' \leq we^{-bt}, \quad w' \leq ax.$$  

Multiplying by factors to equalize the right hand sides, we obtain

$$\frac{d}{dt}(ax^2 + e^{-bt}w^2) \leq 4axwe^{-bt} \leq 2e^{-bt/2}\sqrt{a(ax^2 + e^{-bt}w^2)}.$$  

This immediately implies

$$ax^2 + e^{bt}v^2 \leq \frac{4\sqrt{a}}{b}(ax_0^2 + v_0^2).$$  

We can read off bounds for each parameter individually:

$$x \leq \frac{2}{a^{1/4}b^{1/2}} \sqrt{ax_0^2 + v_0^2}, \quad v \leq e^{-bt/2}\frac{2a^{1/4}}{b^{1/2}} \sqrt{ax_0^2 + v_0^2}.$$  

**Theorem 2.6.** The following bound holds for any pair of solutions to (21) with the same momentum:

$$a\|X - \tilde{X}\|^2 + e^{bt}\|V - \tilde{V}\|^2 \leq \frac{4\sqrt{a}}{b}(a\|X_0 - \tilde{X}_0\|^2 + \|V_0 - \tilde{V}_0\|^2),$$

where $a = 2M\lambda \|\nabla \phi\|_{\infty} A_0$ and $b = \lambda \gamma M$, $\gamma = \max\{\phi(D), \phi(\tilde{D})\}$.

**2.5. Singular kernels and the issue of collisions.** Let us introduce into consideration the singular kernels (5), which will play a crucial role in macroscopic description of the system. Clearly, singularity at the origin emphasizes predominantly local communication as is natural in some applications. However, with singularity we stumble upon the issue of well-posedness of the system (11) in the case when agents encounter collisions. In fact collisions are common in the bounded kernel case.

**Example 2.7.** Let us assume that $\phi = 1$ in a neighborhood of 0. Let us arrange two agents $x = x_1 = -x_2$ with $0 < x(0) = \varepsilon \ll 1$. And let $v_1 = -v_2 < 0$ be very large. Clearly $x(t)$ will remain in the same neighborhood of 0 as where it has started, and so the system reads

$$\frac{d}{dt}x = v, \quad \frac{d}{dt}v = -2v.$$  

Solving it explicitly we can see that the two agents will collide at the origin.

Heuristically, however, strong singularity should prevent such collisions. The induced alignment forces should become strong enough to correct velocities of converging agents before collision happens in the first place. Exactly how singular the kernel should be can be seen from the following example.
Example 2.8. Let the kernel be given by (5) and let us consider the same setup as previously. Then we obtain the system
\[
\frac{d}{dt}x = v, \quad \frac{d}{dt}v = -2\frac{v}{x^\beta}.
\]
This system has a conservation law provided \( \beta < 1 \): \(v + \frac{2^{1-\beta}}{1-\beta} = C_0\). So, if initially \( C_0 \ll 0 \), then \( v < C_0 \ll 0 \) as well. This means that \( x \) will reach the origin in finite time.

This example demonstrates that the threshold singularity necessary to prevent collisions must be non-integrable. It is indeed true as we prove in the following theorem.

**Theorem 2.9.** Under the strong singularity condition (8) the flock experiences no collisions between agents for any non-collisional initial datum. Consequently, any non-collisional initial datum gives rise to a unique global solution.

In view of global existence and absence of collisions the content of Theorem 2.4 holds true as stated provided the kernel is singular (8) condition.

**Proof.** For a given non-collisional initial data \((x_i, v_i)\) let us assume that collision occurs first time at \( t = T^* \). Denote by \( I^* \subset \{1, ..., N\} \) the indexes of agents that collided at one point in space – note that this may not be unique collection. Consequently, there is a \( \delta > 0 \) such that \(|x_{ik}(t)| \geq \delta\) for all \( i \in I^* \) and \( k \in \Omega \setminus I^* \). Let us denote
\[
D^*(t) = \max_{i,j \in I^*} |x_{ij}(t)|, \quad A^*(t) = \max_{i,j \in I^*} |v_{ij}(t)|, \quad t < T^*.
\]
Directly from the characteristic equation we obtain \( \frac{d}{dt} D^* \leq A^* \), and hence
\[
-\frac{d}{dt} D^* \leq A^*.
\]
For velocity variation we obtain, using a maximizing triple \( \ell \in (\mathbb{R}^n)^*, i, j \in I^* \):
\[
\frac{d}{dt} A^* = \frac{1}{N} \sum_{k=1}^{N} \phi_{ik\ell}(v_{ki}) - \phi_{kj\ell}(v_{kj}) = \frac{1}{N} \sum_{k \in I^*} \phi_{ik\ell}(v_{kj} - v_{ij}) + \phi_{kj\ell}(-v_{kj} - v_{ij}) + \frac{1}{N} \sum_{k \not\in I^*} \phi_{ik\ell}(v_{ki}) - \phi_{kj\ell}(v_{kj}).
\]
In the first sum we notice that all terms are negative, so we can pull out the minimal value of the kernel, which is \( \phi(D^*) \). In the second sum, all the distance \(|x_{ik}|, |x_{kj}|\) stay away from zero up to \( T^* \). So, the kernels and the whole sum remains bounded. In summary we obtain
\[
\frac{d}{dt} A^* \leq C_1 - C_2 \phi(D^*) A^*.
\]
Considering the energy functional
\[
E(t) = A^*(t) + C_2 \int_{D^*(t)}^1 \phi(r) \, dr,
\]
we readily find that \( \frac{d}{dt} E \leq C_1 \), hence \( E \) remains bounded up to the critical time. This means that \( D^*(t) \) cannot approach zero value.

The global existence part is now a routine application of the Picard iteration and the standard continuation argument. \( \square \)

When the power of kernel \( \beta \geq 2 \) a quantitative estimate on the minimal distance between agents can be derived. Assuming that the singularity assumption holds the local region \( r < r_0 \):
\[
\phi(r) \gtrsim \frac{1_{r < r_0}}{r^\beta},
\]
let us consider the collision functional

\[ C = \begin{cases} 
\frac{1}{N^2} \sum_{i,j=1}^{N} \frac{1}{(|x_{ij}| \cap r_0)^{\beta-2}}, & \beta > 2 \\
\frac{1}{N^2} \sum_{i,j=1}^{N} \ln(|x_{ij}| \cap r_0), & \beta = 2.
\end{cases} \]

For \( \beta > 2 \) we have for the derivative

\[
\frac{dC}{dt} = \frac{(2-\beta)}{N^2} \sum_{i,j=1}^{N} \frac{|x_{ij}| \cap r_0}{(|x_{ij}| \cap r_0)^{\beta-1}} \leq \frac{|\beta - 2|}{N^2} \sum_{i,j=1}^{N} \frac{1}{(|x_{ij}| \cap r_0)^{\beta-1}} |v_{ij}| I_{|x_{ij}| < r_0} \\
\leq |\beta - 2| \left( \frac{1}{N^2} \sum_{i,j=1}^{N} v_{ij}^2 \frac{1}{|x_{ij}|^\beta} I_{|x_{ij}| < r_0} \right)^{1/2} \left( \frac{1}{N^2} \sum_{i,j=1}^{N} |x_{ij}|^{2-\beta} I_{|x_{ij}| < r_0} \right)^{1/2} \\
\leq C \sqrt{I_2} \sqrt{C}.
\]

This implies

\[ \sqrt{C(t)} \leq \sqrt{C(0)} + C \int_0^t \sqrt{I_2(s)} \, ds, \]

and recalling that \( I_2 \) is integrable on \( \mathbb{R}^+ \) we conclude

\[ C(t) \lesssim t. \]

For \( \beta = 2 \), a similar computation gives \( \frac{dC}{dt} \leq C \sqrt{I_2} \), hence \( C(t) \lesssim \sqrt{t} \). We thus arrive at the following bounds

\[ |x_{ij}(t)| \geq \begin{cases} 
\frac{c}{t^{\beta-2}}, & \beta > 2, \\
\frac{c}{e^{C\sqrt{t}}}, & \beta = 2.
\end{cases} \]

2.6. Degenerate communication. Corrector Method. Let us consider higher order variations

\[ V_p = \frac{1}{pN^2} \sum_{i,j=1}^{N} |v_i - v_j|^p, \quad p \geq 1, \]

\[ I_p = \frac{1}{N^2} \sum_{i,j=1}^{N} \phi(x_i, x_j)|v_i - v_j|^p. \]

We observe that \( V_p \)'s are non-increasing. Indeed,

\[
\frac{d}{dt} V_p = \frac{1}{N^3} \sum_{i,j,k} |v_{ij}|^{p-2} v_{ij} \cdot (v_{ki} \phi_{ki} - v_{kj} \phi_{kj}) = \frac{2}{N^3} \sum_{i,j,k} |v_{ij}|^{p-2} v_{ij} \cdot v_{ki} \phi_{ki} \\
= \frac{1}{N^3} \sum_{i,j,k} (|v_{ij}|^{p-2} v_{ij} - |v_{kj}|^{p-2} v_{kj}) \cdot v_{ki} \phi_{ki} \\
= \frac{1}{N^3} \sum_{i,j,k} (|v_{ij}|^{p-2} v_{ij} - |v_{kj}|^{p-2} v_{kj}) \cdot (v_{kj} - v_{ij}) \phi_{ki},
\]

with the convention that \( |v_{ij}|^{p-2} v_{ij} = 0 \) if \( i = j \). The right hand side is non-positive due to the elementary inequality

\[ (|a|^{p-2}a - |b|^{p-2}b) \cdot (a - b) \geq 0. \]
The two special cases, \( i = j \) and \( k = j \), produce the term \(-|v_{ik}|^p - |v_{ij}|^p\). So, in general we have an \( N \)-dependent inequality

\[
\frac{d}{dt} V_p \leq -\frac{1}{N} \mathcal{I}_p, \quad p \geq 1.
\]

For the case \( p = 2 \) we have the \( N \)-independent energy law:

\[
\frac{d}{dt} V_2 = -\mathcal{I}_2.
\]

**Theorem 2.10.** Suppose the kernel \( \phi \geq 0 \) is either smooth or satisfying (8). Suppose also that it dominates a monotone fat tail: there exists a non-increasing \( \Phi(r) \), such that for some \( r_0 > 0 \)

\[
\phi(r) \geq \Phi(r), \forall r > r_0, \quad \text{and} \quad \int_{r_0}^{\infty} \Phi(r) \, dr = \infty.
\]

Then

(i) Any solution to the discrete system (11) aligns: \( V_2(t) \leq \frac{C}{N} t \), and \( D(t) \leq \mathcal{D}_N \), with constants depending on \( N \).

(ii) If \( \phi \) is smooth, then any solution to the discrete system (11) aligns: \( V_4 \to 0 \), with a rate independent of \( N \).

The advantage of (i) over (ii) is that it gives a faster, although \( N \)-dependent, rate as well as flocking. It also holds for singular kernels satisfying non-collision condition (8). However, it is not extendable to the macroscopic or kinetic case, while (ii) is.

**Proof of Theorem 2.10 (i).** We start by defining the following corrector

\[
\mathcal{G} = \frac{1}{N^2} \sum_{i,j,k=1}^{N} |v_{ij}| \psi(d_{ij}) \chi(|x_{ij}|),
\]

where \( d_{ij} \) is a longitudinal displacement function defined by

\[
d_{ij} = -x_{ij} \cdot \frac{v_{ij}}{|v_{ij}|},
\]

and the two auxiliary functions \( \chi : \mathbb{R}^+ \to \mathbb{R}^+ \) and \( \psi : \mathbb{R} \to \mathbb{R}^+ \) are defined by

\[
\chi(r) = \begin{cases} 
1, & r < r_0, \\
2 - \frac{r}{r_0}, & r_0 \leq r \leq 2r_0, \\
0, & r > 2r_0,
\end{cases}
\quad \text{and} \quad
\psi(d) = \begin{cases} 
0, & d < -r_0, \\
d + r_0, & |d| \leq r_0, \\
2r_0, & d > r_0.
\end{cases}
\]

Let us compute the derivative of the corrector

\[
\frac{d}{dt} \mathcal{G} = -\frac{1}{N^2} \sum_{i,j=1}^{N} |v_{ij}|^2 \mathbb{1}_{|d_{ij}|<r_0} \chi(|x_{ij}|) + \mathcal{R}_1 + \mathcal{R}_2 + \mathcal{R}_3,
\]

where

\[
\mathcal{R}_1 = 2 \frac{3}{N^3} \sum_{i,j,k=1}^{N} \frac{v_{ij}}{|v_{ij}|} \cdot v_{ki} \phi(x_{ik}) \psi(d_{ij}) \chi(|x_{ij}|),
\]

\[
\mathcal{R}_2 = -2 \frac{3}{N^3} \sum_{i,j,k=1}^{N} x_{ij} \cdot \left( \mathbb{1} - \frac{v_{ij} \otimes v_{ij}}{|v_{ij}|^2} \right) v_{ki} \phi(x_{ik}) \mathbb{1}_{|d_{ij}|<r_0} \chi(|x_{ij}|)
\]

\[
\mathcal{R}_3 = \frac{1}{N^2} \sum_{i,j=1}^{N} |v_{ij}| \psi(d_{ij}) \chi'(|x_{ij}|) \frac{x_{ij}}{|x_{ij}|} \cdot v_{ij}.
\]
The first term on the right hand side of (50) patches up communication where it was originally missing – at the close range. Let us address it first. Without loss of generality we can assume that \( \Phi \) is a bounded decreasing function on \( \mathbb{R}^+ \). Hence,

\[
1_{r<r_0} + \phi(r) \geq c \Phi(r),
\]

for all \( r > 0 \) and some \( c > 0 \). Using \( 1_{|d_{ij}|<r_0} \geq 1_{|x_{ij}|<r_0} \), we have

\[
-\frac{1}{N^2} \sum_{i,j} |v_{ij}|^2 1_{|d_{ij}|<r_0} \chi(x_{ij}) \leq -\frac{1}{N^2} \sum_{i,j} |v_{ij}|^2 1_{|x_{ij}|<r_0}
= -\frac{1}{N^2} \sum_{i,j} |v_{ij}|^2 (1_{|x_{ij}|<r_0} + \phi(x_{ij})) + \frac{1}{N^2} \sum_{i,j} |v_{ij}|^2 \phi(x_{ij})
\leq -c\Phi(D)V_2 + I_2.
\]

Proceeding to the error terms, by construction of the auxiliary functions, we have

\[
|R_1|, |R_2| \lesssim I_1,
\]

and

\[
R_3 \leq \frac{1}{N^2} \sum_{i,j} |v_{ij}|^2 |\chi'(x_{ij})| \leq \frac{1}{N^2} \sum_{i,j} |v_{ij}|^2 1_{r_0<|x_{ij}|<2r_0} \lesssim \frac{1}{N^2} \sum_{i,j} |v_{ij}|^2 \phi(x_{ij}) = I_2.
\]

Hence, with constants \( a, b, c \) only depending on \( \phi \), we obtain

\[
\frac{d}{dt} G \leq -c\Phi(D)V_2 + aI_2 + bI_1.
\]

We can now define a new Lyapunov functional:

(51)

\[
\mathcal{L} = G + aV_2 + bNV_1,
\]

\[
\frac{d}{dt} \mathcal{L} \leq -c\Phi(D)V_2.
\]

Using that \( \frac{d}{dt} D \leq C_N \sqrt{V_2} \), we find another Lyapunov functional:

\[
\tilde{\mathcal{L}} = \mathcal{L} + \frac{c}{C_N \sqrt{V_2}} \int_0^D \Phi(r) \, dr.
\]

Consequently, due fat tail on \( \Phi \), we have flocking \( D(t) \leq \overline{D}_N \) for all time and for some \( N \)-dependent \( \overline{D}_N \). Returning to (51) we conclude that

\[
A := \int_0^\infty V_2(t) \, dt < \infty.
\]

So, on every time interval \([T, e^AT] \) there is a \( t \) such that \( V_2(t) \leq \frac{1}{7} \). By monotonicity, this implies a similar bound for all \( t \). \( \square \)

**Proof of Theorem 2.10 (ii).** To achieve \( N \)-independent, although slower, rate we consider a third order corrector given by

\[
\mathcal{G}_3 = \frac{1}{N^2} \sum_{i,j} |v_{ij}|^3 \psi(d_{ij}) \chi(|x_{ij}|),
\]

with \( \psi \) and \( \chi \) defined previously. In this case

\[
\frac{d}{dt} \mathcal{G}_3 = -\frac{1}{N^2} \sum_{i,j=1}^N |v_{ij}|^4 1_{|d_{ij}|<r_0} \chi(|x_{ij}|) + R_1 + R_2 + R_3,
\]
with
\[ R_1 = \frac{6}{N^3} \sum_{i,j,k=1}^{N} |v_{ij}|v_{ij} \cdot v_{ki} \phi_{ik} \psi(d_{ij}) \chi(|x_{ij}|), \]
\[ R_2 = -\frac{2}{N^3} \sum_{i,j,k=1}^{N} |v_{ij}|^2 x_{ij} \cdot \left( I - \frac{v_{ij} \otimes v_{ij}}{|v_{ij}|^2} \right) v_{ki} \phi_{ki} \mathbb{1}_{|d_{ij}| < r_0} \chi(|x_{ij}|) \]
\[ R_3 = \frac{1}{N^2} \sum_{i,j=1}^{N} |v_{ij}|^3 \psi(d_{ij}) \chi'(|x_{ij}|) \frac{x_{ij}}{|x_{ij}|} \cdot v_{ij}. \]

The gain term is estimated as before with the use of a priori uniform bound on velocities $|v_i(t)| \leq |v(0)|_\infty$:

\[ -\frac{1}{N^2} \sum_{i,j=1}^{N} |v_{ij}|^4 \mathbb{1}_{|d_{ij}| < r_0} \chi(|x_{ij}|) \leq -c \Phi(D) V_4 + |v(0)|_\infty I_2. \]  

(52)

Next,
\[ R_3 \lesssim \frac{1}{N^2} \sum_{i,j} |v_{ij}|^3 \phi(|x_{ij}|) \lesssim I_2. \]

By Young’s inequality for a small $\epsilon > 0$ to be specified later,
\[ \frac{\epsilon}{N^2} \sum_{i,j=1}^{N} |v_{ij}|^4 |x_{ij}|^2 \chi(|x_{ij}|)^2 + \frac{1}{\epsilon} |\phi|_\infty \sum_{i,k=1}^{N} |v_{ki}|^2 \phi_{ki}. \]

Note that $\chi(|x_{ij}|) = 0$ if $|x_{ij}| > 2r_0$. So,
\[ R_2 \lesssim \frac{\epsilon}{N^2} \sum_{i,j=1}^{N} |v_{ij}|^4 \mathbb{1}_{|x_{ij}| \leq 2r_0} + \frac{1}{\epsilon N^2} \sum_{i,k=1}^{N} |v_{ki}|^2 \phi_{ki} \]
\[ \leq \frac{\epsilon}{N^2} \sum_{i,j=1}^{N} |v_{ij}|^4 \mathbb{1}_{|x_{ij}| \leq r_0} + \frac{\epsilon}{N^2} \sum_{i,j=1}^{N} |v_{ij}|^4 \mathbb{1}_{r_0 < |x_{ij}| < 2r_0} + \frac{1}{\epsilon N^2} \sum_{i,k=1}^{N} |v_{ki}|^2 \phi_{ki} \]

For $\epsilon$ sufficiently small the first sum gets absorbed into the gain term (52). The second and third sums are dominated by $I_2$. The term $R_1$ can be estimated in exact same manner. We obtain
\[ \frac{d}{dt} G_3 \leq -c \Phi(D) V_4 + a I_2. \]

By analogy with previous proof we define the Lyapunov functional $L = G_3 + a V_2$ and conclude that

\[ \int_0^\infty \Phi(D(t)) V_4(t) dt < \infty, \]

(53)

with the bound being independent of $N$. Furthermore, due to uniform bound on velocity, $D(t) \leq ct + D_0$. Thus,

\[ \int_0^\infty \Phi(ct + D_0) V_4(t) dt < \infty. \]

(54)

By virtue of the fat tail condition on $\Phi$, $V_4$ cannot be bounded away from the zero. And in view of its monotonicity, $V_4 \to 0$. \[\Box\]
Remark 2.11. In either of the results above we can actually extract a specific rate of alignment if we make an explicit assumption about the tail of the kernel. Thus, if
\[ \phi(r) \sim \frac{1}{r^\beta}, \quad \forall r > r_0, \]
and \( \beta \leq 1 \), then from (53),
\[ \int_1^\infty \frac{1}{t^\beta} V_4(t) \, dt < \infty, \]
where \( V \) is the corresponding functional. So, if \( \beta = 1 \), then there exists an \( A > 0 \) such that for any \( T > 1 \) there exists a \( t \in [T, T^A] \) such that \( V_4(t) < \frac{1}{t^A} \). Since \( \ln t \) is proportional to \( \ln T \) for all \( t \in [T, AT] \) this proves the above bound for all large times. If \( \beta < 1 \), then we argue that for some large \( A > 0 \) and all \( T > 0 \) we find \( t \in [T, AT] \) such that \( V_4(t) \leq \frac{1}{t^{1-\beta}} \). But the latter is comparable for all values of \( t \in [T, AT] \). Thus we obtain the power rate as above for all \( t \).

Let us summarize the obtained results.
\[ V_4(t) \lesssim \frac{1}{\ln t}, \quad \beta = 1, \]
\[ V_4(t) \lesssim \frac{1}{t^{1-\beta}}, \quad \beta < 1. \]

Of course the same argument applies to \( V_2 \) under the conditions of part (ii).

2.7. Dynamics under external potential forces: confinement. One of the natural ways to force the Cucker-Smale model is to include a potential confinement to the momentum equation:
\[ \begin{cases} \dot{x}_i = v_i, \\ \dot{v}_i = \frac{1}{N} \sum_{j=1}^N \phi(x_i - x_j)(v_j - v_i) - \nabla U(x_i), \end{cases} \]
where \( U \) is a convex radially increasing potential. The macroscopic version of this system was analyzed by Shu and Tadmor in recent work [11]. The case is a good illustration just how dramatically the behavior of the system adapts to present forces and how crucially it relies on the particular structure of potential. While in general one can perform analysis in the case of a strictly convex potential, the results are not as conclusive as they rely on additional assumptions on the strength of communication. We focus here on one particular case of quadratic \( U \):
\[ U(x) = \frac{1}{2} |x|^2. \]

In this case, the system reads
\[ \begin{cases} \dot{x}_i = v_i, \\ \dot{v}_i = \frac{1}{N} \sum_{j=1}^N \phi(x_i - x_j)(v_j - v_i) - x_i, \end{cases} \]

The natural limiting state of this system is not, in fact, alignment, but rather a harmonic oscillator. Indeed, denoting as before \( \bar{x}, \bar{v} \) the center of mass and momentum, we obtain
\[ \frac{d}{dt} \bar{x} = \bar{v}, \quad \frac{d}{dt} \bar{v} = -\bar{x}. \]

Also due to the linear nature of the forcing we can shift the system of coordinates to \( (\bar{x}, \bar{v}) \) and assume that \( \bar{x} = 0, \bar{v} = 0 \). So, the question becomes to show that the solution tends to zero in the new reference frame (still denoted \( (x, v) \)).
The full energy of the system is given by
\[ \mathcal{E} = \mathcal{K} + \mathcal{P} \]
(60)
\[ \mathcal{K} = \frac{1}{2N} \sum_{i=1}^{N} |v_i|^2, \quad \mathcal{P} = \frac{1}{2N} \sum_{i=1}^{N} |x_i|^2. \]

It will also be important to consider the “particle energy”, i.e. \( L^\infty \)-version of the energy:
\[ \mathcal{E}_\infty = \frac{1}{2} \max_i (|v_i|^2 + |x_i|^2). \]
(61)

**Theorem 2.12.** Suppose that the kernel \( \phi \) is bounded, decreasing, and satisfies the weak fat tail condition
\[ \int_0^\infty r\phi(r) \, dr = \infty. \]
(62)

Then the system (58) settles on its harmonic oscillator (59) exponentially fast, meaning
\[ \max_i (|v_i(t) - \bar{v}(t)|^2 + |x_i(t) - \bar{x}(t)|^2) \leq Ce^{-\delta t}, \]
for some \( \delta > 0 \) and \( C = C(v_0, x_0, \phi) \) independent of \( N \).

**Proof.** The result amounts to establishing an exponential bound on \( \mathcal{E}_\infty \).

Straight from the equations we obtain the energy law
\[ \frac{d}{dt} \mathcal{E} = -I_2 \leq -\phi(D)\mathcal{K}. \]
(64)

Note that the dissipation is not coercive even if we knew a lower bound \( \phi(D) > c_0 \). However, we proceed with \( \mathcal{E}_\infty \):
\[ \frac{d}{dt} \mathcal{E}_\infty \leq \frac{1}{N} \sum_j \phi_{ij}(v_j - v_i) \cdot v_i \leq \frac{1}{2N} \sum_j \phi_{ij}(|v_j|^2 - |v_i|^2) \leq \frac{1}{2N} \sum_j \phi_{ij}|v_j|^2 \leq |\phi|_\infty \mathcal{K}. \]

To combine the two equations into one system, let us note that \( D \leq 4\sqrt{\mathcal{E}_\infty} \). Thus,
\[ \frac{d}{dt} \mathcal{E} \leq -\phi(4\sqrt{\mathcal{E}_\infty})\mathcal{K}, \]
\[ \frac{d}{dt} \mathcal{E}_\infty \leq CK. \]
(65)

Consider the Lyapunov function
\[ \mathcal{L} = \mathcal{E} + \frac{1}{C} \int_0^{\mathcal{E}_\infty} \phi(4\sqrt{r}) \, dr. \]

From our fat tail assumption it follows that \( \mathcal{E}_\infty \) remains bounded, from which we conclude that \( D(t) \leq D_\infty \). Now the energy law (64) reads
\[ \frac{d}{dt} \mathcal{E} \leq -c_0 \mathcal{K}. \]
(66)

Next we proceed with a hypocoercivity argument to restore \( \mathcal{K} \) on the right hand side to the full energy \( \mathcal{E} \). We consider the corrector (longitudinal momentum)
\[ \mathcal{X} = \frac{1}{N} \sum_{i=1}^{N} x_i \cdot v_i, \]
(67)
and note that for $\lambda < \frac{1}{4}$, $\mathcal{E} + \lambda \mathcal{X} \sim \mathcal{E}$. Let us compute the derivative:

$$
\frac{d}{dt} \mathcal{X} = \frac{1}{N^2} \sum_{j,i} \phi_{ij} (v_j - v_i) \cdot x_i + \frac{1}{N} \sum_i (|v_i|^2 - |x_i|^2)
$$

$$
\leq \frac{\|\phi\|_{\infty}}{N^2} \sum_{j,i} (4|v_j|^2 + 4|v_i|^2 + \frac{1}{2}|x_i|^2) + \frac{1}{N} \sum_i (|v_i|^2 - |x_i|^2) \leq C\mathcal{K} - \frac{1}{4} \mathcal{P}.
$$

Choosing $\lambda < c_0/2C$ we obtain

$$
\frac{d}{dt} (\mathcal{E} + \lambda \mathcal{X}) \leq -c_1 \mathcal{E} \sim -(\mathcal{E} + \lambda \mathcal{X}).
$$

This establishes exponential decay of $\mathcal{E}$.

Now that the $L^2$-energy is decaying exponentially, we can extrapolate to obtain exponential decay for $\mathcal{E}_\infty$ as well. The method is similar – we consider an amended version of $\mathcal{E}_\infty$:

$$
\mathcal{E}_\infty^\lambda = \frac{1}{2} \max_i (|v_i|^2 + |x_i|^2 + \lambda x_i \cdot v_i).
$$

Again, for $\lambda < \frac{1}{4}$, this does not alter the particle energy much, $\mathcal{E}_\infty \sim \mathcal{E}_\infty^\lambda$. Differentiating at a point of maximum we obtain

$$
\frac{d}{dt} \mathcal{E}_\infty^\lambda \leq \frac{1}{N} \sum_j \phi_{ij} (v_j - v_i) \cdot (v_i + \lambda x_i) + \lambda |v_i|^2 - \lambda |x_i|^2
$$

$$
\leq \frac{1}{N} \sum_j \phi_{ij} v_j \cdot v_i + \lambda \frac{1}{N} \sum_j \phi_{ij} v_j \cdot x_i - \frac{1}{N} \sum_j \phi_{ij} |v_i|^2 - \lambda \frac{1}{N} \sum_j \phi_{ij} v_i \cdot x_i + \lambda |v_i|^2 - \lambda |x_i|^2.
$$

In the gain term $-\frac{1}{N} \sum_j \phi_{ij} |v_i|^2$ we replace $\phi_{ij}$ by a lower bound $c_0$:

$$
-\frac{1}{N} \sum_j \phi_{ij} |v_i|^2 \leq -c_0 |v_i|^2.
$$

and in the rest we simply use boundedness of the kernel. Hence, if $\lambda$ is small enough, the term $\lambda |v_i|^2$ gets absorbed. We obtain at this point

$$
\frac{d}{dt} \mathcal{E}_\infty^\lambda \lesssim \frac{1}{N} \sum_j |v_j||v_i| + \lambda \frac{1}{N} \sum_j |v_j||x_i| + \lambda \frac{1}{N} \sum_j |v_i||x_i| - c_1 |v_i|^2 - \lambda |x_i|^2
$$

$$
\lesssim 4\mathcal{K} + \frac{c_1}{4} |v_i|^2 + 4\lambda \mathcal{K} + \frac{1}{4} \lambda |x_i|^2 + \lambda |v_i|^2 + \frac{1}{4} \lambda |x_i|^2 - c_1 |v_i|^2 - \lambda |x_i|^2
$$

$$
\lesssim \mathcal{K} - |v_i|^2 - |x_i|^2 \lesssim \mathcal{K} - \mathcal{E}_\infty^\lambda.
$$

Since we already know that $\mathcal{K}$ is exponentially decaying, this establishes a similar bound on $\mathcal{E}_\infty^\lambda$. □

Let us note that for small number of agents, when the dependence on $N$ is not an issue, one can actually obtain a much stronger result: as long as $\phi(r) > 0$ for all $r > 0$, the conclusion of the theorem holds true. Indeed, from the first lines we have established that the total energy is decaying and, hence, bounded. But from the potential part $\mathcal{P}$ this immediately implies that the flock is bounded, with a bound depending on $N$. This sets the rest of the argument to go through.

This observation clearly shows that it is impossible to construct a simple example with a few agents, like we did in Example 2.3, to prove sharpness of the fat tail condition (62).
2.8. **Attraction-Alignment models.** We consider the system with pairwise interactions determined by a radially symmetric smooth potential \( U \in C^2(\mathbb{R}_+) \):

\[
\begin{cases}
\dot{x}_i = v_i, \\
\dot{v}_i = \frac{1}{N} \sum_{j=1}^{N} \phi(x_i - x_j)(v_j - v_i) - \frac{1}{N} \sum_{j=1}^{N} \nabla U(x_i - x_j).
\end{cases}
\]  

(69)

Note that in general it may not be possible to achieve a natural aligned steady state as it may not even exist for the system with repulsion/attraction forces. Indeed all such solutions will need to land on the minimum of \( U \), which is this case is an annulus. In other words all \(|x_{ij}| \in [r_0, r_1]|\). This is only possible with a limited number of agents, \( N \), where \( N \) depends on the width of the annulus. This 3Zone model, however, is very important a mathematical realization of the classical Craig Reynolds’ theory of flock formations, and is an underlying tool in 3D computer animations for picture production. We focus here on two less restrictive cases: repulsion-only or attraction-only potential, the so called 2Zone models.

Notice first that the system (69) preserves momentum and is Galilean invariant. So, we can shift the center of mass and momentum to zero. Let us define the classical energies

\[
E = K + P,
\]

(70)

\[
K = \frac{1}{2N} \sum_{i=1}^{N} |v_i|^2, \quad P = \frac{1}{N^2} \sum_{i,j=1}^{N} U(x_{ij}),
\]

the forces

\[
F_i = \frac{1}{N} \sum_{j=1}^{N} \nabla U(x_i - x_j), \quad F = (F_1, \ldots, F_N), \quad |F|^2 = \frac{1}{N} \sum_{j=1}^{N} |F_j|^2.
\]

The energy satisfies the classical law:

\[
\frac{d}{dt}E = -\frac{1}{N^2} \sum_{i,j=1}^{N} \phi_{ij} |v_{ij}|^2 = -\mathcal{I}.
\]

(71)

In this section we consider an Attraction-Alignment 2Zone model with non-degenerate communication kernel:

\[
\phi'(r) \leq 0, \quad \phi(r) \geq \frac{c_0}{(r)^\gamma}, \quad \text{for } r \geq 0.
\]

(72)

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

- **Support:** \( U \in C^2(\mathbb{R}_+), \quad U(r) = 0, \quad \forall r \leq L \),

- **Growth:** \( U(r) \geq a_0 r^\beta, \quad |U'(r)| \leq a_1 r^{\beta-1}, \quad |U''(r)| \leq a_2 r^{\beta-2}, \quad \forall r > L' \),

- **Convexity:** \( U'(r), U''(r) \geq 0, \quad \forall r > 0 \).

**Theorem 2.13.** Under the assumptions (72) and (73) on the kernel and potential in the range of parameters given by

\[
\gamma < \begin{cases} 
1, & 1 < \beta < \frac{4}{3}, \\
\frac{3}{2} \beta - 1, & \frac{4}{3} \leq \beta < 2, \\
2, & \beta \geq 2,
\end{cases}
\]

(74)

all solutions to the system (69) flock

\[
\mathcal{D}(t) \leq \overline{\mathcal{D}} < \infty,
\]
and align

\begin{equation}
\mathcal{E}(t) \leq \frac{C_\delta}{(t)^{1-\beta}}, \quad \forall \delta > 0.
\end{equation}

**Proof.** We will operate with the particle energy defined similarly to the confinement case:

\begin{equation}
\mathcal{E}_i = \frac{1}{2}|v_i|^2 + \frac{1}{N} \sum_{k=1}^{N} U(x_{ik}), \quad \mathcal{E}_\infty = \max_i \mathcal{E}_i.
\end{equation}

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

\begin{equation}
\mathcal{E}_i \geq U(x_i) \geq (|x_i| - L')^\beta.
\end{equation}

So,

\begin{equation}
\mathcal{D} \leq \mathcal{E}_\infty^{1/\beta} + L'.
\end{equation}

Let us now establish a bound on \(\mathcal{E}_\infty\). For each \(i\) we compute

\begin{equation}
\frac{d}{dt} \mathcal{E}_i = \frac{1}{N} \sum_{k=1}^{N} \phi_{ik} \phi_{ik} \cdot v_i - \frac{1}{N} \sum_{k=1}^{N} \nabla U(x_{ik}) \cdot v_k.
\end{equation}

For the kinetic part we use the identity

\begin{equation}
v_{ki} \cdot v_i = -\frac{1}{2}|v_{ki}|^2 - \frac{1}{2}|v_i|^2 + \frac{1}{2}|v_k|^2.
\end{equation}

Discarding all the negative terms, we bound

\begin{equation}
\frac{1}{N} \sum_{k=1}^{N} \phi_{ik} v_{ki} \cdot v_i \leq |\phi|_\infty K.
\end{equation}

Due to the energy law, \(K\), of course will remain bounded, but we will keep it for now. As to the potential term, there are several ways we can handle it.

For any \(1 \leq \beta \leq 4/3\) we can derive a direct estimate from the first derivative:

\begin{equation}
\left| \frac{1}{N} \sum_{k=1}^{N} \nabla U(x_{ik}) \cdot v_k \right| \leq \sqrt{K} \left( \frac{1}{N} \sum_{k=1}^{N} |\nabla U(x_{ik})|^2 \right)^{1/2} \leq \sqrt{KD}^{\beta-1}.
\end{equation}

Consequently,

\begin{equation}
\frac{d}{dt} \mathcal{E}(i) \leq c_1 K + c_2 \sqrt{KD}^{\beta-1} \leq \sqrt{K}(1 + \mathcal{E}_\infty^{\beta-1}),
\end{equation}

and

\begin{equation}
\frac{d}{dt} \mathcal{E}_\infty \leq c_3 \sqrt{K}(1 + \mathcal{E}_\infty^{\beta-1}) \quad \Rightarrow \quad \mathcal{E}_\infty \lesssim \langle t \rangle^{\beta} \quad \Rightarrow \quad \mathcal{D} \lesssim \langle t \rangle.
\end{equation}

In the range \(4/3 \leq \beta \leq 2\) it is better to make use of the second derivative:

\begin{equation}
\left| \frac{1}{N} \sum_{k=1}^{N} \nabla U(x_{ik}) \cdot v_k \right| = \frac{1}{N} \sum_{k=1}^{N} (\nabla U(x_{ik}) - \nabla U(x_i)) \cdot v_k \leq \|D^2 U\|_\infty \sqrt{K} \left( \frac{1}{N} \sum_{k=1}^{N} |x_k|^2 \right)^{1/2}
\end{equation}

\begin{equation}
\leq c_4 \sqrt{K} \left( \frac{1}{N^2} \sum_{i,j=1}^{N} |x_{ij}|^2 \right)^{1/2}.
\end{equation}
The following inequality will be used repeatedly

\[ \frac{1}{N^2} \sum_{i,j=1}^{N} |x_{ij}|^2 \leq (L')^2 + \frac{1}{N^2} \sum_{i,j=1}^{N} (|x_{ij}| - L')^2 \leq C(1 + D^{(2-\beta) + \mathcal{P}}). \]

Continuing the above,

\[ \left| \frac{1}{N} \sum_{k=1}^{N} \nabla U(x_{ik}) \cdot v_k \right| \leq c_4 \sqrt{K} (1 + D^{2-\beta}) \leq c_5 \sqrt{K} (1 + \mathcal{E}_\infty)^{\frac{2-\beta}{2\beta}}. \]

In this case,

\[ \frac{d}{dt} \mathcal{E}_\infty \leq c_6 \sqrt{K} (1 + \mathcal{E}_\infty)^{\frac{2-\beta}{2\beta}} \implies \mathcal{E}_\infty \lesssim \langle t \rangle^{\frac{2\beta}{2\beta-2}} \implies D \lesssim \langle t \rangle^{\frac{2}{2\beta-2}}. \]

Finally, for \( \beta > 2 \), we argue similarly, using that \( |D^2 U(x_{ik})| \leq D^{\beta-2} \), and (83), to obtain

\[ \left| \frac{1}{N} \sum_{k=1}^{N} \nabla U(x_{ik}) \cdot v_k \right| \leq \sqrt{K} D^{\beta-2}, \]

and hence,

\[ \frac{d}{dt} \mathcal{E}_\infty \leq c_7 \sqrt{K} (1 + \mathcal{E}_\infty)^{\frac{\beta-2}{2\beta}} \implies \mathcal{E}_\infty \lesssim \langle t \rangle^{\frac{\beta}{1}} \implies D \lesssim \langle t \rangle^{\frac{1}{2}}. \]

We have proved the following a priori estimate:

\[ D(t) \lesssim \langle t \rangle^d, \quad \text{where} \quad d = \begin{cases} 1, & 1 \leq \beta < \frac{4}{3}, \\ \frac{2}{3\beta - 2}, & \frac{4}{3} \leq \beta < 2, \\ \frac{1}{2}, & \beta \geq 2. \end{cases} \]

Denote

\[ \zeta(t) = \langle t \rangle^{-\gamma d}. \]

According to the basic energy equation (71) we have

\[ \frac{d}{dt} \mathcal{E} \leq -\frac{1}{2} \mathcal{I} - c \zeta(t) \mathcal{K}. \]
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 (87). We introduce one more auxiliary power function 
\[ \eta(t) = \langle t \rangle^{-\alpha}, \quad \gamma d \leq \alpha < 1. \]
First, we consider the same longitudinal momentum 
\[ \mathcal{X} = \frac{1}{N} \sum_{i=1}^{N} x_i \cdot v_i. \]
It will come with a prefactor \( \varepsilon \eta(t) \), where \( \varepsilon \) is a small parameter. Let us estimate using (83):
\[
\varepsilon \eta(t) |\mathcal{X}| \leq \varepsilon \mathcal{K} + \varepsilon \eta^2(t) \frac{1}{N^2} \sum_{i,j=1}^{N} |x_{ij}|^2 \leq \varepsilon \mathcal{K} + c \varepsilon \eta^2(t) D^{(2-\beta)_+} \mathcal{P}.
\]
The potential term is bounded by \( \varepsilon \mathcal{P} \) as long as 
\[ 2\alpha \geq d(2 - \beta)_+. \]
Hence,
\[(88) \quad \varepsilon \eta(t) |\mathcal{X}| \leq \varepsilon \mathcal{E} + c \eta^2(t). \]
This shows that 
\[ \mathcal{E} + \varepsilon \eta(t) \mathcal{X} + 2c \eta^2(t) \sim \mathcal{E} + c \varepsilon \eta^2(t). \]
Let us now consider the derivative 
\[ \mathcal{X}' = \frac{1}{N^2} \sum_{i=1}^{N} |v_i|^2 + \frac{1}{N^2} \sum_{i,k=1}^{N} x_{ik} \cdot v_{ki} \phi_{ki} - \frac{1}{N^2} \sum_{i,k=1}^{N} x_{ik} \cdot \nabla U(x_{ik}) = \mathcal{K} + A - B. \]
The gain term \( B \), by convexity dominates the potential energy \( B \geq \mathcal{P} \). This is the main reason why we introduced the \( \mathcal{X} \)-corrector. As to \( A \):
\[
|A| \leq \frac{||\phi||_{\infty}}{2\varepsilon^{1/2} \eta(t)} I + \frac{\varepsilon^{1/2} \eta(t)}{2} \frac{1}{N^2} \sum_{i,j=1}^{N} |x_{ij}|^2 \lesssim \frac{1}{\varepsilon^{1/2} \eta(t)} I + \varepsilon^{1/2} \eta(t) + \varepsilon^{1/2} \eta(t) D^{(2-\beta)_+} \mathcal{P}.
\]
By requiring a more stringent assumption on parameters 
\[(89) \quad \alpha \geq d(2 - \beta)_+, \]
we can ensure that the potential term is bounded by \( \sim \varepsilon^{1/2} \mathcal{P} \), which can be absorbed by the gain term. So far, we have obtained
\[(90) \quad \frac{d}{dt}(\mathcal{E} + \varepsilon \eta(t) \mathcal{X} + 2c \eta^2(t)) \leq -c_1 \varepsilon \eta(t) \mathcal{E} + c_2 \eta^2(t) + \varepsilon \eta'(t) \mathcal{X}. \]
In view of (88),
\[
|\varepsilon \eta'(t) \mathcal{X}| \leq \varepsilon \frac{1}{\langle t \rangle} \eta(t) |\mathcal{X}| \leq \varepsilon \frac{1}{\langle t \rangle} \mathcal{E} + \varepsilon \frac{\eta^2(t)}{\langle t \rangle}.
\]
Since \( \alpha < 1 \), the energy term will be absorbed, and the free term is even smaller then \( \eta^2 \). Denoting 
\[ E = \mathcal{E} + \varepsilon \eta(t) \mathcal{X} + 2c \eta^2(t), \]
we obtain
\[
\frac{d}{dt} E \leq -c_1 \eta(t) E + c_2 \eta^2(t).
\]
By Duhamel’s formula,
\[
E(t) \lesssim \exp\{-\langle t \rangle^{1-\alpha} \} + \exp\{-\langle t \rangle^{1-\alpha} \} \int_{0}^{t} \frac{e(s)^{1-\alpha}}{\langle s \rangle^{2\alpha}} ds.
\]
By an elementary asymptotic analysis,
\[ \int_0^t \frac{e^{(s)^{1-\alpha'}}}{(s)^{\alpha''}} \, ds \sim \exp\{ (t)^{1-\alpha'} \} \frac{1}{(t)^{\alpha''-\alpha'}}. \]
Thus, we obtain an algebraic decay rate
\[ E(t) \lesssim \frac{1}{(t)^\alpha}, \quad \forall \alpha < 1, \]
provided
\[ d\gamma < 1 \quad \text{and} \quad d(2 - \alpha)_+ < 1. \]
This translates exactly into the conditions on \( \gamma \) given by (74), and (91) automatically implies (75).

Going back to the estimates (81) and (84), but keeping the kinetic energy with its established decay, we obtain a new decay rate for the diameter
\[ D \leq C_\delta (t)^\frac{2}{3}, \quad \forall \delta > 0. \]

At the next stage we prove flocking: \( D(t) < D_\infty \). In order to achieve this we return again to the particle energy estimates. Let us denote
\[ \mathcal{P}_i = \frac{1}{N} \sum_{k=1}^N U(x_{ik}), \quad \mathcal{I}_i = \frac{1}{N} \sum_{k=1}^N \phi_{ik} |v_{ki}|^2, \quad \mathcal{X}_i = x_i \cdot v_i. \]
Using (79), (80), (82), (83) and the fact that \( D^{2-\beta}_+ \mathcal{P} \) has a negative rate of decrease, we obtain
\[ \frac{d}{dt} \mathcal{E}_i \leq K - \frac{1}{2} \phi(D) |v_i|^2 - \mathcal{I}_i + c\sqrt{K} \lesssim -\frac{1}{2} \phi(D) |v_i|^2 - \mathcal{I}_i + \frac{1}{(t)^\frac{1}{2} - \delta}, \quad \forall \delta > 0. \]
In view of (92), we can pick \( \alpha \) and \( \delta \) such that
\[ \frac{d\gamma}{2} + \delta < \frac{1}{2} - 2\delta < \alpha < \frac{1}{2} - \delta \]
\[ (2 - \alpha)_+ d + 2\delta (2 - \beta)_+ < 2\alpha. \]
We use as before the auxiliary rate function \( \eta(t) = (t)^{-\alpha} \). Let us estimate the corrector
\[ |\varepsilon \eta(t) \mathcal{X}_i| \leq \varepsilon |v_i|^2 + \varepsilon \eta^2(t) |x_i|^2 \leq \varepsilon |v_i|^2 + \varepsilon \eta^2(t) D^{2-\beta}_+ \mathcal{P}_i + L^2 \varepsilon^2 \eta^2(t) \leq \varepsilon |v_i|^2 + c \varepsilon \mathcal{P}_i + L^2 \varepsilon^2 \eta^2(t). \]
So,
\[ E_i := \mathcal{E}_i + \varepsilon \eta(t) \mathcal{X}_i + 2L^2 \varepsilon \eta^2(t) \sim \mathcal{E}_i + L^2 \varepsilon \eta^2(t). \]
Differentiating,
\[ \mathcal{X}_i' = |v_i|^2 + \frac{1}{N} \sum_{k=1}^N x_i \cdot v_{ki} \phi_{ki} - \frac{1}{N} \sum_{k=1}^N x_{ik} \cdot \nabla U(x_{ik}) + \frac{1}{N} \sum_{k=1}^N x_k \cdot (\nabla U(x_{ik}) - \nabla U(x_i)) \]
\[ \leq |v_i|^2 + \varepsilon^{1/2} \eta(t) |x_i|^2 + \frac{1}{\varepsilon^{1/2} \eta(t)} \mathcal{I}_i - \mathcal{P}_i + \frac{1}{N^2} \sum_{l,k=1}^N |x_{kl}|^2 \]
\[ \leq |v_i|^2 + \varepsilon^{1/2} L^2 \eta(t) + \varepsilon^{1/2} D^{2-\beta}_+ \eta(t) \mathcal{P}_i + \frac{1}{\varepsilon^{1/2} \eta(t)} \mathcal{I}_i - \mathcal{P}_i + C \]
in view of (93), \( \varepsilon^{1/2} D^{2-\beta}_+ \eta(t) \leq \varepsilon^{1/2} \), so the potential term is absorbed by \(-\mathcal{P}_i\),
\[ \leq |v_i|^2 + \frac{1}{\eta(t)} \mathcal{I}_i - \frac{1}{2} \mathcal{P}_i + C. \]
Again in view of (93), \( \eta(t) \) decays faster than \( \phi(D) \), so plugging into the energy equation we obtain

\[
\frac{d}{dt} E_i \leq -\varepsilon \eta(t) E_i + \eta(t) + \sqrt{K} + \varepsilon \eta'(t) X_i,
\]

and as before \( \varepsilon \eta'(t) X \) is a lower order term which is absorbed in the negative energy term and \( + \eta^2 \). So,

\[
\frac{d}{dt} E_i \leq -\varepsilon \eta(t) E_i + \eta(t) + \sqrt{K}.
\]

By our choice of constants (93), \( \sqrt{K} \) decays faster than \( \eta(t) \), hence,

\[
\frac{d}{dt} E_i \lesssim -\varepsilon \eta(t) E_i + \eta(t).
\]

This proves boundedness of \( E_i \), and hence that of \( \mathcal{E}_i + L^2 \varepsilon \eta^2(t) \), and hence that of \( \mathcal{E}_i \). In view of (78), this implies flocking:

(94) \[
\mathcal{D}(t) < \mathcal{D}, \quad \forall t > 0.
\]

\[ \square \]

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:

(95) \[
U(r) \geq a_0 r^2, \quad r < L',
\]
then we can establish exponential alignment in terms of the energy \( \mathcal{E} \). Indeed, since we already know that the diameter is bounded, the basic energy equation reads

\[
\frac{d}{dt} \mathcal{E} \leq -c_0 K - \frac{1}{2} \mathcal{I}.
\]

The momentum corrector needs only an \( \varepsilon \)-prefactor to satisfy the bound

\[
|\varepsilon X| \leq \varepsilon K + \varepsilon c \mathcal{P}.
\]

This is due to the assumed quadratic order of the potential near the origin and, again, boundedness of the diameter. Hence, \( \mathcal{E} + \varepsilon X \sim \mathcal{E} \). The rest of the argument is similar to the confinement case. We obtain

\[
\mathcal{X} \lesssim \mathcal{K} + \varepsilon^{1/2} \mathcal{P} + \frac{1}{\varepsilon^{1/2}} \mathcal{I} - \mathcal{P} \leq \mathcal{K} - \frac{1}{2} \mathcal{P} \frac{1}{\varepsilon^{1/2}} \mathcal{I}
\]

Thus,

\[
\frac{d}{dt} (\mathcal{E} + \varepsilon \mathcal{X}) \leq -c_1 \mathcal{E} \sim -c_1 (\mathcal{E} + \varepsilon \mathcal{X}).
\]

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

\[
\frac{d}{dt} \mathcal{E}_i \leq -c_1 |\mathbf{v}_i|^2 - \frac{1}{2} \mathcal{I}_i + \text{Exp}.
\]

In view of \( |x_i|^2 \lesssim \mathcal{P}_i \),

\[
\varepsilon |X_i| \leq \varepsilon |\mathbf{v}_i|^2 + \varepsilon \mathcal{P}_i,
\]

so \( \mathcal{E}_i + \varepsilon X_i \sim \mathcal{E}_i \). Further following the estimates as in the proof,

\[
\mathcal{X}_i \lesssim |\mathbf{v}_i|^2 + \frac{1}{\varepsilon^{1/2}} \mathcal{I}_i - \frac{1}{2} \mathcal{P}_i.
\]

Thus,

\[
\frac{d}{dt} (\mathcal{E}_i + \varepsilon \mathcal{X}_i) \leq -c_1 (\mathcal{E}_i + \varepsilon \mathcal{X}_i) + \text{Exp}.
\]

This establishes exponential decay for \( \mathcal{E}_\infty \), and hence for the individual velocities. This also proves that \( \mathcal{D}(t) = \text{Exp} \). So, the alignment outcome here is exponential shrinking a point.
Theorem 2.14. Let us assume that the support of the potential spans the entire space and (95). Then the solutions flock and align exponentially fast:

\[ \mathcal{D}(t) + |v(t) - \bar{v}|_\infty \leq Ce^{-\delta t}, \]

for some \( C, \delta > 0 \).

2.9. Multi-flocks. It is clear from formulas (29) and (30) that the rate of alignment of a flock is proportional to the mass \( M \). At the same time it is inversely proportional to the diameter of the flock (through the upper bound \( \mathcal{D} \)). This creates a somewhat unrealistic scenario if two well separated flocks, each being massive, are described by the same communication \( \phi \) throughout the domain. In this case the fast alignment which is supposed to happen within each flock gets hijacked by the other flock due to long separation. This situation is better modeled by a system with multi-scaling that incorporates different time scales on which alignment is achieved within and between the flocks in a bigger cluster. We can formulate an even more general multi-scale model where communication between flocks is regulated by a kernel different from internal ones.

To derive such a model, we assume that positions and velocities \( y_{ai} = y_{ai}(t, \tau), u_{ai} = u_{ai}(t, \tau) \), where \( \alpha = 1, \ldots, A \) and \( i = 1, \ldots, N_{\alpha} \), depend on two time parameters, in which \( t \) is a fast time and \( \tau \) is a slow time. We postulate that on the fast time scale the \( \alpha \)-flock does not react on other flocks’ motion, and evolves autonomously:

\[
\begin{align*}
\partial_t y_{ai} &= u_{ai}, \\
\partial_t u_{ai} &= \lambda_{\alpha} \sum_{j=1}^{N_{\alpha}} m_{\alpha j} \phi_{\alpha}(y_{ai}, y_{aj})(u_{aj} - u_{ai}).
\end{align*}
\]

On a slower time scale \( \tau \), however, the agents of the \( \alpha \)-flock are influenced by other flocks via their macroscopic parameters. In other words \( \alpha \)-agents adjust to other flocks’ motion according to their consensus direction:

\[
Y_{\alpha} = \frac{1}{M_{\alpha}} \sum_{i=1}^{N_{\alpha}} m_{\alpha i} y_{ai}, \quad U_{\alpha} = \frac{1}{M_{\alpha}} \sum_{i=1}^{N_{\alpha}} m_{\alpha i} u_{ai}, \quad M_{\alpha} = \sum_{i=1}^{N_{\alpha}} m_{\alpha i}.
\]

Thus,

\[
\begin{align*}
\partial_\tau y_{ai} &= U_{\alpha}, \\
\partial_\tau u_{ai} &= \sum_{\beta \neq \alpha} M_{\beta} \Psi(Y_{\alpha}, Y_{\beta})(U_{\beta} - u_{ai}),
\end{align*}
\]

where \( \Psi \) is an inter-flock communication kernel. Notice that the flock momenta are stationary on fast scale \( \partial_t U_{\alpha} = 0 \). It follows directly from (96). However, on slow scale the macroscopic parameters \((Y_{\alpha}, U_{\alpha})\) satisfy the “up-scaled” Cucker-Smale system:

\[
\begin{align*}
\partial_\tau Y_{\alpha} &= U_{\alpha}, \\
\partial_\tau U_{\alpha} &= \sum_{\beta \neq \alpha} M_{\beta} \Psi(Y_{\alpha}, Y_{\beta})(U_{\beta} - U_{\alpha}).
\end{align*}
\]

The ratio of time scales is given by

\[
\varepsilon := \frac{\tau}{t} << \min_{\alpha} \lambda_{\alpha}.
\]

With \( \varepsilon \) being fixed and small we consider a pair of new variables depending on only one time \( t \):

\[
x_{ai}(t) = y_{ai}(t, \varepsilon t), \quad v_{ai}(t) = u_{ai}(t, \varepsilon t) + \varepsilon U_{\alpha}(\varepsilon t).
\]
Notice that $\dot{x}_{\alpha i} = v_{\alpha i}$, and
\[
\dot{v}_{\alpha i} = \partial_t u_{\alpha i}(t, \varepsilon t) + \varepsilon \partial_t u_{\alpha i}(t, \varepsilon t) + \varepsilon^2 \partial_t U_{\alpha}(\varepsilon t)
\]
\[
= \lambda_{\alpha} \sum_{j=1}^{N_{\alpha}} m_{\alpha j} \phi_{\alpha}(x_{\alpha i}, x_{\alpha j})(u_{\alpha j} - u_{\alpha i}) + \varepsilon \sum_{\beta \neq \alpha} M_{\beta} \Psi(X_{\alpha}, X_{\beta})(U_{\beta} - U_{\alpha})
\]
\[
+ \varepsilon^2 \sum_{\beta \neq \alpha} M_{\beta} \Psi(X_{\alpha}, X_{\beta})(U_{\beta} - U_{\alpha}).
\]
Noting that $u_{\alpha j} - u_{\alpha i} = v_{\alpha j} - v_{\alpha i}$ and $V_{\alpha} = (1 + \varepsilon) U_{\alpha}$ we arrive at the following system:
\[
\begin{align*}
\dot{x}_{\alpha i} &= v_{\alpha i}, \\
\dot{v}_{\alpha i} &= \lambda_{\alpha} \sum_{j=1}^{N_{\alpha}} m_{\alpha j} \phi_{\alpha}(x_{\alpha i} - x_{\alpha j})(v_{\alpha j} - v_{\alpha i}) + \varepsilon \sum_{\beta \neq \alpha} M_{\beta} \Psi(X_{\alpha} - X_{\beta})(V_{\beta} - V_{\alpha}),
\end{align*}
\]
where
\[
X_{\alpha} = \frac{1}{M_{\alpha}} \sum_{i \in \Omega_{\alpha}} m_{\alpha i} x_{\alpha i}, \quad V_{\alpha} = \frac{1}{M_{\alpha}} \sum_{i \in \Omega_{\alpha}} m_{\alpha i} v_{\alpha i}.
\]
It follows from the contraction that the macroscopic variables $X_{\alpha}, V_{\alpha}$ satisfy the upscaled system
\[
\begin{align*}
\dot{X}_{\alpha} &= V_{\alpha}, \\
\dot{V}_{\alpha} &= \varepsilon \sum_{\beta \neq \alpha} M_{\beta} \Psi(X_{\alpha} - X_{\beta})(V_{\beta} - V_{\alpha}).
\end{align*}
\]
Let us consider the following size variables of the system (101) and (102):
\[
D_{\alpha} = \max_{i,j} |x_{\alpha i} - x_{\alpha j}|, \quad D = \max_{\alpha, \beta} |X_{\alpha} - X_{\beta}|
\]
\[
A_{\alpha} = \max_{i,j} |v_{\alpha i} - v_{\alpha j}|, \quad A = \max_{\alpha, \beta} |V_{\alpha} - V_{\beta}|.
\]
The alignment of macroscopic quantities follows from the same system of ODEs as we derived in the classical case:
\[
\begin{align*}
\dot{A} &\leq -\varepsilon M \Psi(D) A \\
\dot{D} &\leq A.
\end{align*}
\]
Thus, under fat tail condition on $\Psi$ the consensus directions $V_{\alpha}$ will in fact align exponentially fast according to Theorem 2.4. To understand the alignment within each flock we consider a maximizing triple $\ell, i, j$ for an $\alpha$-flock and apply computation (26):
\[
\frac{d}{dt} \ell(v_{\alpha i} - v_{\alpha j}) \leq -\lambda_{\alpha} \phi(D_{\alpha}) A_{\alpha} - \varepsilon M \Psi(D) A_{\alpha},
\]
we obtain the system of ODEs:
\[
\begin{align*}
\dot{A}_{\alpha} &\leq -\lambda_{\alpha} M_{\alpha} \phi(D_{\alpha}) A_{\alpha} - \varepsilon M \Psi(D) A_{\alpha} \\
\dot{D}_{\alpha} &\leq A_{\alpha} \\
\dot{A} &\leq -\varepsilon M \Psi(D) A \\
\dot{D} &\leq A
\end{align*}
\]
Ignoring $-\varepsilon M \Psi(D) A_{\alpha}$ in the $A_{\alpha}$ equation for the moment, we see that the $\alpha$-flock decouples from the rest. Consequently, a fast internal alignment insues applying Theorem 2.4.
Theorem 2.15 (Fast internal flocking). If for a given \( \alpha \in \{1, \ldots, A\} \) the kernel \( \phi_\alpha \) has a fat tail, the \( \alpha \)-flock aligns exponentially fast at a rate functionally dependent on \( \phi_\alpha \), initial data, and \( \lambda_\alpha \):

\[
\max_i |v_{\alpha i}(t) - V_\alpha(t)| \lesssim e^{-\delta t}.
\]

It is interesting to note that this alignment process is completely independent from the inter-flock communication. So, long range internal communication leads to local emergence despite potentially destabilizing influence of the outside crowd. On the other hand, if the inter-flock communication \( \Psi \) is global, e.g. satisfies the fat tail condition (7), then the global alignment occurs even if internal communications are weak or even absent. This is clear from (104) if we drop \(-\lambda_\alpha M_\alpha \phi(D_\alpha) V_\alpha\) and conclude boundedness of \( D \) from the last two equations. Alignment rate in this case is global but occurs on the slow time scale.

Theorem 2.16 (Slow global flocking). Assuming that \( \Psi \) has a fat tail and all \( \phi_\alpha \geq 0 \), all solutions to (101) align exponentially fast at a rate functionally dependent on \( \Psi \), initial data, and \( \varepsilon \),

\[
\max_{\alpha,i} |v_{\alpha i}(t) - V| \lesssim e^{-\delta t}.
\]

Asymptotic dependence of the implied alignment rates for small \( \varepsilon \) and large \( \lambda_\alpha \) for the Cucker-Smale kernel can be readily obtained from formulas (29) and (30). Thus, in the context of fast local alignment we obtain \( \delta \sim \lambda_\alpha \) for all \( \beta \leq 1 \), while in the context of slow alignment, \( \delta \sim \varepsilon^{1/\beta} \), for \( \beta < 1 \), and \( \delta \sim \varepsilon^{-1/\varepsilon} \) for \( \beta = 1 \).

It is sometimes convenient to pass to the reference frame evolving with the momentum and center of mass of the flock:

\[
w_{\alpha i} = v_{\alpha i} - V_\alpha, \quad y_{\alpha i} = x_{\alpha i} - X_\alpha.
\]

Using (101) and (102) one readily obtains the system

\[
\begin{cases}
\dot{y}_{\alpha i} = w_{\alpha i}, \\
\dot{w}_{\alpha i} = \lambda_\alpha \sum_{j=1}^{N_\alpha} m_{\alpha j} \phi_{\alpha ij}(w_{\alpha i} - w_{\alpha j}) - \varepsilon R_\alpha(t) w_{\alpha i},
\end{cases}
\]

where

\[
R_\alpha(t) = \sum_{\beta \neq \alpha} M_\beta \Psi(X_\alpha - X_\beta).
\]

We used a shortcut to denote \( \phi_{\alpha ij} = \phi_\alpha(y_{\alpha i} - y_{\alpha j}) \). The following lemma is straightforward.

**Lemma 2.17.** The old set of variables \( (x_{\alpha i}, v_{\alpha i})_{\alpha,i} \) satisfies (101) if and only if the new set \( (y_{\alpha i}, w_{\alpha i})_{\alpha,i} \) satisfies (106) and the macroscopic variables \( (X_\alpha, V_\alpha)_{\alpha} \) satisfy (102).

An immediate consequence of (106) is the maximum principle within each flock relative to its momentum. In particular we obtain a class of solutions, called Mikado flocks, given by

\[
w_{\alpha i}(t) = w_{\alpha i}(t) r_\alpha,
\]

where \( r_\alpha \) are fixed unit vectors. We will address those in more detail in the context of hydrodynamic systems in Section 7.2.

### 3. Kinetic models

In the large crowd systems, where \( N \sim \infty \), it is more efficient to resort to mesoscopic level of description of the Cucker-Smale dynamics. The corresponding kinetic formulation of (11) can be derived formally via the BBGK hierarchy. Looking slightly ahead we seek to derive the following
Vlasov-type model which describes evolution of a mass probability distribution \( f(x, v, t) \) of agents in phase space \((x, v)\):

\[
\partial_t f + v \cdot \nabla_x f + \lambda \nabla_v \cdot [f F(f)] = 0,
\]

where

\[
F(f)(x, v, t) = \int_{\mathbb{R}^{2n}} \phi(x, y)(w - v)f(y, w, t) \, dw \, dy.
\]

3.1. BBGKY hierarchy: formal derivation. Let us consider a probability density \( P^N = P^N(x_1, v_1, ..., x_N, v_N, t) \) of a system of \( N \) agents in the ensemble configuration space 

\[
(x_1, v_1, ..., x_N, v_N) \in \mathbb{R}^{2nN}.
\]

The conservation of mass in the Gibbs ensemble propagated according to the given system (11) leads to the classical Liouville equation:

\[
P_t^N + \sum_{i=1}^N v_i \cdot \nabla_x P^N + \sum_{i=1}^N \nabla v_i \cdot (\partial_t P^N) = 0.
\]

We assume the effective radius of communication between agents remains independent of \( N \), i.e. the kernel \( \phi \) is not rescaled with \( N \). This scaling regime called the mean-field limit. We further assume that the total mass \( M = \sum m_i \) remains constant and \( \max_i m_i \to 0 \). As a result, the agents become more and more indistinguishable, which we reflect in the symmetry condition

\[
P^N(..., x_i, v_i, ..., x_j, v_j, ..., t) = P^N(..., x_j, v_j, ..., x_i, v_i, ..., t).
\]

We seek to derive an equation for the first marginal

\[
P^1_{\lambda,N}(x, v, t) = \int_{\mathbb{R}^{2n(N-1)}} P^N(x, v, \bar{x}, \bar{v}, t) \, d\bar{x} \, d\bar{v},
\]

where \( \bar{x} = (x_2, ..., x_N) \) and \( \bar{v} = (v_2, ..., v_N) \). Thus, integrating in \( \bar{x}, \bar{v} \) in (109) we obtain

\[
P_t^1_{\lambda,N} + v \cdot \nabla_x P^1_{\lambda,N} + \lambda \nabla_v \cdot \int_{\mathbb{R}^{2n(N-1)}} \sum_{j=2}^N m_j \phi(x, x_j)(v_j - v) P^N d\bar{x} d\bar{v} = 0.
\]

In view of the symmetry of \( P^N \), we achieve equality of the integrals in the sum above, and hence,

\[
P_t^1_{\lambda,N} + v \cdot \nabla_x P^1_{\lambda,N} + \lambda (M - m_1) \nabla_v \cdot \int_{\mathbb{R}^{2n}} \phi(x, y)(w - v) P^2_{\lambda,N}(x, v, y, w, t) \, dy \, dw = 0,
\]

where \( P^2_{\lambda,N} \) is the second marginal:

\[
P^2_{\lambda,N}(x, v, y, w, t) = \int_{\mathbb{R}^{2n(N-2)}} P^N(x, v, y, w, z, \bar{u}, t) \, dz \, d\bar{u}.
\]

Denoting the limiting densities by \( P = \lim_{N \to \infty} P^1_{\lambda,N}, Q = \lim_{N \to \infty} P^2_{\lambda,N} \) we obtain

\[
P_t + v \cdot \nabla_x P + \lambda M \nabla_v \cdot \int_{\mathbb{R}^{2n}} \phi(x, y)(w - v) Q(x, v, y, w, t) \, dy \, dw = 0.
\]

We close by making the molecular chaos assumption

\[
Q(x, v, y, w, t) = P(x, v, t) P(y, w, t),
\]

which results in precisely the following Vlasov-type equation (108) for the mass density \( f = MP \).

Formally, the kernel takes macroscopic form of any of the models discussed:

**Motsch-Tadmor:**

\[
\phi(x, y) = \frac{\phi(x - y)}{\phi * \rho(x)};
\]

\[(110)\]
We endow $M$ on continuous bounded function i.e. for every $\varepsilon > 0$, formulation (114) is easily extendable to functions with quadratic growth of the material derivative: to make the right hand side of (114) well defined. From compactly supported test functions, will not detail the statements above here.

Definition 3.1. We say that $\mu_{t} = \sum_{i=1}^{N} m_{i} \delta_{x_{i}(t)} \otimes \delta_{v_{i}(t)}$, which satisfies a weak formulation of (108) if and only if the system $\{x_{i}, v_{i}\}_{i}$ solves the classical Cucker-Smale equations (11). This leads us first to define (108) in a weak sense for a special class of measure-valued solutions. We denote by $M_{+}(\mathbb{R}^{2n})$ the set of non-negative Radon measures on $\mathbb{R}^{2n}$, and

$$M_{+}^p = \{ \mu \in M_{+}(\mathbb{R}^{2n}) : \|\mu\|_{p} = \int_{\mathbb{R}^{2n}} (1 + |v|^p) \, d\mu(x,v) < \infty \}.$$  

We endow $M_{+}$ and well as $M_{+}^p$ with the topology of weak convergence, which means convergence on continuous bounded function $C_{b}(\mathbb{R}^{2n})$. Note that any bounded set in $M_{+}^p$ is automatically tight, i.e. for every $\varepsilon > 0$ there exists a compact set $K_{\varepsilon}$ (a ball in fact) such that $\mu(\mathbb{R}^{2n} \setminus K_{\varepsilon}) < \varepsilon$ for all $\mu$ in the family. So, by Prohorov’s theorem weak convergence on $C_{b}(\mathbb{R}^{2n})$ is equivalent to the classical weak* convergence on $C_{0}(\mathbb{R}^{2n})$, the predual of $M(\mathbb{R}^{2n})$. Therefore all bounded sets in $M_{+}^p$ are weakly precompact. Since we will deal with measures confined to a bounded set anyway, we will not detail the statements above here.

Let us note that for any $\mu \in M_{+}^p$, $p > 1$ the integral

$$F(\mu)(x,v) = \int_{\mathbb{R}^{2n}} \phi(x-y)(w-v) \, d\mu(y,w)$$

defines a $C^{1}$ smooth locally bounded field on $\mathbb{R}^{2n}$. Moreover, for a time-dependent weakly continuous family $\{\mu_{t} \}_{0 \leq t < T} \subset C_{w^{*}}([0,T]; M_{+}^{p}(\mathbb{R}^{2n}))$, $p > 1$, the field $F$ becomes continuous $F(\mu) \in C([0,T] \times \mathbb{R}^{2n})$ and uniformly Lipschitz $F(\mu) \in L^{\infty}([0,T]; \text{Lip}_{\text{loc}}(\mathbb{R}^{2n}))$ with

$$|F(\mu_{t})(x,v)| \leq C(1 + |v|).$$

This is sufficient to define the proper global flow map on $[0,T] \times \mathbb{R}^{2n}$ later.

Definition 3.1. We say that $\{\mu_{t}\}_{0 \leq t < T} \subset C_{w^{*}}([0,T]; M_{+}^{2}(\mathbb{R}^{2n}))$ is a measure-valued solution to (108) with initial condition $\mu_{0}$ if for any test-function $g \in C_{0}^{\infty}([0,T] \times \mathbb{R}^{2n})$ one has for all $0 < t < T$

$$g(t,x,v) \, d\mu_{t}(x,v) = \int_{\mathbb{R}^{2n}} g(0,x,v) \, d\mu_{0}(x,v) +$$

$$+ \int_{0}^{t} \int_{\mathbb{R}^{2n}} (\partial_{s} g + v \cdot \nabla_{x} g + \lambda F(\mu_{s}) \cdot \nabla_{v} g) \, d\mu_{s}(x,v) \, ds.$$  

The choice of $p = 2$ is motivated by our interest in the energy class, however any $p > 1$ is sufficient to make the right hand side of (114) well defined. From compactly supported test functions, formulation (114) is easily extendable to functions with quadratic growth of the material derivative: $|\partial_{s} g| + |v|\nabla_{x} g| + |v|\nabla_{v} g| \lesssim |v|^2$. Immediate consequences are the mass conservation (plugging $g = 1$)

$$\frac{d}{dt} \mu_{t}(\mathbb{R}^{2n}) = 0,$$
conservation of momentum (plugging $g = v_i$),

$$\frac{d}{dt} \int_{\mathbb{R}^{2n}} v \, d\mu_t(x, v) = 0, \quad (116)$$

and the energy law (plugging $g = |v|^2$),

$$\frac{d}{dt} \int_{\mathbb{R}^{2n}} |v|^2 \, d\mu_t(x, v) \leq -\int_{\mathbb{R}^{2n}} \int_{\mathbb{R}^{2n}} \phi(x - y) |w - v|^2 \, d\mu_t(y, w) \, d\mu_t(x, v). \quad (117)$$

One implication of the decaying energy is that any solution will automatically retain uniformly bounded second momentum on its interval of existence:

$$\|\mu_t\|_2 \leq \|\mu_0\|_2. \quad (118)$$

The proof of the following lemma is straightforward.

**Lemma 3.2.** The empirical measure (112) satisfies the weak kinetic formulation (114) if and only if the set of pairs $(x_i, v_i)_i$ is a solution to (11).

To fully exploit the transport structure of the kinetic equation (108) let us consider the characteristic flow of the field $\langle v, \lambda F(\mu_t) \rangle$:

$$\frac{d}{dt} X(t, s, x, v) = V(t, s, x, v), \quad X(s, s, x, v) = x \quad (119)$$

$$\frac{d}{dt} V(t, s, x, v) = \lambda F(\mu_t)(X, V), \quad V(s, s, x, v) = v. \quad (120)$$

We also denote $X(t, 0, x, v) = X(t, x, v), V(t, 0, x, v) = V(t, x, v)$, and sometimes $(x, v) = \omega$. Note that $F(\mu_t)$ is smooth in $\omega$, so the flow is well-defined and smooth. In view of the linear bound (113) we also conclude that the ODE (119)–(120) is wellposed on the entire existence time interval of the solution $\mu$. Using the test-function $g(s, \omega) = h(X(t, s, \omega), V(t, s, \omega))$ in (114), for some $h \in C^\infty_0(\mathbb{R}^{2n})$, we have

$$\partial_s g + v \cdot \nabla_x g + \lambda F(\mu_t) \cdot \nabla_v g = 0.$$ 

So, (114) reads

$$\int_{\mathbb{R}^{2n}} h(\omega) \, d\mu_t(\omega) = \int_{\mathbb{R}^{2n}} h(X(t, \omega), V(t, \omega)) \, d\mu_0(\omega). \quad (121)$$

We say that $\mu_t$ is a push-forward of the measure $\mu_0$ under the flow-map $(X, V)$, $\mu_t = (X, V) \# \mu_0$. This will be the key formula to deduce contractivity of the solution map to (108) and to study the mean-field limit.

Due to finite second momenta of the measures, (121) is extendable to $|h| \lesssim |v|^2$ (note here a rough estimate on the velocity flow $|V| \leq C(t, s)(1 + |v|)$). Hence, we can apply (121) to the $V$-equation (120) to rewrite it completely in Lagrangian coordinates

$$\frac{d}{dt} V(t, \omega) = \int_{\mathbb{R}^{2n}} \phi(X(t, \omega) - X(t, \omega'))(V(t, \omega') - V(t, \omega)) \, d\mu_0(\omega'). \quad (122)$$

In this form it presents a direct analogue to the discrete system (11), and as a result we can carry out all the basic flocking results in a way very similar to the discrete case.

To finish general discussion, we note that the momentum

$$\overline{V} = \frac{1}{M} \int_{\mathbb{R}^{2n}} v \, d\mu_t(x, v).$$

determines the average direction of the flock centered around the mass center

$$\overline{X} = \int_{\mathbb{R}^{2n}} x \, d\mu_t(x, v), \quad \frac{d}{dt} \overline{X} = \overline{V}.$$
Due to Galilean invariance of the equation (108)
\[ \mu_t \rightarrow \tilde{\mu}_t, \quad \int g(t, x, v)\tilde{\mu}_t(x, v) = \int g(t, x + X(t), v + V) \, d\mu_t(x, v) \]
we can always assume that \( V = X = 0 \).

3.3. Kinetic maximum principle and flocking. It is clear from (122) that the velocity characteristics are concentrating towards the mean value conserved in time. To quantify the rate of convergence stumbles upon the need for global communication in a way similar to the discrete case. Let us assume for now that we have a general non-negative kernel \( \phi = \phi(x - y) \) and work out a system of equation for the kinetic flock parameters. We define the amplitude and diameter over a general compact domain \( \Omega \) containing the initial flock, \( \text{Supp} \mu_0 \subset \Omega \). This will serve multiple purposes – to show alignment on an arbitrarily wide domain, and to provide a tool to compare two different solutions. So, let us assume that we are working with a given solution \( \mu \in C_{\omega^*}([0, T]; \mathcal{M}^2_{+}(\mathbb{R}^{2n})) \) with finite initial support. We define
\[
D_{\Omega}(t) = \max_{\omega', \omega'' \in \Omega} |X(t, \omega') - X(t, \omega'')|,
\]
\[
A_{\Omega}(t) = \max_{\omega', \omega'' \in \Omega} |V(t, \omega') - V(t, \omega'')|.
\]
We will perform a computation similar in the spirit to the discrete system (26), but minding the wider range of the flock parameters. So, let us pick a triple \( \ell \in (\mathbb{R}^{2n})^*, |\ell| = 1, \omega', \omega'' \in \Omega \) which maximizes \( A_{\Omega}(t) \),
\[ A_{\Omega}(t) = \ell(V(t, \omega'') - V(t, \omega')). \]
We abbreviate \( V(t, \omega) = V, V(t, \omega') = V', V(t, \omega'') = V'' \), etc. Then
\[
\frac{d}{dt} A_{\Omega}(t) = \lambda \int_{\mathbb{R}^{2n}} \phi(X - X'') \ell(V - V'') \, d\mu_0(\omega) - \int_{\mathbb{R}^{2n}} \phi(X - X') \ell(V - V') \, d\mu_0(\omega)
\]
\[
= \lambda \int_{\mathbb{R}^{2n}} \phi(X - X'') \ell((V - V') - (V'' - V')) \, d\mu_0(\omega)
\]
\[
+ \lambda \int_{\mathbb{R}^{2n}} \phi(X - X') \ell((V'' - V) - (V'' - V')) \, d\mu_0(\omega)
\]
using that \( \text{Supp} \mu_0 \subset \Omega \) and maximality of \( \ell(V'' - V') \),
\[
\leq \lambda \phi(D_{\Omega}) \int_{\mathbb{R}^{2n}} \ell((V - V') - (V'' - V') + (V'' - V) - (V'' - V')) \, d\mu_0(\omega)
\]
\[
= -\lambda M \phi(D_{\Omega}) A_{\Omega}(t).
\]
We also have trivially
\[
\frac{d}{dt} D_{\Omega} \leq A_{\Omega}.
\]
So, we recover the same system of ODEs as in the discrete case (35). For general kernels, it simply implies that \( A_{\Omega} \) is a decreasing quantity, and hence \( D(t) \leq t \). This implies a priori control on the growth of the radius of support of \( \mu_t \).
\[
\text{Supp} \mu_t \subset B_{R_1 + tR_2}(0).
\]
In the case of fat tail kernel we obtain the full analogue of Theorem 2.4 in terms of kinetic parameters.
**Theorem 3.3** (Alignment for kinetic model). Let \(\mu \in C_\star^w(\mathbb{R}^+; \mathcal{M}_+^2(\mathbb{R}^{2n}))\) be a given solution with compact initial support and let \(\Omega\) be a compact domain with \(\text{Supp } \mu_0 \subset \Omega\). Suppose \(\mathcal{D}\) is a solution to

\[
\int_{\mathcal{D}(0)} \phi(r) \, dr = \frac{A_\Omega(0)}{\lambda M}.
\]

Then the solution \(\mu\) flocks exponentially fast according to

\[
\sup_{t \geq 0} \mathcal{D}(t) \leq \mathcal{D}, \quad A_\Omega(t) \leq A_\Omega(0) e^{-t\lambda M \phi(\mathcal{D})}.
\]

Besides flocking behavior estimates (127) imply that in the fat kernel case the supports of \(\mu_t\) remain a priori uniformly bounded

\[
\text{Supp } \mu_t \subset B_{\mathcal{D}}(0), \quad \forall t \geq 0.
\]

Further refinement of flocking behavior can be provided by establishing estimates on the deformation tensor of the flow-map \((X, V)\). Straight from the Lagrangian formulation we obtain the system for all \(t \geq 0, \omega \in \Omega:\)

\[
\partial_t \nabla X(t, \omega) = \nabla V(t, \omega)
\]

\[
\partial_t \nabla V(t, \omega) = \lambda \int_{\mathbb{R}^{2n}} \nabla^T X(t, \omega) \nabla \phi(X(t, \omega) - X(t, \omega')) \otimes (V(t, \omega') - V(t, \omega)) \, d\mu_0(\omega')
\]

\[
- \nabla V(t, \omega') \lambda \int_{\mathbb{R}^{2n}} \phi(X(t, \omega) - X(t, \omega')) \, d\mu_0(\omega').
\]

Thus,

\[
\frac{d}{dt} \| \nabla X \|_{L^\infty(\Omega)} \leq \| \nabla V \|_{L^\infty(\Omega)},
\]

and again with the use of a functional which maximizes \(\ell(\nabla V(t, \omega)), \omega \in \Omega,\)

\[
\frac{d}{dt} \| \nabla V \|_{L^\infty(\Omega)} \leq \lambda |\nabla \phi|_{\infty} \| \nabla X \|_{L^\infty(\Omega)} A_\Omega(t) - \lambda M \phi(\mathcal{D}(t)) \| \nabla V \|_{L^\infty(\Omega)}.
\]

For general kernels, we simply know the bound \(A_\Omega(t) \leq A_\Omega(0)\) and so the above calculation implies an exponential bound

\[
L(t) \leq C_1 e^{C_2 t}.
\]

For fat tail kernels, with the use of (127) we obtain a system of type (36)

\[
\frac{d}{dt} \| \nabla X \|_{L^\infty(\Omega)} \leq \| \nabla V \|_{L^\infty(\Omega)}
\]

\[
\frac{d}{dt} \| \nabla V \|_{L^\infty(\Omega)} \leq \lambda |\nabla \phi|_{\infty} \| \nabla X \|_{L^\infty(\Omega)} A_\Omega(t) e^{-\lambda M \phi(\mathcal{D})} - \lambda M \phi(\mathcal{D}) \| \nabla V \|_{L^\infty(\Omega)}.
\]

The resulting estimate (37) implies, noting that initially \(\nabla(X, V) = \text{Id},\)

\[
a \| \nabla X(t) \|^2_{L^\infty(\Omega)} + e^{bt} \| \nabla V(t) \|^2_{L^\infty(\Omega)} \leq \frac{4\sqrt{a}}{b} (a + 1)
\]

where \(a = \lambda M A_\Omega(0), b = \lambda M \phi(\mathcal{D}).\) In particular, \(\| \nabla V(t) \|_{L^\infty(\Omega)}\) is exponentially decaying.
3.4. Stability. **Kantorovich-Rubinstein metric. Contractivity.** In this section we provide an analogue of stability property of solutions to (108) similar to one obtained in Section 2.4. On the kinetic level of description analytically suitable way to measure closeness of two flocks is via use of Kantorovich-Rubinstein metric (a particular case of Wasserstein distance), which is compatible with weak topology on $\mathcal{M}_+$. For two measures of equal mass $\mu, \nu \in \mathcal{M}_+$, we define

$$d(\mu, \nu) = \sup_{\text{Lip}(g) \leq 1} \left| \int_{\mathbb{R}^{2n}} g(\omega) \, d\mu(\omega) - \int_{\mathbb{R}^{2n}} g(\omega) \, d\nu(\omega) \right|.$$ 

It follows directly from the definition that if $\text{Lip}(g) \leq L$, then

$$\left| \int_{\mathbb{R}^{2n}} g(\omega) \, d\mu(\omega) - \int_{\mathbb{R}^{2n}} g(\omega) \, d\nu(\omega) \right| \leq L d(\mu, \nu).$$

**Lemma 3.4.** For a sequence of measures with $\text{Supp} \, \mu_n \subset B_R(0)$, $d(\mu_n, \mu) \to 0$ if and only if $\mu_n \to \mu$ weakly.

**Proof.** Clearly if $d(\mu_n, \mu) \to 0$ then $\int_{\mathbb{R}^{2n}} g(\omega) \, d\mu_n(\omega) \to \int_{\mathbb{R}^{2n}} g(\omega) \, d\mu(\omega)$ for every Lipschitz function $g$. However, Lipschitz functions are dense in $C(B_2R(0))$, hence $\mu_n \to \mu$ weakly. On the other hand, if $\mu_n \to \mu$ weakly, then $\mu_n \to \mu$ uniformly on any precompact subset of $C(B_2R(0))$, which, in particular, is the set of all $g$ with $\text{Lip}(g) \leq 1$. \hfill \Box

So, let us consider two solutions on a common interval of existence $\mu, \nu \in C_w^\ast([0,T]; \mathcal{M}_+)$ with compact initial supports which we confine into a fixed convex compact domain $\Omega$

$$\text{Supp} \, \mu_0 \cup \text{Supp} \, \nu_0 \subset \Omega.$$ 

We also assume that the solutions have equal masses $M_\mu = M_\nu$ and momenta $\nabla \mu = \nabla \nu$. Clearly,

$$\frac{d}{dt} \|X_\mu(t) - X_\nu(t)\|_{L^\infty(\Omega)} \leq \|V_\mu(t) - V_\nu(t)\|_{L^\infty(\Omega)}.$$ 

For the velocities we apply the same strategy as usual by fixing a maximizing functional $\ell$ and computing

$$\frac{d}{dt} \|V_\mu(t) - V_\nu(t)\|_{L^\infty(\Omega)} \leq \lambda \int_{\mathbb{R}^{2n}} \phi(X_\mu - X'_\mu) \ell(V'_\mu - V_\mu) \, d\mu_0(\omega')$$

$$- \lambda \int_{\mathbb{R}^{2n}} \phi(X_\nu - X'_\nu) \ell(V'_\nu - V_\nu) \, d\nu_0(\omega')$$

$$= \lambda \int_{\mathbb{R}^{2n}} \phi(X_\mu - X'_\mu) \ell(V'_\mu - V_\mu) \, d\mu_0(\omega') - \lambda \int_{\mathbb{R}^{2n}} \phi(X_\nu - X'_\nu) \ell(V'_\nu - V_\nu) \, d\nu_0(\omega')$$

$$+ \lambda \int_{\mathbb{R}^{2n}} (\phi(X_\mu - X'_\mu) - \phi(X_\nu - X'_\nu)) \ell(V'_\mu - V_\mu - V_\nu) \, d\nu_0(\omega')$$

$$+ \lambda \int_{\mathbb{R}^{2n}} \phi(X_\nu - X'_\nu) \ell((V'_\mu - V_\nu) - (V_\mu - V_\nu)) \, d\nu_0(\omega').$$

There are three terms to estimate on the right hand side. For the first we use the KR-distance. Since $\Omega$ is a convex domain, it is bounded by

$$\lambda \|\phi\|_{W^{1,\infty}} \left( \|\nabla X_\mu\|_{L^\infty(\Omega)} A_{\mu,\Omega}(t) + \|\nabla V_\mu\|_{L^\infty(\Omega)} \right) d(\mu_0, \nu_0).$$

The second term is bounded by

$$2\lambda \|\nabla \phi\|_{\infty,M} \|X_\mu(t) - X_\nu(t)\|_{L^\infty(\Omega)} A_{\mu,\Omega}(t).$$
For the last term we use maximality of $\ell(V_\mu - V_\nu)$ and pull out the kernel first
\[
\lambda \int_{\mathbb{R}^2^n} \phi(X_\nu - X'_\nu) \ell((V'_\nu - V'_\nu) - (V_\mu - V_\nu)) \, d\nu_0(\omega') \leq \lambda \phi(D_\nu,\Omega) \int_{\mathbb{R}^2^n} \ell((V'_\nu - V'_\nu) - (V_\mu - V_\nu)) \, d\nu_0(\omega')
\]
\[
= \lambda \phi(D_\nu,\Omega) \ell \left( \int_{\mathbb{R}^2^n} (V'_\nu - V'_\nu) \, d\nu_0(\omega') \right) - \lambda \phi(D_\nu,\Omega) M \|V_\mu(t) - V_\nu(t)\|_{L^\infty(\Omega)}
\]
\[
= \lambda \phi(D_\nu,\Omega) \ell \left( \int_{\mathbb{R}^2^n} V'_\mu \, d\nu_0(\omega') - \nabla_\nu \right) - \lambda \phi(D_\nu,\Omega) M \|V_\mu(t) - V_\nu(t)\|_{L^\infty(\Omega)}
\]
\[
= \lambda \phi(D_\nu,\Omega) \ell \left( \int_{\mathbb{R}^2^n} V'_\mu \, d\nu_0(\omega') - \int_{\mathbb{R}^2^n} V'_\nu \, d\mu_0(\omega') \right) - \lambda \phi(D_\nu,\Omega) M \|V_\mu(t) - V_\nu(t)\|_{L^\infty(\Omega)}
\]
where in the last step we used equality of momenta. Continuing we obtain
\[
\leq \lambda |\phi|_\infty \|\nabla V_\mu\|_{L^\infty(\Omega)} d(\mu_0, \nu_0) - \lambda \phi(D_\nu,\Omega) M \|V_\mu(t) - V_\nu(t)\|_{L^\infty(\Omega)}.
\]
Putting all the estimates together we obtain the system
\[
\frac{d}{dt} \|X_\mu(t) - X_\nu(t)\|_{L^\infty(\Omega)} \leq \|V_\mu(t) - V_\nu(t)\|_{L^\infty(\Omega)}
\]
\[
\frac{d}{dt} \|V_\mu(t) - V_\nu(t)\|_{L^\infty(\Omega)} \leq \lambda |\phi|_{W^{1,\infty}} (\|\nabla X_\mu\|_{L^\infty(\Omega)} + \|\nabla V_\mu\|_{L^\infty(\Omega)}) d(\mu_0, \nu_0)
\]
\[
+ 2\lambda |\nabla \phi|_\infty M \|X_\mu(t) - X_\nu(t)\|_{L^\infty(\Omega)} A_{\nu,\Omega}(t)
\]
\[
- \lambda \phi(D_\nu,\Omega) M \|V_\mu(t) - V_\nu(t)\|_{L^\infty(\Omega)}.
\]

Two conclusions follow as before from the system above. For general kernels $\phi$ we can use the bound (128) to estimate
\[
\frac{d}{dt} \|X_\mu(t) - X_\nu(t)\|_{L^\infty(\Omega)} + \|V_\mu(t) - V_\nu(t)\|_{L^\infty(\Omega)} \leq C_1 e^{C_2 t} d(\mu_0, \nu_0) + C_3 \|X_\mu(t) - X_\nu(t)\|_{L^\infty(\Omega)}.
\]
Since initially $\|X_\mu(0) - X_\nu(0)\|_{L^\infty(\Omega)} + \|V_\mu(0) - V_\nu(0)\|_{L^\infty(\Omega)} = 0$, we obtain
\[
\|X_\mu(t) - X_\nu(t)\|_{L^\infty(\Omega)} + \|V_\mu(t) - V_\nu(t)\|_{L^\infty(\Omega)} \leq C e^{C t} d(\mu_0, \nu_0),
\]
for some $C > 0$ which depends only on initial condition and the kernel.

For fat tail kernels, we use more robust estimate on the deformation tensor (129) and on the diameter and amplitude (127) to conclude
\[
\frac{d}{dt} \|V_\mu(t) - V_\nu(t)\|_{L^\infty(\Omega)} \leq ae^{-bt} [d(\mu_0, \nu_0) + \|X_\mu(t) - X_\nu(t)\|_{L^\infty(\Omega)}] - b \|V_\mu(t) - V_\nu(t)\|_{L^\infty(\Omega)}
\]
So, we obtain the same system (36) but for the new pair
\[
\frac{d}{dt} \|V_\mu(t) - V_\nu(t)\|_{L^\infty(\Omega)} \leq ae^{-bt} [d(\mu_0, \nu_0)] - b \|V_\mu(t) - V_\nu(t)\|_{L^\infty(\Omega)}.
\]
Noting that $x(0) = d(\mu_0, \nu_0)$, and $v(0) = 0$, we obtain from (37)
\[
\|X_\mu(t) - X_\nu(t)\|_{L^\infty(\Omega)} \leq C d(\mu_0, \nu_0), \quad \|V_\mu(t) - V_\nu(t)\|_{L^\infty(\Omega)} \leq C e^{-ct} d(\mu_0, \nu_0),
\]
for all time $t > 0$, where $C, c > 0$ depend only on the initial kinetic diameter of the flocks, mass, and the kernel.

Although (130) and (131) by themselves express characteristics stability of the flock, the ultimate application lies in estimating the KR-distance $d(\mu_1, \nu_1)$ and establishing contractivity of the kinetic dynamics. So, let us assume that on a given time interval $[0, T)$ we have two solutions $\mu, \nu \in C_w([0, T); M_+)$ with bounded supports $\text{Supp} \mu_0 \cup \text{Supp} \nu_0 \in \Omega$, where $\Omega$ is convex and compact,
for example we can take convex hull of $\text{Supp}{\mu}_0 \cup \text{Supp}{\nu}_0$. Let us fix a function $h$ with $\text{Lip}(h) \leq 1$, and use conservation law (121)

$$
\int_{\mathbb{R}^2^N} h(\omega) \, d\mu - \int_{\mathbb{R}^2^N} h(\omega) \, d\nu = \int_{\mathbb{R}^2^N} h(X_\mu, V_\mu) \, d\mu_0 - \int_{\mathbb{R}^2^N} h(X_\nu, V_\nu) \, d\nu_0
$$

$$
= \int_{\mathbb{R}^2^N} h(X_\mu, V_\mu) \,(d\mu_0 - d\nu_0) + \int_{\mathbb{R}^2^N} [h(X_\mu, V_\mu) - h(X_\nu, V_\nu)] \, d\nu_0
$$

$$
\leq \text{Lip}_\Omega(h(X_\mu, V_\mu)) d(\mu_0, \nu_0) + M(\|X_\mu - X_\nu\|_{L^\infty(\Omega)} + \|V_\mu - V_\nu\|_{L^\infty(\Omega)}),
$$

Using that

$$
\text{Lip}_\Omega(h(X_\mu, V_\mu)) \leq \|\nabla V_\mu\|_{L^\infty(\Omega)} + \|\nabla X_\mu\|_{L^\infty(\Omega)},
$$

and applying the stability and deformation estimates (128), (129), (130), (131) we conclude the following bounds

(132) \qquad d(\mu_t, \nu_t) \leq C e^{Ct} d(\mu_0, \nu_0), \quad \text{general kernels},

(133) \qquad d(\mu_t, \nu_t) \leq C \text{d} \delta(\mu_0, \nu_0), \quad \text{fat tail kernels}.

3.5. **Mean-field limit.** The mean-field limit refers to passage from discrete to kinetic system as $N \to \infty$ in the scaling regime the range of the interactions remains independent of $N$, i.e. $\phi(x, y)$ remains unrescaled.

The analysis of the previous section makes passing to the limit $N \to 0$ almost trivial. Let us start with a given measure $\mu_0 \in M_+$ with compact support. We discretize it in the classical atomic approximation. We consider a box $Q$ of side length $L$ containing $\text{Supp}{\mu}_0$, and decompose it into $N^n$ subboxes of side length $L/N$, denote them $Q_k$, $k = 1, \ldots, N^n$. Next we find a $\omega_k = (v_k, x_k)$ such that

$$
v_k = \frac{1}{\mu_0(Q_k)} \int_{Q_k} v \, d\mu_0(x, v).
$$

We dismiss those $Q_k$ which have no mass. Finally, we define

$$
\mu_0^N = \sum_{k=1}^{N^n} \mu_0(Q_k) \delta_{\omega_k}.
$$

Clearly, all $\mu_0^N$’s have the same momentum and mass. Moreover, $\mu_0^N \to \mu_0$ weakly, and all supports $\text{Supp}{\mu}_0^N$ are confined to the same box $Q$. Let us define the empirical measure

$$
\mu_t^N = \sum_{k=1}^{N^n} \mu_0(Q_k) \delta_{\omega_k(t)},
$$

where $\omega_k(t) = (x_k(t), v_k(t))$ solves the Cucker-Smale system (11) with masses $m_k = \mu_0(Q_k)$. These measures will have uniformly bounded supports on any given time interval $[0, T]$, and in fact on all $\mathbb{R}^+$ if $\phi$ has fat tail. Hence, estimates (132), (133) apply to give us

$$
d(\mu_t^N, \mu_t^M) \leq C(T) d(\mu_0^N, \mu_0^M), \quad \forall t < T,
$$

and

$$
d(\mu_t^N, \mu_t^M) \leq C d(\mu_0^N, \mu_0^M), \quad \forall t > 0,
$$

respectively. This means at any time $t$, $\mu_t^N$ is a Cauchy sequence in $M_+$, hence there exists a weak limit $\mu_t^N \to \mu_t$, uniform on any $[0, T]$, and $\mathbb{R}^+$, respectively. The following lemma concludes the passage.
Lemma 3.5 (Stability under weak limits). Suppose a sequence of solutions \( \mu^n \in C_w^\infty([0,T]; M_+) \) with \( \text{Supp} \mu^n_t \subseteq B_R(0) \), for all \( t < T \) and \( n \in \mathbb{N} \) converges weakly pointwise, i.e. \( \mu^n_t \to \mu_t \) for all \( 0 \leq t < T \). Then \( \mu \in C_w^\infty([0,T]; M_+) \) is a weak solution to (108).

Proof. Weak continuity will follow immediately from (114) once it is established. It is clear that all the linear terms in (114) converge to the natural limits. As to \( F(\mu^n_t) \), note that the family of functions \( \{ \phi(x-\cdot)(\cdot-v) \}_{x,v} \in B_R(0) \) is uniformly Lipschitz on \( B_R(0) \), hence precompact in \( C(B_R(0)) \). So, \( F(\mu^n_t)(x,v) \to F(\mu_t)(x,v) \) converges uniformly on \( B_R(0) \). This implies

\[
\int_0^t \int_{\mathbb{R}^2n} F(\mu^n_s)(x,v) \cdot \nabla_v g(x,v) \, d\mu^n_s(x,v) \, ds \to \int_0^t \int_{\mathbb{R}^2n} F(\mu_s)(x,v) \cdot \nabla_v g(x,v) \, d\mu_s(x,v) \, ds.
\]

Indeed, adding and subtracting cross-terms we obtain trivially

\[
\int_0^t \int_{\mathbb{R}^2n} F(\mu^n_s)(x,v) \cdot \nabla_v g(x,v) \, d\mu^n_s(x,v) \, ds \to 0,
\]

and

\[
\left| \int_0^t \int_{\mathbb{R}^2n} [F(\mu^n_s)(x,v) - F(\mu_s)(x,v)] \cdot \nabla_v g(x,v) \, d\mu^n_s(x,v) \, ds \right|
\leq |\nabla g|_\infty \int_0^t \int_{\mathbb{R}^2n} \| F(\mu^n_s) - F(\mu_s) \|_{L^\infty(B_R(0))} \, d\mu^n_s(x,v) \, ds
\]

\[
= CM_0^n \int_0^t \int_{\mathbb{R}^2n} \| F(\mu^n_s) - F(\mu_s) \|_{L^\infty(B_R(0))} \, ds \to 0.
\]

If the initial measure \( \mu_0 \) is in fact absolutely continuous \( d\mu_0(\omega) = f_0(\omega) \, d\omega \), then the resulting solution will give density \( f(t, \omega) \) obtained by the transport along characteristic flow according to (121). Explicitly we have the Liouville formula for the Jacobian of the flow:

\[
\det \nabla_\omega (X,V)(\omega, t) = \exp \left\{-\lambda n \int_0^t \phi * \rho(X(\omega, s), s) \, ds \right\},
\]

where \( \rho \) is the macroscopic density measure \( d\rho(x,t) = \int_{\mathbb{R}^n} \mu_t(x,v) \, dv \), i.e. first marginal of \( \mu_t \). Note that \( \phi * \rho \in C^\infty \). Then

\[
f(X(\omega,t), V(\omega,t), t) = f_0(\omega) \exp \left\{ \lambda n \int_0^t \phi * \rho(X(\omega, s), s) \, ds \right\}.
\]

Inverting the flow we recover \( f(t) \). It is clear that \( f(t) \) inherits smoothness of the initial condition as well, given the smoothness of the flow map. At the \( C^1 \) level it is seen from (128). Further regularity can be deduced by proving higher regularity of the flow map.

3.6. Macroscopic description. Hydrodynamic limit. By taking \( v \)-moments of (108) we can read off the system for macroscopic density and momentum

\[
\rho(x,t) = \int_{\mathbb{R}^n} f(x,v,t) \, dv, \quad \rho u = \int_{\mathbb{R}^n} vf(x,v,t) \, dv,
\]

\[
\rho_t + \nabla \cdot (\rho u) = 0,
\]

\[
(\rho u)_t + \nabla_x \cdot (\rho u \otimes u + \mathcal{R}) = \int_{\mathbb{R}^n} \rho(x)\rho(y)(u(y) - u(x))\phi(x,y) \, dy,
\]

where \( \mathcal{R} \) is the Reynolds stress tensor

\[
\mathcal{R}(x,t) = \int_{\mathbb{R}^n} (v - u(x,t)) \otimes (v - u(x,t)) f(x,v,t) \, dv.
\]
One can formally close the system by considering a monokinetic density ansatz concentrated at the macroscopic velocity \( \mathbf{u} \):

\[
\begin{align*}
  f(x,v,t) &= \rho(x,t)\delta(v - \mathbf{u}(x,t)).
\end{align*}
\]

Clearly, such an ansatz creates zero stress \( R = 0 \), and hence we obtain the system, which we write for the \((\rho, \mathbf{u})\)-pair

\[
\begin{align*}
  \rho_t + \nabla \cdot (\rho \mathbf{u}) &= 0, \\
  \mathbf{u}_t + \mathbf{u} \cdot \nabla \mathbf{u} &= \int_{\mathbb{R}^n} \rho(y)(\mathbf{u}(y) - \mathbf{u}(x))\phi(x, y) \, dy.
\end{align*}
\]

4. Euler Alignment System

\[
\begin{align*}
  \rho_t + \nabla \cdot (\rho \mathbf{u}) &= 0, \\
  \mathbf{u}_t + \mathbf{u} \cdot \nabla \mathbf{u} &= \int_{\Omega^n} \phi(x, y)(\mathbf{u}(y) - \mathbf{u}(x))\rho(y) \, dy \quad (x, t) \in \Omega^n \times \mathbb{R}_+.
\end{align*}
\]

It is sometimes advantageous to write the velocity equation in the conservative form

\[
\begin{align*}
  (\rho \mathbf{u})_t + \nabla \cdot (\rho \mathbf{u} \otimes \mathbf{u}) &= \int_{\Omega^n} \phi(x, y)(\mathbf{u}(y) - \mathbf{u}(x))\rho(y)\rho(x) \, dy.
\end{align*}
\]

Consequently, just as in the discrete case, the system (138) conserves mass and momentum:

\[
\begin{align*}
  \frac{d}{dt} \int_{\Omega^n} \rho \, dx, \quad \int_{\Omega^n} \rho(x, t) \, dx.
\end{align*}
\]

This allows to predict the limiting alignment velocity to be \( \mathbf{u} = \frac{1}{\rho} \int_{\Omega^n} \rho \, dx \). For convolution kernels, \( \phi(x, y) = \phi(x - y) \), the system is also Galilean invariant. If this is the case we can always assume that \( \mathbf{u} = 0 \).

The crucial feature of the alignment term in (138) is its commutator representation, which is given by

\[
\begin{align*}
  \mathcal{L}_\phi(\mathbf{u}, \rho) = \mathcal{L}_\phi(\rho \mathbf{u}) - \mathcal{L}_\phi(\rho) \mathbf{u},
\end{align*}
\]

where \( \mathcal{L} \) can have two different forms

\[
\begin{align*}
  \mathcal{L}_\phi(f)(x) &= \int_{\Omega^n} \phi(x, y)f(y) \, dy, \\
  \mathcal{L}_\phi(f)(x) &= \int_{\Omega^n} \phi(x, y)(f(y) - f(x)) \, dy.
\end{align*}
\]

more suitable for smooth kernels, and

\[
\begin{align*}
  \mathcal{L}_\phi(f)(x) &= \int_{\Omega^n} \phi(x, y)(f(y) - f(x)) \, dy,
\end{align*}
\]

suitable for singular kernels. Note that for smooth kernels of convolution type we have \( \mathcal{L}_\phi f = \phi \ast f \).

The kinetic energy is given by

\[
\begin{align*}
  \mathcal{E} = \frac{1}{2} \int_{\Omega^n} \rho |\mathbf{u}|^2 \, dx.
\end{align*}
\]

If \( \mathbf{u} = 0 \), the energy becomes a measure of alignment due to its relation to the \( L^2 \)-variation given by

\[
\begin{align*}
  \mathcal{E} = \frac{1}{4M} \int_{\Omega^n} |\mathbf{u}(x) - \mathbf{u}(y)|^2 \rho(x)\rho(y) \, dy \, dx.
\end{align*}
\]

However, as in the discrete case we mostly work with the variation functions to make the computations not dependent on Galilean invariance of the models. So, let us introduce the variations:

\[
\begin{align*}
  \mathcal{V}_p(t) = \int_{\Omega^n} \rho(x, t) \, dx.
\end{align*}
\]
The following energy equation is obtained directly by testing the momentum equation:

\[ \frac{d}{dt} \mathcal{V}_2 = -M \mathcal{I}_2, \quad \mathcal{I}_2 = \int_{\Omega^n \times \Omega^n} \phi(x, y) |u(y) - u(x)|^2 \rho(y) \rho(x) \, dy. \]

Another fundamental property of the system is the maximum principle for each scalar velocity component \( \ell(u), \ell \in (\mathbb{R}^n)^* \), provided the maxima are achieved, which is course true on periodic domain or if \( \bar{u} = 0 \) and the solution decays at infinity on \( \mathbb{R}^n \).

Since all the macroscopic quantities appear to be measured per mass it is natural to introduce the density measure

\[ d m_t = \rho(x, t) \, dx. \]

In view of the transport nature of the continuity equation, this measure is transported along the flow of \( u \). Namely, if

\[ \frac{d}{dt} X(\alpha, t, t_0) = u(X(\alpha, t, t_0), t), \quad t > t_0 \]
\[ X(\alpha, t_0, t_0) = \alpha, \]

then \( dm_t \) is a push-forward of \( dm_{t_0} \) under \( X(\cdot, t, t_0) \):

\[ dm_t = X(\cdot, t, t_0) \# dm_{t_0}. \]

In other words, for any \( g \),

\[ \int_{\mathbb{R}^n} g(X(\alpha, t, t_0)) \, dm_{t_0}(\alpha) = \int_{\mathbb{R}^n} g(x) \, dm_t(x). \]

We also denote \( X(\cdot, t, 0) = X(\cdot, t) \). In particular, the density support is transported by the flow:

\[ \text{Supp } \rho(t) = X(\text{Supp } \rho_0, t). \]

Denoting the Lagrangian velocity by

\[ v(\alpha, t) = u(X(\alpha, t), t) \]

for short, and denoting

\[ v_{\alpha\beta} = v(\alpha, t) - v(\beta, t), \quad \phi_{\alpha\beta} = \phi(X(\alpha, t) - X(\beta, t)), \]

we can rewrite the velocity equation as

\[ \frac{d}{dt} v(\alpha, t) = \int_{\Omega^n} \phi_{\alpha\beta} [v(\beta, t) - v(\alpha, t)] \, dm_0(\beta). \]

The variation becomes

\[ \mathcal{V}_p(t) = \frac{1}{p} \int_{\Omega^n \times \Omega^n} |v_{\alpha\beta}|^p \, dm_0(\alpha, \beta), \]

where \( dm_0(\alpha, \beta) = dm_0(\alpha) \times dm_0(\beta) \). Thus, in Lagrangian coordinates some computations become very similar to the discrete case. For example, it is straightforward to see that all \( \mathcal{V}_p \)'s are decaying in time and (46) translates into just \( \frac{d}{dt} \mathcal{V}_p \leq 0. \)

4.1. **Hydrodynamic flocking and stability.** We assume here that the domain is \( \mathbb{R}^n \) as the discussion carries over to periodic settings directly, and \( \phi \) is of convolution type unless noted otherwise.

Let us consider a flock of finite diameter and a compact domain \( \Omega \) containing \( \text{Supp } \rho_0 \). We define the flock parameters as follows

\[ D_\Omega(t) = \max_{\alpha, \beta \in \Omega} |X(\alpha, t) - X(\beta, t)|, \quad A_\Omega(t) = \max_{\alpha, \beta \in \Omega} |v(\alpha, t) - v(\beta, t)|. \]

Just as in kinetic case it is important to consider this more general setup in order to have a capability to compare flocks and study stability. Since the domain \( \Omega \) is fixed for all time, one can apply the
Rademacher Lemma 2.5 and carry out the same argument as discrete and kinetic cases. Thus, for
general non-negative kernels we have the maximum principle
\begin{equation}
\frac{d}{dt} A_\Omega(t) \leq 0.
\end{equation}
For fat kernels we have the following result whose proof will be omitted.

**Theorem 4.1.** The system (138) aligns and flocks exponentially fast provided \( \phi \) is a non-increasing, everywhere positive, and satisfying the large tail condition (126):
\begin{equation}
\sup_{t \geq 0} D_\Omega(t) = \overline{D}, \quad A_\Omega(t) \leq A_\Omega(0)e^{-\lambda M \phi(\overline{D}) t}.
\end{equation}

The corresponding Motsch-Tadmor and topological versions (31), (32) carry over to macroscopic context ad verbatim. For systems with degenerate communication the corrector method yields the same statement as in Theorem 2.10(ii), which was proved independent of the number of agents. Here one makes use of macroscopic quantities throughout:
\[ d_{\alpha\beta} = -x_{\alpha\beta} \cdot \frac{v_{\alpha\beta}}{|v_{\alpha\beta}|}, \quad G_3 = \int_{\mathbb{R}^n} |v_{\alpha\beta}|^3 \psi(d_{\alpha\beta}) \chi(|x_{\alpha\beta}|) \, dm_0(\alpha, \beta), \quad \text{etc.} \]

The Lagrangian formulation of the Euler Alignment system given by
\begin{align}
\frac{d}{dt} X(\alpha, t) &= v(\alpha, t) \\
\frac{d}{dt} v(\alpha, t) &= \lambda \int_{\mathbb{R}^n} \phi(X(\alpha, t) - X(\beta, t)) [v(\beta, t) - v(\alpha, t)] \rho_0(\beta) \, d\beta
\end{align}
has almost identical structure to its kinetic counterpart (122). This allows to obtain similar stability and regularity estimates for smooth type communication kernels. We start with estimates on deformation and it will be useful to reproduce them pointwise. So, we have
\[ \partial_t \nabla X(\alpha, t) = \nabla v(\alpha, t) \]
\[ \partial_t \nabla v(\alpha, t) = \lambda \int_{\mathbb{R}^n} \nabla^T X(\alpha, t) \nabla \phi(X(\alpha, t) - X(\beta, t)) \otimes (v(\beta, t) - v(\alpha, t)) \rho_0(\beta) \, d\beta \]
\[ - \lambda \nabla v(\alpha, t) \int_{\mathbb{R}^n} \phi(X(\alpha, t) - X(\beta, t)) \rho_0(\beta) \, d\beta. \]

In the case of local kernels, we do not hope for good long time control. So, we simply note that the amplitude \( A_{\text{Supp} \rho_0 \cup \{\alpha\}}(t) \) is bounded by initial condition due to (149), and the kernel is bounded. This implies
\begin{equation}
\|\nabla v(t)\|_{L^\infty(\mathbb{R}^n)} + \|\nabla X(t)\|_{L^\infty(\mathbb{R}^n)} \leq C_1 e^{C_2 t},
\end{equation}
For fat tail kernels, we can deduce better long time estimate on every fixed compact domain \( \Omega \) containing \( \text{Supp} \rho_0 \). So, let us fix such \( \Omega \) and deduce as in kinetic case
\[ \frac{d}{dt} \|\nabla X\|_{L^\infty(\Omega)} \leq \|\nabla v\|_{L^\infty(\Omega)} \]
\[ \frac{d}{dt} \|\nabla v\|_{L^\infty(\Omega)} \leq \lambda \|\phi\|_{\infty} \|\nabla X\|_{L^\infty(\Omega)} A_\Omega(0) e^{-t \lambda M \phi(\overline{D}_\Omega)} - \lambda M \phi(\overline{D}_\Omega) \|\nabla v\|_{L^\infty(\Omega)}. \]
The resulting estimate (37) implies, noting that initially \( \nabla X = \text{Id}, \nabla v = \nabla u_0 \),
\begin{equation}
\alpha \|\nabla X(t)\|_{L^\infty(\Omega)}^2 + \beta e^{t} \|\nabla v(t)\|_{L^2(\Omega)}^2 \leq \frac{4 \sqrt{\alpha}}{b} (\alpha + \|\nabla u_0\|_{L^2(\Omega)}^2) \]
where \( \alpha = \lambda M A_\Omega(0), \beta = \lambda M \phi(\overline{D}_\Omega) \). In particular, we can see that \( \|\nabla v(t)\|_{L^\infty(\Omega)}^2 \) decays exponentially fast.
Suppose now we have two flocks with initial densities \( \rho'_0, \rho''_0 \). It turns out our kinetic result can be carried out almost ad verbatim in the hydrodynamic setting with the use of Kantorovich-Rubinstein distance on \( \mathcal{M}_+ (\mathbb{R}^n) \). So, if \( \Omega \) is a convex compact domain with

\[ \text{Supp} \rho'_0 \cup \text{Supp} \rho''_0 \subset \Omega, \]

and the masses and momenta are equal, \( M' = M'', \varphi' = \varphi'' \), then we obtain

\[
\begin{align*}
\frac{d}{dt} \| X'(t) - X''(t) \|_{L^\infty(\Omega)} & \leq \| \varphi'(t) - \varphi''(t) \|_{L^\infty(\Omega)} \\
\frac{d}{dt} \| \varphi'(t) - \varphi''(t) \|_{L^\infty(\Omega)} & \leq \lambda \| \varphi \|_{W^{1,1}} (\| \nabla X'(t) \|_{L^\infty(\Omega)} A'_1(t) + \| \nabla \varphi' \|_{L^\infty(\Omega)})d(\rho'_0, \rho''_0) \\
& \quad + 2\lambda \| \nabla \varphi \|_{L^\infty(\Omega)} (\| X'(t) - X''(t) \|_{L^\infty(\Omega)} A'_1(t) \\
& \quad - \lambda \| \varphi \|_{L^\infty(\Omega)}^2 (\| \varphi'(t) - \varphi''(t) \|_{L^\infty(\Omega)}).
\end{align*}
\]

Subsequently, with the use of (153) and (154), and recalling that characteristics start from the same values and \( \| \varphi'(0) - \varphi''(0) \|_{L^\infty(\Omega)} = \| u'_0 - u''_0 \|_{L^\infty(\Omega)} \), we obtain for general kernels

\[
\| X'(t) - X''(t) \|_{L^\infty(\Omega)} + \| \varphi'(t) - \varphi''(t) \|_{L^\infty(\Omega)} \leq C e^{C t}[d(\rho'_0, \rho''_0) + \| u'_0 - u''_0 \|_{L^\infty(\Omega)}]
\]

while for kernels with fat tail,

\[
\begin{align*}
\| X'(t) - X''(t) \|_{L^\infty(\Omega)} & \leq C[d(\rho'_0, \rho''_0) + \| u'_0 - u''_0 \|_{L^\infty(\Omega)}], \\
\| \varphi'(t) - \varphi''(t) \|_{L^\infty(\Omega)} & \leq C e^{-C t}[d(\rho'_0, \rho''_0) + \| u'_0 - u''_0 \|_{L^\infty(\Omega)}].
\end{align*}
\]

Continuing with the same KR-distance computation leading to (132) - (133) we obtain

\[
\begin{align*}
\| X'(t) - X''(t) \|_{L^\infty(\Omega)} & \leq C[d(\rho'_0, \rho''_0) + \| u'_0 - u''_0 \|_{L^\infty(\Omega)}], \\
\| \varphi'(t) - \varphi''(t) \|_{L^\infty(\Omega)} & \leq C d(\rho'_0, \rho''_0) + \| u'_0 - u''_0 \|_{L^\infty(\Omega)},
\end{align*}
\]

fat tail kernels.

Remark 4.2. On the real line \( \mathbb{R} \) the KR-distance between two densities \( \rho', \rho'' \) is equal to \( L^1 \)-norm between the corresponding cumulative distribution functions \( F'(x) = \int_{-\infty}^x \rho'(y) \, dy, F''(x) = \int_{-\infty}^x \rho''(y) \, dy \):

\[ d(\rho', \rho'') = \| F' - F'' \|_{L^1}. \]

See [20] for details.

4.2. Spectral method. Hydrodynamic connectivity. When communication is strictly local, the role of connectivity becomes more prominent and the analysis of flocking behavior is more amenable to the periodic settings. Indeed, like in the discrete case one can easily arrange to flocks heading in opposite directions which will never align. In hydrodynamic context connectivity is encoded in no-vacuum condition: \( \rho > 0 \). The lower bound on the density quantifies connectivity in a way similar to the weighted Fiedler number. In this section we make this quantification more precise and apply the spectral method to develop a conditional alignment criterion for singular local kernels of rather general nature: \( \phi(x, y) = \phi(y, x) \) and satisfying

\[
\lambda \frac{1_{|x-y|<r_0}}{|x-y|^{n+\alpha}} \leq \phi(x, y, t) \leq \frac{\Lambda(t)}{|x-y|^{n+\alpha}}, \quad 0 < \alpha < 2,
\]

here \( \Lambda(t) \) is simply assumed finite for any \( t \geq 0 \). Since the compactness of embedding \( H^s \hookrightarrow L^2 \) is crucial to this discussion, as well as a uniform lower bound on the density, which is consistent with finite ness of mass only on a compact environment \( \Omega^n \), we restrict ourselves to the periodic domain \( \Omega^n = \mathbb{T}^n \).

A quantitative expression of coercivity of the commutator (140) relies on lower bounds on the density. Precisely how large that lower should be in order to ensure alignment is investigated in the following proposition.
Proposition 4.3. Let $\phi$ be a symmetric, local, singular kernel satisfying (160) and let $(\rho, u)$ be a global strong solution to (138). Assume that
\begin{equation}
 C \geq \rho(t, x) \geq \frac{c}{\sqrt{1 + t}}, \quad C > c > 0. 
\end{equation}
Then the solution aligns at an algebraic rate. Namely, there exist $\eta > 0$ such that
\begin{equation}
 \int_{T^d} |u(t, x) - \bar{u}|^2 \rho(t, x) \, dx \leq \frac{1}{2M} \eta^\nu. 
\end{equation}

Proof. We consider the family of eigenvalue problems parameterized by time:
\begin{equation}
 \int_{T^n} \phi(x, y)(u(x) - u(y)) \, dm_t(y) = \kappa(t)u(x), \quad u \in H^\alpha. 
\end{equation}
We seek the minimal eigenvalue, which is of course 0 corresponding to the constant eigenfunction. To remove this trivial solution we restrict it to the time-dependent 1-codimensional subspace
\begin{equation}
 H^\alpha_0 = \left\{ u \in H^\alpha : \int_{T^n} u \, dm_t = 0 \right\}. 
\end{equation}
The key issue here is that $H^\alpha_0$ depends on time. We will return to this later. So, we seek the second minimal eigenvalue of (163) restricted to $H^\alpha_0$, as a solution to the variational problem
\begin{equation}
 \kappa_2(t) = 2M \inf_{u \in H^\alpha_0} \frac{\int_{T^n} \phi(x, y)|u(y) - u(x)|^2 \rho(t, y) \rho(t, x) \, dx \, dy}{\int_{T^n} |u(x) - u(y)|^2 \rho(t, x) \rho(t, y) \, dx \, dy}. 
\end{equation}
In view of (160), and the assumed bounds on the density (161), the upper norm is equivalent for the $H^{\alpha/2}$, and the lower to $L^2$, so the existence follows classically by compactness. The number $\kappa_2(t)$ bears complete resemblance with the discrete Fiedler number, discussed in Section 2.2. In terms of this Fiedler number the energy equation (144) take form
\begin{equation}
 \frac{d}{dt} \mathcal{V}_2 \leq -\kappa_2(t)\mathcal{V}_2. 
\end{equation}
Consequently,
\begin{equation}
 \mathcal{V}_2(t) \leq \mathcal{V}_2(0) \exp \left\{ -\int_0^t \kappa_2(s) \, ds \right\}. 
\end{equation}
We will derive now the lower bound $\kappa_2(t) \geq c/(1 + t)$ which clearly implies the statement of the proposition. Using the bounds on the density (161), the mean-zero condition on $u$, and the lower bound of the kernel (160) we obtain
\begin{align*}
 \int_{T^n} |u(x) - u(y)|^2 \rho(t, y) \rho(t, x) \, dx \, dy &= 2M \int_{T^n} |u(x)|^2 \rho(t, x) \, dx \leq C\|u\|_2^2, \\
 \int_{T^n} \phi(x, y)|u(y) - u(x)|^2 \rho(t, y) \rho(t, x) \, dx \, dy &\geq \frac{c}{t} \int_{|x - y| < \rho_0} \frac{|u(x) - u(y)|^2}{|x - y|^{n+\alpha}} \, dx \, dy.
\end{align*}
\begin{equation}
 \kappa_2(t) \geq \frac{c}{t} \inf_{u \in H^\alpha_0} \frac{\int_{|x - y| < \rho_0} \frac{|u(x) - u(y)|^2}{|x - y|^{n+\alpha}} \, dx \, dy}{\|u\|_2^2}. 
\end{equation}
Technically, the infimum still depends on time since the mean-zero condition is. So, the last piece to show is that this infimum stays bounded away from zero. We argue by contradiction. Suppose there is a sequence of times $t_k > 0$, and $u_k \in H^{\alpha/2}$ with $\int u_k \, dm_{t_k} = 0$ such that $\|u_k\|_2 = 1$ and
\begin{equation}
 \int_{|x - y| < \rho_0} \frac{|u_k(x) - u_k(y)|^2}{|x - y|^{n+\alpha}} \, dx \, dy \to 0. 
\end{equation}
We now reach a contradiction if we prove the limit we conclude that \( \|u\|_{H^{\alpha/2}} = 0 \), and hence \( u \) is a constant field, with \( |u| = 1 \) due to \( \|u_k\|_2 \to \|u\|_2 \).

At the same time, since \( \int \rho(t_k, x) \, dx = M \), there exists a weak* limit of a further subsequence \( m_{t_k} \to m \), where \( m \) is a positive Radon measure on \( \mathbb{T}^d \) with non-trivial total mass \( m(\mathbb{T}^d) = M \). We now reach a contradiction if we prove the limit

\[
0 = \int_{\mathbb{T}^d} u_k(x) \rho(t_k, x) \, dx \to M u \neq 0.
\]

To prove the claimed limit note that the assumed uniform upper-bound of the density implies

\[
\int_{\mathbb{T}^d} u_k(x) \rho(t_k, x) \, dx - M u = \int_{\mathbb{T}^d} (u_k(x) - u) \rho(t_k, x) \, dx,
\]

and the latter is clearly bounded by \( C \|u_k - u\|_2 \to 0 \).

This proves that \( \kappa_2(t) \geq c/t \), and the result follows. \( \square \)

In the course of the proof we essentially established a statement analogous to Lemma 2.1 in the discrete case. Let us state it separately.

**Lemma 4.4.** Let \( \kappa_2(t) \) be the weighted Fiedler number defined by (165), and suppose that

\[
\int_0^\infty \kappa_2(s) \, ds = \infty.
\]

Then the solution aligns: \( \mathcal{V}_2 \to 0 \).

We can see now that unconditional flocking is generally achieved under the lower bound on the density, \( \rho(t, \cdot) \gtrsim (1 + t)^{-1/2} \). The difficulty is that this lower bound is too restrictive and is not given a priori for any strong solution. Situation improves considerably for the topological models which yield unconditional flocking under more accessible assumption on the density. We discuss those next.

### 4.3. Topological models. Adaptive diffusion.

Technically, topological models belong to a more general class of systems with kernels depending on all agent positions \( x = (x_1, \ldots, x_N) \). So, we start with the general formulation

\[
\begin{align*}
\dot{x}_i &= v_i, \\
\dot{v}_i &= \lambda \sum_{j: |x_i - x_j| < r_0} m_j \phi_{ij}(x)(v_j - v_i).
\end{align*}
\]

(170)

Here, we interpret the mass \( m_j \) as a parameter that quantifies power of the agent \( x_j \) to influence others. Thus, the bigger \( m_j \) is the more direct influence \( x_j \) exerts on others. At the same time, it is natural to assume that the “massive” agents are more resistant to the influence by other agents. This latter property will be encoded into the adjacency matrix \( \{\phi_{ij}(x)\}_{i,j=1}^N \) which reflects a chosen communication protocol. Our particular choice of \( \phi_{ij} \) will involve both geometric and topological features, where agent \( x_i \) senses proximity to others not only according to the Euclidean distance but also by the power of their influence. The model we have in mind is based on the following three principles:

1. Every agent \( x_i \) has a finite influence range, which is a Euclidean ball of radius \( r_0 \) centered at \( x_i \), denoted \( B(x_i, r_0) \).
2. Agent \( x_i \) exerts influence on another agent \( x_j \) via a transfer of information through a symmetric communication domain between the two agents, denoted \( \Omega(x_i, x_j) \), where \( x_i, x_j \in \partial \Omega(x_i, x_j) \).
3. The quantity

\[ d_{ij} = \left[ \sum_{k: x_k \in \Omega_{ij}} m_k \right]^{\frac{1}{n}} \]

is a measure of collective power of the intermediaries. We assume that the strength of communication is inversely proportional to \( d_{ij} \).

Based on the outlined principles, we make the following choice:

\[ \phi_{ij}(x) = \frac{1}{d_{ij}} \psi(|x_i - x_j|), \]

where \( \psi \) is a non-negative function supported in the ball of radius \( r_0 \), and \( \tau > 0 \) is a parameter. The kernel \( \psi \) encapsulates the metric component of the kernel and is limited to communication cutoff scale \( r_0 \), while \( \tau \) gauges presence of topological effects.

The choice of the domain \( \Omega(x, y) \) can be rather flexible, as long as it satisfies two basic requirements (refer to Figure 5): a) the region is a subset of the ball determined by \([x, y]\) as its diameter chord, and two cones of opening \(< \pi\) at vertices \( x \) and \( y \), b) \( \Omega(x, y) = \Omega(y, x) \), and the boundary of the region is smooth everywhere except for \( x, y \). For simplicity we also assume that topological communication is homogeneous and isotropic, i.e. \( \Omega(x, y) \) is constructed by a shift, rotation, and dilation of a basic domain \( \Omega(-e_1, e_1) \). In order to insure symmetry b) we assume that \( \Omega(-e_1, e_1) = -\Omega(-e_1, e_1) \).

Note that in the macroscopic limit \( N \to \infty \), the natural interpretation of the collective powers \( d_{ij} \) is given by

\[ \text{d}(x, y) = \left[ \int_{\Omega(x,y)} \rho(t, z) \, dz \right]^{\frac{1}{n}}. \]

Clearly, \( \text{d}(x, y) \) scales like a distance between \( x \) and \( y \), that is why we call it a “topological distance”. In fact in 1D, where

\[ \text{d}(x, y) = \int_x^y \rho(t, z) \, dz, \]

this does define a proper metric. As to the choice of metric component, it will play a more important role when we discuss regularity of the models, but for now it suffices to assume that

\[ \psi(r) \geq \lambda \mathbb{1}_{r < r_0}. \]
The corresponding hydrodynamic system is given by

$$\begin{cases}
\rho_t + \nabla \cdot (\rho u) = 0, \\
u_t + u \cdot \nabla u = \int_{\Omega^n} \phi(x,y)(u(y) - u(x))\rho(y)\,dy, \quad \phi(x,y) = \frac{\psi(|x-y|)}{d^n(x,y)}.
\end{cases}$$

We note that a proper care has to be given in order to properly define the singular integral operator $L\phi \ast f$ and the commutator $C\phi$ in this case. These issues are discussed in [13].

**Theorem 4.5.** Let $(\rho, u)$ be a global smooth solution of the topological model (173) on the torus $\mathbb{T}^n$, with $\tau \geq n$. Assume that the density satisfies the lower bound

$$\rho(t, x) \geq \frac{c}{1+t}, \quad \forall t > 0, x \in \mathbb{T}^n.$$  

Then the solution aligns with a logarithmic rate given by

$$|u(t) - \bar{u}|_\infty \leq \frac{c}{\sqrt{\ln t}}.$$  

**Remark 4.6.** The assumption on the density (174) holds automatically in 1D case, as we will see in Section 6.

**Proof.** We aim to prove (175) for each component of velocity $u_i$. Let us denote $u = u_i$ for short. By the Galilean invariance of the system we can add a constant to $u$ if necessary and assume that $u(t) > 0$. By the maximum principle the extrema of $u(t)$, denoted $u_+(t)$ and $u_-(t)$, are monotone.

**Step 1: Alignment near extremes.**

Denote by $x_+(t)$ a point where the maximum of $u(t, \cdot)$ is achieved, and by $x_-(t)$ for minimum. Let us fix a time-dependent $\delta(t) > 0$ to be specified later. Consider the sets

$$G_+^{\delta}(t) = \{u < u_+(t)(1 - \delta(t))\}, \quad G_-^{\delta}(t) = \{u > u_-(t)(1 + \delta(t))\}.$$  

The amount flattening near extreme values will be quantified in terms of conditional expectations of the above sets relative to the local balls $B(x_{\pm}(t), r_0)$. We denote such expectations by

$$\mathbb{E}_t[A|B] = \frac{m_t(A \cap B)}{m_t(B)}.$$  

First, let us show that

$$\int_0^\infty \mathbb{E}_t[G_+^{\delta}(t)|B(x_{\pm}(t), r_0)]\,dt < \infty.$$  

Figure 6. Initial flattening near points of maximum and minimum.
To this purpose let us compute the equation at \((x_+(t), t)\) (with the use of Rademacher’s lemma 2.5)

\[
\frac{d}{dt} u_+ = \int_{\Omega} \phi(x_+, y)(u(y) - u_+) \rho(y) \, dy.
\]

We now use our assumption (172) and the fact that \(\tau \geq n\) to obtain the bound

\[
(177) \quad \frac{\lambda \int_{r<n_0} (|x-y|)}{d^n(x, y)} \leq \frac{\psi(x-y)d^{\tau-n}(x, y)}{d^n(x, y)} \leq M^{\tau-n} \phi(x, y).
\]

Thus, we have

\[
-\frac{d}{dt} u_+ = \int \phi(x_+, y)(u_+ - u(y)) \rho(y) \, dy \\
\geq \int_{B(x_+, r_0)} \frac{1}{d(x_+, y)} (u_+(t) - u(y)) \rho(y) \, dy \\
\geq \frac{1}{m_t(B(x_+, r_0))} \int_{G_+^1(t) \cap B(x_+, r_0)} (u_+ - u(y)) \rho(y) \, dy \quad \text{(since } \Omega(x_+, y) \subset B(x_+, 0)) \\
\geq \frac{\delta(t) u_+}{m_t(B(x_+, r_0))} \int_{G_+^1(t) \cap B(x_+, r_0)} \rho(y) \, dy \\
= \delta(t) u_+ \mathbb{E}_t[G_+^1(t)|B(x_+, r_0)].
\]

Integrating we obtain

\[
\int_0^\infty \delta(t) \mathbb{E}_t[G_+^1(t)|B(x_+(t), r_0)] \, dt \lesssim \ln \frac{u_+(0)}{\lim_{t \to -\infty} u_+(t)} \leq \ln \frac{u_+(0)}{u_-(0)}.
\]

**STEP 2: USE OF CAMPANATO-MORREY NORM.** On this step we show that \(u\) does not deviate much from its averages over local balls. We express such measure of deviation in terms of a Campanato-Morrey metric to be specified later. Let us denote

\[
u_{x,r} = \frac{1}{m_t(B(x, r))} \int_{B(x, r)} u(t, z) \, dm_t(z).
\]

Recall that the communication domains \(\Omega(x, x')\) are confined to cones intersected with the diametric ball. As a result the following geometric observation is true, see Figure 7.

*Claim 4.7.* There exists a \(c_0 > 0\) depending only on the opening angle of the cones such that for any \(r > 0\) and any triple of points \((x, x', x^*)\) with \(|x - x^*| < c_0 r\) and \(|x' - x^*| < r\), we have \(\Omega(x, x') \subset B(x^*, r)\).

Let us fix an arbitrary \(x^* \in \mathbb{T}^n\). By Hölder inequality, we have the following estimate for any \(r < r_0/2\):

\[
\int_{|x-x^*|<c_0r} |u(x) - u_{x^*, r}|^2 \rho(x) \, dx \lesssim \int_{|x-x^*|<c_0r} \frac{1}{m_t(B(x^*, r))} |u(x) - u(x')|^2 \rho(x) \rho(x') \, dx' \, dx \\
\]

using that \(m_t(B(x^*, r)) \geq m_t(\Omega(x, x')) = d^n(x, x')\)

\[
\leq \int_{|x-x'|<1+c_0r} \frac{1}{d^n(x, x')} |u(x) - u(x')|^2 \rho(x) \rho(x') \, dx' \, dx \\
\leq C \int_{\mathbb{T}^n} \phi(x, x') |u(x) - u(x')|^2 \rho(x) \rho(x') \, dx' \, dx.
\]
From the energy equation (144), the right hand side is globally integrable on $\mathbb{R}_+$. Consequently, we obtain a global time integrability for the following Campanato-Morrey semi-norm,

$$(178) \quad \int_0^\infty [u(t)]^2_{r_0} \, dt < \infty, \quad [u(t)]^2_{r_0} := \sup_{x^* \in \mathbb{T}^n, r < \frac{r_0}{2}} \int_{|x-x^*| < c_0 r} |u(x) - u_{x^*, r}|^2 \rho(x) \, dx.$$  

In combination with (176) we obtain

$$I = \int_0^\infty \left( \delta(t)\mathbb{E}_t[G^+_\delta(t)|B(x_+(t), r_0)] + [u(t)]^2_{r_0} \right) \, dt < \infty.$$ 

Denoting $A = e^{2I}$ we have

$$\int_T^{T^A} \frac{dt}{t \ln t} = 2I \quad \text{for all} \quad T > 0.$$ 

Consequently, for any $T > 0$ there exists a $t \in [T, T^A]$ such that

$$(179) \quad [u(t)]^2_{r_0} < \frac{1}{t \ln t} \quad \mathbb{E}_t[G^+_\delta(t)|B(x_+(t), r_0)] + \mathbb{E}_t[G^-_\delta(t)|B(x_-(t), r_0)] < \frac{1}{\delta(t)t \ln t}$$

By virtue of the lower bound on the density (174) this implies that

$$(180) \quad \sup_{x^*, r < \frac{r_0}{2}} \int_{|x-x^*| < c_0 r} |u(x) - u_{x^*, r}|^2 \, dx \leq \frac{1}{\ln t}.$$ 

**Step 3: Sliding averages.** Let $t \in [T, T^A]$ be the moment of time fixed above. Let us fix $r = \frac{1}{3}r_0$, which lies within the reach of the Campanato metric. We will connect the two averages $u_{x_+, r}$ and $u_{x_-, r}$ sliding along the line connecting $x_+$ and $x_-$, and show that the fluctuation of those averages is small.

To this end, consider the direction vector $\mathbf{n} = \frac{x_+ - x_-}{|x_+ - x_-|}$ and define a sequence of balls, $B_k = B(x_k, c_0 r)$, $k = 0, \ldots, K$, with centers given by $x_0 = x_-$ and defined recursively by $x_{k+1} = x_k + c_0 r \mathbf{n}$ up to $k = K - 1$ and ending with $x_K = x_+$. The point is that the balls overlap significantly: $|B_k \cap B_{k+1}| \geq c_1 r_0^n$. 

**Figure 7.** $\Omega(x, x')$ is trapped in the outer ball if $x$ is close to the center $x^*$. 
By the Chebychev inequality, followed by (180) applied to the ball centered at \( x_0 \), yields
\[
|\{ x \in B_0 \cap B_1 : |u(x) - u_{x_0,r}| > \eta \}| \leq \frac{1}{\eta^2} \int_{B_0} |u(x) - u_{x_0,r}|^2 \, dx \leq \frac{1}{\eta^2 \ln t}.
\]
Let us set \( \eta = \frac{2}{\sqrt{c_1 r_0^m \ln t}} \) so that
\[
|\{ x \in B_0 \cap B_1 : |u(x) - u_{x_0,r}| > \eta \}| \leq \frac{1}{4} |B_0 \cap B_1|.
\]
By the same argument applied to the fluctuation around the averaged value \( u_{x_1,r} \) we obtain
\[
|\{ x \in B_0 \cap B_1 : |u(x) - u_{x_1,r}| > \eta \}| \leq \frac{1}{4} |B_0 \cap B_1|.
\]
Hence, the complements of the two sets must have a common point in the intersection \( B_0 \cap B_1 \):
\[
\{ x \in B_0 \cap B_1 : |u(x) - u_{x_0,r}| \leq \eta \} \cap \{ x \in B_0 \cap B_1 : |u(x) - u_{x_1,r}| \leq \eta \} \neq \emptyset.
\]
This implies
\[
|u_{x_0,r} - u_{x_1,r}| \leq 2\eta.
\]
Continuing in similar fashion we recover the same bound for all consecutive pairs of averages:
\[
|u_{x_k,r} - u_{x_{k+1},r}| \leq 2\eta.
\]
Hence,
\[
(181) \quad |u_{x_-,r} - u_{x_+,r}| \leq 2K\eta \lesssim \frac{1}{\sqrt{\ln t}}.
\]
Notice the absolute bound on \( K \lesssim 1/r_0 \). Furthermore, in view of (179), the follows estimate holds
\[
u_{x_+,r} \geq \frac{1}{m_t(B(x_+,r))} \int_{B(x_+,r) \setminus G_\delta} u_+(t)(1 - \delta(t)) \, dm_t
\]
\[
\geq u_+(t)(1 - \delta(t))(1 - E_t[G_\delta^+(t) \mid \delta(t) R_0]) \geq u_+(t)(1 - \delta(t)) \left( 1 - \frac{1}{\delta(t) t \ln t} \right).
\]
Consequently,
\[
u_+(t) - u_{x_+,r}(t) \lesssim \delta(t) + \frac{1}{\delta(t) t \ln t} \lesssim \frac{1}{\sqrt{t \ln t}}
\]
provided we make the following choice of \( \delta \)
\[
\delta(t) = \frac{1}{\sqrt{\ln t}}.
\]
Exact same argument can be make to estimate the bottom average. In combination with (181) these imply
\[
|u_+(t) - u_-(t)| \lesssim \frac{1}{\sqrt{\ln t}}.
\]
To conclude we notice that the maximum principle implies
\[
|u_+(T^A) - u_-(T^A)| \lesssim \frac{1}{\sqrt{\ln t}} \sim \frac{1}{\sqrt{\ln(T^A)}}.
\]
Since \( T \) is arbitrary the proof is complete.
Remark 4.8. Theorem 4.5 allows for an extension which improves upon the rate of alignment under more restrictive bound from below on the density. Specifically, the following statement can be proved along the lines of the above argument: suppose
\begin{equation}
\rho(t, x) \geq \frac{c}{(1 + t)^\gamma}, \quad 0 \leq \gamma \leq 1,
\end{equation}
then the solution aligns with the following algebraic rate
\begin{equation}
\|u(t) - \bar{u}\|_\infty = \frac{o(1)}{t^{\frac{1}{2}(1-\gamma)}}.
\end{equation}

5. Regularity I: Local well-posedness and continuation criteria

The regularity theory for models with smooth communication is understandably quite different from singular models – the former is essentially Burgers’ equation with a dumping mechanism, while the latter is a degenerate fractional parabolic system with dissipation in the momentum equation. In this section we will go through a rather elementary but necessary for future exposition first step – proving local existence of classical solutions. We are not seeking the sharpest spaces to keep exposition simple. However, we do require our local solutions to have certain level of regularity for various phenomena to remain classically verifiable, such as mass conservation, and existence of characteristic flow. So, the results presented below will respect such requirements. We will also limit ourselves to the metric models, where the estimates are not excessively contaminated with density dependent coefficients.

5.1. Smooth models. Let us assume throughout that \( \phi \) is sufficiently smooth to take as many derivatives as necessary in the course of our arguments below. We also assume that \( \phi = \phi(x - y) \) is of convolution type and that the environment domain is \( \mathbb{R}^n \). The exact same results will carry over to \( \mathbb{T}^n \) with slight modifications.

Using the commutator structure of the alignment term and (141) we write the system (138) as
\begin{equation}
\begin{cases}
\rho_t + \nabla \cdot (\rho u) = 0, \\
u_t + u \cdot \nabla u = \phi \ast (u \rho) - u \phi \ast \rho.
\end{cases}
\end{equation}

Suppose we would like to prove local existence of solutions in Sobolev class \( u \in H^m, \rho \in H^k \cap L^1_+ \), where \( L^1_+ \) denotes the set of non-negative functions in \( L^1 \). Note that the Sobolev embedding does not guarantee that \( \rho \in L^1_+ \) if it is in a higher Sobolev class, yet \( u \in L^\infty \) automatically if \( m > \frac{n}{2} \). Both conditions are natural quantities to include in the class as they are controled by dynamics a priori.

One can obtain local existence rather easily for a viscous regularization:
\begin{equation}
\begin{cases}
\rho_t + \nabla \cdot (\rho u) = \varepsilon \Delta \rho, \\
u_t + u \cdot \nabla u = \phi \ast (u \rho) - u \phi \ast \rho + \varepsilon \Delta u.
\end{cases}
\end{equation}

Indeed, we are going to denote the grand quantity \( Z = (u, \rho) \) and consider the equivalent mild formulation of (184):

\[ Z(t) = e^{\varepsilon t \Delta} Z_0 + \int_0^t e^{\varepsilon (t-s) \Delta} \mathcal{N}(Z(s)) \, ds, \]

where \( \mathcal{N}(Z) \) denotes all the non-linear terms in (184). The argument goes by the standard contractivity argument. Let us fix \( Z_0 \in H^m \times (H^k \cap L^1_+) \) and consider the map

\[ T[Z](t) = e^{\varepsilon t \Delta} Z_0 + \int_0^t e^{\varepsilon (t-s) \Delta} \mathcal{N}(Z(s)) \, ds. \]
We need to show that for some small $T$ this maps is a contraction on $C([0, T); B_1(Z_0))$, where $B_1$ is understood in metric of $X = H^m \times (H^k \cap L^1)$. Let us invariance, while contractivity following similarly. First, by continuity of the heat semigroup,

$$\|Z_0 - e^{\varepsilon t \Delta}Z_0\|_X \leq \frac{1}{2}^\varepsilon t$$

for small $t$. To estimate the integral we first recall analyticity property of the heat semigroup:

$$\|\nabla e^{\varepsilon t \Delta} f\|_{L^p} \leq \frac{1}{\sqrt{\varepsilon t}} \|f\|_{L^p}, \quad 1 \leq p \leq \infty.$$  

So, considering the $\rho$-component we obtain

$$\left\| \int_0^t e^{\varepsilon (t-s) \Delta} \cdot (\rho u) \, ds \right\|_{L^1} \leq \int_0^t \frac{1}{\sqrt{\varepsilon (t-s)}} \|u \rho (s)\|_{L^1} \, ds$$

$$\leq \frac{1}{\varepsilon t^{1/2}} \sup_s \|u(s)\|_\infty\|\rho(s)\|_1 \leq \frac{T^{1/2}}{\varepsilon t^{1/2}} (\|Z_0\|^2 + 1) < \frac{1}{2},$$

for small $T$. For $H^k$ we argue similarly for $L^2$. So, let $\partial^k$ be a multiindex partial derivative of order $k$. We will use the product estimate

$$\|\partial^k (u \rho)(s)\|_{L^2} \leq \|u\|_\infty \|\rho\|_{H^k} + \|u\|_{H^k} \|\rho\|_\infty.$$  

It is clear at this point that in order to close the estimates in $X$ we need to assume that $\frac{n}{2} < k \leq m$, in which case

$$\|\partial^k (u \rho)(s)\|_{L^2} \leq \|Z(s)\|_{X}^2.$$  

So, by the same argument we obtain

$$\left\| \partial^k \int_0^t e^{\varepsilon (t-s) \Delta} \cdot (\rho u) \, ds \right\|_{L^2} \leq \int_0^t \frac{1}{\sqrt{\varepsilon (t-s)}} \|\partial^k(u \rho)(s)\|_{L^2} \, ds < \frac{1}{2}.$$  

On the velocity side there are two terms to handle. For the transport part, the $L^2$-estimate is straightforward, and

$$\left\| \partial^m \int_0^t e^{\varepsilon (t-s) \Delta} u \cdot \nabla u \, ds \right\|_{L^2} \leq \int_0^t \frac{1}{\sqrt{\varepsilon (t-s)}} \|\partial^{m-1}(u \cdot \nabla u)(s)\|_{L^2} \, ds$$

$$\leq \int_0^t \frac{1}{\sqrt{\varepsilon (t-s)}} (\|u\|_{H^{m-1}} \|\nabla u\|_\infty + \|u\|_{H^m} \|u\|_\infty) \, ds < \frac{1}{4},$$

provided $m > \frac{n}{2} + 1$ to ensure embedding of $W^{1, \infty}$ into $H^m$. The commutator term $\phi \ast (u \rho) - u \phi \ast \rho$ is even easier, since the derivatives are absorbed by the kernel except when all fall on $u$, which results in the same estimate.

We have shown that $T : C([0, T); B_1(Z_0)) \rightarrow C([0, T); B_1(Z_0))$ is a contraction, and so, we obtain a local solution on a time interval dependent on $\varepsilon$. Denoting $T^*$ the maximal time of existence in $C([0, T); X)$ we show that $T^*$ depends only on the $X$-norm of the initial condition. We do it by establishing a priori estimates that are independent of $\varepsilon$ and which will allow us to pass to the limit of vanishing viscosity. So, the grand quantity we are trying to control is

$$Y_{m,k} = \|u\|^2_{H^m} + \|\rho\|^2_{H^k} + |\rho|^2_{1}.$$  

To start, we write the continuity equation as

$$\rho_t + u \cdot \nabla \rho + (\nabla \cdot u) \rho = 0.$$  

So, testing with $\partial^{2k} \rho$ we obtain

$$\frac{d}{dt} \|\rho\|^2_{H^k} = \int (\nabla \cdot u |\partial^k \rho|^2 \, dx - \int (\partial^k (u \cdot \nabla \rho) - u \cdot \nabla \partial^k \rho) \partial^k \rho \, dx - \int \partial^k ((\nabla \cdot u) \rho) \partial^k \rho \, dx - \varepsilon \|\rho\|^2_{H^{k+1}}.$$
We dismiss the last term. Recalling the classical commutator estimate
\begin{equation}
\|\partial^k (fg) - f \partial^k g \|_2 \leq |\nabla f| |g| \|\hat{H}^{k-1} + \|f\| \|\hat{H}^k| g\|_\infty,
\end{equation}
we obtain
\[
\frac{d}{dt} \|\rho\|_{\hat{H}^k}^2 \leq |\nabla u|_\infty \|\rho\|_{\hat{H}^k}^2 + \|u\|_{\hat{H}^k} \|\rho\|_{\hat{H}^k} |\nabla \rho|_\infty + \|u\|_{\hat{H}^{k+1}} \|\rho\|_{\hat{H}^k} |\rho|_\infty
\leq C (|\nabla u|_\infty + |\nabla \rho|_\infty + |\rho|_\infty) Y_{m,k},
\]
provided \( m \geq k + 1 \). The \( L^2 \) norm of \( \rho \) obeys a similar estimate trivially, and the \( L^1 \)-norm is conserved.

For the velocity equation we apply the same commutator estimate for the material derivative part:
\[
\int \partial^m (u \cdot \nabla u) \partial^m u \, dx = - \int \nabla \cdot u |\partial^m u|^2 \, dx + \int [\partial^m (u \cdot \nabla u) - (u \cdot \nabla \partial^m u)] \partial^m u \, dx
\lesssim |\nabla u|_\infty \|u\|_{\hat{H}^m}^2.
\]
For the alignment term, we can put all the derivatives onto the kernel whenever possible and the only term that is left out is \( |\partial^m u|^2 \phi \star \rho \) with \( \phi \star \rho \) clearly bounded by \( |\phi|_\infty M \), a priori conserved quantity. So, we obtain
\[
\frac{d}{dt} \|u\|_{\hat{H}^m}^2 \leq (|\nabla u|_\infty + C(\|\phi\|_{C^m}, M)) \|u\|_{\hat{H}^m}^2.
\]
The similar bound for \( \frac{d}{dt} \|u\|_{\hat{H}^2}^2 \) is derived trivially. So, we obtain
\begin{equation}
\frac{d}{dt} \|u\|_{\hat{H}^m}^2 \leq (|\nabla u|_\infty + C(\|\phi\|_{C^m}, M)) \|u\|_{\hat{H}^m}^2.
\end{equation}
It is important to note that this bound is independent of the higher norms of the density. Combining the two equations we obtain
\begin{equation}
\frac{d}{dt} Y_{m,k} \leq C (|\nabla u|_\infty + |\nabla \rho|_\infty + |\rho|_\infty + C(\|\phi\|_{C^m}, M)) Y_{m,k}.
\end{equation}
Of course \( |\nabla u|_\infty + |\nabla \rho|_\infty + |\rho|_\infty \leq Y_{m,k} \) provided \( k > \frac{n}{2} + 1 \), which adds the last restriction on the exponents for the argument to work. So, if \( m \geq k + 1 > \frac{n}{2} + 2 \), then
\[
\frac{d}{dt} Y_{m,k} \leq C_1 Y_{m,k} + C_2 Y_{m,k}^2.
\]
Solving the Riccati equation gives a uniform bound on the \( X \)-norm on a time interval inversely proportional to \( \|Z_0\|_X \), but independent of \( \varepsilon \). Thus, solutions to (185) with the same initial data exist on a common time interval \([0, T_0]\) where they are uniformly bounded in \( C([0, T_0]; X) \).

Let us also note that keeping the dissipative terms in the estimates above also shows that
\[
\varepsilon \int_0^{T_0} (\|\rho(s)\|_{\hat{H}^{k+1}}^2 + \|u(s)\|_{\hat{H}^{m+1}}^2) \, ds < C,
\]
where \( C \) is independent of \( \varepsilon \). Then
\[
\|Z_t\|_{L^2} \leq \|Z\|_{X}^2 + \varepsilon \|Z\|_{\hat{H}^2} \leq \|Z\|_{X}^2 + \varepsilon \|Z\|_{\hat{H}^{k+1} \times \hat{H}^{m+1}}.
\]
So, \( Z_t \in L^2([0, T_0]; L^2) \). Passing to a subsequence we find a weak limit \( Z_\varepsilon \to Z \) in \( L^\infty([0, T_0]; X) \) and \( (Z_\varepsilon)_t \to Z_t \) in \( L^2([0, T_0]; L^2) \) (technically, a limit in \( L^1 \) may end up being a measure of bounded variation, however as a member of \( H^k \) it is absolutely continuous, hence in \( L^1 \)). Since \( Z_t \in L^2([0, T_0]; L^2) \), \( Z \) is weakly continuous with values in \( L^2 \). Since \( L^2 \) is dense in \( H^{-m} \) and \( H^{-k} \) this
implies weak continuity $Z \in C_w([0,T_0]; H^m \times H^k)$. Strong continuity of the density follows from the equations itself:

$$\|\rho_t\|_{L^1} \leq \|\rho \nabla u\|_1 + \|u \nabla \rho\|_1 \leq \|Z\|_{X}^2 < C.$$ 

Further regularity in time $Z_t$ follows from measuring smoothness of the system one level down and performing similar product estimates as above.

Having established local existence in $X$ let us come back to (188) and notice that this solution can in fact be extended beyond $T_0$ if we know that

$$\int_{T_0}^T \|\nabla u\|_{\infty} \, dt < \infty.$$ 

Indeed, $|\rho|_{\infty}$ can be bounded by solving the continuity equation along characteristics

$$\rho(X(t, \alpha), t) = \rho(\alpha, 0) \exp \left\{ - \int_0^t \nabla \cdot u(X(s, \alpha), s) \, ds \right\}.$$ 

Using (187) we see that $\|u\|_{H^m}^2$ is also bounded uniformly, and hence so is $|\nabla^2 u|_{\infty}$ since $m > \frac{n}{2} + 2$. Bootstrapping further by differentiating the continuity equation we bound $|\nabla \rho|_{\infty}$ in a similar fashion.

Having this continuation criterion at hand we can further improve the local existence result by establishing control over $|\nabla u|_{\infty}$ directly. First, by the maximum principle, $|u(t)|_{\infty} \leq |u_0|_{\infty}$. Writing equation for one component $\partial_i u_j$ we have

$$\partial_t \partial_i u_j + u \cdot \nabla \partial_i u_j + \partial_i u \cdot \nabla u_j = (\partial_i \phi) * (u_j \rho) - \partial_j u \cdot \partial_i \phi * \rho - u_j (\partial_i \phi) * \rho.$$ 

Evaluating at the maximum and minimum and adding up over $i, j$ we obtain

$$\frac{d}{dt} |\nabla u|_{\infty} \leq |\nabla u|_{\infty}^2 + CM |u|_{\infty} + M |\nabla u|_{\infty}.$$ 

Hence, $|\nabla u|_{\infty}$ is uniformly bounded a priori on a time interval depending only on $|\nabla u_0|_{\infty}^{-1}$. So, the continuation criterion allows to extend our local solution to that time interval.

Using the transport structure of the momentum equation we can further relax (189) to a condition on divergence of $u$, which will be extremely useful in the future. Let us recall that we have uniform bounds on the deformation tensor of the characteristic flow map given by (153). To translate this to a bound on $|\nabla u|_{\infty}$ in Eulerian coordinates we invert the flow map

$$\nabla u(x, t) = \nabla^{-\top} X(X^{-1}(x, t), t) \nabla \phi(X^{-1}(x, t), t).$$ 

Using that

$$|\nabla^{-\top} X(\alpha, t)| \leq \frac{C_1}{|\det \nabla X(\alpha, t)|} e^{C_2 t}$$ 

we recall the Liouville formula for the Jacobian

$$\det \nabla X(\alpha, t) = \exp \left\{ \int_0^t \nabla \cdot u(X(\alpha, t), t) \, dt \right\}.$$ 

So, as long as

$$\int_{T_0}^T \inf_{x \in \mathbb{R}^n} \nabla \cdot u(t, x) \, dt > -\infty,$$ 

this guarantees that the Jacobian does not vanish. This establishes uniform bound on $|\nabla u(t)|_{\infty}$ on time interval $[0, T_0]$.

Let us record the obtained results in the following theorem.
Theorem 5.1 (Local existence of classical solutions). Suppose $m \geq k + 1 > \frac{n}{2} + 2$, and $(u_0, \rho_0) \in H^m \times (H^k \cap L^1_+)$. Then there exists time $T_0 = T_0(|\nabla u_0|^{-1}_\infty, M)$ and a unique solution to (184) on time interval $[0, T_0)$ in the class

$$\tag{192} (u, \rho) \in C_w([0, T_0); H^m \times (H^k \cap L^1_+) \cap \text{Lip}([0, T_0); H^{m-1} \times (H^{k-1} \cap L^1_+))$$

satisfying the given initial condition. Moreover, any classical local solution on $[0, T_0)$ in class (192) and satisfying (191) can be extended beyond $T_0$.

Theorem 5.1 can be used as a stepping stone to obtain solutions with less smoothness, especially for $\rho$, as long as regularity of $u$ permits to define smooth characteristic flow map with sufficient compactness properties.

So, let us first assume $u_0 \in H^m$, $m > \frac{n}{2} + 1$ and $\rho_0 \in L^1_+$, the most basic assumption on the density. Mollifying the data $((u_0)_\varepsilon, (\rho_0)_\varepsilon)$, due to Theorem 5.1, we obtain a family of local solutions on a common time interval $T_0$, since $H^m \subset W^{1,\infty}$, and $|\nabla (u_0)_\varepsilon|_\infty \leq |\nabla u_0|_\infty$. We also note that the estimate (187) holds for any integer $m$, hence $u_\varepsilon \in L^{\infty}([0, T_0); H^m)$ uniformly. We also have $\partial_t u_\varepsilon \in L^{\infty}([0, T_0); H^{m-1}) \subset L^{\infty}([0, T_0); L^\infty)$.

Let us note that $H^m \subset W^{1+\delta,\infty}$ for some $\delta > 0$, and the embedding is compact on any bounded set, and of course $W^{1+\delta,\infty} \subset L^\infty$. So, the Aubin-Lions-Simon Lemma implies that the family is compact in $C([0, T_0); W^{1+\delta,\infty})$ on any bounded set. Passing to a subsequence we obtain a weak limit $u$ in $L^{\infty}([0, T_0); H^m)$ and strong in $C([0, T_0); W^{1+\delta,\infty})$ on any bounded set. The velocity also belongs to $C_w([0, T_0); H^m)$ as a consequence of the two memberships. Similarly, the family of flows $X_\varepsilon$ belongs to $L^{\infty}([0, T_0); W^{1,\infty}) \cap \text{Lip}([0, T_0); L^\infty)$ so is compact in $C([0, T_0); C^1)$, $\delta < 1$, on any bounded set. We can thus claim strong uniform convergence of the flow maps as well. Considering the solution to the continuity equation

$$\tag{193} \rho_\varepsilon(X_\varepsilon(t, \alpha), t) = \rho_\varepsilon(\alpha, 0) \exp\left\{-\int_0^t \nabla \cdot u_\varepsilon(X_\varepsilon(s, \alpha), s) \, ds\right\},$$

we can clearly pass to the strong limit in $L^1$ and the limit satisfies (190).

Note that the density essentially plays the role of a passive scalar. In particular, if $\rho_0 \in L^1_+ \cap L^\infty$ initially, then by formula (190) it will remain in the same class on the entire time interval.

The argument above can be elevated to any higher smoothness $H^m \times (L^1_+ \cap W^{1,\infty})$ as long as $m > \frac{n}{2} + k + 1$. It goes by differentiating the continuity equation $k$ times and running the same compactness procedure. Weak continuity of the density in $L^1_+ \cap W^{k,\infty}$ follows from the established regularity properties of the velocity field and the corresponding formula for the solution of $\partial^k \rho$. The continuation criterion (191) remains valid in this setting as well.

Theorem 5.2 (Local existence of strong solutions). Suppose $m > \frac{n}{2} + k + 1$, $k = 0, 1, \ldots$ and $(u_0, \rho_0) \in H^m \times (L^1_+ \cap W^{k,\infty})$. Then there exists time $T_0 = T_0(|\nabla u_0|^{-1}_\infty, M)$ and a unique solution to (184) on time interval $[0, T_0)$ in the class

$$\tag{194} (u, \rho) \in C_w([0, T_0); H^m \times (L^1_+ \cap W^{k,\infty}))$$. Moreover, any such solution satisfying (191) can be extended beyond $T_0$.

The case $k = 1$ will already be interesting in dimension 1, where the given regularity is sufficient to justify pointwise evaluation of $u_\varepsilon$ and to study strong flocking.

5.2. Singular models. For models with singular kernels given by (5), $\beta = n + \alpha$, $0 < \alpha < 2$, the operator $L_\phi$ becomes the fractional Laplacian, although we also consider local kernels. So we set

$$\phi(r) = \frac{h(r)}{r^{n+\alpha}}.$$
and consider the model on periodic domain

\[
\begin{aligned}
\rho_t + \nabla \cdot (\rho u) &= 0, \\
u_t + u \cdot \nabla u &= \int_{\mathbb{T}^n} \phi(x-y)(u(y)-u(x))\rho(y)\,dy \\
(x,t) &\in \mathbb{T}^n \times \mathbb{R}_+.
\end{aligned}
\]

The important fact is that we have a coercivity bound

\[
c_1\|f\|_{\dot{H}^\alpha} - c_2\|f\|_2 \leq \|L_\phi f\|_2 \leq c_1\|f\|_{\dot{H}^\alpha} + c_3\|f\|_2.
\]

We will be casting our regularity theory for singular models on the periodic domain \(\mathbb{T}^n\) and for non-vacuous solutions only. This is motivated by technical reasons rather than applications, although one can argue that periodic conditions are suitable for studying flocks in the bulk. The primary reason is that we require uniform parabolicity of the commutator \((140)\) for estimates to go through. Such parabolicity depends on the pointwise bound \(\rho > c_0 > 0\), which is consistent with finite mass of the flock only on bounded domains.

Furthermore, we can easily construct a blowing up solution with vacuum and with a local kernel. Indeed, consider a local kernel \(\text{Supp } \phi \subset B_1(0)\). Let initial density \(\text{Supp } \rho_0 \subset B_{\varepsilon}(0)\), while \(u_0 = 1\) on \(B_{10}(0)\), \(v_0 = 0\) on \(B_{10+\varepsilon}(0)\) and smooth in between. Then the density will remain in \(B_2(0)\) for a time period of at least \(t < 1\), due to \(u \leq 1\). During this time the momentum equation will remain pure Burgers, hence the solution will evolve into a shock at a time \(t \sim \varepsilon < 1\). A similar argument can be done even for a global singular kernel on the periodic domain and this was formalized in \([18]\).

Performing energy estimates in the same fashion as for smooth models will inevitably create a derivative overload on the density. Instead we consider another "almost conserved" quantity

\[
e = \nabla \cdot u + L_\phi \rho,
\]

which satisfies the equation

\[
e_t + \nabla \cdot (ue) = (\nabla \cdot u)^2 - \text{Tr}(\nabla u)^2.
\]

Let us derive it in general for the sake of completeness.

We have

\[
\partial_t e + \nabla \cdot [ue] = (\nabla \cdot u)^2 - \text{Tr}(\nabla u)^2.
\]

Since \(\phi\) is a convolution kernel, we have that

\[
\partial_t L_\phi + \nabla \cdot L_\phi (\rho u) = 0.
\]

Taking the divergence of the velocity equation, we obtain

\[
\partial_t (\nabla \cdot u) + \nabla \cdot [u \cdot \nabla u] = \nabla \cdot L_\phi (\rho u) - \nabla \cdot [u L_\phi \rho]
\]

with

\[
\nabla \cdot [u L_\phi \rho] = L_\phi \rho \nabla \cdot u + u \cdot \nabla L_\phi \rho
\]

and

\[
\nabla \cdot [u \cdot \nabla u] = \text{Tr}(\nabla u)^2 + u \cdot \nabla (\nabla \cdot u).
\]

On one hand, combining \((200)\) and \((201)\), we obtain that

\[
\partial_t e + L_\phi \rho \nabla \cdot u + u \cdot \nabla e + \text{Tr}(\nabla u)^2 = 0.
\]

Adding and subtracting now \((\nabla \cdot u)^2\) produces \((213)\).

From the order of terms that enter into the formula for \(e\), it is clear that the natural correspondence in regularity for state variables involved is \((u \in H^{m+1}) \sim (\rho \in H^{m+\alpha})\).

Note that in 1D the right hand side vanishes and we have a perfect conservation law. This case will be discussed at length in Section 6.
The grand quantity to be estimated is
\[ Y_m = \|u\|_{H^{m+1}}^2 + \|\varepsilon\|_{H^m}^2 + |\varepsilon|_\infty + |\rho|_1 + |\rho^{-1}|_\infty, \]
which is equivalent to \( Y_m \sim \|u\|_{H^{m+1}}^2 + \|\rho\|_{H^{m+\alpha}} + |\rho^{-1}|_\infty \) in view of (196).

Our strategy will be very similar to the smooth case, where we obtain local solutions via viscous regularization, and prove a continuation criterion via a priori estimates on \( Y_m \). We assume throughout that \( m > \frac{5}{2} + 1 \) and \( 0 < \alpha < 2 \).

So, let us start with (185) and consider the mild formulation
\[
\begin{align*}
\rho(t) &= e^{\varepsilon t \Delta} \rho_0 - \int_0^t e^{\varepsilon (t-s) \Delta} \nabla \cdot (u\rho)(s) \, ds \\
u(t) &= e^{\varepsilon t \Delta} u_0 - \int_0^t e^{\varepsilon (t-s) \Delta} u \cdot \nabla u(s) \, ds + \int_0^t e^{\varepsilon (t-s) \Delta} \mathcal{C}_\phi(u,\rho)(s) \, ds.
\end{align*}
\]
Let us denote as before \( Z = (\rho, u) \) and by \( T[Z](t) \) the right hand side of the mild formulation. In order to apply the standard fixed point argument we have to show that \( T \) leaves the set \( C([0,T_R]; B_\delta(Z_0)) \) invariant, where \( B_\delta(Z_0) \) is the ball of radius \( \varepsilon \) around initial condition \( Z_0 \), and that it is a contraction. We limit ourselves to showing details for invariance as the estimates involved in proving Lipschitzness are similar.

First we assume that \( \rho \) has no vacuum: \( \rho_0(x) \geq c_0 > 0 \). Since the metric we are using for \( \rho \in H^{m+\alpha} \) controls \( L^\infty \) norm, if \( \delta > 0 \) is small enough then for any \( \|\rho - \rho_0\|_{H^{m+\alpha}} < \delta \) one obtains
\[
\rho(x) > \frac{1}{2} c_0.
\]
So, let us assume that \( Z \in C([0,T]; B_\delta(Z_0)) \). It is clear that \( \|e^{\varepsilon t \Delta} Z_0 - Z_0\| < \delta \) provided time \( t \) is short enough. The \( Z \) has some bound \( \|Z\| \leq C \). Using that let us estimate the norms under the integrals. First, recall that \( \| \Lambda_\alpha e^{\varepsilon t \Delta} \|_{L^2 \rightarrow L^2} \lesssim \frac{1}{(1 + t)^{n/2}} \). In the case \( \alpha \geq 1 \), we have
\[
\left\| \partial^m \Lambda_\alpha \int_0^t e^{\varepsilon (t-s) \Delta} \nabla \cdot (u\rho)(s) \, ds \right\|_2 \lesssim \int_0^t \frac{1}{(t-s)^{\alpha/2}} \| \partial^{m+1} (u\rho)(s) \|_2 \, ds \\
\leq \int_0^t \frac{1}{(t-s)^{\alpha/2}} \| u \|_{H^{m+1}} \| \rho \|_{H^{m+\alpha}} \, ds \lesssim C^2 t^{1-\alpha/2} < \frac{\delta}{2},
\]
provided \( T = T(\delta,\varepsilon) \) is small enough. In the case \( \alpha < 1 \), we combine instead one full derivatives with the heat semigroup, and the rest \( \partial^{m+\alpha} \) gets applied to \( u\rho \), which produces a similar bound.

Moving on to the \( u \)-equation, we have
\[
\left\| \partial^{m+1} \int_0^t e^{\varepsilon (t-s) \Delta} u \cdot \nabla u(s) \, ds \right\|_2 \lesssim \int_0^t \frac{1}{(t-s)^{1/2}} \| \partial^m (u \cdot \nabla u)(s) \|_2 \, ds \\
\leq \int_0^t \frac{1}{(t-s)^{\alpha/2}} \| u \|_{H^{m+1}} \| u \|_{H^m} \, ds \lesssim C^2 t^{1/2} < \frac{\delta}{4}.
\]
As to the commutator form, for \( \alpha \leq 1 \) the computation is very similar: we combine one derivative with the heat semigroup and for the rest we use (196):
\[
\| \partial^m \mathcal{C}_\phi(u,\rho) \|_2 \lesssim \| u \|_{m+\alpha} \| \rho \|_{m+\alpha} < C^2,
\]
and the rest follows as before. When \( \alpha > 1 \) we combine \( \alpha \) derivatives with the semigroup, and the rest follows as before.

We have proved that \( \|T[Z](t) - Z_0\| < \delta \), for a short time and hence, \( T \) leaves \( C([0,T(\delta,\varepsilon)); B_\delta(Z_0)) \) invariant.

Now let us make a priori estimates for viscous solutions independent of \( \varepsilon \). Note that the dissipation terms in all the following computations are negative and as such will be ignored.
First, evaluating the continuity equation at a point of minimum $x_-$ and denoting $\rho_- = \min \rho$ we readily obtain
\[
\frac{d}{dt}\rho_- = -\rho_- \nabla u + \varepsilon \Delta \rho(x_-) \geq -\rho_- |\nabla u|_\infty.
\]
Hence,
\[
\frac{d}{dt}|\rho^{-1}|_\infty \leq |\rho^{-1}|_\infty |\nabla u|_\infty \leq |\nabla u|_\infty Y_m.
\]
Furthermore,
\[
(205) \quad \frac{d}{dt}|e|_\infty \leq |\nabla u|_\infty |e|_\infty + |\nabla u|_\infty^2 \leq |\nabla u|_\infty Y_m.
\]
Let us continue with estimates on the $e$-quantity. We have (dropping integral signs)
\[
\frac{d}{dt}\|e\|^2_{H_m} \leq \partial^m e \cdot \nabla \partial^m e + \partial^m e [\partial^m (u \cdot \nabla e) - u \cdot \nabla \partial^m e] + \partial^m e \partial^m (e \cdot \nabla u) + \partial^m e [((\nabla \cdot u)^2 - \text{Tr}(\nabla u)^2]
\]
In the first term we integrate by part and estimate
\[
|\partial^m e \cdot \nabla \partial^m e| \leq \|e\|^2_{H_m} |\nabla u|_\infty.
\]
For the next commutator term we use (186)
\[
|\partial^m e [\partial^m (u \cdot \nabla e) - u \cdot \nabla \partial^m e]| \leq \|e\|^2_{H_m} |\nabla u|_\infty + \|e\|_{H_m} \|u\|_{H_m} |\nabla e|_\infty.
\]
Using Gagliardo-Nirenberg inequality we estimate the latter term as
\[
\|e\|_{H_m} \|u\|_{H_m} |\nabla e|_\infty \leq \|e\|^2_{H_m} |\nabla u|_\infty \leq \|e\|_{H_m} \|u\|_{H_m} |\nabla e|_\infty.
\]
where $\theta_1 = \frac{n-2(m-1)}{n-2m}$ and $\theta_2 = \frac{2}{n-2m}$. The two exponents add up to 1, so by the generalized Young inequality,
\[
\leq (\|e\|^2_{H_m} + \|u\|^2_{H_m+1}/(|e|_\infty + |\nabla u|_\infty) \leq (|e|_\infty + |\nabla u|_\infty)Y_m.
\]
Next term in the $e$-equation is estimated by the product formula
\[
(206) \quad \|\partial^m (fg)\|_2 \leq \|f\|_{H^m} g|_\infty + |f|_\infty \|g\|_{H^m}.
\]
So, we have
\[
|\partial^m e \partial^m (e \nabla u)| \leq \|e\|^2_{H_m} |\nabla u|_\infty + \|e\|_{H_m} \|e\|_{H_m+1} \|\nabla u|_{H_m+1} \leq (|e|_\infty + |\nabla u|_\infty)Y_m.
\]
Finally,
\[
|\partial^m e [(\nabla \cdot u)^2 - \text{Tr}(\nabla u)^2]| \leq \|e\|_{H_m} \|u\|_{H_m+1} |\nabla u|_\infty \leq |\nabla u|_\infty Y_m.
\]
Thus,
\[
(207) \quad \frac{d}{dt}\|e\|^2_{H_m} \leq (|e|_\infty + |\nabla u|_\infty)Y_m.
\]
Next perform the main technical estimate on the velocity equation. We have
\[
\partial_t \|u\|^2_{H_{m+1}} = -\partial^{m+1}(u \cdot \nabla u) \cdot \partial^{m+1} u + \partial^{m+1} C_{\phi}(u, \rho) \cdot \partial^{m+1} u.
\]
The transport term is estimated using the classical commutator estimate
\[
\partial^{m+1}(u \cdot \nabla u) \cdot \partial^{m+1} u = \nabla [(\partial^{m+1} u) \cdot \partial^{m+1} u]
\]
Then
\[
u \cdot (\nabla [\partial^{m+1} u] \cdot \partial^{m+1} u) = -\frac{1}{2} (\nabla \cdot u) |\partial^{m+1} u|^2 \leq |\nabla u|_\infty \|u\|^2_{H_{m+1}},
\]
and using (186) we obtain
\[
[(\partial^{m+1}, u) |\nabla u \cdot \partial^{m+1} u| \leq |\nabla u|_\infty \|u\|^2_{H_{m+1}}.
\]
Thus,
\[
\partial_t \|u\|^2_{H_{m+1}} \leq |\nabla u|_\infty Y_m + \partial^{m+1} C_{\phi}(u, \rho) \cdot \partial^{m+1} u.
\]
Let us expand the commutator
\[ \partial^{m+1} C_\phi(u, \rho) = \sum_{l=0}^{m+1} \binom{m+1}{l} C_\phi(\partial^l u, \partial^{m+1-l} \rho). \]

One end-point case, \( l = m + 1 \), gives rise to a dissipative term:
\[ \int_{\mathbb{R}^n} C_\phi(\partial^{m+1} u, \rho) \cdot \partial^{m+1} u \, dx = -\frac{1}{2} \int_{\mathbb{R}^2n} \phi(z) |\delta_z \partial^{m+1} u(x)|^2 \, dz \, dx \\
- \frac{1}{2} \int_{\mathbb{R}^2n} \phi(z) \delta_z \partial^{m+1} u(x) \delta_z \rho(x) \, dz \, dx. \]

The first term is bounded by
\[ -\rho_- \int_{\mathbb{R}^2n} \phi(z) |\delta_z \partial^{m+1} u(x)|^2 \, dz \, dx \sim -\rho_- \|u\|_{H^{m+1+\frac{\alpha}{2}}}^2, \]
which is the main dissipation term. The second is estimated as follows. Let us pick an \( \varepsilon > 0 \) so small that \( 1 + \frac{\alpha}{2} > \alpha + \varepsilon \). Then
\[ \left| \int_{\mathbb{R}^2n} \phi(z) \delta_z \partial^{m+1} u(x) \partial^{m+1} u(x) \delta_z \rho(x) \, dz \, dx \right| \leq |\nabla \rho|_{\infty} \int_{\mathbb{R}^2n} |\partial^{m+1} \delta_z u(x)| |\partial^{m+1} u(x)| \, dz \, dx \\
\leq |\nabla \rho|_{\infty} \|u\|_{H^{m+\alpha+\varepsilon}} \leq |\nabla \rho|_{\infty} \|u\|_{H^{m+\alpha}} \|u\|_{H^{m+1+\frac{\alpha}{2}}} \leq \frac{1}{2} \rho_- \|u\|_{H^{m+1+\frac{\alpha}{2}}}^2 + \rho_- |\nabla \rho|_{\infty}^2 Y_m, \]
where the first term is absorbed into dissipation. So,
\[ \int_{\mathbb{R}^2n} C_\phi(\partial^{m+1} u, \rho) \cdot \partial^{m+1} u \, dx \lesssim -\rho_- \|u\|_{H^{m+1+\frac{\alpha}{2}}}^2 + \rho_- |\nabla \rho|_{\infty}^2 Y_m. \]

Let us consider first the other end-point case of \( l = 0 \). In this case the density suffers a derivative overload. We apply the following “easing” technique:
\[ \int_{\mathbb{R}^n} C_\phi(u, \partial^{m+1} \rho) \cdot \partial^{m+1} u \, dx = \int_{\mathbb{R}^2n} \phi(z) \delta_z u(x) \partial^{m+1} \rho(x + z) \partial^{m+1} u(x) \, dz \, dx. \]

Observe that
\[ \partial^{m+1} \rho(x + z) = \partial_z \partial_x^m \rho(x + z) = \partial_z (\partial_x^m \rho(x + z) - \partial_x^m \rho(x)) = \partial_z \partial_x^m \rho(x). \]

Let us now integrate by parts in \( z \):
\[ \int_{\mathbb{R}^n} C_\phi(u, \partial^{m+1} \rho) \cdot \partial^{m+1} u \, dx = \int_{\mathbb{R}^2n} \partial_z \phi(z) \delta_z u(x) \delta_z \partial_x^m \rho(x) \partial^{m+1} u(x) \, dz \, dx + \\
+ \int_{\mathbb{R}^2n} \phi(z) \partial u(x + z) \delta_z \partial_x^m \rho(x) \partial^{m+1} u(x) \, dz \, dx := J_1 + J_2. \]

Let us start with the \( J_2 \) first. By symmetrization,
\[ J_2 = \int_{\mathbb{R}^2n} \delta_z \partial u(x) \delta_z \partial_x^m \rho(x) \partial^{m+1} u(x) \phi(z) \, dz \, dx - \int_{\mathbb{R}^2n} \partial u(x) \delta_z \partial_x^m \rho(x) \delta_z \partial^{m+1} u(x) \phi(z) \, dz \, dx := J_{2,1} + J_{2,2}. \]

Term \( J_{2,1} \) will appear in a series of similar terms that we will estimate systematically below. The bound for \( J_{2,2} \) is rather elementary:
\[ J_{2,2} \leq |\nabla u|_{\infty} \|u\|_{H^{m+1+\alpha}/2} + \|\rho\|_{H^{m+\alpha}/2} \leq \varepsilon \rho_- \|u\|_{H^{m+1+\alpha}/2}^2 + \rho_- |\nabla u|_{\infty}^2 Y_m. \]

Similar computation can be made for \( J_1 \). Indeed, using that \( \partial_z \phi(z) \) is odd, by symmetrization, we have
\[ J_1 = \frac{1}{2} \int_{\mathbb{R}^2n} \partial_z \phi(z) \delta_z u(x) \delta_z \partial_x^m \rho(x) \delta_z \partial^{m+1} u(x) \, dz \, dx. \]
Replacing \(|\delta \mathbf{u}(x)| \leq |z| |\mathbf{u}|_\infty\), the rest of the term is estimated exactly as \(J_{2,2}\).

To summarize, we have obtained the bound
\[
\int_{\mathbb{T}^n} C_\phi(\mathbf{u}, \partial^{m+1} \rho) \cdot \partial^{m+1} \mathbf{u} \, dx \leq \varepsilon \rho_0 \||\mathbf{u}|^2_{H^{m+1+\alpha/2}} + \rho_0^{-1} |\nabla \mathbf{u}|^2_Y Y_m.
\]

Let us now examine the rest of the commutators \(C_\phi(\partial^l \mathbf{u}, \partial^{m+1-l} \rho)\) for \(l = 1, \ldots, m\). After symmetrization we obtain
\[
\int_{\mathbb{T}^n} C_\phi(\partial^l \mathbf{u}, \partial^{m+1-l} \rho) \cdot \partial^{m+1} \mathbf{u} \, dx = \frac{1}{2} \int_{\mathbb{T}^n} \delta_\phi \partial^l \mathbf{u}(x) \delta_\phi \partial^{m+1-l} \rho(x) \partial^{m+1} \mathbf{u}(x) \phi(z) \, dz \, dx + \frac{1}{2} \int_{\mathbb{T}^n} \delta_\phi \partial^l \mathbf{u}(x) \partial^{m+1-l} \rho(x) \delta_\phi \partial^{m+1} \mathbf{u}(x) \phi(z) \, dz \, dx := J_1 + J_2.
\]

Estimates on the new terms, \(J_1, J_2\) are a little more sophisticated as we seek to optimize distribution of \(L^p\)-norms inside their components. Notice that the case \(l = 1\) corresponds to the previously appeared term \(J_{2,1}\).

So, let us assume that \(l = 1, \ldots, m\). We will distribute the parameters in \(J_1\) as follows
\[
J_1 = \int_{\mathbb{T}^n} \frac{\delta_\phi \partial^l \mathbf{u}(x)}{|z|^{\frac{q}{2} + \frac{\alpha}{2} + 2\delta}} \frac{\delta_\phi \partial^{m+1-l} \rho(x)}{|z|^{\frac{q}{2} + \frac{\alpha}{2} - \frac{n}{p}}} \frac{\partial^{m+1} \mathbf{u}(x)}{|z|^{\frac{q}{2} - \delta}} \frac{1}{|z|^{\frac{q}{2} - \delta}} \, dz \, dx,
\]
where \(\delta > 0\) is a small parameter to be determined later, and \((2, p, q, r)\) is a Hölder quadruple defined by
\[
p = 2 \frac{m + \frac{q}{2}}{l - 1 + \frac{\alpha}{2}}, \quad q = 2 \frac{m + \alpha - 1}{m - l + \frac{\alpha}{2}}, \quad r = 1 - \frac{1}{2} - \frac{1}{p} - \frac{1}{q}.
\]

The existence of finite \(r\) is warranted by the strict inequality which is verified directly:
\[
\frac{1}{2} + \frac{1}{p} + \frac{1}{q} < 1.
\]

By the Hölder inequality,
\[
J_1 \leq \||\mathbf{u}|_{H^{l+\frac{q}{2} + 2\delta, p}}\| \||\rho|_{W^{m+1-l+\frac{q}{2}, q}}\||\mathbf{u}|_{H^{m+1}}.
\]

Let us apply the following Gagliardo-Nirenberg inequalities to all the terms
\[
\||\mathbf{u}|_{H^{m+1}} \leq \||\mathbf{u}|_{H^{m+1+\alpha/2}}^{2m} \||\nabla \mathbf{u}|_{H^{m+1+\alpha/2}}^{\alpha} \leq \||\mathbf{u}|_{H^{m+1+\alpha/2}}^{2m} \||\nabla \mathbf{u}|_{\infty}^{\alpha / 2m},
\]
\[
\||\mathbf{u}|_{H^{l+\frac{q}{2} + 2\delta, p}} \leq \||\mathbf{u}|_{H^{m+1+\alpha/2}}^{\theta_1} \||\nabla \mathbf{u}|_{\infty}^{1 - \theta_1},
\]
\[
\||\rho|_{H^{m+1-l+\frac{q}{2}, q}} \leq \||\rho|_{H^{m+1+\alpha/2}}^{\theta_2} \||\nabla \rho|_{\infty}^{1 - \theta_2},
\]
where
\[
\theta_1 = \frac{l - 1 + \frac{\alpha}{2} - \frac{n}{p} + 2\delta}{m + \frac{\alpha}{2} - \frac{n}{p}}, \quad \theta_2 = \frac{m - l + \frac{\alpha}{2} - \frac{n}{q}}{m + \alpha - 1 - \frac{n}{q}}.
\]

The exponents satisfy the necessary requirements
\[
1 \geq \theta_1 \geq \frac{l - 1 + \frac{\alpha}{2} + 2\delta}{m + \frac{\alpha}{2}}, \quad 1 \geq \theta_2 = \frac{m - l + \frac{\alpha}{2}}{m + \alpha - 1},
\]
and in fact,
\[
\theta_1 = \frac{l - 1 + \frac{\alpha}{2}}{m + \frac{\alpha}{2}} + O(\delta).
\]

Now, we have
\[
J_1 \leq \||\mathbf{u}|_{H^{l+\frac{q}{2} + \theta_1, 2}}^{2m+\theta_1} \||\rho|_{H^{m+\alpha}}^{\theta_2} \||\nabla \mathbf{u}|_{\infty}^{\frac{\alpha}{2m+\alpha} + 1 - \theta_1} \||\nabla \rho|_{\infty}^{1 - \theta_2}.
\]
By generalized Young,
\[ J_1 \leq \varepsilon \rho_- \| u \|_{H^{m+\alpha/2}}^2 + \rho_-^{-1} \| \rho \|_{H^{m+\alpha}}^{\theta_2 Q} (|\nabla u|_\infty^{\alpha m + \alpha + 1 - \theta_1} |\nabla \rho|_\infty^{1 - \theta_2})^Q, \]
where \( Q \) is the conjugate to \( \frac{2m}{2m + \alpha} + \theta_1 \). We have \( \theta_2 Q < 2 \) as long as
\[ \theta_1 + \theta_2 < 2 - \frac{2m}{2m + \alpha}. \]

We in fact have even stronger inequality, \( \theta_1 + \theta_2 < 1 \) provided \( \delta \) is small enough. So, we arrived at
\[ J_1 \leq \varepsilon \rho_- \| u \|_{H^{m+\alpha/2}}^2 + \rho_-^{-1} p_N (|\nabla \rho|_\infty, |\nabla u|_\infty) Y_m, \]
for some polynomial \( p_N \).

Finally, moving on to \( J_2 \), we distribute the exponents as follows
\[ J_2 \leq \int_{\mathbb{T}^n} \left| \frac{\delta_x \partial_t u(x)}{|x|^{\frac{p}{2} + \frac{\alpha}{2} + \frac{\theta_1}{2}}} \right| \frac{\left| \delta_x \partial_t^m \rho(x) \right|}{|x|^{\frac{p}{2} - \theta_1}} \frac{1}{|x|^{\frac{\alpha}{2} + \frac{\theta_2}{2}}} \, dx \leq \| u \|_{W^{l+\delta, \frac{p}{2} + \frac{\theta_1}{2}}} \| \rho \|_{W^{m+\alpha, \frac{p}{2} + \theta_2}} \| u \|_{H^{m+\alpha/2}}. \]

Here we choose \((r, p, q, \delta)\) as follows
\[ q = 2 \frac{m + \alpha - 1}{m - l}, \quad p = 2 \frac{m + \frac{\alpha}{2}}{l - 1 + \frac{\alpha}{2}}, \quad \frac{1}{r} = 1 - \frac{1}{2} - \frac{1}{p} - \frac{1}{q}, \]
and \( \delta \) is small. With these choices we proceed with the Gagliardo-Nirenberg inequalities
\[ \| u \|_{W^{l+\delta, \frac{p}{2} + \frac{\theta_1}{2}}} \leq \| u \|_{H^{m+\alpha/2}}^{\theta_1} |\nabla u|_\infty^{1 - \theta_1}, \]
\[ \| \rho \|_{W^{m+\alpha, \frac{p}{2} + \theta_2}} \leq \| \rho \|_{H^{m+\alpha}}^{\theta_2} |\nabla \rho|_\infty^{1 - \theta_2}, \]
where
\[ \theta_1 = \frac{l - 1 + \frac{\alpha}{2} + 2\delta}{m + \frac{\alpha}{2} - \frac{\theta_1}{2}} = \frac{l - 1 + \frac{\alpha}{2} + O(\delta)}{m + \frac{\alpha}{2}} + O(\delta), \quad \theta_2 = \frac{m - l}{m + \alpha - 1}. \]

Now to achieve the bound
\[ J_2 \leq \varepsilon \rho_- \| u \|_{H^{m+\alpha/2}}^2 + \rho_-^{-1} p_N (|\nabla \rho|_\infty, |\nabla u|_\infty) Y_m, \]
we have to make sure that \( \theta_1 + \theta_2 \leq 1 \), which is true for small \( \delta \).

We have proved the following a priori bound on \( u \):
\[ \partial_t \| u \|_{H^{m+1}}^2 \leq -\frac{1}{2} \rho_- \| u \|_{H^{m+\alpha/2}}^2 + \rho_-^{-1} p_N (|\nabla \rho|_\infty, |\nabla u|_\infty) Y_m. \]

Together with the previously established bounds we obtain
\[ \frac{d}{dt} Y_m \leq -\frac{1}{2} \rho_- \| u \|_{H^{m+\alpha/2}}^2 + \rho_-^{-1} p_N (|\nabla \rho|_\infty, |\nabla u|_\infty, |e|_\infty) Y_m. \]

This of course implies a Riccati inequality, provided \( m > \frac{\alpha}{2} + 1 \):
\[ \frac{d}{dt} Y_m \leq CY_m^N, \]
and provides a priori bound independent of the viscosity coefficient. Thus, we can extend it to an interval independent of \( \varepsilon \) as well. By the compactness argument similar to the smooth kernel case, we obtain a local solution in the same class as initial data and \( u \in L^2 H^{m+\alpha/2} \). In addition, we obtain a continuation criterion – as long as \( |\nabla \rho|_\infty, |\nabla u|_\infty, |e|_\infty \) remain bounded on \([0, T_0])\) the solutions can be extended beyond \( T_0 \). However everything is reduced to a control over the first two quantities, because \( |e|_\infty \) remains bounded as long as \( |\nabla u|_\infty \) is in view of (205).
Theorem 5.3 (Local existence of classical solutions). Suppose \(m > \frac{n}{2} + 1\), \(0 < \alpha < 2\), and
\[
(u_0, \rho_0) \in H^{m+1}(\mathbb{T}^n) \times H^{m+\alpha}(\mathbb{T}^n),
\]
and \(\rho_0(x) > 0\) for all \(x \in \mathbb{T}^n\). Then there exists time \(T_0 > 0\) and a unique non-vacuous solution to (195) on time interval \([0, T_0]\) in the class
\[
(209) \quad u \in C_w([0, T_0); H^{m+1}(\mathbb{T}^n)) \cap L^2([0, T_0); \dot{H}^{m+1+\alpha/2}), \quad \rho \in C_w([0, T_0); H^{m+\alpha}).
\]
Moreover, any such solution satisfying
\[
(210) \quad \sup_{t \in [0, T_0)} (|\nabla \rho(t)|_\infty + |\nabla u(t)|_\infty) < \infty
\]
can be extended beyond \(T_0\).

It is clear from the proof that (210) can be replaced with an integrability condition with some high power depending on \(m, n, \alpha\).

6. One-dimensional theory

In dimension 1, most complete regularity theory of alignment models is available. In this chapter we discuss some of its highlights.

In 1D the system is given by
\[
(211) \begin{cases}
\rho_t + (\rho u)_x = 0, \\
u_t + uu_x = \int_\Omega \phi(x-y)(u(y) - u(x))\rho(y) \, dy
\end{cases} \quad (x,t) \in \Omega \times \mathbb{R}_+.
\]
Here \(\Omega\) is either \(\mathbb{R}\) or periodic circle \(\mathbb{T}\). The underlying theme in this chapter will be to relate regularity and flocking behavior of the system to the new conserved quantity which is available in dimension 1 and a few other exceptional multi-D cases:
\[
(212) \quad e = u_x + L_\phi \rho,
\]
where \(L_\phi\) takes form of either of the integral representations (141) or (142) depending on the context.

6.1. Smooth kernels: threshold for GWP and stability. The system (211) possesses an extra conservation law provided \(\phi\) is a convolution kernel, \(\phi = \phi(x - y)\). To see this we note that the alignment term in the velocity equation is given by the commutator
\[
C_\phi(u, \rho) = \phi * (u \rho) - u \phi * \rho.
\]
Using this commutator structure and by elementary manipulation with the equations we obtain
\[
(213) \quad e_t + (ue)_x = 0, \quad e = u_x + \phi * \rho.
\]
This quantity, whose physical role will be illuminated later in Section 6.2, plays a crucial role in regularity theory. Note that the Burgers’ part of the velocity equation tends to create shocks, while the alignment part would counterforce it to smooth out the solution. One would expect then that a threshold condition would separate singular behavior from regular. For pure Burgers equation such condition is provided by positive initial slope \(u_x \geq 0\). In view of the regularizing effect of alignment such condition can be relaxed to \(e_0 \geq \frac{1}{2}\). To see that let us rewrite the \(e\)-equation as a non-autonomous logistic ODE along characteristics:
\[
(214) \quad \dot{e} = e(\phi * \rho - e)
\]
It is clear that the sign of \(e\) will be preserved pointwise. So, if \(e_0(x_0) < 0\) at some point \(x_0\), then \(\dot{e} < -e^2\), and consequently the solution blows up in finite time. On the other hand, supposing \(e_0 \geq \frac{1}{2}\) we have \(e(t) \geq \frac{1}{2}\) for all times, and \(e\) remains a priori bounded since \(\phi * \rho \leq CM\). This in
particular implies that $u_x \in L^\infty_t L^1_x$. In view of Theorem 5.1, this ensures global existence of solutions. Moreover, we can also obtain a hydrodynamic version of strong flocking. Let us make this precise.

**Theorem 6.1** (Threshold for global existence). Consider the system (211) on $\mathbb{R}$ or $\mathbb{T}$ with smooth kernel. For any initial condition $(u_0, \rho_0) \in H^m \times (L^1_x \cap W^{1,\infty})$, $m > \frac{n}{2} + 2$, which satisfies the threshold condition $e_0 \geq 0$ there exists a unique global solution $(u, \rho) \in C_w([0, \infty); H^m \times (L^1_x \cap W^{1,\infty})$. If $e_0(x) < 0$ at some point, then the solution blows up in finite time.

Furthermore, suppose that $\phi$ has a fat tail. Let the initial flock has compact support $\text{Supp} \rho_0$ and $e_0 \geq 0$. Then the solution flock strongly in the following sense: there exists $C, \delta > 0$ depending on $\phi$, and initial data, such that the velocity satisfies

\begin{equation}
\sup_{x \in \text{Supp} \rho(t)} |u(x, t) - \bar{u}| + |u_x(x, t)| + |u_{xx}(x, t)| \leq Ce^{-\delta t}.
\end{equation}

and the density $\rho$ converges to a traveling wave $\bar{\rho}$ in the metric of $C^\gamma$ for any $0 < \gamma < 1$:

\begin{equation}
\|\rho(\cdot, t) - \bar{\rho}(\cdot - \bar{u}t)\|_{C^\gamma} \leq Ce^{-\delta t}, \quad t > 0.
\end{equation}

**Proof.** The global existence has already been proved in the preceding remarks.

Let us now assume that we have a global solution with $e_0 \geq 0$ and $\phi$ has a fat tail. We know from Theorem 4.1 that the diameter of the flock will remain finite, $\mathcal{D}$. Then we can estimate the convolution from below: for any $x \in \text{Supp} \rho(t)$:

\[ \phi \ast \rho(t, x) \geq \phi(\mathcal{D})M := c_0. \]

Solving the logistic ODE: $\dot{e} \geq c_0(1 - e)$ we find that starting from some time $t_0$ for all $t > t_0$ and $x \in \text{Supp} \rho(t)$ we have $e(x, t) \geq c_0/2$. With this in mind let us write the equation for $u_x$ as follows

\begin{equation}
\frac{D}{Dt} u_x = \int_{\mathbb{R}} \phi'(x - y)(u(y) - u(x))\rho(y) \, dy - e(x)u_x(x).
\end{equation}

We already know from Theorem 4.1 that the velocity fluctuations are decaying with exponential rate. Hence, the integral above will be bounded by $|\phi'|_\infty ME(t)$, where we denote by $E$ any quantity that shows exponential decay. Thus, multiplying (217) with $u_x$ and evaluating at the maximum over $\text{Supp} \rho_0$ we obtain

\begin{equation}
\frac{d}{dt}\|u_x\|_{L^\infty(\text{Supp} \rho(t))} = E(t) - \frac{1}{2}c_0\|u_x\|_{L^\infty(\text{Supp} \rho(t))}.
\end{equation}

This implies the desired result by integration. In view of (190) the density enjoys a pointwise global bound

\begin{equation}
\sup_{t > 0} \|\rho(\cdot, t)\|_{\infty} = R < \infty.
\end{equation}

For the second derivative, we have

\begin{equation}
\frac{D}{Dt} u_{xx} + 2u_x u_{xx} = \int_{\mathbb{R}} \phi''(x - y)(u(y) - u(x))\rho(y) \, dy - 2u_x \phi' \ast \rho - eu_{xx}.
\end{equation}

We see that the integral term as well as $u_{xx}$ is of type $E(t)$ in view of the previously established bounds. Note also that $|\phi_x \ast \rho| \leq |\phi_x|_\infty M$. So, we obtain

\begin{equation}
\frac{d}{dt}\|u_{xx}\|_{L^\infty(\text{Supp} \rho(t))} = E(t) - \frac{1}{2}c_0\|u_{xx}\|_{L^\infty(\text{Supp} \rho(t))}.
\end{equation}

This implies exponential decay. Moving to the density, we have

\begin{equation}
\partial_t \rho_x + u \rho_{xx} = -2u_x \rho_x - u_{xx} \rho = E \rho_x + E.
\end{equation}

This shows that $\rho_x$ remains uniformly bounded. Now, to establish strong flocking we pass to the moving frame $x - \bar{u}t$ and write the continuity equation in new coordinates

\begin{equation}
\frac{D}{Dt} \rho = -(u - \bar{u}) \rho_x - u_x \rho = E.
\end{equation}
This shows that $\rho(t)$ is Cauchy in $t$ in the metric of $L^\infty$. Hence, there exists $\bar{\rho} \in L^\infty$ such that $\|\rho(t) - \bar{\rho}\|_\infty = E(t)$. Since $\rho'$ is uniformly bounded, this also shows that $\bar{\rho}$ is Lipschitz. Convergence in $C^\gamma$, $\gamma < 1$, follows by interpolation.

Using stability estimate (159) we can easily conclude stability of the limiting flock distributions in the sense of KR-distance. Indeed, if $\rho \to \bar{\rho}$ in $C^\gamma$ then certainly $d(\rho, \bar{\rho}) \to 0$. So, if we start from two flocks with the same mass and momentum and bounded supports, then using (159) we obtain (by translation invariance of the KR-metric)

$$d(\rho', \bar{\rho}'') \leq C[d(\rho_0', \rho_0'') + \|u_0' - u_0''\|_{L^\infty(\Omega)}],$$

where $\Omega = \operatorname{conv}\{\operatorname{Supp} \rho_0', \operatorname{Supp} \rho_0''\}$. This establishes a direct stability control of limiting flocks with respect to initial perturbation. To bootstrap stability estimate to higher regularity class, let us note that in the course of proof of Theorem 6.1 we established global bound on $\rho_x$ by a constant depending on initial data. Thus, $\bar{\rho}$ will remain in $W^{1,\infty}$ with similar bound. Interpolating between $W^{-1,\infty}$ (which is equivalent to KR-distance, see Remark 4.2) and $W^{1,\infty}$ gives a bound in H"older class $C^\gamma$, $\gamma < 1$:

$$\|\rho' - \rho''\|_{C^\gamma} \lesssim [d(\rho_0', \rho_0'') + \|u_0' - u_0''\|_{L^\infty(\Omega)}]^{\frac{1+\gamma}{2}}.$$


Although in the case of symmetric communication the limiting velocity is determined from the initial condition by use of momentum conservation, the limiting shape of the density profile $\bar{\rho}$ is an emergent quantity which is not known a priori. Yet it is easy to notice that the $e$-quantity is somehow involved in determining $\bar{\rho}$. Let us assume in this section that $e = u_x + \mathcal{L}_\phi \rho$, where $\mathcal{L}_\phi$ is defined by (142) in either smooth or singular case. Let us write the continuity equation with the use of $e$:

$$\rho_t + u\rho_x + e\rho = \rho \mathcal{L}_\phi \rho.$$

Let us assume for a moment that $\phi$ is a smooth absolute kernel on $T$. Suppose that $\epsilon_0 = 0$, and hence $e(t) = 0$ for all time. Note that in this case $u_x + \phi * \rho \geq \rho(x)|\phi|_1 \geq 0$, so the global solution exists. Then

$$\rho_t + u\rho_x = \rho \mathcal{L}_\phi \rho,$$

and so, $\rho$ obeys the maximum principle. Let us assume that there is no vacuum $\rho_-(0) > 0$. Let us then write the equation for the new quantity $\ln \rho$:

$$(\ln \rho)_t + u(\ln \rho)_x = \mathcal{L}_\phi \rho.$$

Evaluating at a point of minimum we obtain

$$(\ln \rho_-)_t = \int \phi(x, y)(\rho(y) - \rho_-) \, dy \geq c_0(M - 2\pi \rho_-),$$

and at the maximum

$$(\ln \rho_+)_t = \int \phi(x, y)(\rho(y) - \rho_+) \, dy \leq c_0(M - 2\pi \rho_+).$$

Subtracting the two we obtain

$$\frac{d}{dt} \ln \frac{\rho_+}{\rho_-} \leq -2\pi c_0(\rho_+ - \rho_-) \leq -2\pi c_0 \rho_- (0) \left( \frac{\rho_+}{\rho_-} - 1 \right) \leq -2\pi c_0 \rho_- (0) \ln \frac{\rho_+}{\rho_-}.$$

We conclude that

$$\ln \frac{\rho_+}{\rho_-} \leq c_1 e^{-c_2 t}.$$

Since the maximum also stays bounded we have inequality

$$\ln \frac{\rho_+}{\rho_-} \geq c \left( \frac{\rho_+}{\rho_-} - 1 \right),$$

where $c > 0$ and $c_1, c_2, c_3$ are constants depending on $e_0$.
So, we conclude that the density flattens out exponentially fast to a uniformly distributed state \( \hat{\rho} = \frac{1}{2\pi} M \).

In view of this computation we can see that \( e \) is directly responsible for the flattening of the density. It turns out that, first, a similar result is true for local kernels and even vacuous solutions. And second, in general the size of \( e \) per mass, i.e. the quotient \( q = \frac{e}{\hat{\rho}} \), measures how far \( \hat{\rho} \) is from the uniform distribution. Thus, \( e \) plays the role of a topological entropy of the flock – a measure of disorder. We will address this interpretation in the next theorem.

**Theorem 6.2.** Let \((\rho, u)\) be a smooth solution to the system (211) on the 1D torus \( \mathbb{T} \), and \( \phi \) is a smooth local kernel:

\[
\phi(r) \geq \lambda \mathbf{1}_{r < r_0}.
\]

If \( e_0 = 0 \), then

\[
\|\rho(t) - \bar{\rho}\|_1 \leq c_1(\|\rho_0\|_2)e^{-c_2(\lambda, r_0, M, \|\rho_0\|_\infty)t},
\]

where \( \bar{\rho} = \frac{1}{2\pi} M \).

In general, provided \( \|q_0\|_\infty < \|\phi\|_1 \), one has

\[
\limsup_{t \to \infty} \|\rho(:, t) - \bar{\rho}\|_1 \leq \frac{M\|q_0\|_\infty\|\phi\|_\infty}{\lambda c(r_0)(\|\phi\|_1 - \|q_0\|_\infty)}.
\]

Let us note that the dependence on \( \|q_0\|_\infty \) is linear for small values. At the same time, the bound is inversely proportional to the strength \( \lambda \), which shows the stabilizing effect of communication on the structure of the flock. Let us note that \( q \) satisfies the transport equation

\[
d\frac{d}{dt} q + u q_x = 0, \quad q = \frac{e}{\rho},
\]

and hence the value of \( \|q\|_{L^\infty} \) is preserved for all time.

The proof will be split in the following few subsections. First, let us recall a very useful tool – The Csiszár-Kullback inequality.

### 6.2.1. The Csiszár-Kullback inequality

Let us consider two functions \( f \geq 0, g > 0 \) on a measure space \((\Omega, \Sigma, \mu)\), such that

\[
\int_\Omega f(x) \, d\mu(x) = \int_\Omega g(x) \, d\mu(x) = 0.
\]

Then

\[
\int_\Omega |f - g|^2 \frac{d\mu}{g} \geq \int_\Omega f \log \frac{f}{g} \, d\mu \geq \frac{1}{8} \|f - g\|_{L^1}^2.
\]

**Proof.** Let us start from an elementary inequality:

\[x(x - 1) \geq x \log x \geq (x - 1) + \frac{1}{2}(x - 1)^2 \mathbf{1}_{\{x < 1\}}.\]

The upper inequality is elementary:

\[
\int_\Omega f \log \frac{f}{g} \, d\mu \leq \int_\Omega f(\frac{f}{g} - 1) \, d\mu = \int_\Omega (f - g)(\frac{f}{g} - 1) \, d\mu = \int_\Omega |f - g|^2 \frac{d\mu}{g}.
\]

Let us prove the lower inequality. On the one end,

\[
\int_\Omega f \log \frac{f}{g} \, d\mu(x) \geq \int_\Omega (f - g) \, d\mu(x) + \frac{1}{2} \int_{f < g} g(f/g - 1)^2 \, d\mu(x),
\]

and on the other end,

\[
\|f - g\|_{L^1} = \int_{f < g} (g - f) \, d\mu + \int_{f \leq g} (f - g) \, d\mu = \int_{f < g} (g - f) \, d\mu - \int_{f < g} (f - g) \, d\mu = 2 \int_{f < g} (g - f) \, d\mu.
\]
Considering $g \, d\mu$ as a probability measure, we use the Hölder inequality:
\[
\int_{f<g} (g - f) \, d\mu = \int_{f<g} (1 - f/g) g \, d\mu \leq \left( \int_{f<g} |1 - f/g|^2 g \, d\mu \right)^{\frac{1}{2}}.
\]
Connecting the two ends produces lower inequality in (229).

The object to our study will be the relative entropy defined by
\[
\mathcal{H} = \int_{\mathbb{T}} \rho \log \frac{\rho}{\bar{\rho}} \, dx = \int_{\mathbb{T}} \rho \log \rho \, dx - M \log \bar{\rho},
\]
By rescaling the Csiszár-Kullback inequality applied to $f = \rho/M$, $g = 1/2\pi$ on $\mathbb{T}$ we obtain
\[
\frac{1}{16\pi} \|\rho - \bar{\rho}\|_{L^1}^2 \leq \mathcal{H} \leq \|\rho - \bar{\rho}\|_{L^2}^2.
\]

### 6.2.2. Evolution of the entropy.

At the heart of the argument is the equation for the entropy (230) which one obtains testing the continuity equation with $\log \rho + 1$:
\[
(\rho \log \rho)_t = \rho_t (\log \rho + 1) = - (\rho u)' (\log \rho + 1)
\]
\[
= - \rho' (\log \rho + 1) u - \rho u' (\log \rho + 1)
\]
\[
= - (\rho \log \rho)' u - (\rho \log \rho) u' - \rho u'
\]
\[
= - [u(\rho \log \rho)]' - \rho u' = - [u(\rho \log \rho)]' - \rho^2 q + \mathcal{L}_\psi \rho.
\]
Therefore,
\[
\frac{d \mathcal{H}}{dt} = \frac{d}{dt} \int_{\mathbb{T}} \rho \log \rho \, dx = - \int_{\mathbb{T}} \rho^2 q \, dx - \int_{\mathbb{T}^2} \phi(x-y)(\rho(x) - \rho(y)) \rho(x) \, dx \, dy.
\]
Noting that $\int_{\mathbb{T}} \rho q \, dx = \int_{\mathbb{T}} \epsilon \, dx = 0$, we can subtract $\bar{\rho}$ from one density in the first integral on the left hand side. After additionally symmetrizing the last integral we obtain
\[
\frac{d \mathcal{H}}{dt} = - \int_{\mathbb{T}} (\rho - \bar{\rho}) \rho q \, dx - \frac{1}{2} \int_{\mathbb{T}^2} \phi(x-y) |\rho(x) - \rho(y)|^2 \, dx \, dy.
\]

### 6.2.3. Bounds on the dissipation.

If our kernel was absolute, it would be easy to get a positive lower bound on the dissipation term:
\[
\int_{\mathbb{T}^2} \phi(x-y)|\rho(x) - \rho(y)|^2 \, dx \, dy \geq (\inf \phi) \int_{\mathbb{T}^2} |\rho(x) - \rho(y)|^2 \, dx \, dy = 2(\inf \phi) \|\rho - \bar{\rho}\|_{L^2}^2.
\]
Since we have a non-trivial lower bound on the kernel only near the diagonal $\{(x,y) \in \mathbb{T}^2 : |x-y| < r_0\}$, we need a substitute for (234) stated in the following Lemma.

**Lemma 6.3.** The following inequality holds:
\[
\int_{|x-y|<r_0} |\rho(x) - \rho(y)|^2 \, dy \, dx \geq c(r_0) \|\rho - \bar{\rho}\|_{L^2}^2.
\]

**Proof.** Denote by $\chi$ be any nonnegative bump function supported on $B_{r_0}(0)$, constant on $B_{r_0/2}$ and $\int \chi(r) \, dr = 1$. Then on the Fourier side, $\hat{\chi}(0) = 1$ and $|\hat{\chi}(k)| < 1$ for all $k \in \mathbb{Z}\setminus\{0\}$. On the other hand, by the Riemann-Lebesgue Lemma, $\hat{\chi}(k) \to 0$ as $k \to \infty$. Therefore, $|\hat{\chi}(k)| \leq 1 - \epsilon$ for some $\epsilon > 0$ depending only on $r_0$ ($k \neq 0$). Define $\tilde{\rho}_{r_0}(x) = \chi * \rho(x)$, so that
\[
(\rho - \tilde{\rho}_{r_0}) \hat{\chi}(k) = (1 - \hat{\chi}(k)) \hat{\rho}(k).
\]
Hence,
\[
|\rho - \tilde{\rho}_{r_0}|(k) \geq \epsilon \hat{\rho}(k), \quad k \in \mathbb{Z}, \ k \neq 0,
\]
and \( \bar{\rho}(0) = \tilde{\rho}_{r_0}(0) \). Consequently,
\[
\|\rho - \bar{\rho}\|_2^2 = \sum_{k \in 2\mathbb{N}(0)} |\tilde{\rho}(k)|^2 \leq \varepsilon^{-2} \sum_{k \in \mathbb{Z}} |(\rho - \bar{\rho}_{r_0})\hat{k}|^2 = \varepsilon^{-2} \|\rho - \bar{\rho}_{r_0}\|_2^2.
\]

By \( \int_T \chi = 1 \) and the Minkowski inequality,
\[
\|\rho - \bar{\rho}_{r_0}\|_2^2 = \left\| \int T \chi(y)(\rho(\cdot) - \rho(\cdot - y)) \, dy \right\|_2^2 \leq \int_{|y| < r_0} \|\rho(\cdot) - \rho(\cdot - y)\|_2^2 \, dy
= \int_T \int_{|z| < r_0} |\rho(x) - \rho(x + z)|^2 \, dz \, dx.
\]

Combining the above we obtain
\[
\|\rho - \bar{\rho}\|_2^2 \leq \varepsilon^{-2} \int_T \int_{|z| < r_0} |\rho(x) - \rho(x + z)|^2 \, dz \, dx.
\]
Choosing \( c(r_0) = \varepsilon^2 \) concludes the proof.

By virtue of the lemma, the dissipation term has the following lower bound
\[
\frac{1}{2} \int_{T^2} \phi(x - y)|\rho(x) - \rho(y)|^2 \, dx \, dy \geq c\|\rho - \bar{\rho}\|_2^2.
\]

6.2.4. The entropy equation revisited. We now return to the entropy equation (233). We have
\[
\dot{\mathcal{H}}(t) \leq \|\rho(t)\|_{L^\infty}\|q_0\|_{L^\infty}\|\rho(\cdot, t) - \bar{\rho}\|_{L^1} - c\|\rho(\cdot, t) - \bar{\rho}\|_{L^2}^2
\leq \|\rho(t)\|_{L^\infty}\|q_0\|_{L^\infty}\sqrt{16\pi\bar{\rho}\mathcal{H}(t)} - c\bar{\rho}\mathcal{H}(t).
\]

Setting \( Y = \sqrt{\mathcal{H}} \), we obtain
\[
\dot{Y}(t) \leq \|\rho(t)\|_{L^\infty}\|q_0\|_{L^\infty}\sqrt{\pi\bar{\rho}} - c\bar{\rho}Y(t).
\]

By Grönwall’s lemma we arrive at
\[
Y(t) \leq Y_0 e^{-c\bar{\rho}t} + \sqrt{\pi\bar{\rho}}\|q_0\|_{L^\infty}\int_0^t \|\rho(s)\|_{L^\infty} e^{-c\bar{\rho}(t-s)} \, ds.
\]

It is now easy to reach the conclusion of Theorem 6.2. If \( \varepsilon_0 \equiv 0 \), then the second term in (237) drops out completely and (231) closes this particular case.

For general \( \varepsilon \), we have
\[
\limsup_{t \to \infty} \|\rho(\cdot, t) - \bar{\rho}\|_{L^1} \leq M\|q_0\|_{L^\infty}\limsup_{t \to \infty} \int_0^t \|\rho(s)\|_{L^\infty} e^{-c\bar{\rho}(t-s)} \, ds
\leq \frac{\|q_0\|_{L^\infty}}{\lambda c(r_0)} \limsup_{t \to \infty} \|\rho(t)\|_{L^\infty}.
\]

Thus the proof of Theorem 6.2 is reduced to estimating the density amplitude.

6.2.5. Bounds on the Density Amplitude. First let us observe that if \( \dot{X}(t) \leq AX(t)[B - X(t)] \), where \( A \) and \( B \) are positive constants and \( X(t) \) is a positive function, then
\[
X(t) \leq \frac{BX(0)}{X(0) + (B - X(0)) \exp(-ABt)}.
\]

In particular, \( \limsup_{t \to \infty} X(t) \leq B \).
Let $\rho_+(t)$ denote the maximum $\rho$ at time $t$, and $x_+$ be a point where the maximum is achieved. Then if $\|q_0\|_{L^\infty} < \|\phi\|_{L^1}$, one can get an upper bound on $\|\rho(t)\|_{L^\infty}$ by integrating the differential inequality derived below.

\[
\frac{d}{dt} \rho_+(t) = -\rho_+(t)u'(x_+, t) = -\rho_+(t)^2 q(x_+, t) + \rho_+(t) \int_T \phi(x_+ - y)(\rho(y, t) - \rho_+(t)) \, dy \\
\leq (\|q_0\|_{\infty} - \|\phi\|_{1}) \rho_+(t)^2 + M\|\phi\|_{\infty} \rho_+(t) \\
= (\|\phi\|_{1} - \|q_0\|_{\infty}) \rho_+(t) \left[ \frac{M\|\phi\|_{\infty}}{\|\phi\|_{1} - \|q_0\|_{\infty}} - \rho_+(t) \right].
\]

In view of (239) we obtain

\[
\limsup_{t \to \infty} \|\rho(t)\|_{\infty} \leq \frac{M\|\phi\|_{\infty}}{\|\phi\|_{1} - \|q_0\|_{\infty}}.
\]

Plugging into (238) we conclude

\[
\limsup_{t \to \infty} \|\rho(\cdot, t) - \bar{\rho}\|_{L^1} \leq \frac{C|q_0|\|\phi\|_{\infty}}{\lambda c(r_0)(\|\phi\|_{1} - \|q_0\|_{\infty})}.
\]

6.3. **Alignment on $\mathbb{T}$ with degenerate kernel.** The corrector method. On the periodic circle, there is a mechanism for alignment with local communication even for vacuous solutions. It is clear already on the level of particle dynamics – if two agents have not yet aligned and are past their communication range, they would meet at the opposite end of the circle and reestablish communication. The alignment can be established in this case with an adaptation of the corrector method we introduced in Section 2.6.

**Theorem 6.4.** For any solution of either the discrete or hydrodynamic system on $\mathbb{T}$ the following holds

(i) For sub-quadratic communication

\[
\lambda \mathbf{1}_{r < r_0} \leq \phi(r) \leq \frac{\Lambda}{r^2},
\]

for some $\Lambda > 0$. One has

\[
V_2(t) \leq C \ln \frac{t}{r^2},
\]

as $t \to \infty$, where $C$ depends only on the initial condition.

(ii) If the kernel satisfies the more singular assumption

\[
\mathbf{1}_{r < r_0} \frac{\lambda}{r^\beta} \leq \phi(r) \leq \frac{\Lambda}{r^\beta}, \quad \beta > 2
\]

then we can only conclude $V_2(t) \to 0$, $t \to \infty$.

Let us note that this result does not improve the rate in the topological models since the variations of the density preclude us from making assumption (240).

**Proof.** The proof will be carried out in the discrete case only since the continuous version is entirely similar. We will limit ourselves to providing necessary modifications in that case.

We go back to the basic energy law (47) and construct a corrector functional $G$ which serves to compensate for the missing interactions. To do that we first define a periodic analogue of the directed distance:

\[
d_{ij}(t) = -x_{ij} \text{sgn}(v_{ij}) \mod 2\pi,
\]
where \( x_i, x_j \in [0, 2\pi) \) are viewed on the same coordinate chart. The distance picks up the length of the arch between \( x_i \) and \( x_j \) which contracts under the evolution of the agents. The distance undergoes jump discontinuities at \( x_i = x_j \) and \( v_i = v_j \). At any other point, we have

\[ \frac{d}{dt} d_{ij} = -|v_{ij}|. \]

Next, we define a slope kernel \( \psi \geq 0 \) as follows (see Figure 8)

\[
\psi(x) = \begin{cases} 
-x + r_0, & -r_0 \leq x \leq r_0 \\
\frac{r_0}{\pi - r_0} x - \frac{r_0^2}{\pi - r_0}, & r_0 < x < 2\pi - r_0,
\end{cases}
\]

extended periodically on \( \mathbb{R} \). Finally, we define the corrector

\[ \mathcal{G}(t) = \frac{1}{N^2} \sum_{i,j=1}^{N} |v_{ij}| \psi(d_{ij}). \]

Let us look into differentiability of \( \mathcal{G} \):

\[
\frac{d}{dt} \mathcal{G} = -\frac{1}{N^2} \sum_{i,j=1}^{N} |v_{ij}|^2 \psi'(d_{ij}) + \frac{2}{N^3} \sum_{i,j=1}^{N} \psi(d_{ij}) \text{sgn}(v_{ij}) \sum_{k=1}^{N} v_{ki} \phi_{ki}.
\]

The formula can be justified classically, at those times when there is no jump, i.e. \( x_i \neq x_j \) and \( v_i \neq v_j \), due to (243). When two agents pass each other \( x_i = x_j \) we use periodicity of \( \psi \), and when \( v_{ij} = 0 \), the factor \(|v_{ij}|\) vanishes.

We continue

\[
\frac{d}{dt} \mathcal{G}(t) = \frac{1}{N^2} \sum_{i,j=1}^{N} |v_{ij}|^2 1_{|x_{ij}| \leq r_0} - \frac{r_0}{\pi - r_0} \frac{1}{N^2} \sum_{i,j=1}^{N} |v_{ij}|^2 1_{|x_{ij}| \geq r_0} + \mathcal{R},
\]

where

\[ \mathcal{R} = \frac{2}{N^3} \sum_{i,j,k=1}^{N} \psi(d_{ij}) \text{sgn}(v_{ij}) v_{ki} \phi_{ki}. \]

Symmetrizing over \( i, k \), we obtain

\[ \mathcal{R} = \frac{1}{N^3} \sum_{i,j,k=1}^{N} (\psi(d_{ij}) \text{sgn}(v_{ij}) - \psi(d_{kj}) \text{sgn}(v_{kj})) v_{ki} \phi_{ki}. \]
In the case $v_i \geq v_j \geq v_k$ or $v_i \leq v_j \leq v_k$, the summand is negative, and so we can neglect it. Continuing,

$$
R \leq \frac{1}{N^3} \sum_{i,j,k=1}^{N} \left( \psi(d_{ij}) \text{sgn}(v_{ij}) - \psi(d_{kj}) \text{sgn}(v_{kj}) \right) v_{ki} \phi_{ki} 1_{v_j > \max(v_i, v_k)}
$$

$$
+ \frac{1}{N^3} \sum_{i,j,k=1}^{N} \left( \psi(d_{ij}) \text{sgn}(v_{ij}) - \psi(d_{kj}) \text{sgn}(v_{kj}) \right) v_{ki} \phi_{ki} 1_{v_j < \min(v_i, v_k)}
$$

$$
= \frac{1}{N^3} \sum_{i,j,k=1}^{N} \left( \psi(d_{kj}) - \psi(d_{ij}) \right) v_{ki} \phi_{ki} 1_{v_j > \max(v_i, v_k)}
$$

$$
+ \frac{1}{N^3} \sum_{i,j,k=1}^{N} \left( \psi(d_{ij}) - \psi(d_{kj}) \right) v_{ki} \phi_{ki} 1_{v_j < \min(v_i, v_k)}.
$$

In the cases $v_j > \max(v_i, v_k)$ and $v_j < \min(v_i, v_k)$, we see that $v_i - v_j$ and $v_k - v_j$ have the same sign so $d_{ij}$ and $d_{kj}$ are computed in the same direction. So, by the Lipschitz continuity of $\psi$ and by the triangle inequality we find that $|\psi(d_{ij}) - \psi(d_{kj})| \leq C|x_i - x_k|$. Therefore,

$$
R \leq \frac{C}{N^3} \sum_{i,j,k=1}^{N} |x_{ik}| v_{ik} \phi_{ki} = \frac{C}{N^2} \sum_{i,k=1}^{N} |x_{ik}| v_{ik} \phi_{ki} \leq \frac{1}{t} \frac{C}{bN^2} \sum_{i,k=1}^{N} |x_{ik}|^2 \phi_{ki} + t \frac{b}{N^2} \sum_{i,k=1}^{N} v_{ik}^2 \phi_{ki}.
$$

Let us proceed now under the assumption of (i). Here we obtain

$$
R \leq \frac{c}{t} + bt I_2.
$$

Then the corrector equation becomes

$$
\frac{d}{dt} G(t) \leq a I_2 + b(t I_2 - V_2) + \frac{c}{t}.
$$

Let us form form another functional:

$$
L = G + bt V_2 + a V_2.
$$

It satisfies inequality $\frac{d}{dt} L \leq \frac{c}{t}$. Thus, $L(t) \lesssim \ln t$, and the resulting bound follows.

For part (ii) we use the collision potential (41) with precomputed bound in (42),

$$
R \leq C \left( \frac{1}{N^2} \sum_{i,k=1}^{N} |x_{ik}|^{2-\beta} |x_{ik}|^\beta \phi_{ki} \right)^{1/2} \sqrt{I_2} \lesssim \sqrt{I_2} \sqrt{C} \leq c_1 \sqrt{I_2(t)} + c_2 \sqrt{I_2(t)} \int_0^t \sqrt{I_2(s)} \, ds.
$$

We can replace by the generalized Young inequality,

$$
c_1 \sqrt{I_2(t)} \leq \frac{c_3}{t} + bt I_2(t),
$$

and obtain

$$
\frac{d}{dt} G(t) \leq a I_2 + b(t I_2 - V_2) + \frac{c_3}{t} + c_2 \sqrt{I_2(t)} \int_0^t \sqrt{I_2(s)} \, ds.
$$

With $L$ being defined as before, we continue

$$
\frac{d}{dt} L \lesssim \frac{c_3}{t} + c_2 \sqrt{I_2(t)} \int_0^t \sqrt{I_2(s)} \, ds.
$$

Integrating over $[0, T]$,

$$
L(T) \lesssim L(0) + \ln T + \left( \int_0^T \sqrt{I_2(s)} \, ds \right)^2.
$$
Thus,
\[ \mathcal{V}_2(T) \leq \frac{1}{T} \mathcal{L}(T) \lesssim \frac{\ln T}{T} + \frac{1}{T} \left( \int_0^T \sqrt{\mathcal{I}_2(s)} \, ds \right)^2. \]

The right hand side tends to zero which can be readily seen by splitting the integral into \((0, T')\) and \((T', T)\), where \(T'\) is large.

We already noted that the hydrodynamic version of the result is identical. Let us make some remarks about the proof. We work with the Lagrangian formulation (148). As in the discrete case we define the directed distance
\[ d_{\alpha\beta}(t) = (x(\alpha, t) - x(\beta, t)) \text{sgn}(v(\beta, t) - v(\alpha, t)) \mod 2\pi, \]
and the corrector with \(\psi\) as before:
\[ \mathcal{G} = \int_{T^2} |u_{\alpha\beta}| \psi(d_{\alpha\beta}) \, d\mu_0(\alpha, \beta). \]

We calculate the derivative of \(\mathcal{G}\):
\[ \frac{d}{dt} \mathcal{G} = -\int_{T^2} |u_{\alpha\beta}|^2 \psi'(d_{\alpha\beta}) \, d\mu_0(\alpha, \beta) + \int_{T^2} \text{sgn}(u_{\alpha\beta}) \psi(d_{\alpha\beta}) \phi_{\alpha\gamma} u_{\gamma\alpha} \, d\mu_0(\alpha, \beta, \gamma) \]
\[ \leq a\mathcal{I}_2 - b\mathcal{V}_2 + \mathcal{R}. \]

Here,
\[ \mathcal{R} = \int_{T^3} \text{sgn}(u_{\alpha\beta}) \psi(d_{\alpha\beta}) \phi_{\alpha\gamma} u_{\gamma\alpha} \, d\mu_0(\alpha, \beta, \gamma) \]
\[ = \frac{1}{2} \int_{T^3} [\text{sgn}(u_{\alpha\beta}) \psi(d_{\alpha\beta}) - \text{sgn}(u_{\gamma\beta}) \psi(d_{\gamma\beta})] \phi_{\alpha\gamma} u_{\gamma\alpha} \, d\mu_0(\alpha, \beta, \gamma) \]
\[ \leq \int_{T^3} (\psi(d_{\alpha\beta}) - \psi(d_{\gamma\beta})) \phi_{\alpha\gamma} u_{\gamma\alpha} 1_{u(\beta) < \min\{u(\alpha), u(\gamma)\}} \, d\mu_0(\alpha, \beta, \gamma) \]
\[ + \int_{T^3} (\psi(d_{\gamma\beta}) - \psi(d_{\alpha\beta})) \phi_{\alpha\gamma} u_{\gamma\alpha} 1_{u(\beta) > \max\{u(\alpha), u(\gamma)\}} \, d\mu_0(\alpha, \beta, \gamma) \]
\[ \leq \int_{T^2} |x_{\alpha\gamma}| u_{\gamma\alpha} \phi_{\alpha\gamma} \, d\mu_0(\alpha, \gamma). \]

In case (i) we obtain
\[ \mathcal{R} \leq \frac{c}{t} + bt\mathcal{I}_2, \]
and the proof concludes as in the agent-based settings. In case (ii) we consider the collisional potential
\[ \mathcal{C} = \int_{T^2} \frac{d\mu_0(\alpha, \beta)}{(|x_{\alpha\beta}| \wedge r_0)^{\beta - 2}}. \]

It is well-posed for \(\beta < 3\) (in view of also the fact that the density is bounded for regular solutions). A similar computation establishes (42), and from this point on the proof is ad verbatim.

In hydrodynamic settings the \(L^2\)-based alignment result does not provide sufficient information for pointwise behavior. So, it is desirable to obtain \(L^\infty\)-based alignment statement in this context. The mechanism for such alignment comes from considering regions where the density is non-negligible – here the alignment term works faster than the transport to avoid agent collisions. At the same time if density is thin the equation acts as classical Burgers’ equation. So, in order to avoid a blowup it must have low velocity fluctuations. In other words, it has to be aligned sufficiently well.

\[ \square \]
Theorem 6.5. Consider the system (211) on $\mathbb{T}$ with a smooth non-trivial non-negative kernel. Then any global classical solution aligns:

\begin{equation}
\sup_x |u(x, t) - \bar{u}| \leq C \left( \frac{\ln t}{t} \right)^{\frac{1}{5}}.
\end{equation}

Proof. By the Galilean invariance we can assume throughout that $\bar{u} = 0$. As a consequence of the energy equality and (241) we obtain

\[ \int_T^\infty \int_{\mathbb{T}^2} \phi_{\alpha\beta} |v(\alpha, t) - v(\beta, t)|^2 \, dm_0(\alpha, \beta) \, dt \leq C \ln T \frac{T}{T} := \varepsilon. \]

Here we passed to Lagrangian coordinates $v(\alpha, t) = u(x(\alpha, t), t)$. Denote

\[ I_2(\alpha, T) = \int_T^\infty \int_{\mathbb{T}^2} \phi_{\alpha\beta} |v(\alpha, t) - v(\beta, t)|^2 \, dm_0(\beta) \, dt. \]

So, we have

\[ \int_{\mathbb{T}} I_2(\alpha, T) \, dm_0(\alpha) \leq \varepsilon. \]

Let us fix another small parameter $\delta > 0$ and define the “good set”:

\[ G_\delta(T) = \{ \alpha : I_2(\alpha, T) \leq \delta \}. \]

We denote by $G^c_\delta$ the complement of $G_\delta$, so that $m_0(G^c_\delta) = M - m_0(G_\delta)$ (recall that $M$ is the total mass of the flock). By the Chebychev inequality,

\begin{equation}
\frac{m_0(G^c_\delta)}{m_0(G_\delta)} < \varepsilon. \tag{246}
\end{equation}

Thus, the good set occupies almost all of the domain provided $\varepsilon \ll \delta$. We now proceed by proving that alignment occurs first on the good set identified above, and then on the rest of the torus later in time within a controlled time scale.

Lemma 6.6 (Alignment on $G_\delta$). We have

\[ \sup_{\alpha_1, \alpha_2 \in G_\delta(T), t \geq T} |v(\alpha_1, t) - v(\alpha_2, t)| \lesssim \delta^{2/3}. \]

Proof. It suffices to establish alignment at time $T$ only because of monotonicity of the our $I_2$-function:

\[ I_2(\alpha, t) \leq I_2(\alpha, T), \quad t > T, \]

which in particular implies that the good sets are increasing in time, $G_\delta(T) \subset G_\delta(t)$.

Integrating the equation Euler-Alignment system

\[ \frac{d}{dt} v(\alpha, t) = \int_{\mathbb{T}} \phi_{\alpha\beta} v_{\alpha\beta} \, dm_0(\beta) \]

over $[T, t]$ for any $\alpha \in G_\delta$ we obtain

\begin{equation}
|v(\alpha, t) - v(\alpha, T)| \leq \int_T^t \int_{\mathbb{T}} \phi_{\alpha\beta} |v_{\alpha\beta}| \, dm_0(\beta) \lesssim \delta \sqrt{t - T}. \tag{247}
\end{equation}

Assume that for some $\alpha_1, \alpha_2 \in G_\delta$ we have

\[ v(\alpha_1, T) - v(\alpha_2, T) > U, \]

where $U$ to be determined later. Then in view of (247),

\[ v(\alpha_1, t) - v(\alpha_2, t) > \frac{U}{2}. \]
so long as
\[ t - T \lesssim \frac{U^2}{\delta^2}. \]
During this time interval the corresponding characteristics will undergo a significant displacement
\[ x(\alpha_1, t) - x(\alpha_2, t) \geq x(\alpha_1, T) - x(\alpha_2, T) + \frac{1}{2} U(t - T) \mod 2\pi, \]
where \( \frac{1}{2} U(t - T) > 4\pi \) as long as \( t - T \gtrsim \frac{1}{U} \). If this is allowed to happen, then the characteristics will find themselves at the separation distance equal to \( 2\pi = 0 \) at some point in time, which means they would collapse. We then obtain
\[ \frac{1}{U} \gtrsim \frac{U^2}{\delta^2}, \]
which gives \( U \lesssim \delta^{2/3} \) as claimed. \[ \square \]

On the next step we show that our solution aligns at a certain not too remote later time \( t > T \).

**Lemma 6.7** (Alignment outside \( G_\delta \)). For all \( t \gtrsim T + \frac{1}{\delta^{1/3} + (\varepsilon/\delta)^{1/2}} \) we have
\[ \sup_{\alpha \in \mathcal{T}, \gamma \in G_\delta(T)} |v(\alpha, t) - v(\gamma, t)| \lesssim \delta^{1/3} + (\varepsilon/\delta)^{1/2}. \]

**Proof.** Fix \( \alpha \in \mathcal{T} \) and \( \gamma \in G_\delta(T) \). Let us write
\[
\frac{d}{dt} v(\alpha, t) = \int_T v_{\beta \alpha} \phi_{\alpha \beta} \, dm_0(\beta) = \int_T (v_{\beta \gamma} + v_{\gamma \alpha}) \phi_{\alpha \beta} \, dm_0(\beta)
= (\phi \ast \rho)(x(\alpha, t), t)v_{\gamma \alpha} + \int_T v_{\beta \gamma} \phi_{\alpha \beta} \, dm_0(\beta).
\]
The integral term on the right hand side above will remain small for all \( t \geq T \), by virtue of Lemma 6.6 and (246). Indeed,
\[
\left| \int_T v_{\beta \gamma}(t) \phi_{\alpha \beta} \, dm_0(\beta) \right| = \left| \int_{G_\delta(T)} v_{\beta \gamma}(t) \phi_{\alpha \beta} \, dm_0(\beta) \right| + \left| \int_{G_\delta(T)^c} v_{\beta \gamma}(t) \phi_{\alpha \beta} \, dm_0(\beta) \right|
\lesssim \delta^{2/3} + \frac{\varepsilon}{\delta}.
\]
Thus,
\[ (\phi \ast \rho)v_{\gamma \alpha} - \delta^{2/3} - \frac{\varepsilon}{\delta} \leq \frac{d}{dt} v(\alpha, t) \leq (\phi \ast \rho)v_{\gamma \alpha} + \delta^{2/3} + \frac{\varepsilon}{\delta}. \]
Let us consider a fixed time \( t \gtrsim T + \frac{1}{\delta^{1/3} + (\varepsilon/\delta)^{1/2}} \), and assume that \( v_{\alpha \gamma}(t) = U > 0 \) for some \( U \) to be determined later. Let us now reverse the dynamics backwards in time from the moment \( t \). For a time period \([s, t]\), where \( T < s < t \), the difference will remain positive \( v_{\alpha \gamma}(s) > 0 \). On that time period, the right hand side of (250) implies
\[ \frac{d}{dt} v \leq \delta^{2/3} + \frac{\varepsilon}{\delta} \]
and hence,
\[ v(\alpha, t) - \left( \delta^{2/3} + \frac{\varepsilon}{\delta} \right)(t - s) \leq v(\alpha, s). \]
Simultaneously, by (247) applied for \( \gamma \in G_\delta \), we obtain
\[ |v(\gamma, t) - v(\gamma, s)| \leq \delta(t - s)^{1/2}. \]
In combination with the previous, this implies
\[
U - \left( \delta^{2/3} + \frac{\varepsilon}{\delta} \right) (t - s) - \delta (t - s)^{1/2} = v_{\alpha \gamma}(t) - \left( \delta^{2/3} + \frac{\varepsilon}{\delta} \right) (t - s) - \delta(t - s)^{1/2} \leq v_{\alpha \gamma}(s).
\]
We find that
\[
v_{\alpha \gamma}(s) \geq \frac{U}{2},
\]
as long as \((t - s) \lesssim \frac{U}{\delta^{2/3} + \frac{\varepsilon}{\delta}}\) and \((t - s) \lesssim \frac{U^2}{\delta^2}\). The former condition is more restrictive, unless \(U \lesssim \delta^{4/3}\), in which case we have reached our goal. Arguing as in Lemma 6.6 we obtain collision backwards in time, provided \((t - s) \sim 1/U\). This becomes possible if \(U \gtrsim \delta^{1/3} + (\varepsilon/\delta)^{1/2}\) on the time interval of length \(t - T \gtrsim 1/U\), which is true under the assumption.

Arguing similarly from the opposite end, \(v_{\alpha \gamma}(t) = -U < 0\), we obtain the bound from below. □

Lemma 6.7 implies the following quantified global alignment starting from \(t \gtrsim T + \frac{1}{\delta^{1/3} + (\varepsilon/\delta)^{1/2}}\)
\[
\sup_{\alpha, \gamma \in \mathbb{T}} |v(\alpha, t) - v(\gamma, t)| \lesssim \delta^{1/3} + (\varepsilon/\delta)^{1/2}.
\]
Optimization over \(\delta\), produces the choice \(\delta = \varepsilon^{3/5}\). Recalling that \(\varepsilon = \ln T/T\), we obtain
\[
\sup_{\alpha, \gamma \in \mathbb{T}} |v(\alpha, t) - v(\gamma, t)| \lesssim \left( \frac{\ln T}{T} \right)^{1/5},
\]
for \(t \sim T + \left( \frac{T}{\ln T} \right)^{1/5} \sim T\). This concludes the proof. □

6.4. Singular models: GWP and strong flocking. In this section we establish global well-posedness of solutions to the Euler alignment system \((211)\) on the torus \(\mathbb{T}\) for the case of singular kernel \(\phi\). More specifically, we assume that \(\phi\) is the kernel of the classical fractional Laplacian \(\Lambda_\alpha\):
\[
\phi(z) = \sum_{k \in \mathbb{Z}} \frac{1}{|z + 2\pi k|^{1+\alpha}}, \quad 0 < \alpha < 2.
\]
Very often we refer to \(\phi\) as the kernel over the line \(\phi(z) = \frac{1}{|z|^{1+\alpha}}\) which is justified by extending corresponding functions periodically to the whole line:
\[
\Lambda_\alpha f(x) = p.v. \int_{\mathbb{T}} \phi(z) \delta_z f(x) \, dz = p.v. \int_{\mathbb{R}} \delta_z f(x) \frac{dz}{|z|^{1+\alpha}}.
\]
The analysis can be carried out for local metric kernels as well in a similar fashion.

As always in 1D the corresponding entropy will play a key role in establishing regularity of solutions to \((211)\):
\[
(251) \quad e = u_x + \Lambda_\alpha \rho.
\]
The main result is the following.

**Theorem 6.8.** Suppose \(m \geq 3\) and \(0 < \alpha < 2\). Let \((u_0, \rho_0) \in H^{m+1}(\mathbb{T}) \times H^{m+\alpha}(\mathbb{T})\), and \(\rho_0(x) > 0\) for all \(x \in \mathbb{T}^m\). Then there exists a unique non-vacuous global in time solution to \((211)\) in the class
\[
(252) \quad u \in C_w([0, \infty); H^{m+1})(0, \infty); \dot{H}^{m+1+\alpha/2}), \quad \rho \in C_w([0, \infty); H^{m+\alpha}).
\]
Moreover, the solution obeys uniform bounds on the density
\[
(253) \quad c_0 \leq \rho(x, t) \leq C_0, \quad t \geq 0,
\]
and strong flocking: \(\tilde{\rho} \in H^{m+\alpha}\) such that
\[
(254) \quad \|u(t) - \bar{u}\|_{W^{2, \infty}} + \|\rho(\cdot, t) - \tilde{\rho}(\cdot - \bar{u}t)\|_{C^{\gamma}} \leq Ce^{-\delta t} \quad t > 0, \quad 0 < \gamma < 1.
\]
According to our local well-posedness Theorem 5.3 we already have a local solution \((u, \rho)\) on time interval \([0, T_0]\). We proceed in several steps. First, we establish uniform bounds (253) on the density which depend only on the initial conditions. So, such bounds hold uniformly on the available time interval \([0, T_0]\). Next, we invoke results from the theory fractional parabolic equations to conclude that our solution gains Hölder regularity after a short period of time, and the Hölder exponent as well as the bound on the Hölder norm depend on the \(L^\infty\) bound of the solution. Finally, we establish a continuation criterion much weaker than that of Theorem 5.3 – claiming that any Hölder regularity of the density propels higher order norms beyond \(T_0\). Here the case \(\alpha = \) turns out to be more challenging than the rest of the range.

Paired with the mass equation we find that the ratio \(q = \frac{e}{\rho}\) satisfies the transport equation

\[
\frac{D}{Dt} q = q_t + u q_x = 0.
\]

Starting from sufficiently smooth initial condition with \(\rho_0\) away from vacuum we can assume that

\[
Q = |q(t)|_\infty = |q_0|_\infty < \infty.
\]

**Step 1: Bounds on the density.** We start by establishing (253) on the given time interval.

First, recall that \(q = \frac{e}{\rho}\) is transported, see (228), and hence is bounded for all time with its initial value \(|q_0|_\infty\). So, we can write the continuity equation as

\[
\rho_t + u \rho_x = -q \rho^2 + \rho \Lambda_\alpha(\rho).
\]

Let us evaluate at a point \(x_+\) where the maximum of \(\rho\), denoted \(\rho_+\), is reached. We obtain

\[
\begin{align*}
\frac{d}{dt}\rho_+ &= -q(x_+, t) \rho_+^2 + \rho_+ \int \phi(|z|)(\rho(x_+ + z, t) - \rho_+) \, dz \\
&\leq |q_0|_\infty \rho_+^2 + \rho_+ \int_{|z|<r} \phi(|z|)(\rho(x_+ + z, t) - \rho_+) \, dz \\
&\leq |q_0|_\infty \rho_+^2 + \frac{1}{\alpha} \rho_+ (M - 2 \rho_+) = |q_0|_\infty \rho_+^2 + \frac{1}{\alpha} M \rho_+ - \frac{2}{\alpha} \rho_+^2.
\end{align*}
\]

Let us pick \(r\) large enough so that \(\frac{2}{\alpha} > |q_0|_\infty + 1\). Then

\[
\frac{d}{dt}\rho_+ \leq -\rho_+^2 + C(M, r) \rho_+,
\]

which establishes the upper bound by integration.

As to the lower bound we argue similarly. Let \(\rho_-\) be the minimum value of \(\rho\) and \(x_-\) a point where such value is achieved. We have

\[
\begin{align*}
\frac{d}{dt}\rho_- &\geq -|q_0|_\infty \rho_-^2 + \rho_- \int \phi(|z|)(\rho(x_- + z, t) - \rho_-) \, dz \\
&\geq -|q_0|_\infty \rho_-^2 + \phi_- \rho_-(M - 2 \rho_-) = -c_1 \rho_-^2 + c_2 \rho_-.
\end{align*}
\]

This readily implies the bound from below. Note that at this point the global communication of the model is crucial: \(\phi_- > 0\).

As a consequence of the lower bound on the density we have a global bound on the entropy:

\[
\sup_{t \in [0, T_0]} |e(t)|_\infty < \infty.
\]

**Step 2: Hölder regularization.** The representation of continuity equation in the form (257) puts it into the class of forced fractional parabolic equations with bounded drift and force:

\[
\partial_t v + u \cdot \nabla v = L[v] + f,
\]
where $L$ has kernel

$$K(x, z, t) = \rho(x)\frac{1}{|z|^{1+\alpha}},$$

which is even with respect to $z$. The bounds on the density provide uniform ellipticity bounds on the kernel $\frac{1}{|z|^{1+\alpha}} \lesssim K(x, z, t) \lesssim \frac{1}{|z|^{1+\alpha}}$.

With these ingredients at hand, the case $\alpha = 1$ falls under the assumptions of Silverstre’s results [17] which provides Hölder regularization bound given by

$$|\rho|_{C^\gamma(\mathbb{T} \times [T_0/2, T_0])} \leq C(|\rho|_{L^\infty(\mathbb{T} \times [0, T_0])} + |\rho e|_{L^\infty(\mathbb{T} \times [0, T_0])}),$$

for some $\gamma > 0$.

The case $\alpha < 1$ falls under the same result provided $u \in L^\infty([0, T_0); C^{1-\alpha})$. This is indeed the case as follows from

$$\Lambda_\alpha^{-1} \partial_x u = \Lambda_\alpha^{-1} e - \rho \in L^\infty_{t,x}.$$

Note that $\Lambda_\alpha^{-1} \partial_x$ a $(1 - \alpha)$-order differential operator.

Finally, for $\alpha > 1$ the Hölder continuity follows from a similar identity for $\rho$:

$$\Lambda_{\alpha-1} \rho = \Lambda_1^{-1} e - \mathcal{H} u,$$

where $\mathcal{H}$ is the Hilbert transform. Note that it sends functions in $L^\infty$ to $B^0_{\infty,\infty}$. Hence, $\rho \in B^{\alpha-1}_{\infty,\infty} = C^{\alpha-1}$.

**Step 3: Continuation Criterion.** Last step is to show that if the density is bounded in $C^\gamma$ on time interval $[T_0/2, T_0)$ then the solution remains uniformly in $W^{1,\infty}$, and hence the continuation criterion of Theorem 5.3 applies. While doing that we will keep track of estimates on the $W^{1,\infty}$ with the purpose to obtaining long time asymptotics.

**Step 3A: Control over $\rho'$.** So, let us start with $\rho'$:

$$\partial_t \rho' + u \rho'' + u' \rho' + e' \rho + e \rho' = \rho' \Lambda_\alpha \rho + \rho \Lambda_\alpha \rho'. \tag{260}$$

Using again $u' = e - \Lambda_\alpha \rho$ we rewrite

$$\partial_t \rho' + u \rho'' + u' \rho' + 2e \rho' = 2 \rho' \Lambda_\alpha \rho + \rho \Lambda_\alpha \rho'. \tag{261}$$

Evaluating at the maximum of $|\rho'|$ and multiplying by $\rho'$ we obtain

$$\partial_t |\rho'|^2 + e' \rho |\rho'| + 2e |\rho'|^2 = 2 |\rho'|^2 \Lambda_\alpha \rho + \rho \Lambda_\alpha \rho'. \tag{262}$$

Let us note that $q'$ satisfies the continuity equation, and consequently, $\frac{q'}{\rho}$ is transported. So, $|q'| \leq C \rho$ pointwise. For the e-quantity itself this implies pointwise bound

$$|e'(x, t)| \leq C(|\rho'(x, t)| + \rho(x, t)).$$

Let us note that in order to make pointwise evaluation possible in (262) one has to assume regularity $e' \in C(\mathbb{T})$ which guaranteed provided $m \geq 2$. With this at hand, and in view of (253) and (258) we can bound

$$|e' \rho \rho' + 2e |\rho'|^2| \leq C(|\rho'|^2 + |\rho'|).$$

Thus,

$$\partial_t |\rho'|^2 = C(|\rho'|^2 + |\rho'|) + 2 |\rho'|^2 \Lambda_\alpha \rho + \rho \Lambda_\alpha \rho'. \tag{263}$$

Due to the bound from below on $\rho$, we estimate

$$\rho \Lambda_\alpha \rho' \leq c_1 \int_\mathbb{R} \frac{(\rho'(x + z) - \rho'(x))\rho'(x + z)}{|z|^{1+\alpha}} \, dz \leq -c_2 D\alpha \rho'(x). \tag{264}$$

where

$$D\alpha \rho'(x) = \int_\mathbb{R} |\rho'(x) - \rho'(x + z)|^2 \, dz.$$
Lemma 6.9 (Non-local maximum principle). The following pointwise bound holds
\begin{equation}
D_\alpha \rho'(x) \geq \epsilon \frac{\rho'(x)^{2+\alpha}}{|\rho|^{\alpha}}.
\end{equation}

Proof. Fix an $r > 0$ to be determined later. We write
\begin{equation}
D_\alpha \rho'(x) \geq \int_{|z| > r} \frac{\rho'(x) - \rho'(x + z)^2}{|z|^{1+\alpha}} \, dz \geq \int_{|z| > r} \frac{\rho'(x)^2 - 2\rho'(x + z)\rho'(x)}{|z|^{1+\alpha}} \, dz
= \frac{\rho'(x)^2}{r^{\alpha}} - 2\rho'(r) \int_{|z| > r} \rho'(x + z) \, dz.
\end{equation}
Integrating by parts in the last integral we further estimate
\begin{equation}
D_\alpha \rho'(x) \geq \frac{\rho'(x)^2}{r^{\alpha}} - c_\rho \rho'(x) |\rho|_{\infty} \frac{1}{r^{1+\alpha}}.
\end{equation}
Choosing $r = C \frac{\rho|_{\infty}}{|\rho'(x)|}$, where $C$ is large proves the estimate. \hfill \square

In view of density bounds we have a priori (253), the non-local maximum principle yields the following non-linear bound
\begin{equation}
D_\alpha \rho'(x) \geq c |\rho'(x)|^{2+\alpha}.
\end{equation}
We arrive at
\begin{equation}
\partial_t |\rho'|^2 = C(|\rho'|^2 + |\rho'|) + 2 |\rho'|^2 \Lambda_\alpha \rho - c |\rho'|^{2+\alpha} - \frac{1}{2} D_\alpha \rho'(x).
\end{equation}
The lower order terms $|\rho'|^2 + |\rho'|$ can be absorbed into dissipation by the generalized Young inequality:
\begin{equation}
|\rho'|^2 + |\rho'| \leq c_\epsilon + \epsilon |\rho'|^{2+\alpha},
\end{equation}
for $\epsilon > 0$ small. So, it remains to obtain estimate on the remaining term $|\rho'|^2 \Lambda_\alpha \rho$.

To do that we fix a scale parameter $1 > r > 0$ to be determined later, and split the integral representation of the fractional Laplacian into three parts: short-range, mid-range, and long-range
\begin{equation}
\Lambda_\alpha \rho(x) = \int_{|z| < r} [\delta_\alpha \rho(x) - \rho'(x) z] \, dz + \int_{r < |z| < 1} \delta_\alpha \rho(x) \, dz + \int_{|z| > 1} \delta_\alpha \rho(x) \, dz
:= S + M + L.
\end{equation}
For the short-range we use the dissipation directly:
\begin{equation}
|\rho(x + z) - \rho(x) - \rho'(x) z| = \left| \int_0^z (\rho'(x + w) - \rho'(x)) \, dw \right| \leq \sqrt{D_\alpha \rho'(x)} |z|^{1+\frac{\alpha}{2}},
\end{equation}
so,
\begin{equation}
|S| \leq r^{1-\alpha/2} \sqrt{D_\alpha \rho'(x)}.
\end{equation}
In the mid-range we use the available Hölder continuity (here we can assume without loss of generality that $\gamma < \alpha$):
\begin{equation}
|M| \leq |\rho|_{C^{\gamma}} r^{\gamma-\alpha} \lesssim r^{\gamma-\alpha}.
\end{equation}
And finally, for the long-range we simply use the boundedness of $\rho$:
\begin{equation}
|L| \lesssim |\rho|_{\infty}.
\end{equation}
The competition occurs only between the short- and mid-range terms. Optimizing over $r$ we set $r = (D_\alpha \rho'(x))^{-\frac{1}{2+\alpha+2\gamma}}$ unless such expression is $> 1$, in which case we have an absolute bound on the dissipation and the proof proceeds trivially. With the established bounds we obtain the following pointwise estimate
\begin{equation}
|\Lambda_\alpha \rho(x)| \lesssim c_1 + c_2(D_\alpha \rho'(x))^{\frac{\alpha-\gamma}{1+\alpha+\gamma}}.
\end{equation}
Note that $\frac{\alpha-\gamma}{2+\alpha+2\gamma} < \frac{\alpha}{2+\alpha}$. So, we can use the generalized Young inequality we obtain

$$|\rho'|^2 |\Lambda_\alpha \rho| \lesssim c_\varepsilon + \varepsilon |\rho'|^{2+\alpha} + \varepsilon D_\alpha \rho'(x).$$

Plugging this into (266) we arrive at

$$\partial_t |\rho'|^2 \leq c_1 - c_2 |\rho'|^{2+\alpha}. \tag{268}$$

This concludes the proof of uniform bound $\rho \in L^\infty([0, T_0); W^{1,\infty})$.

**Step 3b: control over $u'$.** Note that for the case $0 < \alpha < 1$ this bound is straightforward from the $e$-quantity. Indeed, the $e$-quantity is uniformly bounded by (258), while $\Lambda_\alpha \rho \in L^\infty$ simply by $|\Lambda_\alpha \rho|_\infty \leq |\rho'|_\infty$. However, we will seek more precise estimates with the view towards long time behavior. So, we will revisit this case also in the course of completing this step.

So, let’s write the equation for $u'$, evaluated at maximum of $|u'|$ and multiplied by $u'$:

$$\frac{d}{dt} |u'|^2 \leq |u'|^3 + u'(x) \int_\mathbb{R} \delta_z u'(x) \rho(x + z) \frac{dz}{|z|^{1+\alpha}} + u'(x) \int_\mathbb{R} \delta_z u(x) \rho'(x + z) \frac{dz}{|z|^{1+\alpha}}. \tag{270}$$

Note that the last term is bounded by $C|u'(x)|A(t)$, and we know that a priori $A(t)$ is an exponentially decaying quantity. The dissipation term is bounded, as before by

$$u'(x) \int_\mathbb{R} \delta_z u'(x) \rho(x + z) \frac{dz}{|z|^{1+\alpha}} \leq -cD_\alpha u'(x).$$

The dissipation term obeys another non-local maximum principle similar to (265) where instead of $u'$ we replace it with $(u - \bar{u})'$, thus the denominator contain the amplitude $A$ rather than $|u|_\infty$:

$$D_\alpha u'(x) \geq c \frac{|u'(x)|^{2+\alpha}}{A^{\alpha}(t)} \tag{269}$$

We continue

$$\frac{d}{dt} |u'|^2 \leq |u'|^3 - c\frac{|u'(x)|^{2+\alpha}}{A^{\alpha}(t)} + C|u'(x)|A(t). \tag{270}$$

If $\alpha > 1$, we absorb the cubic and linear terms by

$$|u'|^3 + C|u'(x)|A(t) \leq \varepsilon \frac{|u'(x)|^{2+\alpha}}{A^{\alpha}(t)} + c_\varepsilon A^{\frac{3\alpha}{\alpha-1}}(t) + c_\varepsilon A^2(t).$$

So,

$$\frac{d}{dt} |u'|^2 \leq E(t) - c\frac{|u'(x)|^{2+\alpha}}{A^{\alpha}(t)}, \tag{271}$$

where $E$ denotes an exponentially decaying quantity. In particular this establishes uniform control over $|u'|_\infty$.

If $0 < \alpha < 1$, we already know from the remark in the beginning of this step that $|u'|$ is uniformly bounded. So, we estimate $|u'|^3 \lesssim |u'|^2$, and the rest goes as before to obtain (271).

Let us investigate the remaining critical case $\alpha = 1$. This is when the dissipation is not strong enough to control nonlinearity, and at the same time $\Lambda_\alpha \rho$ is not automatically bounded either. We will do it on the next step.

**Step 3c: control over $\Lambda_1 \rho$.** Let us assume that $\alpha = 1$ and denote $\Lambda_1 = \Lambda$. We obtain an estimate on $\Lambda \rho$ indirectly, by establishing an energy-type bound on $\rho''$. So, assuming we have proved that $|\rho''|_2 \in L^\infty([0, T_0])$, control over $\Lambda \rho$ goes as follows:

$$\Lambda \rho(x) = \int_{|z|<1} [\delta_z \rho(x) - \rho'(x)z] \frac{dz}{|z|^2} + \int_{1<|z|} \delta_z \rho(x) \frac{dz}{|z|^2}.$$
The second integral is clearly bounded uniformly. Next,
\[ |\delta_z \rho(x) - \rho'(x)z| \leq \int_0^z \int_0^y \rho''(x+y) \, dy \leq |\rho''|_2 |z|^{3/2}.\]
So, the first integral is bounded by a constant multiple of $|\rho''|_2$. This shows that $\Lambda \rho \in L^\infty([0,T_0); L^\infty)$.

So, let us write the second derivative of density:
\begin{equation}
\partial_t \rho'' + u \rho''' + u' \rho'' + c' \rho' + 3c' \rho' + 2c \rho'' = 2 \rho' \Lambda \rho + 3 \rho' \Lambda \rho' + \rho \Lambda \rho'.
\end{equation}

Let us apply the test-function $\rho''/\rho$. Via routine computation with the use of the density equation, one can observe that
\[ \left\langle \partial_t \rho'' + u \rho''' + u' \rho'' + \frac{c'' \rho'}{\rho}, \rho'' \right\rangle = \frac{1}{2} \partial_t \int \frac{1}{\rho} |\rho''|^2 \, dx. \]
In view of the bounds on the density we note that $\int \frac{1}{\rho} |\rho''|^2 \, dx \sim |\rho''|_2^2$. So, it is sufficient to bound the rest of the terms in terms of $|\rho''|_2^2$. Considering the last three terms on the left hand side, let us make one observation: since $q'/\rho$ is transported, then $(q'/\rho)'$ satisfied the continuity equation, and hence $(q'/\rho)/\rho$ is transported again. Solving for $\rho''$ in this expression results in piecewise bound
\begin{equation}
|e''(x,t)| \leq C(|\rho''(x,t)| + |\rho'(x,t)| + \rho(x,t)).
\end{equation}
In order for this bound to make sense we require $m \geq 3$. With the use of a priori estimates established so far,
\[ \left\langle e'' \rho + 3e' \rho' + 2e \rho'', \frac{\rho''}{\rho} \right\rangle \lesssim 1 + |\rho''|_2^2. \]
At this point we have (dropping $\rho, \rho'$ that are already bounded)
\begin{equation}
\partial_t \int \frac{1}{\rho} |\rho''|^2 \, dx \lesssim 1 + |\rho''|_2^2 + \int |\rho''|^2 |\Lambda \rho| \, dx + I_3.
\end{equation}
Clearly, the last term $I_3$ is dissipative:
\[ I_3 \lesssim - \int D_{\alpha} \rho''(x) \, dx - \frac{1}{|\rho''|_\infty} \int \rho'' |\Lambda \rho| \, dx = -\|\rho''|_H^{1/2} - |\rho''|_3^2, \]
where in the latter we dropped $\frac{1}{|\rho''|_\infty}$ from inside the integral since this term is bounded from below.

To tackle $I_1$ we fix $\varepsilon > 0$ small and split the fractional Laplacian:
\[ \Lambda \rho(x) = \int_{|z| < \varepsilon} [\delta_z \rho(x) - \rho'(x)z] \frac{dz}{|z|^2} + \int_{\varepsilon < |z|} \delta_z \rho(x) \frac{dz}{|z|^2} \leq \varepsilon^{1/2} |\rho''|_2 + c_\varepsilon. \]
So,
\[ I_1 \leq \varepsilon^{1/2} |\rho''|_2 + c_\varepsilon |\rho''|_2 \leq \varepsilon^{1/2} |\rho''|_3 + c_\varepsilon |\rho''|_2. \]
The cubic term gets absorbed by dissipation for small $\varepsilon$.

For $I_2$ we simply use the Hölder inequality:
\[ |I_2| \leq |\rho''|_2 |\Lambda \rho'|_2 \lesssim |\rho''|_2. \]
So,
\begin{equation}
\partial_t \int \frac{1}{\rho} |\rho''|^2 \, dx \lesssim 1 + |\rho''|_2^2 - |\rho''|_3^2.
\end{equation}
This finishes the step.
Coming back to Step 3b, we conclude that \( u' = e - \lambda \rho \) remains bounded. This concludes the proof of global existence.

**STEP 4: FLOCKING.** In the cases when \( \alpha \neq 1 \), the estimates above come with a good a priori bound (271), which shows that \( |u'|_\infty \leq E(t) \), an exponentially decaying quantity. In the exceptional case \( \alpha = 1 \) we go back to (270) to obtain

\[
\frac{d}{dt} |u'|^2 \leq |u'|^3 - c \frac{|u'(x)|^3}{A(t)} + A^2(t).
\]

Clearly, in the long run the dissipation term overtakes \( |u'|^3 \) and we arrive at the same conclusion.

Since on Step 3a we showed that \( \rho \) is uniformly bounded \( W^{1,\infty} \), the proof of strong flocking for the density, \( \rho \rightarrow \rho(-\dot{u}t) \), follows along the lines of Theorem 6.1.

Lastly, showing exponential decay of \( |u''|_\infty \) follows similar estimates on the evolution of the norm \( |u''|_\infty^2 \), and will not be presented here for the sake of brevity. We refer to [15] for full details.

### 7. Multi-dimensional systems

#### 7.1. Unidirectional flocks.

One class of solutions that behaves like 1D is the class of unidirectional oriented flows. These are given by

\[
\mathbf{u}(x, t) = u(x, t) \mathbf{d}, \quad \mathbf{d} \in S^{n-1}, \quad u : \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}.
\]

The same conservation law holds for the entropy

\[
e = \mathbf{d} \cdot \nabla u + \phi * \rho, \quad \partial_t e + \nabla \cdot (e \mathbf{u}) = 0,
\]

although in this case the entropy does not control the full gradient of the velocity field. Nonetheless, one can develop a theory fully analogous to the 1D in the smooth communication case, which is what we will cover in this section.

First of all by the maximum principle applied in any direction perpendicular to \( \mathbf{d} \) one can see that the ansatz (276) is preserved in time. Second, in view of rotational invariance of the Euler Alignment System, we can assume that \( \mathbf{d} \) points in the direction of the \( x_1 \)-axis. So, we can assume

\[
\mathbf{u}(x, t) = \langle u(x, t), 0, \ldots, 0 \rangle \quad \text{for} \quad u : \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}.
\]

The full system (138) takes for of a system of scalar conservation laws:

\[
(x, t) \in \mathbb{R}^n \times \mathbb{R}^+ \quad \begin{cases}
\partial_t \rho + \partial_1 (\rho u) = 0, \\
\partial_t u + \frac{1}{2} \partial_1 (u^2) = \phi * (\rho u) - u \phi * \rho.
\end{cases}
\]

The entropy takes form

\[
e := \partial_1 u + \phi * \rho, \quad \partial_t e + \partial_1 (ue) = 0.
\]

As a result global well-posedness follows in the same way as in 1D via the threshold condition \( e_0 \geq 0 \). We can prove the full analogue of the one dimensional Theorem 6.1:

\[
\mathcal{A}(t) + \|\nabla u(t)\|_{L^\infty(\text{Supp}(\rho(t)))} + \|\nabla^2 u(t)\|_{L^\infty(\text{Supp}(\rho(t)))} \leq C e^{-\delta t},
\]

together with strong flocking (216).

We start as before by noting that the diameter of the flock \( D(t) \) remains bounded. So, there exists a time \( t^* > 0 \) such that \( e(x, t) \geq \frac{1}{2} \phi(D) M \) for all \( x \in \text{Supp} \rho(\cdot, t) \) and \( t > t^* \). Let us write an equation for \( \partial_t u \) in the following form

\[
\partial_t \partial_t u + \partial_1 u \partial_1 u + u \partial_1 \partial_1 u = \partial_1 \phi * (\rho u) - \partial_1 u \phi * (\rho) - u \partial_1 \phi * (\rho),
\]
or along characteristics

\[ \frac{d}{dt} \partial_i u = C_{i,\phi}(u, \rho) - e \partial_i u. \]

Recall from Theorem 4.1 that the velocity fluctuations \( A(t) \) are exponentially decaying. Hence, the integral above will be bounded by \( |\partial_i \phi|_\infty ME(t) \) where we denote by \( E(t) \) a generic exponentially decaying quantity. Evaluating (281) at the maximum over \( \text{Supp} \rho(\cdot, t) \) we obtain

\[ \partial_t \|\partial_i u\|_{L^\infty(\text{Supp} \rho(\cdot, t))} \leq E(t) - \frac{1}{2} \phi(\mathcal{D}) M \|\partial_i u\|_{L^\infty(\text{Supp} \rho(\cdot, t))}. \]

This readily implies the exponential bound on \( \|\partial_i u\|_{L^\infty(\text{Supp} \rho(\cdot, t))} \) for \( i = 1, \ldots, n \).

Moving on to the second order derivatives, we write the equations along characteristics

\[ \frac{d}{dt} \partial_j \partial_l u = \int_\Omega \partial_j \partial_l \phi(|x - y|) (u(y) - u(x)) \rho(y) \, dy - e \partial_j \partial_l u - \partial_j \partial_l e \rho \partial_j \partial_l u, \]

where

\[ \partial_j e = \partial_j \partial_l u + \partial_j \phi \ast \rho \quad \text{and} \quad \partial_l e = \partial_l \partial_j u + \partial_l \phi \ast \rho. \]

We prove exponential decay by bootstrapping information from the partial \( \partial_i \partial_1 u \), then \( \partial_i \partial_j u \), and then general \( \partial_i \partial_j \). So, first, we consider the case \( i = j = 1 \). Note that for this particular case, we have

\[ \frac{d}{dt} \partial_1^2 u = \int_\Omega \partial_1^2 \phi(|x - y|) (u(y) - u(x)) \rho(y) \, dy - e \partial_1^2 u - 2 \partial_1 e \partial_1 u. \]

Using that \( \partial_1 e = \partial_1^2 u + \partial_1 \phi \ast \rho \), we arrive at

\[ \frac{d}{dt} \partial_1^2 u \leq E(t) - \partial_1^2 u(e - E(t)). \]

Now, as \( e(x, t) \geq \frac{1}{2} \phi(\mathcal{D}) M > 0 \) for all \( x \in \text{Supp} \rho(\cdot, t) \) and \( t > t^* \) we have that

\[ \frac{d}{dt} \partial_1^2 u \leq E(t) - \partial_1^2 u \left(\frac{1}{2} \phi(\mathcal{D}) M - E(t)\right) \quad \text{for} \ t > t^*. \]

Since \( E(t) \) decays exponentially fast, there must exists \( t^{**} > t^* \) such that

\[ \frac{d}{dt} \partial_1^2 u \leq E(t) - \frac{1}{4} \phi(\mathcal{D}) M \partial_1^2 u \quad \text{for} \ t > t^{**}. \]

Then, evaluating the previous inequality at the maximum over \( \text{Supp} \rho(\cdot, t) \) we have obtained the desired result by integration.

Second, we consider the case \( i = 1 \) and \( j \neq 1 \). In this case, we have

\[ \frac{d}{dt} \partial_j \partial_1 u = \int_\Omega \partial_j \partial_1 \phi(|x - y|) (u(y) - u(x)) \rho(y) \, dy - e \partial_j \partial_1 u - \partial_j e \partial_1 u - \partial_1 e \partial_j u, \]

where

\[ \partial_j e = \partial_j \partial_1 u + \partial_j \phi \ast (\rho) \quad \text{and} \quad \partial_1 e = \partial_1^2 u + \partial_1 \phi \ast (\rho). \]

Using that \( \|\nabla u\|_{L^\infty(\text{Supp} \rho(\cdot, t))} \leq E(t) \) and the fact that \( \|\partial_1^2 u\|_{L^\infty(\text{Supp} \rho(\cdot, t))} \leq E(t) \) we get

\[ \frac{d}{dt} \partial_j \partial_1 u \leq E(t) - \partial_j \partial_1 u (e - E(t)) \]

and doing the same as before we obtain that \( \|\partial_j \partial_1 u\|_{L^\infty(\text{Supp} \rho(\cdot, t))} \leq E(t) \) for \( j \neq 1 \).

Finally, the case \( i, j \neq 1 \) relies on the previous in a similar manner. We get

\[ \frac{d}{dt} \partial_j \partial_i u \leq E(t) - \partial_j \partial_i u \]

and hence \( \|\partial_j \partial_i u\|_{L^\infty} \leq E(t) \).

The same argument as in 1D shows that \( \|\nabla \rho\|_{L^\infty} \) remains uniformly bounded, and with the exponential decay of velocity this implies strong flocking by the same argument as in 1D.
7.2. Mikado clusters and hydrodynamic multi-scale model. The macroscopic counterpart of the discrete multi-flock system (101) introduced in Section 2.9 can be derived in a similar fashion to the mono-flock case. Letting macroscopic flock variables denoted by \((\rho_\alpha, u_\alpha)\) for \(\alpha = 1, \ldots, A\) and global flock parameters by

\[
X_\alpha(t) := \frac{1}{M_\alpha} \int_{\mathbb{R}^n} x \rho_\alpha(x,t) \, dx, \quad M_\alpha := \int_{\mathbb{R}^n} \rho_\alpha(x,t) \, dx, \quad V_\alpha(t) := \frac{1}{M_\alpha} \int_{\mathbb{R}^n} u_\alpha(x,t) \rho_\alpha(x,t) \, dx,
\]

the multiflock Euler Alignment system is given by

\[
\begin{aligned}
\frac{\partial}{\partial t} \rho_\alpha + \nabla \cdot (\rho_\alpha u_\alpha) &= 0, \\
\frac{\partial}{\partial t} u_\alpha + u_\alpha \cdot \nabla u_\alpha &= \lambda_\alpha \left[ \phi_\alpha * (\rho_\alpha u_\alpha) - u_\alpha (\phi_\alpha * \rho_\alpha) \right] + \varepsilon \sum_{\beta \neq \alpha} M_\beta \Psi(X_\alpha - X_\beta) \left( V_\beta - u_\alpha \right),
\end{aligned}
\]

where as before communication between flocks is assumed to be weaker than communication inside each of the flocks \(\varepsilon \ll \min_\alpha \lambda_\alpha\).

One can develop a similar regularity theory in 1D as for mono-flock case and show analogues of Theorem 2.15 and Theorem 2.16 together with strong flocking statements. Here the \(\alpha\)-entropies \(e_\alpha = \partial_x u_\alpha + \lambda_\alpha \phi_\alpha * \rho_\alpha\) determine the threshold condition \(e_\alpha \geq 0\) for global existence, see [12].

Now, by analogy with the discrete system we can pass the new system of variables tied to the reference frame of each flock

\[
v_\alpha(x,t) := u_\alpha(x - X_\alpha(t),t) - V_\alpha(t) \quad \text{and} \quad \varrho_\alpha(x,t) := \rho_\alpha(x - X_\alpha(t),t).
\]

The new system reads

\[
\begin{aligned}
\frac{\partial}{\partial t} \varrho_\alpha + \nabla \cdot (\varrho_\alpha v_\alpha) &= 0, \\
\frac{\partial}{\partial t} v_\alpha + v_\alpha \cdot \nabla v_\alpha &= \lambda_\alpha \left[ \phi_\alpha * (\varrho_\alpha v_\alpha) - v_\alpha (\phi_\alpha * \varrho_\alpha) \right] + \varepsilon R_\alpha(t) v_\alpha,
\end{aligned}
\]

where

\[
R_\alpha(t) := \sum_{\beta \neq \alpha} M_\beta \Psi(X_\alpha(t) - X_\beta(t)).
\]

Now, each flock satisfies the maximum principle and we can study unidirectional configurations:

\[
v_\alpha(x,t) = v_\alpha(x,t) r_\alpha \quad \text{for} \quad v_\alpha : \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}, \quad r_\alpha \in S^{n-1},
\]

We call these solutions Mikado clusters, see Figure 10 – by analogy with Mikado solutions to the 3D incompressible Euler equation which played crucial role in resolution of the celebrated Onsager conjecture, [4, 7].
Theorem 7.1. Consider initial Mikado cluster (284) with \((v_\alpha(0), \rho_\alpha(0)) \in H^m \times (L^1_+ \cap W^{1,\infty})\) with \(m > \frac{d}{2} + 2\), satisfying the threshold condition \(e_\alpha(0) \geq 0\) for all \(\alpha = 1, \ldots, A\). Then there exists a global in time unique solution to system (282) which retains the same form (284) and satisfies \((v_\alpha, \rho_\alpha) \in C^0([0, \infty)); H^m \times (L^1_+ \cap W^{1,\infty})\). Moreover,

- **[Fast local flocking]**. Assuming that for a given \(\alpha \in \{1, \ldots, A\}\) the \(\alpha\)-flock has compact support and the internal kernel \(\phi_\alpha\) has a fat tail, then there exists \(\delta_\alpha(\lambda, \phi_\alpha, \rho_\alpha(0), u_\alpha(0))\) such that
  \[
  \sup_{x \in \text{Supp} \rho_\alpha(x,t)} \left[ |u_\alpha(x,t) - V_\alpha| + |\nabla u_\alpha(x,t)| + |\nabla^2 u_\alpha(x,t)| \right] \lesssim e^{-\delta_\alpha t},
  \]
  \[
  |\rho_\alpha(\cdot, t) - \bar{\rho}_\alpha(\cdot - X_\alpha(t))|_{C^\gamma} \lesssim e^{-\delta_\alpha t} \quad (0 < \gamma < 1).
  \]

- **[Slow global flocking]**. Suppose the inter-flock kernel \(\Psi\) has a fat tail and the internal kernels \(\phi_\alpha \geq 0\) are arbitrary. If the multi-flock has a finite diameter initially, then global alignment occurs at a rate \(\delta(\Psi, \varepsilon, \rho_\alpha(0), u_\alpha(0))\) such that
  \[
  \sup_{x \in \text{Supp} \rho_\alpha(x,t)} \left[ |u_\alpha(x,t) - V| + |\nabla u_\alpha(x,t)| + |\nabla^2 u_\alpha(x,t)| \right] \lesssim e^{-\delta t},
  \]
  \[
  |\rho_\alpha(\cdot, t) - \bar{\rho}_\alpha(\cdot - tV)|_{C^\gamma} \lesssim e^{-\delta t} \quad (0 < \gamma < 1),
  \]
  where \(V = \frac{1}{M} \sum_{\alpha = 1}^{A} M_\alpha V_\alpha\) is the global momentum.

As in the monoflock case the entropy plays a crucial role,

\[
(285) \quad e_\alpha := r_\alpha \cdot \nabla v_\alpha + \lambda_\alpha \phi_\alpha \ast \varrho_\alpha,
\]

which satisfies

\[
\partial_t e_\alpha + \nabla \cdot (v_\alpha \varrho_\alpha) = -\varepsilon R_\alpha(t) (\nabla \cdot v_\alpha),
\]

or equivalently along characteristics

\[
\frac{d}{dt} e_\alpha = (\varepsilon R_\alpha(t) + e_\alpha)(\lambda_\alpha \phi_\alpha \ast \varrho_\alpha - e_\alpha).
\]

Since \(R_\alpha \geq 0\), the initial positive entropy \(e_\alpha \geq 0\) will preserve its sign, and also be globally bounded. Thus, \(\nabla \cdot u_\alpha\) is bounded, and hence we obtain global existence by Theorem 5.1.

It was already shown in [12] that any classical solution to a multi-flock aligns exponentially fast. To prove strong flocking we simply observe that the scalar pair \((v_\alpha, \varrho_\alpha)\) satisfies

\[
(286) \quad \begin{cases}
\partial_t \varrho_\alpha + \nabla \cdot (\varrho_\alpha v_\alpha r_\alpha) = 0, \\
\partial_t v_\alpha + (r_\alpha \cdot \nabla v_\alpha) v_\alpha = \lambda_\alpha [\phi_\alpha \ast (\varrho_\alpha v_\alpha) - v_\alpha (\phi_\alpha \ast \varrho_\alpha)] + \varepsilon R_\alpha(t) v_\alpha,
\end{cases}
\]

which is similar to (278) with the exception of the extra term \(\varepsilon R_\alpha(t) v_\alpha\), which is a damping term since \(R_\alpha \geq 0\). So, the same analysis as in the previous section applies.

7.3. **Spectral dynamics approach.** In dimension 2 one can obtain an alternative threshold condition based on spectral dynamics approach. Let us assume that the kernel \(\phi\) is smooth, of convolution type, and satisfies the usual fat tail condition (7). Recall the entropy

\[
(287) \quad e = \nabla \cdot u + \phi \ast \rho,
\]

satisfying the equation

\[
(288) \quad e_t + \nabla \cdot (ue) = (\nabla \cdot u)^2 - \text{Tr} (\nabla u)^2.
\]

In 2D the right hand side is equal exactly to \(2 \det(\nabla u)\). So, if we attempt to appeal as in 1D to the logistic nature of the equation we write

\[
\frac{d}{dt} e = e (\phi \ast \rho - e) + 2 \det(\nabla u).
\]
So, the residual term $\det(\nabla \mathbf{u})$ gets in the way of controlling the growth or sign of $e$. It is difficult however to track down dynamics of $\det(\nabla \mathbf{u})$ since $\nabla \mathbf{u}$ is non-symmetric. Instead, one can track the dynamics of the symmetric part of $\nabla \mathbf{u}$, and in particular the eigenvalues of $S = \frac{1}{2}(\nabla \mathbf{u} + \nabla^\top \mathbf{u})$. In order to see exactly what we are aiming for, let us note that

$$\det(\nabla \mathbf{u}) = \det S + \omega^2,$$

where $\omega = \frac{1}{2}(\partial_1 u_2 - \partial_2 u_1)$ is the scalar vorticity of the field. Denote by $\mu_1, \mu_2$ the eigenvalues of $S$. Then

$$\det S = \mu_1 \mu_2.$$

The issue now reduces to whether we can control the spectral gap $\eta$ and vorticity $\omega$. It turns out that evolution of both quantities can be read off easily from the equation for $\nabla \mathbf{u}$. Indeed, let us write the full matrix equation first:

$$\partial_t \nabla \mathbf{u} + \mathbf{u} \cdot \nabla (\nabla \mathbf{u}) + (\nabla \mathbf{u})^2 = -\nabla \phi \cdot \nabla \mathbf{u} + E,$$

where $E$ is an exponentially decaying quantity. To be precise,

$$E = C\nabla \phi(\mathbf{u}, \rho),$$

and according to (25),

$$|E| \leq A_0 e^{-\lambda M \phi(\mathbf{u}) |\nabla \phi|_\infty M},$$

where $\mathbf{u}$ is determined solely from initial condition by equation (28). We decompose $\nabla \mathbf{u}$ into symmetric and skew-symmetric parts

$$\nabla \mathbf{u} = S + \Omega, \quad S = \frac{1}{2}(\nabla \mathbf{u} + \nabla^\top \mathbf{u}), \quad \Omega = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix}.$$

And we have decomposition of the square matrix

$$(\nabla \mathbf{u})^2 = \begin{pmatrix} S \otimes S + \Omega \Omega \end{pmatrix}_{\text{sym}} + \begin{pmatrix} \Omega \Omega \end{pmatrix}_{\text{skew-sym}} = S^2 - \omega^2 \mathbb{I}_{2 \times 2} + \Omega \nabla \cdot \mathbf{u}.$$
due to orthogonality $s_i \cdot \frac{d}{dt}s_i = 0$. So, multiplying the $S$-equation by $s_i$ from both sides, we obtain a system for spectral dynamics of the eigenvalues
$$\frac{d}{dt} \mu_i + \mu_i^2 = \omega^2 - (\phi \ast \rho) \mu_i + E.$$ 
Lastly, taking the difference and using that $\mu_1^2 - \mu_2^2 = \eta (\nabla \cdot u)$ we obtain
$$\frac{d}{dt} \eta + e \eta = E.$$ 
Collecting together the equations we have obtained the system
\begin{align*}
2 \dot{e} + e^2 &= (\phi \ast \rho)^2 + 4 \omega^2 - \eta^2 \\
\dot{\omega} + e \omega &= E \\
\dot{\eta} + e \eta &= E.
\end{align*}
(293)
Let us note in passing that the bound on $E$'s in (291) is still valid up to an absolute constant due to algebraic manipulations above.
So, let us now fix an initial condition $(u_0, \rho_0) \in H^m \times (L^1_+ \cap W^{k, \infty})$ and assume that $e_0(x) > 0$ for every $x \in \mathbb{R}^2$. According to Theorem 5.2 we have a local solution on a maximal time interval $[0, T_0)$. By continuity, $e(X(t, x), t) > 0$ for some short time $t < T(x)$. On that time interval the spectral gap solution reads
$$\eta(t) = \eta_0 \exp \left\{ -\int_0^t e(s) \, ds \right\} + \int_0^t \exp \left\{ -\int_s^t e(\tau) \, d\tau \right\} E(s) \, ds.$$ 
So,
$$|\eta(t)| \leq |\eta_0| + c A_0 \frac{||\nabla \phi||_\infty}{\phi(D)}.$$ 
Using this we obtain from the $e$-equation
$$2 \dot{e} \geq \phi^2(D) M^2 - \left( |\eta_0| + c A_0 \frac{||\nabla \phi||_\infty}{\phi(D)} \right)^2 - e^2.$$ 
Assuming now the small gap and small amplitude condition
$$|\eta_0| < \frac{1}{4} \phi(D) M, \quad A_0 < \frac{1}{4 c} \frac{\phi^2(D) M}{||\nabla \phi||_\infty}$$
we find that
$$2 \dot{e} \geq \frac{1}{2} \phi^2(D) M^2 - e^2.$$ 
This shows that $e$ will remain positive on the entire interval $[0, T_0)$, which implies that $\nabla \cdot u$ remains bounded from below. The continuation criterion (191) proves global existence.

**Theorem 7.2.** Suppose $(u_0, \rho_0) \in H^m \times (L^1_+ \cap W^{1, \infty})$ and assume that $e_0(x) > 0$ for every $x \in \mathbb{R}^2$. Assume also the smallness conditions
$$|\eta_0| < \frac{1}{4} \phi(D) M, \quad A_0 < \frac{\phi^2(D) M}{c||\nabla \phi||_\infty}.$$ 
Then there exists a global solution with this initial condition.

It is interesting to note that the small amplitude condition alone would guarantee global existence for singular models due to additional dissipation enhancement effect, see Section 7.4.

Let us observe that in the case of unidirectional flocks described in the previous section, $|\eta_0| = |\nabla u_0|$. So, the spectral gap not always applies to those flows. Yet, we know they are globally well-posed. It would be interesting to bridge the gap between the two classes of solutions.
7.4. GWP of nearly aligned flocks: small initial data. Lack of control on $e$ in multi dimensional case is part of the reason why the model has no well developed regularity theory. For singular models, however, dissipation provided by the alignment may be a tool to preserve regularity in some multi dimensional situations. The nonlinear bound (269) gives a good indication how the alignment amplifies diffusion in the velocity equation, which if initially small remains so for all time. This gives a mechanism to control the entropy and subsequently obtain a global existence of solutions with small initial fluctuation of velocity field. We carry out this idea in the case of $n$-dimensional torus for non-vacuous data.

To fix the notation, we denote by $[\cdot]_s$ the metric of the homogeneous Hölder class $C^s(\mathbb{T}^n)$. For higher $s$, we will resort to a finite difference definition of $[\cdot]_s$ stated as follows. First we denote

$$\delta_h f(x) = f(x + h) - f(x), \quad \tau_z f(x) = f(x + z)$$

$$\delta^2_h f = \delta_h(\delta_h f), \quad \delta^3_h f = \delta_h(\delta_h(\delta_h f)).$$

We then define, for $0 < \gamma < 1$:

$$[f]_{2+\gamma} = \sup_{x, h \in \mathbb{T}^n} \frac{|\delta^2_h f(x)|}{|h|^{2+\gamma}}.$$

The equivalence of (294) to the classical norm $[\nabla^2 f]_\gamma$ is a well known result in approximation theory, see [19]. For integer values of smoothness parameter $k \in \mathbb{N}$ we use classical homogeneous metric $[f]_{k} = \|\nabla^k f\|_\infty$.

We consider the global singular communication given by kernel of the classical fractional Laplacian $\Lambda_\alpha$:

$$\phi(z) = \sum_{k \in \mathbb{Z}^n} \frac{1}{|z + 2\pi k|^{n+\alpha}}, \quad 0 < \alpha < 2.$$

Like in 1D it is sometimes convenient to use both open space and periodic representations of $\Lambda_\alpha$:

$$\Lambda_\alpha f(x) = p.v. \int_{\mathbb{T}^n} \phi(z) \delta_z f(x) \, dz = p.v. \int_{\mathbb{R}^n} \delta_z f(x) \frac{dz}{|z|^{n+\alpha}}.$$

Let us state the result.

**Theorem 7.3.** Consider the Euler Alignment System (195) on the torus $\mathbb{T}^n$ with kernel given by (295). There exists an $N \in \mathbb{N}$ such that for any sufficiently large $R > 0$, depending only on $\alpha$ and dimension $n$, any initial condition $(u_0, \rho_0) \in H^{m+1}(\mathbb{T}^n) \times H^{m+\alpha}(\mathbb{T}^n)$, $m > \frac{n}{2} + 3$, satisfying

$$|\rho_0|_\infty, |\rho_0^{-1}|_\infty, [u_0]_3, [\rho_0]_3 \leq R,$$

$$A_0 \leq \frac{1}{R^N},$$

gives rise to a unique global solution in class $C_w([0, \infty); H^{m+1} \times H^{m+\alpha})$. Moreover, the solution converges to a flocking state exponentially fast:

$$A(t) + [u(t)]_1 + [u(t)]_2 + ||\rho(t) - \bar{\rho}(t)||_{C^1} < C e^{-\delta t}.$$

in the course of the proof of Theorem 7.3 we establish a uniform control on $C^2$-norm of $u$ and the distance between the initial density $\rho_0$ and its final profile $\bar{\rho}$. As a byproduct, we obtain the following stability result for flocking states.

**Theorem 7.4.** Let $(\bar{u}, \bar{\rho})$ be a traveling wave, where $\bar{\rho}(x, t) = \bar{\rho}(x - t \bar{u})$, and let $(v_0, r_0)$ be an initial data satisfying the conditions of Theorem 7.3. Suppose $|v_0 - \bar{u}|_\infty + |r_0 - \bar{\rho}|_\infty < \varepsilon$. Then the solution will converge to another flock $\bar{r}$ with $||\bar{r} - \bar{\rho}_0||_\infty < \varepsilon^\theta$, where $\theta \in (0, 1)$ depends only on $\alpha$. 

The idea of the proof is to establish control over a higher H"older norm $|u|_{2+\gamma}$. This serves multiple purposes. First, it automatically shows boundedness of the gradients $|\nabla u|_\infty$ and $|\nabla \rho|_\infty$, where for the latter we need to control $|\nabla^2 u|_\infty$. So, we fulfill the continuation criterion of Theorem 5.3 and conclude global existence as stated. Second, with $C^{2+\gamma}$ norm uniformly bounded, we obtain exponential decay of $|u(t)|_1 + |u(t)|_2$ simply by interpolation with $A$, which readily implies strong flocking as in Theorem 6.1.

From now on we will fix an exponent $0 < \gamma < 1$ to be identified later but dependent only on $\alpha$.

The proof will be structured in several steps.

**Step 1: Breakthrough scenario.** According to Theorem 5.3 we have a local solution $(u, \rho) \in C_w((0, T_0) : H^{m+1} \times H^{m+\alpha})$ satisfying the assumptions of Theorem 7.3. Note that in view of the smallness assumption on $A_0$, the norm $|u(t)|_{2+\gamma}$ will remain smaller than 1 at least for a short period of time. If the solution can be extended beyond $T_0$ there exists a possible critical time $t^* < T_0$ at which the solution reaches size $R$ for the first time:

$$ [u(t^*)]_{2+\gamma} = R, \quad |u(t)|_{2+\gamma} < R, \quad t < t^*. $$

A contradiction is be achieved when we show that $\partial_t |u(t^*)|_{2+\gamma} < 0$. This establishes the bound $|u(t)|_{2+\gamma} < R$ on the entire interval $[0, T_0)$, and hence, continuation of the solution beyond $T_0$ by Theorem 5.3. In the course of the argument we pick $\gamma$ based on several occurring restrictions, but ultimately depending only on $\alpha$.

**Step 2: Preliminary estimates on $[0, t^*]$.** We will make a few preliminary estimate on various H"older norms of the data. We fix $R$ and $N$ are sufficiently large, and $N$ depending only on $\alpha$, for all the arguments below to go through.

First, we notice two direct bounds:

$$ [u(t)]_1, [u(t)]_2 < R^{-\frac{6}{\alpha}} e^{-\frac{6t}{R}}, \quad \text{for all } t \leq t^*. $$

Indeed, by interpolation and in view of (297),

$$ [u]_1 \leq A_0^{\frac{1+\alpha}{2+\gamma}} [u]_{2+\gamma}^{\frac{1}{2+\gamma}} \leq R^{1-N/2} e^{-\alpha t/R} < R^{-\frac{6}{\alpha}} e^{-\frac{6t}{R}}, $$

and similarly,

$$ [u]_2 \leq A_0^{\frac{2+\gamma}{2+\gamma}} [u]_{2+\gamma}^{\frac{2}{2+\gamma}} e^{-\alpha t/R} \leq R^{-\frac{6}{\alpha}} e^{-\alpha t/R}. $$

Next, we consider density. Let us denote $\rho$ and $\overline{\rho}$ the minimum and maximum of $\rho$, respectively. Denote $d = \nabla \cdot u$. From (190) we conclude the bounds

$$ \rho_0 \exp \left\{ - \int_0^t |d(s)|_\infty \, ds \right\} \leq \rho(t), \quad \overline{\rho}(t) \leq \overline{\rho}_0 \exp \left\{ \int_0^t |d(s)|_\infty \, ds \right\}. $$

By (299), $|d(s)|_\infty \leq R^{-3} e^{-\alpha s/R}$. Consequently,

$$ \int_0^t |d(s)|_\infty \, ds \leq cR^{-2} \leq \ln 2, $$

We have arrived at

$$ \frac{1}{2R} \leq \rho(t), \quad \overline{\rho}(t) \leq 2R. $$

Next we obtain higher order bounds on $\rho$ with the help of the $e$-quantity. Note that the right hand side of the $e$-equation (199) is bounded by

$$ |(\nabla \cdot u)^2 - \mathrm{Tr}(\nabla u)^2| \leq c|u|^2_1 \lesssim R^{-6} e^{-\alpha t/R}. $$

So, from (199) we obtain

$$ \frac{d}{dt} |e|_\infty \leq R^{-3} e^{-\alpha t/R} |e|_\infty + R^{-6} e^{-\alpha t/R}. $$


By the Grönwall inequality, and using that $|\epsilon_0|_\infty < R$, we conclude
\begin{equation}
|\epsilon(t)|_\infty \leq 2R, \quad t < t^*.
\end{equation}
By a similar computation for $\nabla \epsilon$ we obtain
\begin{equation}
\frac{d}{dt} [\epsilon ]_1 \leq [u]_1 [\epsilon ]_1 + [u]_2 [\epsilon ]_1 + c[u]_1 [u]_2,
\end{equation}
and using (299),
\begin{equation}
\frac{d}{dt} [\epsilon ]_1 \leq R^{-3} e^{-\alpha t/R} [\epsilon ]_1 + 2R^{-2} e^{-\alpha t/R} + cR^{-3} e^{-\alpha t/R}.
\end{equation}
Since initially $|\epsilon_0|_1 \leq [u]_2 + [\rho_0]_3 < 2R$, again by the Grönwall inequality,
\begin{equation}
[\epsilon ]_1 \leq 4R.
\end{equation}
This implies that $[A_\alpha \rho]_1 < 5R$. So, if $\alpha \neq 1$, this translates directly into the Hölder norm and we obtain
\begin{equation}
[\rho]_{1+\alpha} \leq c_0 R,
\end{equation}
while for $\alpha = 1$, however it implies bounds in other border-line classes, and as a consequence,
\begin{equation}
[\rho]_{2-\gamma} \leq c_0 R.
\end{equation}

**Step 3: Higher Order Non-Local Maximum Principle.** Here we adapt the argument of Lemma 6.9 to obtain the non-local maximum principle for higher order finite differences. As before we denote
\begin{equation}
D_\alpha f(x) = \int_{\mathbb{R}^n} |f(x+z) - f(x)|^2 \frac{dz}{|z|^{n+\alpha}}.
\end{equation}

**Lemma 7.5.** There is an absolute constant $c_0 > 0$ such that
\begin{equation}
D_\alpha [\delta_h^3 f](x) \geq c_0 \frac{|\delta_h^3 f(x)|^{2+\alpha}}{|f|_2^2 |h|^{3\alpha}}.
\end{equation}

**Proof.** Fix a smooth cut-off function $\chi$, and $r > 0$ to be specified later. We obtain
\begin{align*}
D_\alpha [\delta_h^3 f](x) & \geq \int_{\mathbb{R}^n} |\delta_z \delta_h^3 f(x)|^2 \frac{1 - \chi(z/r)}{|z|^{n+\alpha}} \, dz \\
& \geq \int_{\mathbb{R}^n} (|\delta_h^3 f(x)|^2 - 2\delta_h^3 f(x) \delta_h^3 f(x + z)) \frac{1 - \chi(z/r)}{|z|^{n+\alpha}} \, dz \\
& \geq |\delta_h^3 f(x)|^2 \frac{1}{r^\alpha} - 2\delta_h^3 f(x) \int_{\mathbb{R}^n} \delta_h^3 f(x + z) \frac{1 - \chi(z/r)}{|z|^{n+\alpha}} \, dz.
\end{align*}
Note the Taylor residue formula
\begin{align*}
\delta_h^3 f(x + z) = \int_0^1 \int_0^1 \nabla_z^3 f(x + z + (\theta_1 + \theta_2 + \theta_3)h)(h, h, h) \, d\theta_1 \, d\theta_2 \, d\theta_3.
\end{align*}
Integrating by parts in $z$, and using the bound
\begin{equation}
\left| \nabla_z \left( \frac{1 - \chi(z/r)}{|z|^{n+\alpha}} \right) \right| \leq \frac{c}{|z|^{n+\alpha+1}} \chi(|z| > r),
\end{equation}
we obtain
\begin{equation}
\left| \int_{\mathbb{R}^n} \delta_h^3 f(x + z) \frac{1 - \chi(z/r)}{|z|^{n+\alpha}} \, dz \right| \leq C[f]_2 \frac{|h|^3}{r^{\alpha+1}}.
\end{equation}
Continuing with main estimate we obtain
\begin{equation}
D_\alpha [\delta_h^3 f](x) \geq |\delta_h^3 f(x)|^2 \frac{1}{r^\alpha} - C[f]_2 |\delta_h^3 f(x)| \frac{|h|^3}{r^{\alpha+1}}.
\end{equation}
Choosing \( r \sim \frac{|f_x|\|\delta_h^3 f(x)\| |h|^{3}}{\|\delta_h^3 f(x)\|^2} \) produces (305).

**Step 4: Main estimates.** We are now in a position to use the velocity equation to make estimates on the derivative of \([u]_{2+\gamma}\). Let us fix a pair \((x, h) \in \mathbb{T}^n\) which maximizes (294), and we have at time \(t^*\)
\[
[u]_{2+\gamma} = \frac{\|\delta_h^3 u(x)\|}{|h|^{2+\gamma}} = R.
\]
Consequently, at this time, using (305)
\[
\frac{1}{|h|^{4+2\gamma}} D_\alpha \delta_h^3 u(x) \geq \frac{R^{8+\alpha}}{|h|^\alpha(1-\gamma)}.
\]
Let us now write the equation for the third finite difference:
\[
\partial_t \delta_h^3 u + \delta_h^3 (u \cdot \nabla u) = \int_{\mathbb{R}^n} \delta_h^3 [\rho(x+z)\delta_z u(x)] \frac{dz}{|z|^{n+\alpha}}.
\]
Let us denote the transport and alignment terms by
\[
B = \delta_h^3 (u \cdot \nabla u),
\]
\[
I = \int_{\mathbb{R}^n} \delta_h^3 [\rho(x+z)\delta_z u(x)] \frac{dz}{|z|^{n+\alpha}}.
\]
Let use the test-function \(\delta_h^3 u(x)/|h|^{4+2\gamma}\) and evaluate at the maximizing pair at which point we also have
\[
\frac{\delta_h^3 u(x)}{|h|^{4+2\gamma}} = \frac{[u]_{2+\gamma}}{|h|^{2+\gamma}}.
\]
So we obtain from the equation,
\[
\partial_t [u]_{2+\gamma} + \frac{[u]_{2+\gamma}}{|h|^{2+\gamma}} B = \frac{\delta_h^3 u(x)}{|h|^{4+2\gamma}} I.
\]
Let us first estimate the transport term. By the product formula
\[
\delta_h^3 (f g) = \delta_h^3 f \tau_{3h} g + 3\delta_h^2 f \delta_h \tau_{2h} g + 3\delta_h f \delta_h^2 \tau_{h} g + f \delta_h^3 g.
\]
Thus,
\[
B = \delta_h^3 u \cdot \tau_{3h} \nabla u + 3\delta_h^2 u \cdot \delta_h \tau_{2h} \nabla u + 3\delta_h u \cdot \delta_h^2 \tau_{h} \nabla u + u \cdot \nabla \delta_h^3 u.
\]
Note that the last term vanishes due to criticality. Consequently,
\[
\frac{1}{|h|^{2+\gamma}} |B| \leq |u|_{2+\gamma} |u|_1 + 3|h|^{1-\gamma} |u|_2^3 + 3|u|_1 |u|_{2+\gamma} \lesssim |u|_{2+\gamma} |u|_1 + |h|^{-1-\gamma} |u|_2^3.
\]
Multiplying by another \(|u|_{2+\gamma} \leq R\) and using (299) we obtain
\[
\frac{|u|_{2+\gamma}}{|h|^{2+\gamma}} |B| \lesssim R^{-1} + R^{-5} < 1.
\]
We now turn to the dissipation term. The integrand is given by \(\delta_h^3 [\tau_z \rho \delta_z u]\). So, we expand by the product rule and using commutativity \(\delta_h \delta_z = \delta_z \delta_h\):
\[
\delta_h^3 [\tau_z \rho \delta_z u] = \delta_h^3 \tau_z \rho \tau_{3h} \delta_z u + 3\delta_h^2 \tau_z \rho \tau_{2h} \delta_h \delta_z u + 3\delta_h \tau_z \rho \tau_{h} \delta_h^2 \delta_z u + \tau_z \rho \delta_z \delta_h^3 u.
\]
Multiplying by \(\delta_h^3 u\) the last term provides necessary dissipation:
\[
\tau_z \rho \delta_z \delta_h^3 u \delta_h^3 u \leq -\frac{1}{2} \rho |\delta_z \delta_h^3 u|^2.
\]
Dividing by $|h|^{4+2\gamma}$ and using (306) we obtain the lower bound

\begin{equation}
\frac{1}{2|h|^{4+2\gamma}}D\alpha \delta_h^3 u(x) \geq \frac{R^7}{|h|^{\alpha(1-\gamma)}},
\end{equation}

In particular we can see that the entire transport term estimated in (309) is absorbed by the dissipation at the time $t^*$. As a result we obtain the equation

$$
\partial_t [u]_{2+\gamma}^2 \leq -\frac{R^7}{|h|^{\alpha(1-\gamma)}} + \frac{\delta_h^3 u(x)}{|h|^{4+2\gamma}} J,
$$

where $J$ contains all the remaining three terms of $I$:

$$
J = \int_{\mathbb{R}^n} \left[ \delta_h^3 \tau_z \rho \tau_3 \delta_z u + 3 \delta_h^3 \tau_z \rho \tau_2 \delta_z \delta_z u + 3 \delta_h \tau_z \rho \tau_h \delta_z \delta_z \delta_z u \right] \frac{dz}{|z|^{\gamma+n+\alpha}} = J_1 + 3J_2 + 3J_3.
$$

For the remainder of the proof we provide estimates for each of the $J_i$ terms with the common goal to obtain the bound

\begin{equation}
\frac{1}{|h|^{2+\gamma}} |J_i| \lesssim \frac{|h|^\epsilon}{|h|^{\alpha(1-\gamma)}},
\end{equation}

for some $\epsilon > 0$ and $\gamma$ is sufficiently small. If this is achieved, then the dissipation absorbs all these remaining $J$-terms and we conclude that

$$
\partial_t [u(t^*)]_{2+\gamma}^2 < 0,
$$

which would finish the proof.

So, let us begin with $J_1$. Symmetrizing in $z$ we obtain

\begin{equation}
J_1 = \frac{1}{2} \int_{\mathbb{R}^n} \left[ \delta_h^3 (\tau_z \rho - \tau_{-z} \rho) \tau_3 \delta_z u + \delta_h^3 \tau_z \rho \tau_3 (\delta_z u + \delta_{-z} u) \right] \frac{dz}{|z|^{\gamma+n+\alpha}}.
\end{equation}

For the first summand we use

$$
|\delta_h^3 (\tau_z \rho - \tau_{-z} \rho) \tau_3 \delta_z u| \leq |h|^{\alpha-\gamma} \min\{|z|^2, |h|\}.
$$

For the second summand,

$$
|\delta_h^3 \tau_z \rho \tau_3 (\delta_z u + \delta_{-z} u)| \leq |h|^{1+\alpha-\gamma} \min\{|z|^2, 1\}.
$$

Plugging into (313) and integrating we obtain the desired (312), but the computation extends only up to $\alpha > \frac{1}{2}$. The problem is that the density receives all the $\delta_h$’s and not fully utilizes them. At the same time, $u$ can no longer directly contribute powers of $h$. So, we switch one $h$-difference back onto $u$. So, let us fix $0 < \alpha \leq \frac{1}{2}$. We start from the original formula

$$
J_1 = \int_{\mathbb{R}^n} \delta_h^3 \tau_z \rho(x) \tau_3 \delta_z u(x) \frac{dz}{|z|^{\gamma+n+\alpha}}.
$$

Over the domain $|z| < 10|h|$ we estimate using a cut-off function $\chi$ as before,

$$
\int_{\mathbb{R}^n} |\delta_h^3 \tau_z \rho(x) \tau_3 \delta_z u(x)| \chi \left( \frac{z}{10|h|} \right) \frac{dz}{|z|^{\gamma+n+\alpha}} \lesssim \int_{|z| < 10|h|} |h|^{1+\alpha} \frac{dz}{|z|^{\gamma+n+\alpha-1}} \lesssim |h|^2.
$$
This culminates into (312). For the remaining part, denote for clarity \( f = \delta_1^2 \rho \). So, \( \delta_1^2 \tau_z \rho(x) = f(x + h + z) - f(x + z) \). Let us write

\[
\int_{\mathbb{R}^n} (f(x + h + z) - f(x + z)) \tau_{3h} \delta_z u(x) \frac{(1 - \chi(\frac{z}{10|h|}))}{|z|^{n+\alpha}} \, dz
\]

\[
= \int_{\mathbb{R}^n} f(x + z) \left( \tau_{3h} \delta_z - h u(x) \frac{(1 - \chi(\frac{z-h}{10|h|}))}{|z-h|^{n+\alpha}} - \tau_{3h} \delta_z u(x) \frac{(1 - \chi(\frac{z}{10|h|}))}{|z|^{n+\alpha}} \right) \, dz
\]

\[
= \int_{\mathbb{R}^n} f(x + z) \tau_{3h} (\delta_z - h u(x)) \frac{(1 - \chi(\frac{z-h}{10|h|}))}{|z-h|^{n+\alpha}} \, dz
\]

\[
- \int_{\mathbb{R}^n} f(x + z) \tau_{3h} \delta_z u(x) \frac{(1 - \chi(\frac{z}{10|h|}))}{|z|^{n+\alpha}} \, dz
\]

All the integrals are supported on \(|z| > 9|h|\), where \(|z - h| \sim |z|\). Estimating the first one we use

\[
|\delta_z - h u(x)| = |u(x + z - h) - u(x + z)| \leq |h|
\]

\[
|f(x + z)| \leq |h|^{1+\alpha}.
\]

Consequently,

\[
\left| \int_{\mathbb{R}^n} f(x + z) \tau_{3h} (\delta_z - h u(x)) \frac{(1 - \psi(\frac{z-h}{10|h|}))}{|z-h|^{n+\alpha}} \, dz \right| \leq |h|^{2+\alpha} \int_{|z| \geq |h|} \frac{dz}{|z|^{n+\alpha}} \leq |h|^2,
\]

which implies (312). For the second integral we use

\[
\left| \frac{(1 - \psi(\frac{z-h}{10|h|}))}{|z-h|^{n+\alpha}} - \frac{(1 - \psi(\frac{z}{10|h|}))}{|z|^{n+\alpha}} \right| \leq |h| \int_{|z| \geq |h|} \frac{dz}{|z-h|^{n+\alpha+1}} \leq \frac{|h| |z| \min\{1, 1\}}{R h^{n+\alpha+1}},
\]

and

\[
|f(x + z) \tau_{3h} \delta_z u(x)| \leq |h|^{1+\alpha} |z|.
\]

Integration reproduces the same bound as for the first part.

Next, \( J_2 \). For \( \alpha < 1 \), let us use (299) and (302) to deduce

\[
|\delta_1^2 \tau_z \rho| \leq \rho_{1+\alpha} |h|^{1+\alpha} \lesssim R |h|^{1+\alpha}
\]

\[
|\tau_{2h} \delta_h \delta_z u| \leq |u|_2 |h| \min\{|z|, 1\} \lesssim R^{-1} |h| |z| \min\{|z|, 1\}.
\]

The singularity is now removed and we get

\[
\frac{1}{|h|^{2+\gamma}} |J_2| \lesssim |h|^{\alpha - \gamma},
\]

which implies (312) for sufficiently small \( \gamma \). For \( \alpha \geq 1 \) we first symmetrize

\[
J_2 = \frac{1}{2} \int_{\mathbb{R}^n} \left[ \delta_1^2 (\tau_z - \tau_z) \rho \tau_{2h} \delta_h \delta_z u + \delta_1^2 \tau_z \rho \tau_{2h} \delta_h (\delta_z + \delta_z) u \right] \frac{dz}{|z|^{n+\alpha}}.
\]

Here for the first summand we use (303):

\[
|\delta_1^2 (\tau_z - \tau_z) \rho| \leq R \min\{|h|^{1+\alpha-\gamma}, |h|^{\alpha-\gamma} |z|\}
\]

\[
|\tau_{2h} \delta_h \delta_z u| \leq R^{-1} |h| \min\{|z|, 1\}.
\]

With this at hand we proceed

\[
\frac{1}{|h|^{2+\gamma}} \int_{\mathbb{R}^n} \left| \delta_1^2 (\tau_z - \tau_z) \rho \tau_{2h} \delta_h \delta_z u \right| \frac{dz}{|z|^{n+\alpha}} \leq \frac{|h|^{1+\alpha-\gamma}}{|h|^{2+\gamma}} \leq \frac{|h|^{\alpha-2\gamma-1+\alpha(1-\gamma)}}{|h|^{\alpha(1-\gamma)}} \leq \frac{|h|^\varepsilon}{|h|^{\alpha(1-\gamma)}},
\]
since clearly, \(\alpha - 2\gamma - 1 + \alpha(1 - \gamma) > 0\) for small \(g\). In the second summand, using that \((\delta_z + \delta_{-z})u\)
is the second order finite difference,
\[
|\delta^2_h \tau_z \rho \tau_{2h} \delta_h (\delta_z + \delta_{-z})u| \leq |h|^{2-\gamma} \min\{|z|^2, 1\}
\]
we obtain
\[
\frac{1}{|h|^{2+\gamma}} \int_{\mathbb{R}^n} |\delta^2_h \tau_z \rho \tau_{2h} \delta_h (\delta_z + \delta_{-z})u| \frac{dz}{|z|^{n+\alpha}} \leq \frac{|h|^{2-\gamma}}{|h|^{2+\gamma}} \leq \frac{h^{\alpha(1-\gamma)-2\gamma}}{|h|^{\alpha(1-\gamma)}}.
\]
This finishes the bound on \(J_2\).
Finally, for \(J_3\) we proceed similarly. For \(\alpha < 1\), we use
\[
|\delta_h \tau_z \rho \tau_{h} \delta^2_z u| \leq |h|^2 \min\{|z|, 1\}.
\]
Hence,
\[
\frac{1}{|h|^{2+\gamma}} |J_3| \leq \frac{|h|^2}{|h|^{2+\gamma}} \leq \frac{h^{\alpha(1-\gamma)-\gamma}}{|h|^{\alpha(1-\gamma)}} \leq \frac{h^{\epsilon}}{|h|^{\alpha(1-\gamma)}}.
\]
For \(\alpha \geq 1\), we again symmetrize first
\[
J_3 = \frac{1}{2} \int_{\mathbb{R}^n} \left[ \delta_h (\tau_z - \tau_{-z}) \rho \tau_{2h} \delta^2_h (\delta_z + \delta_{-z})u + \delta_h \tau_z \rho \tau_{h} \delta^2_h (\delta_z + \delta_{-z})u \right] \frac{dz}{|z|^{n+\alpha}},
\]
and using
\[
|\delta_h (\tau_z - \tau_{-z}) \rho \tau_{h} \delta^2_h (\delta_z + \delta_{-z})u| \leq \min\{|h|^{3+\alpha-\gamma}, |h|^{1+\alpha}|z|^2\},
\]
\[
|\delta_h \tau_z \rho \tau_{h} \delta^2_h (\delta_z + \delta_{-z})u| \leq \min\{|h|^3, |h||z|^2\},
\]
integration implies (312).
We have established that \(\partial_t [u(t^*)]^2_{z+\gamma} < 0\) at the critical time. This means that such time \(t^*\)
does not exist and which finishes the proof of existence part.

**Step 5: Flocking and Stability.** As we noted in the beginning, exponential decay of \([u]^2\)
implies uniform control over \(|\nabla \rho|_\infty\).

Arguing as in the proof of Theorem 6.1 we can slightly improve the space in which string flocking
occurs. This is due to (302) - (303) bounds, which imply that \(\bar{\rho} \in W^{1+\alpha-\gamma,\infty}\) by compactness. Using
again (303) and by interpolation we have convergence in the \(W^{1,\infty}\)-metric as well:
\[
|\rho(\cdot, t) - \bar{\rho}(\cdot, t)|_1 < C_2 e^{-\delta t}.
\]
As far stability is concerned, the computation above shows that in fact the limiting profile \(\bar{r}\)
differs little from initial density \(\tau_0\) under the conditions of Theorem 7.4. Indeed, setting \(R\) such that \(\varepsilon = 1/R^N\) (here \(\varepsilon > 0\) is small), we obtain via (299),
\[
|\partial_t r|_\infty \leq CR^{-2} e^{-\alpha t/R}.
\]
Hence, \(|\bar{r} - r_0|_\infty \leq \frac{C}{\alpha R} = e^\theta\). Since \(|r_0 - \bar{r}_0| < \varepsilon\), this finishes the result.

7.5. **Notes and remarks.** The two notable exceptions are the result of Ha, et al [5] demonstrating
global existence in the case of smooth communication kernel \(\phi\) with small initial data in higher
order Sobolev spaces, \(\|u\|_{H^{\gamma+1}} < \varepsilon_0\), where \(\varepsilon_0\) depends on \(\|\rho_0\|_{H^2}\);
8. Pressured Euler Alignment system

We consider a class of compressible fluid models in one space dimension with periodic boundary conditions

\begin{align}
\tag{314}
\partial_t \rho + \partial_x (u \rho) &= 0, \\
\tag{315}
\partial_t (pu) + \partial_x (pu^2) &= -\partial_x p(\rho) + D(u, \rho), \\
\tag{316}
(\rho, u)|_{t=0} &= (\rho_0, u_0).
\end{align}

Here \( f \in L^\infty(\mathbb{T} \times \mathbb{R}^+) \) is a bounded external force, \( D(u, \rho) \) is a diffusion operator and the pressure \( p := p(\rho) \) is a given function of density. The central feature of dissipation operators \( D \) considered here is the existence of another quantity, denoted \( Q \), which allows one to express the term \( \rho^{-1}D \) as a transport of \( Q \):

\begin{equation}
\tag{317}
D_t Q := Q_t + u Q_x = \rho^{-1}D(u, \rho).
\end{equation}

One classical example is the local dissipation in divergence form

\begin{equation}
\tag{318}
D(u, \rho) = (\mu(\rho)u_x)_x.
\end{equation}

This example embodies a broad family of models that appear in various physical phenomena such as barotropic compressible fluids, slender jets, shallow water waves, etc. Here, \( \mu(\rho) \) designates the dynamic viscosity of the fluid which is typically given by the constitutive power law

\begin{equation}
\tag{319}
\mu(\rho) = c_\mu \rho^\alpha, \quad c_\mu > 0, \quad \alpha \geq 0.
\end{equation}

In this case,

\begin{equation}
\tag{320}
Q = \frac{\mu(\rho) \rho_x}{\rho^2}.
\end{equation}

In compressible barotropic fluid models, the pressure equation of state is of the form

\begin{equation}
\tag{321}
p(\rho) = c_p \rho^\gamma, \quad c_p > 0, \quad \gamma > 0.
\end{equation}

Some special cases include a barotropic monatomic gas with \( \alpha = 1/3 \) and \( \gamma = 5/3 \), shallow water waves with \( \alpha = 1 \) and \( \gamma = 2 \) and viscous slender jet dynamics with \( \alpha = 1 \) and \( \gamma = 1/2 \) (but with negative pressure \( c_p < 0 \)). See [14] and [17] for recent studies and further discussion.

Recently, another class of examples, referred to as singular Euler-Alignment models, with non-local density-dependent dissipation appeared in the context of collective behavior [14]:

\begin{equation}
D(u, \rho) = p[L^s(\rho u) - u L^s(\rho)].
\end{equation}

Here \( L^s := -(-\partial_{xx})^{s/2} \) for \( s \in (0, 2) \) is the fractional Laplacian given by the integral representation

\begin{equation}
L^s f(x) = \int_{\mathbb{T}} \phi_s(x - y)(f(y) - f(x)) dy, \quad \phi_s(x) = c \sum_{k \in \mathbb{Z}} \frac{1}{|x + 2\pi k|^{1+s}}.
\end{equation}

The transport representation of the dissipation comes as a consequence of the commutator structure in this case, where one finds

\begin{equation}
\tag{322}
Q = \partial_x^{-1}L^s \rho.
\end{equation}

With regard to the pressure law, two cases arise naturally from the corresponding agent-based Cucker-Smale system introduced in [2, 3]. One is the pressureless case, \( p = 0 \), which presents itself as a limit to monokinetic concentration \( f \to \rho \delta_{v=u} \) in kinetic formulation with a strong local alignment forcing, see [17, 18] for rigorous study. Second is an isentropic pressure given by

\begin{equation}
\tag{323}
p(\rho) = c_p \rho, \quad c_p \geq 0,
\end{equation}

which arises from stochastically forced systems. In the strong dissipation limit, the probability density \( f \) converges to a Maxwellian distribution, see [17, 18, 19]. While the pressureless case is well
understood by now and is covered in extensive studies \cite{14, 15, 16, 13, ?}, the pressured case has virtually been omitted in the literature, except for smooth communication \cite{?}. Both cases fall under the general class covered in our present study.

Since the relation (317) is linear, one can access a family of hybrid local/non-local dissipation models as well:

\begin{equation}
\partial_t (\rho u) + \partial_x (\rho u^2) = -\partial_x p(\rho) + c_{nl} D_{nl}(u, \rho) + c_{loc} D_{loc}(u, \rho) + \rho f
\end{equation}

(324)

\[ D_{loc}(u, \rho) = (\mu(\rho) u_x)_x, \quad D_{nl}(u, \rho) = \rho [L^s(pu) - u L^s(\rho)] \]

In the context of collective behavior these encompass multi-scale alignment models – classical power law at large scales, and strong singular alignment at local small scales. Although multi-scaling has already appeared on the kinetic level in the analysis of hydrodynamic limit performed in \cite{?}, the net effect of the local alignment considered there averages down to zero in the macroscopic formulation. We argue, however, that local dissipation, along with the plethora of constitutive laws (319), appears naturally as a singular limit \( s \to 2 \) of so-called topological model introduced in \cite{13}.

Let us recall the construction. It is observed in many biological behavioral studies, \cite{?, ?}, that “agents”, such as birds or fish, probe local environment sensing only a fixed number of other agents around them. Consequently, the actual communication neighborhood is determined by topologies determined by the density of the flock rather than the classical Euclidean one. The topological density-dependent distance can be defined by the mass of intermediate segment between agents (see \cite{13} for multi-dimensional construction):

\[ d(x, y) = \left| \int_x^y \rho(z, t) \, dz \right| \]

Since communication in dense areas progresses slower, the mass-distance should decrease effective viscosity of the alignment, leading one to consider a kernel inversely dependent on \( d \):

\[ \phi(x, y) = \frac{h(x - y)}{|x - y|^{1+s-\tau} d^\tau(x, y)} \]

where \( \tau > 0 \) is a parameter that gauges contribution of the topological part and \( h \) is a local cut-off function. The corresponding operator is given by

\[ L^{s, \tau} f(x) = \int_T \phi(x, y)(f(y) - f(x)) \, dy. \]

As \( s \to 2 \), the normalized operator formally converges to the local elliptic operator in divergence form:

\[ (2 - s) L^{s, \tau} f \to (\rho^{2-\tau} f_x)_x. \]

Consequently, the alignment term converges to

\begin{equation}
(2 - s) \rho [L^{s, \tau}(u \rho) - u L^{s, \tau} \rho] \to (\rho^{2-\tau} u_x)_x.
\end{equation}

(325)

So, we obtain an example of the local operator (318) with topological viscosity given by \( \mu(\rho) = \rho^{2-\tau} \). To summarize, in the context of flocking, the hybrid model (324) describes a flock driven by a strong local topological alignment and global power law communication.

Returning to the general discussion, we make a key observation – once the transport quantity \( Q \) is identified, one can rewrite the entire system as a system of conservation laws with the momentum equation given by the transport equation

\begin{equation}
D_t X = -h_x(\rho) + f, \quad h'(r) := \frac{p'(r)}{r},
\end{equation}

(326)

for the new quantity

\[ X = u + Q. \]
We will exploit this structure to prescribe an algorithm of constructing a hierarchy of entropy-like quantities which are extremely useful in studying regularity of such systems. The first member in the hierarchy is given by the well known Bresch-Desjardins entropy [7]

\[ \mathcal{H}_0 = \frac{1}{2} \int_T \rho X^2 \, dx + \int_T \pi_0(\rho) \, dx, \]

where \( \pi_0 \) is the pressure potential given by

\[ \pi_0(\rho) = \rho \int_{\tilde{\rho}}^{\rho} \frac{p(s)}{s^2} \, ds, \quad \text{for some} \quad \tilde{\rho} > 0. \]

The algorithm, described in detail later in Section 8.1, gives rise to higher order entropies, \( \mathcal{H}_1, \mathcal{H}_2, \ldots \), each controlling corresponding higher order derivatives \( X, X_{xx}, \rho_x, \rho_{xx}, \) etc. Note that in the pressureless case, in particular, \( X_x \) is precisely the “\( \epsilon \)-quantity” discovered in [1], which determines a threshold for regularity of solutions in the bounded kernel case.

With the use of this method we present, in a unified way, a range of global existence and continuation results for local, non-local, or hybrid models. Let us state our main results now.

**Theorem 8.1** (Continuation Criterion). Consider the system (314)–(316) with

\[ D(u, \rho) := c_{nl} D_{nl}(u, \rho) + c_{loc} D_{loc}(u, \rho), \]

with equations of state given by (319), (321), and \( f \in L^\infty(R^+; C^n) \). Consider the cases

1. purely non-local: \( c_{loc} = 0, c_{nl} > 0 \), in which case we require \( s \in (\frac{5}{3}, 2), \gamma > 0 \).
2. purely local: \( c_{loc} > 0, c_{nl} = 0 \), in which case we require \( (\alpha \geq 0, \gamma > 1) \) or \( (\alpha > \frac{1}{2}, \gamma > 0) \).
3. mixed: \( c_{loc} > 0, c_{nl} > 0 \), in which case we require \( \alpha \geq 0, \gamma > 1, s \in (0, 2) \) or \( \alpha > \frac{1}{2}, \gamma > 0, s \in (0, 2) \) or \( \alpha \geq 0, \gamma > 0, s \in (\frac{3}{2}, 2) \).

Suppose \( (u, \rho) \in H^{m+1-\sigma} \times H^m, m \geq 2 \), is a local solution on time interval \((0, T)\). Suppose also that

\[ \rho := \inf_{t \in [0, T^*]} \min_{x \in \mathbb{T}} \rho(x, t) > 0. \]

Then the solution belongs to the class

\[ u \in L^\infty(0, T; H^{m+1-\sigma}) \cap L^2(0, T; H^m) \]
\[ \rho \in L^\infty(0, T; H^m) \cap L^2(0, T; H^{\sigma/2+m}) \]

where \( \sigma \) is the order of the operator \( D(u, \rho) \), and hence can be extended locally beyond \( T_* \).

**Remark 8.2.** We note that the correspondence

\[ (\rho \in H^m) \sim (u \in H^{m+1-\sigma}) \]

is natural for reasons to be clarified later.

The statement about purely local models in our Theorem 8.1 provides an alternate proof (and extension) of Theorem 1.1 in [7]. In that paper, an “active potential” \( u \) which satisfied a less-degenerate parabolic equation was used to propagate higher regularity provided the density nowhere vanished. In our work, we establish the same result by analyzing the entropy hierarchy.

Our next result establishes global well-posedness for a class of hybrid models, which is proved by propagating a lower bound on the density and appealing to Theorem 8.1.

**Theorem 8.3** (Global Existence). Assume \( f \in L^\infty(R^+; C^n), \rho \in C^{n+1}_{loc}(R^+), \) with \( p'(r) > 0 \) for any \( r > 0 \) and

\[ c_{nl} \geq 0, c_{loc} > 0, \quad \alpha \in (0, 1/2). \]
Then any given local classical solution with non-vacuous initial data enjoys a priori lower bound (363) on its interval of existence. Consequently, any non-vacuous initial condition \((u_0, \rho_0) \in H^{m+1-\sigma} \times H^m, m \geq 2\), gives rise to a unique global solution in the range of parameters stated in Theorem 8.1.

In the purely local case, Theorem 8.3 provides an alternate proof to that of Mellet and Vasseur [?] who proved global well-posedness in the parameter range \(\alpha < 1/2\) and \(\gamma > 1\). As another application, we obtain global existence for collective behavior models.

**Corollary 8.4.** Any collective behavior model, \(\gamma = 1\), with multi-scale diffusion \(c_{nl}, c_{loc} > 0\) in the range of parameters \(\alpha \in (0, 1/2), s \in (3/2, 2)\) is globally well-posed for initial data in the class \((u_0, \rho_0) \in H^{m-1} \times H^m, m \geq 2\).

Our final result concerns the long-time behavior of the velocity and density fields in models which possess a non-local dissipation component. In particular, we show that the energy inequality together with the non-local analogue of Bresch-Desjardins entropy imply flocking in an \(L^2\)-sense.

**Theorem 8.5 (Non-local “Second Law” Implies Flocking).** Consider the forceless system (314)–(316) with \(c_{nl} > 0, c_{loc} \geq 0\) and pressure law given by (323). Then any classical solution undergoes flocking behavior in the weighted \(L^2\)-sense:

\[
\int_{T \times \mathbb{T}} (|u(x) - u(y)|^2 \rho(x) \rho(y) \, dx \, dy + \|\rho(t) - \bar{\rho}\|_{L^1(T)}^2) \leq \frac{\ln t}{t}.
\]

We note that that in the pressurized case, as opposed to pressureless, the density always converges to a uniform state selected by its average. It is an indirect consequence of stochastic diffusion that leads to persistent mixing and eventual homogenization of the flock density. Analogous behavior was also observed in the study of Choi [?], under the assumption of globally bounded velocity field \(u\). Note, however, that such assumption is not guaranteed a priori due to lack of the maximum principle in the pressured system.

8.1. **Hierarchy of Entropies Method.** As observed in the Introduction, due to the transport formulation of the dissipation (317) one can rewrite the entire momentum equation (315) as a transport equation for the new quantity

\[
X = u + Q,
\]

\[
D_t X = -h_x(\rho) + f, \quad h'(r) := \frac{p'(r)}{r}.
\]

Now (314)-(315) becomes as a system of conservation law for the new pair of unknowns \((\rho, \rho X)\). We will exploit this structure to prescribe an algorithm of constructing a hierarchy of entropy-like quantities. First, let us make a general observation – if we have two quantities, \(X\) and \(\rho\), one is transported and the other is conserved

\[
D_t X = 0, \quad \rho_t + (u\rho)_x = 0,
\]

then the “energy” given by \(\frac{1}{2} \int_T \rho X^2 \, dx\) is preserved for all time. In the presence of the pressure such conservation is destroyed,

\[
\frac{d}{dt} \frac{1}{2} \int_T \rho X^2 \, dx = - \int_T X p_x(\rho) \, dx + \int_T \rho f X \, dx,
\]

and the pressure term splits into two elements coming from \(X\)

\[
X p_x(\rho) = u p_x(\rho) + Q p_x(\rho).
\]
It turns out that the $Q$-term is in fact dissipative in all the examples we considered so far. So the only term that needs to be eliminated is $up(\rho)_x$. This is done with the use of the pressure potential given by

$$(337) \quad \pi_0(\rho) = \rho \int_{\rho}^{\tilde{\rho}} \frac{p(s)}{s^2} \, ds, \quad \text{for some } \tilde{\rho} > 0.$$ 

Indeed,

$$(338) \quad \frac{d}{dt} \int_{\mathbb{T}} \pi_0(\rho) \, dx = \int_{\mathbb{T}} up_x(\rho) \, dx.$$ 

We thus recover what is known as Bresch-Desjardins’s entropy

$$(339) \quad H_0 = \frac{1}{2} \int_{\mathbb{T}} \rho X^2 \, dx + \int_{\mathbb{T}} \pi_0(\rho) \, dx.$$ 

According to the computations above we obtain the following balance relation

$$(340) \quad \frac{d}{dt} H_0 = \int_{\mathbb{T}} Q_x p(\rho) \, dx + \int_{\mathbb{T}} \rho f X \, dx.$$ 

In parallel, due to (338), we obtain an energy balance relation for the energy of the system given by

$$(341) \quad E = \frac{1}{2} \int_{\mathbb{T}} \rho |u|^2 \, dx + \int_{\mathbb{T}} \pi_0(\rho) \, dx,$$

$$(342) \quad \frac{d}{dt} E = \int_{\mathbb{T}} uD(u, \rho) \, dx + \int_{\mathbb{T}} \rho uf \, dx.$$ 

In all cases of interest the pressure term in (340) is sign-definite provided $p'(r) \geq 0$. Indeed, in the non-local case (322) we obtain

$$(343) \quad \int_{\mathbb{T}} Q_x p(\rho) \, dx = \int_{\mathbb{T}} p(\rho) \mathcal{L}^s \rho \, dx = - \frac{1}{2} \int_{\mathbb{T}^2} \phi_{s,\rho}(x, y) (\rho(y) - \rho(x))^2 \, dx \, dy \leq 0,$$

where

$$\phi_{s,\rho}(x, y) = \phi_s(x, y) \int_0^1 p'(\theta \rho(x) + (1 - \theta) \rho(y)) \, d\theta.$$ 

In the local case (320) we find

$$(344) \quad \int_{\mathbb{T}} Q_x p(\rho) \, dx = - \int_{\mathbb{T}} \frac{\mu(\rho)\rho^2 p'(\rho)}{\rho^2} \, dx \leq 0.$$ 

Consequently, the initial entropy $H_0$ gives control over $\rho X^2 \in L^\infty_t L^1_x$. Together with the energy conservation, this in turn controls the solo-density term:

$$\int_{\mathbb{T}} \rho |Q|^2 \, dx \leq \int_{\mathbb{T}} \rho |u|^2 \, dx + \int_{\mathbb{T}} \rho X^2 \, dx$$

which will be used to extract initial regularity information on the density in each of local and non-local cases separately.

As to the hybrid case, we have $Q = Q_{nl} + Q_{loc}$. The key observation is that we can extract control on each of the terms separately. Indeed, writing

$$\int_{\mathbb{T}} \rho |Q_{nl} + Q_{loc}|^2 \, dx = \int_{\mathbb{T}} \rho |Q_{nl}|^2 \, dx + \int_{\mathbb{T}} \rho |Q_{loc}|^2 \, dx + 2 \int_{\mathbb{T}} \rho Q_{nl} Q_{loc} \, dx,$$

we observe that the integral of the cross-dissipation is non-negative:

$$(345) \quad \int_{\mathbb{T}} \rho Q_{nl} Q_{loc} \, dx = \int_{\mathbb{T}} \partial_x^{-1} \mathcal{L}^s \rho \frac{\mu(\rho)}{\rho} \, dx = \int_{\mathbb{T}} \partial_x^{-1} \mathcal{L}^s \rho \psi(\rho)_x \, dx = - \int_{\mathbb{T}} \psi(\rho) \mathcal{L}^s \rho \, dx \geq 0.$$
where $\psi'(r) = \mu(r)/r$. The non-negativity holds, in fact, provided $\mu(r) \geq 0$, which is manifestly true for positive viscosities.

Coming back to the entropy construction we now present the next step in the hierarchy. Noting that in the pressureless case the quantity $X_x$ would have satisfied the continuity equation, and hence, in combination with the density the new variable $Y = \frac{1}{\rho}X_x$ would have been transported. By analogy with the previous, we would then start construction with the pressureless term $\rho Y^2 = \rho^{-1}X_x^2$. Note that $X_x$ satisfies

\begin{equation}
\partial_t X_x + (u X_x)_x = -h_{xx}(\rho) + f_x.
\end{equation}

Hence, in conjunction with mass conservation,

\begin{equation}
D_t Y = -\rho^{-1}h_{xx}(\rho) + \rho^{-1}f_x.
\end{equation}

We thus obtain

\begin{equation}
\frac{d}{dt} \frac{1}{2} \int_T \rho Y^2 \, dx = - \int_T h_{xx}(\rho) Y \, dx + \int_T f_x Y \, dx.
\end{equation}

The appropriate next order pressure potential that eliminates the first term on the right hand side is given by

\begin{equation}
\pi_1(\rho, \rho_x) = \frac{1}{2} h'(\rho) \frac{\rho_x^2}{\rho^2}.
\end{equation}

The details of this computation will be provided in the sections below. We thus arrive at the next entropy

\begin{equation}
\mathcal{H}_1 = \frac{1}{2} \int_T \rho Y^2 \, dx + \int_T \pi_1(\rho, \rho_x) \, dx.
\end{equation}

the algorithm is now clear. For the second order entropy we denote $Z = \frac{1}{\rho} Y_x$ and define

\begin{equation}
\pi_2 = \frac{1}{2} h'(\rho) \frac{\rho_{xx}^2}{\rho^4}.
\end{equation}

Continuing in the same fashion we can design an entropy-like quantity of any order, where the pressure potential is given by

\begin{equation}
\pi_n = \frac{1}{2} h'(\rho) \frac{(\partial^n_x \rho)^2}{\rho^{2n}},
\end{equation}

while the kinetic term is constructed inductively,

\begin{equation}
X_n = \frac{1}{\rho} \partial_x X_{n-1}.
\end{equation}

We form the $n$-th entropy accordingly,

\begin{equation}
\mathcal{H}_n = \frac{1}{2} \int_T \rho X_n^2 \, dx + \int_T \pi_n \, dx.
\end{equation}

With the help of this hierarchy we establish a direct control over any higher order regularity of solution, consistent with that of the initial datum, where the relative smoothness of $u$ and $\rho$ mentioned in (333) naturally equilibrates the order of and the velocity and density terms as they enter into an expression for $X_n$. It should be noted however that these entropies do not decay precisely as the first element $\mathcal{H}_0$. Instead, they satisfy ODEs with residual terms. Controlling those residual terms presents the main technical component of the method.
8.2. Global Existence and Flocking. We start by presenting less technical proofs of Theorem 8.3 and 8.5. We begin by remarking that the proof of Theorem 8.3 along with the continuation criterion Thm 8.1 require local well-posedness of the model equations:

Proposition 8.6 (Local Well-Posedness). Let $c_{\text{loc}} \geq 0$ and $c_{\text{nl}} \geq 0$. Assume that $p : \mathbb{R}^+ \to \mathbb{R}$ and $\mu : \mathbb{R}^+ \to \mathbb{R}^+$ are $C^\infty$ functions away from zero. Assume $s \in (0, 2)$ and let $\sigma := 2$ if $c_{\text{loc}} > 0$ and $\sigma := s$ if $c_{\text{nl}} > 0$ and $c_{\text{loc}} = 0$. Let $(u_0, \rho_0) \in H^{m+1-\sigma} \times H^m$, $m \geq 2$, such that $r_0 := \min_{x \in \mathbb{T}} \rho_0 > 0$. Suppose that for all $T > 0$

$$f \in L^2(0, T; H^{m-\sigma}(\mathbb{T})).$$

Then, there exists a time $T_0 > 0$ depending only on $\|(\rho_0, u_0)\|_{H^{m}(\mathbb{T}) \times H^{m+1-\sigma}(\mathbb{T})}$, $r_0$ and $f$, and a unique strong solution $(\rho, u)$ to (314)-(316) on $[0, T_0]$ with data $(\rho_0, u_0)$ such that

$$u \in C(0, T_0; H^{m+1-\sigma}(\mathbb{T})) \cap L^2(0, T_0; H^m(\mathbb{T}))$$

$$\rho \in C(0, T_0; H^m(\mathbb{T})) \cap L^2(0, T_0; H^{\sigma/2+m}(\mathbb{T}))$$

and $\rho(x, t) > \frac{m}{T}$ for all $(x, t) \in \mathbb{T} \times [0, T_0]$.

It should be remarked that this local well-posedness result covers all cases discussed in Theorems 8.1 and 8.3 but it holds in far greater generality. In particular, we do not require power-law forms for the pressure and viscosity constitutive laws. We do not produce a proof of Proposition 8.6 here, which is standard. In fact, the purely local case was established in Appendix II of [7]. On the other hand, the purely nonlocal case follows from a general proof which works in higher dimensions and is provided in Appendix A of [7], see also [14] for singular kernel pressureless case. The mixed case is a routine exercise, so is omitted. With local well-posedness in hand, we proceed with the proof of global well-posedness.

Proof of Theorem 8.3. By Prop. 8.6, we have a local strong solution on some interval $[0, T_0]$. We aim to show that $T_0$ may be taken infinite by establishing a lower bound on the density and appealing to the no-vacuum continuation criteria established in Theorem 8.1.

Let us recall from the previous section that either in the local-only or in hybrid cases we establish control over the local term $\int_\mathbb{T}^T \rho |Q_{\text{loc}}|^2 \, dx$ the entropy $\mathcal{H}_0$ and energy $\mathcal{E}$. Both are bounded uniformly in time due to (340) and (342), where we can estimate the force term by

$$\left| \int_\mathbb{T} f \rho u \, dx \right| \leq |f|_\infty \left( \mathcal{M} + \int_\mathbb{T} \rho |u|^2 \, dx \right) \leq C_1 + C_2 \mathcal{E},$$

$$\left| \int_\mathbb{T} f \rho X \, dx \right| \leq |f|_\infty \left( \mathcal{M} + \int_\mathbb{T} \rho |X|^2 \, dx \right) \leq C_1 + C_2 \mathcal{H}_0.$$ 

For $\mu(r) = r^\alpha$ with $\alpha < 1/2$ this implies

$$\int_\mathbb{T} \rho |Q_{\text{loc}}|^2 \, dx = \int_\mathbb{T} \rho |\rho^{\alpha-2} p_x|^2 \, dx = \int_\mathbb{T} |(\rho^{\alpha-\frac{1}{2}}) x|^2 \, dx = \|\rho^{\alpha-\frac{1}{2}}\|_{H^1} < \infty.$$ 

To establish pointwise bound we simply recall that the density has a conserved finite mass $\|\rho(t)\|_1 = \mathcal{M} > 0$. So, for each time $t$ there exist a point $x_0(t)$ such that $\rho(x_0(t), t) > \mathcal{M}/2$. We find

$$\int_{x_0(t)}^x (\rho^{\alpha-1/2})_x \, dx = \rho^{\alpha-1/2}(x, t) - \rho^{\alpha-1/2}(x_0(t)) \geq \rho^{\alpha-1/2}(x, t) - (\mathcal{M}/2)^{\alpha-1/2}.$$ 

It follows that $\rho^{\alpha-1/2} \in L_1^\infty L_x^\infty$ which implies $1/\rho \in L_1^\infty L_x^\infty$ since $\alpha < 1/2$. This finishes the proof. 

\square
Proof of Theorem 8.5. Due to the energy-entropy law elucidated in the previous section, (340), (342), and the specific form of enstrophy coming from the non-local dissipation, we obtain

\[
\frac{d}{dt} \mathcal{H}_0 \leq -c_1 \int_{T \times T} \phi_s(x-y)(\rho(y) - \rho(x))^2 \, dx \, dy
\]

(355)

\[
\frac{d}{dt} \mathcal{E} \leq -c_2 \int_{T \times T} \phi_s(x-y)|u(x) - u(y)|^2 \rho(x)\rho(y) \, dy \, dx.
\]

Note that under the linear pressure law (323), \(\phi_s = \phi_{s,\rho}\). Moreover, by the Gallilean invariance of the system and conservation of momentum, we may assume that the total momentum remains 0:

\[
\int_T \rho u \, dx = 0.
\]

Let us also assume \(\mathcal{M} = 1\). From the entropy dissipation we have

\[
\int_{T \times T} \phi_s(x-y)(\rho(y) - \rho(x))^2 \, dx \, dy \geq c_0 |\rho - 1|^2 \geq c_0 \int_T \rho \log \rho \, dx \geq c_0 \int_T \pi_0(\rho) \, dx.
\]

This shows that \(\int_0^\infty \int_T \pi_0 \, dx \, dt < \infty\). Then

\[
\frac{1}{2} \int_{T \times T} |u(x) - u(y)|^2 \rho(x)\rho(y) \, dy \, dx = \mathcal{M} \int_T \rho |u|^2 \, dx = c_1 \mathcal{E} - c_2 \int_T \pi_0 \, dx.
\]

Hence,

\[
\frac{d}{dt} \mathcal{E} \leq -c_1 \mathcal{E} + F(t),
\]

where \(F \in L^1(\mathbb{R}_+)\). By Duhamel, we obtain

\[
\mathcal{E}(t) \leq e^{-c_1 t} \mathcal{E}_0 + \int_0^t e^{-c_1 (t-s)} F(s) \, ds.
\]

The asymptotics of convergence \(\mathcal{E} \to 0\) is based on the convolution integral. We can estimate it as follows. Let us define a sequence of times by

\[
t_{m+1} = t_m + \frac{\lambda}{c_1} \ln m, \text{ for some } \lambda > 1.
\]

Then \(t_m \sim \frac{\lambda}{c_1} \ln(m!)\), and by Stirling approximation, \(t_m \sim m \ln m\). Let \(K = \int_0^\infty F(s) \, ds\). Then for every natural \(n \in \mathbb{N}\) there exists an \(m \in [n, (K+1)n]\) such that

\[
\int_{t_{m-1}}^{t_m} F(s) \, ds \leq \frac{1}{n}.
\]

Indeed, otherwise, \(\int F \, ds > K\). At time \(t_m\) we then have an estimate

\[
\mathcal{E}(t_m) \leq \frac{1}{m c_1 m} + \int_0^{t_{m-1}} e^{-c_1 (t_{m-s})} F(s) \, ds + \int_{t_{m-1}}^{t_m} e^{-c_1 (t_{m-s})} F(s) \, ds \leq \frac{1}{m c_1 m} + \frac{K}{m^\lambda} + \frac{1}{n}.
\]

But \(1/n \sim 1/m\) which appears to be the leading order term. Recalling that \(t_m \sim m \ln m\) we conclude that \(1/m \lesssim \ln t_m / t_m\). So, we have

\[
\mathcal{E}(t_m) \lesssim \frac{\ln t_m}{t_m}.
\]

For any other \(t_m < t < t_{m+1}\), we have by monotonicity of the energy

\[
\mathcal{E}(t) \leq \mathcal{E}(t_m) \lesssim \frac{\ln t_m}{t_m} \sim \frac{\ln t_{m+1}}{t_{m+1}} \leq \frac{\ln t}{t}.
\]

This establishes the desired asymptotic.
Finally, by the Csiszar-Kullback inequality, \( \int \pi_0 \, dx \geq |\rho - 1|^2 \), and (356), we obtain a flocking statement in the weighted \( L^2 \)-sense:

\[
\int_{T \times T} |u(x) - u(y)|^2 \rho(x) \rho(y) \, dy \, dx + |\rho - 1|^2 \lesssim \frac{\ln t}{t}.
\]

\[
\square
\]

8.3. **Continuation of non-vacuous solutions.** This section contains main technical ingredients of the hierarchy method, and provides the proof of Theorem 8.1.

Consider the evolution equations

\[
\begin{aligned}
\partial_t \rho + \partial_x (u \rho) &= 0, \\
\partial_t (\rho u) + \partial_x (\rho u^2) &= -\partial_x p(\rho) + c_{nl} D_{nl}(u, \rho) + c_{loc} D_{loc}(u, \rho) + \rho f
\end{aligned}
\]

where, recall, the nonlocal and local dissipative operators are defined as

\[
D_{nl}(u, \rho) := \rho [\mathcal{L}^s (\rho u) - u \mathcal{L}^s (\rho)], \quad D_{loc}(u, \rho) = (\mu(\rho) u_x)_x.
\]

We will be considering \( c_{nl} \geq 0 \) and \( c_{loc} \geq 0 \). The discussion of various cases requires us to break up our argument. In all cases, we consider the power-law for the pressure, i.e. \( p(\rho) = c_p \rho^\gamma \) for some \( \gamma > 0 \) and \( c_p > 0 \). In this case we have the explicit formula

\[
\pi_0(\rho) = c_p \rho \int_\rho^\bar{\rho} s^{\gamma - 2} \, ds = \begin{cases} \\
\frac{c_p}{\gamma - 1} \rho^{\gamma - 1} & \gamma > 1, \ \bar{\rho} = 0, \\
\frac{c_p}{\gamma - 1} \rho^{\gamma - 1} (1 - \rho^\gamma) & \gamma < 1, \ \bar{\rho} = 1. 
\end{cases}
\]

Thus, if \( \gamma > 1 \) then \( \pi_0(\rho) \geq 0 \) is non-negative pointwise. If \( \gamma = 1 \), which is of particular relevance in the context of flocking (323), then upon spatial integration it is non-negative by the Csiszar-Kullback inequality, i.e. \( \int \pi_0 \, dx \geq |\rho - \bar{\rho}|^2 \). This non-negativity will be repeatedly used for extracting information from the a priori estimates arising from the Bresch-Desjardins entropy balance. When \( \gamma < 1 \), we simply note that \( \int \pi_0 \, ds \) is bounded by the mass.

Assume that \((u, \rho)\) is a smooth solution on the time interval \([0, T^*)\) such that for any \( m \geq 2 \)

\[
\begin{aligned}
&u \in L^\infty(0, T; H^{m+1-s}) \cap L^2(0, T; H^m) \\
&\rho \in L^\infty(0, T; H^{m}) \cap L^2(0, T; H^{\sigma/2 + m})
\end{aligned}
\]

for any \( T < T^* \), where we have introduced

\[
\sigma = \begin{cases} \\
2 & c_{nl} \geq 0, c_{loc} > 0, \\
2 s & c_{nl} > 0, c_{loc} = 0.
\end{cases}
\]

Assume also that no vacuum has appeared

\[
\bar{\rho} := \inf_{t \in [0, T^*)} \min_{x \in \mathbb{T}} \rho(x, t) > 0.
\]

8.4. **Mass, Energy and \( \mathcal{H}_0 \)-entropy.** First, from the continuity equation (314) total mass is conserved

\[
\mathcal{M} = \|\rho(\cdot, t)\|_{L^1(\mathbb{T})} = \|\rho_0\|_{L^1(\mathbb{T})}.
\]

Recall next the basic energy balance (342), which in this case reads

\[
\frac{d}{dt} \mathcal{E} = -c_{nl} \int_{T \times T} \frac{1}{2} \phi_s(x - y) |u(x) - u(y)|^2 \rho(x) \rho(y) \, dy \, dx - c_{loc} \int_{T} \mu(\rho) |u_x|^2 \, dx + \int_{T} f \rho u \, dx.
\]

holds for any \( t \in [0, T^*) \). In the view of the estimate (353) this establishes uniform control in the energy space \( u \in L^\infty L^2 \cap L^2 H^{\sigma/2} \) on the given time interval.
The Bresch-Desjardins entropy (339) satisfies the balance (340), (343),

\[
\frac{d}{dt} H_0 = - c_{nl} \int_{\mathbb{T}^2} \frac{1}{2} \phi_s \rho(x, y) (\rho(y) - \rho(x))^2 \, dx \, dy - c_{loc} \int_{\mathbb{T}^2} \frac{\mu(\rho)p'(\rho)}{\rho} |\rho_x|^2 \, dx + \int_{\mathbb{T}} \rho f X \, dx.
\]

holds for any \( t \in [0, T^*] \). Again, using (353), we find that \( H_0 \) remains bounded. We now derive conclusions from these balance in two cases:

**Purely non-local dissipation:** Given the finite energy and controllability of \( \int \pi_0 \, dx \) discussed above, we obtain an \( L^2 \) bound on \( Q = \partial_x^{-1} L^s \rho \), hence, \( \rho \in L^\infty H^{s-1} \). Further condition coming from the dissipation term \( \rho \in L^2 H^{s/2} \) follows from the lower bound on the density and \( p' \) under the assumptions of Theorem 8.1. Provided \( s > 3/2 \), we have the embedding \( L^1 \cap \dot{H}^{s-1} = H^{s-1} \subset L^\infty \), and so, the density is uniformly bounded: \( \rho \in L^\infty(0, T; L^\infty) \).

**Purely local or mixed dissipation:** When the local or both components are active, the uniform bound on the density comes from several sources. Indeed, in light the non-negativity of the dissipation term \( \rho \in L^2 H^{s/2} \) follows from the lower bound on the density and \( p' \) under the assumptions of Theorem 8.3, we find

\[
\int_{\mathbb{T}} \rho |Q_{loc}|^2 \, dx = \|\rho^{\alpha - \frac{1}{2}}\|_{\dot{H}^{1}} < \infty.
\]

**Purely local or mixed dissipation:** When the local or both components are active, the uniform bound on the density comes from several sources. Indeed, in light the non-negativity of the mixed term (345) discussed in the introduction, we have control over the two terms \( \int \rho Q_{nl}^2 \, dx \) and \( \int \rho Q_{loc}^2 \, dx \) separately from the Bresch-Desjardins entropy. Boundedness of the density is then established in the following parameter regimes.

1. \( c_{nl} > 0 \) and \( c_{loc} \geq 0 \) with \( \gamma > 0 \), \( s > 3/2 \) and \( \alpha \geq 0 \). Then \( \rho \in L^\infty(0, T; L^\infty) \) by the non-local argument above.
2. \( c_{nl} \geq 0 \) and \( c_{loc} > 0 \) with \( \gamma > 0 \), \( s \in (0, 2) \) and \( \alpha > 1/2 \). Following the proof of Theorem 8.3, we find

\[
\int_{\mathbb{T}} \rho |Q_{loc}|^2 \, dx = \|\rho^{\alpha - \frac{1}{2}}\|_{\dot{H}^{1}} < \infty.
\]

Boundedness of the density follows, as in the Theorem, from the conserved finite mass.

3. \( c_{nl} \geq 0 \) and \( c_{loc} > 0 \) with \( \gamma > 1 \), \( s \in (0, 2) \) and \( \alpha > 0 \). This follows from the bound

\[
\int |\partial_x (\rho^{\frac{\gamma}{2}})| \, dx \leq \|\rho\|_{L^\gamma}^\gamma \left( \int \rho^{-3} |\rho_x|^2 \, dx \right)^{1/2} \leq c \|\rho\|_{L^\gamma}^\gamma \left( \int \rho Q_{loc}^2 \, dx \right)^{1/2} < \infty
\]

with a constant \( c := c(\rho) \). Thus, since \( \rho \in L^\gamma \) it follows that \( \rho^{\frac{\gamma}{2}} \in L^1 \) so that combined with the above we have \( \rho^{\frac{\gamma}{2}} \in L^\infty(0, T; W^{1,1}(\mathbb{T})) \). Boundedness follows from Sobolev embedding.

Once an upper bound on the density is established in any of the above cases, then the local part of the BD entropy gives the control \( \rho \in L^\infty(0, T; H^1) \). This control trumps what can be obtained from the dissipation of the BD entropy.

All the a priori bounds we have obtained so far in either of the cases can be summarized as follows:

\[
(368) \quad u \in L^\infty(0, T; L^2) \cap L^2(0, T; H^{s/2})
\]

\[
(369) \quad \rho \in L^\infty(0, T; L^\infty) \cap L^\infty(0, T; H^{s-1}) \cap L^2(0, T; H^{s/2})
\]

with norms depending on \( \rho \).

### 8.4.1. \( H_1 \)-entropy and its consequences.

We now work out a detailed balance relation for the \( H_1 \)-entropy. Let us recall the definitions

\[
\mathcal{H}_1 = \frac{1}{2} \int_{\mathbb{T}} \rho Y^2 \, dx + \int_{\mathbb{T}} \pi_1(\rho, \rho_x) \, dx
\]

\[
Y = \frac{1}{\rho} (u_x + Q_x), \quad \pi_1(\rho, \rho_x) = \frac{1}{2} h'(\rho) \rho_x^2.
\]
According to (347), (349) we have

\begin{equation}
\frac{d}{dt} H_1 = - \int_T h_{xx}(\rho) Y \, dx + \frac{d}{dt} \int_T \pi_1(\rho, \rho_x) \, dx + \int_T f_x Y \, dx.
\end{equation}

Looking ahead at the argument below, we remark that expansion of the pressure potential term \(\pi_1\) on the right hand side of this enstrophy budget will produce a dissipation term:

\[-c_m \|\rho\|_{H^{\sigma + 1}}^2 - c_{loc} \|\rho\|_{H^2}^2.\]

We will be using it repeatedly to absorb various residual terms by interpolation using uniform bounds on the density from below, above, and in \(H^{\sigma - 1}\) by (369).

So, first, let us estimate the forcing

\[\left| \int_T f_x Y \, dx \right| \leq c(\|f\|_{C^1}) + H_1.\]

Next, we expand the \(Y\)-term:

\begin{equation}
- \int_T h_{xx}(\rho) Y \, dx = - \int_T h'(\rho) \rho_{xx} Y \, dx - \int_T h''(\rho) \rho_x^2 Y \, dx.
\end{equation}

The second term can be estimated by

\begin{equation}
\left| \int_T h''(\rho) \rho_x^2 Y \, dx \right| \leq H_1^{1/2} |\rho_x|^2, \quad \text{and using that } |\rho_x| \leq \|\rho\|_{H^{\sigma/2+1}}, \quad \text{and } \frac{9-4\sigma}{4-\sigma} \leq 1 \text{ as long as } \sigma \geq \frac{5}{7}, \text{ we obtain}
\end{equation}

\begin{equation}
\left| \int_T h''(\rho) \rho_x^2 Y \, dx \right| \leq c(\varepsilon) H_1 + \varepsilon \|\rho\|_{H^{\sigma/2+1}}^2.
\end{equation}

In view of the remark above we can absorb the term \(\varepsilon \|\rho\|_{H^{\sigma/2+1}}^2\) into the upcoming dissipative contribution from the pressure potential. The remaining residual term \(- \int_T h'(\rho) \rho_{xx} Y \, dx\) will in fact be completely canceled out by another contribution of the pressure potential on which we focus for the remainder of the proof. First, we introduce a couple of shortcuts that greatly simplify the exposition.

- Throughout this proof, we will routinely drop integral signs for brevity. All equalities are intended to hold only upon spatially integrating over \(T\).
- We denote \(g(r) = \frac{1}{2} h'(r) / r^2\) so that \(\pi_1 = g(\rho) \rho_x^2\).

Let us compute the potential now

\[\frac{d}{dt} \pi_1 = -g' \rho_x^2 (u \rho)_x - 2g \rho_x (u \rho)_{xx} = g' \rho_x^2 (u \rho)_x + 2g \rho_{xx} (u \rho)_x \]

\[= 2g \rho \rho_{xx} u_x + g' \rho_x^2 u_x + g' \rho_x^3 u + 2g \rho_{xx} \rho_x u\]

integrating by parts in the last term

\[= 2g \rho \rho_{xx} u_x + g' \rho_x^2 u_x + g' \rho_x^3 u - g' \rho_x^3 u - g' \rho_x^2 u_x\]

\[= 2g \rho \rho_{xx} u_x + g' \rho_x^2 u_x - g' \rho_x^2 u_x.\]

(374)

The last two terms are of the form \(q(\rho) \rho_x^2 u_x\), where \(q\) is a smooth function on \(\mathbb{R}^+\). We can estimate any such term by replacing

\begin{equation}
\left| u_x \right| = \rho Y - c_{nl} L^s \rho - c_{loc}(\mu(\rho) \rho_x^2 / \rho^2)_x.
\end{equation}
The residual term \( q(\rho) \rho \rho_s^2 Y \) enjoys the same estimate as in (373). The local term breaks up into two: \( q(\rho) \rho_s^2 \rho_{xx} \) and \( q(\rho) \rho_s^3 \) for smooth \( q \) (which we redefine line by line). Interpolating between \( H^2 \) and \( H^1 \) we obtain

\[
\int_T q(\rho) \rho_s^2 \rho_{xx} \, dx \leq |\rho_x| \|\rho\|_{H^1} \|\rho\|_{H^2} \leq \varepsilon \|\rho\|_{H^2}^2 + c(\varepsilon),
\]

and using \( |\rho_x| \leq \|\rho\|_H^2 \) we have

\[
\int_T q(\rho) \rho_s^4 \, dx \leq |\rho_x| \|\rho\|_{H^2} \leq \varepsilon \|\rho\|_{H^2}^2 + c(\varepsilon).
\]

To estimate the non-local part \( q(\rho) \rho_s^2 \mathcal{L}^s \rho \), we symmetrize in the integral representation of \( \mathcal{L}^s \) and estimate according to the following

\[
\int_T q(\rho) \rho_s^2 \mathcal{L}^s \rho \, dx \leq |\rho_x| \|\rho\|_{H^{s/2}} + \|\rho_x| \|_{H^{s/2+1}} \|\rho\|_{H^{s/2}}.
\]

By the Gagliardo-Nirenberg inequality,

\[
|\rho_x| \leq \|\rho\|_{H^{s/2+1}} \|\rho\|_{H^{s/2-1}}, \quad \|\rho\|_{H^{s/2}} \leq \|\rho\|_{H^{s/2+1}} \|\rho\|_{H^{s/2-1}}.
\]

Thus,

\[
\int_T q(\rho) \rho_s^2 \mathcal{L}^s \rho \, dx \leq \|\rho\|_{H^{s/2+1}}^2 + \|\rho\|_{H^{s/2}}^2 \leq \varepsilon \|\rho\|_{H^{s/2}}^2 + c(\varepsilon),
\]

due to both powers being less than or equal to 2 as long as \( s > \frac{3}{2} \).

Finally, for the first term on the right hand side of (374) we have

\[
2g \rho \rho_{xx} u_x = 2g \rho^2 \rho_{xx} Y - 2c_{\text{nl}} g \rho \rho_{xx} \mathcal{L}^s \rho - 2c_{\text{loc}} g \rho \rho_{xx} (\mu(\rho) \rho_s / \rho^2)_x.
\]

Notice that \( 2g \rho^2 \rho_{xx} Y = h'(\rho) \rho_{xx} Y \), which cancels with the first term on the right hand side of (371). The main contribution of the last two terms is dissipation. Indeed, omitting constants and integrating by parts,

\[
-g \rho \rho_{xx} \mathcal{L}^s \rho = g \rho_s^2 \mathcal{L}^s \rho + g \rho_s \mathcal{L}^s \rho_x.
\]

The first one we already estimated. The second, after symmetrization, is bounded by

\[
g \rho_s \mathcal{L}^s \rho_x \leq -c_0 \|\rho\|_{H^{s/2+1}}^2 + |\rho_x| \|\rho\|_{H^{s/2+1}} \|\rho\|_{H^{s/2}},
\]

with the latter already being treated in (378). The last local term splits into

\[
-g \rho \rho_{xx} (\mu(\rho) \rho_s / \rho^2)_x = -q_1 |\rho_{xx}|^2 + q_2 |\rho_x|^2 \rho_{xx}, \quad q_1 > 0.
\]

With the use of (376), we estimate it by

\[
-g \rho \rho_{xx} (\mu(\rho) \rho_s / \rho^2)_x \leq -c \|\rho\|_{H^2}^2 + c(\varepsilon).
\]

Collecting the estimates we obtain

\[
\frac{d}{dt} \mathcal{H}_1 \leq c_1 \mathcal{H}_1 + c_2 - c_3 \|\rho\|_{H^{s/2+1}}^2.
\]

This implies a uniform bound on \( \mathcal{H}_1 \) on the time interval at question along with integrability of \( \|\rho\|_{H^{s/2+1}}^2 \). As a consequence of the positivity of \( \pi_1 \), we obtain \( \rho \in L^\infty H^1 \), and

\[
\rho Y = u_x + c_{\text{nl}} \mathcal{L}^s \rho + c_{\text{loc}} (\mu(\rho) \rho_s / \rho^2)_x \in L^\infty L^2.
\]

In the purely non-local case (\( c_{\text{loc}} = 0 \), if we apply \( 1 - s \) derivatives on this expression we still obtain a function in \( L^2 \), yet \( \rho_x \in L^2 \) by the previous. This places \( u \) into \( H^{2-s} \) uniformly. However
the \( L^2 \)-in-time class improves only to \( H^1 \). In the mixed or local cases, no further information is extracted from this computation. We obtain another series of a priori bounds:

\[
\begin{align*}
    u & \in L^\infty(0, T; H^{2-\sigma}) \cap L^2(0, T; H^1) \\
    \rho & \in L^\infty(0, T; H^1) \cap L^2(0, T; H^{\sigma/2+1})
\end{align*}
\]

where we identify \( L^2 = H^0 \). In addition, we record

\[
X \in L^\infty(0, T; H^1).
\]

Notice that in the local/mixed case, the only improvement coming from the \( H_1 \)-entropy at the level of the \( L^2 \)-in-time for \( \rho \), as well as the boundedness of the \( X \) quantity (384). The latter point is important in continuing our procedure.

### 8.4.2. \( \mathcal{H}_2 \)-entropy and its consequences.

Let us denote \( Z = \rho^{-1}Y_x \). Then \( Z \) satisfies

\[
D_t Z = -\rho^{-1}(\rho^{-1}h_{xx}(\rho))_x + \rho^{-1}(\rho^{-1}f_x)_x.
\]

We thus define our next entropy by

\[
\mathcal{H}_2 = \frac{1}{2} \int_T \rho Z^2 \, dx + \int_T \pi_2 \, dx
\]

\[
\pi_2 = \frac{1}{2} h'(\rho) \rho_{xx}^2.
\]

Note that

\[
\frac{d}{dt} \int_T \rho Z^2 \, dx = - \int_T (\rho^{-1}h_{xx})_x Z \, dx + \int_T (\rho^{-1}f_x)_x Z \, dx.
\]

The main term we need to eliminate is in fact \( \rho^{-1}h_{xx}Z \). The remaining ones coming from \( h_{xx} \) and \( \rho^{-1} \) end up being bounded by \( |\rho_x|^2 + |\rho_x||\rho_{xx}| \). This can be estimated by

\[
\int_T |Z|( |\rho_x|^2 + |\rho_x||\rho_{xx}|) \, dx \leq \mathcal{H}_2^{1/2}(|\rho_x|_{\infty} |\rho_{xx}|_2 + |\rho_x|^3).
\]

Keeping in mind that at this stage the dissipation term will be in \( H^{2+\sigma/2} \) (see discussion at the start of section of \( \mathcal{H}_1 \)-entropy). Moreover, the density is uniform in \( H^1 \), we obtain, by interpolation between \( H^1 \) and \( H^{2+\sigma/2} \),

\[
|\rho_x|_{\infty} |\rho_{xx}|_2 + |\rho_x|^3 \lesssim \|\rho\|_{H^{2+\sigma/2}}^3 + \|\rho\|^2_{H^{2+\sigma/2}}.
\]

Here, obviously, \( \frac{3}{\sigma+2} \leq 1 \) if \( \sigma = 2 \) or \( \sigma = s \) in our range of \( s \). Hence the term can be hidden in the dissipation,

\[
\int_T |Z|( |\rho_x|^2 + |\rho_x||\rho_{xx}|) \, dx \leq c(\varepsilon) \mathcal{H}_2 + \varepsilon \|\rho\|_{H^{2+\sigma/2}}^2.
\]

In the main term \( \rho^{-1}h_{xx}Z \) the worst part comes when all derivatives fall on the density:

\[
h_{xxx} = h'\rho_{xxx} + 3h''\rho_{x} + h'''\rho_x^3.
\]

Indeed, the term in the middle repeats (389), while the last one admits

\[
\int_T |Z||\rho_x|^3 \, dx \leq \mathcal{H}_2^{1/2} |\rho_x|^3 \leq \mathcal{H}_2^{1/2} \|\rho\|_{H^{2+\sigma/2}}^2 \leq c(\varepsilon) \mathcal{H}_2 + \varepsilon \|\rho\|_{H^{2+\sigma/2}}^2.
\]

Here and throughout we repeatedly use interpolation between \( H^1 \) and \( H^{2+\sigma/2} \) for the density terms. Thus the worst part of the original term \( - \int_T (\rho^{-1}h_{xx})_x Z \, dx \) gets reduced to just

\[
- \int_T \rho^{-1} h'(\rho) \rho_{xxx} Z \, dx.
\]
As on the $\mathcal{H}_1$-step we expect this term to be canceled by a contribution coming from the pressure potential $\pi_2$. Let us examine it next.

It will be convenient to denote $g(r) = \frac{1}{2} h'(r) / r^4$, and simply write $\pi_2 = g(\rho)\rho_x^2$. Integrating by parts we obtain

$$\frac{d}{dt} \int_T \pi_2 \, dx = g p x x \rho_{x x} u_{x x} + g' \rho_{x x} \rho_x u_{x x} - 2 g \rho_{x x} \rho_x u_{x x} - \frac{1}{2} g' \rho_x^2 u_{x x} - \frac{5}{2} g \rho_{x x}^2 u_x.$$  (392)

In the course of subsequent computations we will encounter a number of similar terms. They can be sorted into two groups – local ones involving $u$ and $X$, and non-local involving operator $\mathcal{L}^6$. All terms come with a prefactor of the form $q(\rho)$ for smooth $q$ which can be ignored. The local ones are

$$\rho_{x x}^2 u_x, \quad \rho_{x x} \rho_x^2 u_x, \quad \rho_{x x} \rho_x^2 X_x, \quad \rho_x^2 X_x;$$  (393)

$$\rho_x^2 \rho_{x x}^2, \quad \rho_x \rho_x^4, \quad \rho_{x x}^3.$$  (394)

the non-locals are $q(\rho)$-multiples of

$$\rho_{x x}^2 \mathcal{L}^6 \rho, \quad \rho_{x x} \rho_x^2 \mathcal{L}^6 \rho, \quad \rho_{x x} \mathcal{L}^6 \rho.$$  (395)

Substituting $u_x = \rho Y - c_m \mathcal{L}^6 \rho - c_m (\mu(\rho) \rho_x / \rho^2)_x$ in the first two local terms, we reduce it to the next locals and non-local ones. We now use interpolation and boundedness of $X_x$ in $L^2$ uniformly, to estimate the local terms as follows

$$\int_T |\rho_{x x}^2 X_x| \, dx \leq |X_x|^2 |\rho_{x x}|^2 \leq c_1 \|\rho\|_{H^{2+\sigma/2}}^{10} \leq c_2 + \varepsilon \|\rho\|_{H^{2+\sigma/2}}^2$$  (396)

$$\int_T |\rho_{x x} \rho_x^2 X_x| \, dx \leq |X_x|^2 |\rho_{x x}|^2 |\rho_x|_\infty \leq c_3 \|\rho\|_{H^{2+\sigma/2}}^{\frac{7}{2}} \leq c_4 + \varepsilon \|\rho\|_{H^{2+\sigma/2}}^2.$$  (397)

The other local terms are

$$\left| \int_T \rho_{x x} \rho_x^3 \, dx \right| \leq |\rho_{x x}|_{\infty} \|\rho\|_{H^1} \|\rho\|_{H^1} + |\rho_{x x}|_{\infty} \|\rho_x^2 \rho_x^2 \|_{H^1} \leq c_5 \|\rho\|_{H^3} + c_6 \|\rho\|_{H^3} \leq c_7 + \varepsilon \|\rho\|_{H^3}.$$  (398)

while which follows after noting that $\|\rho\|_{H^1}$ is uniformly bounded and $|\rho_{x x}|_{\infty} \leq c \|\rho\|_{H^3}$. In the next two terms we simply use Hölder inequality:

$$\left| \int_T \rho_{x x} \rho_x^2 \, dx \right| \leq |\rho_{x x}|_{\infty} \|\rho\|_{H^2} \|\rho\|_{H^2} = \|\rho\|_{H^3} \leq c_7 + \varepsilon \|\rho\|_{H^3}.$$  (399)

$$\left| \int_T \rho_{x x} \rho_x^3 \, dx \right| \leq |\rho_{x x}|_{\infty} \|\rho\|_{H^1} \|\rho\|_{H^1} \leq \|\rho\|_{H^3} \|\rho\|_{H^3} = \|\rho\|_{H^3} \leq c_7 + \varepsilon \|\rho\|_{H^3}.$$  (400)

Let us turn to non-local ones. In the first one we symmetrize in the operator $\mathcal{L}^6$, which produces increments of the other factors that come with it. Thus, we have

$$\int_T q(\rho) \rho_{x x}^2 \mathcal{L}^6 \rho \, dx \leq |\rho_{x x}|_{\infty} \|\rho\|_{H^{1/2+\sigma/2}} \|\rho\|_{H^{1/2+\sigma/2}} + |\rho_{x x}|_{\infty} \|\rho\|_{H^{1/2+\sigma/2}}^2$$  (401)

and noting that $\|\rho\|_{H^{1/2+\sigma/2}}$ is uniformly bounded and $|\rho_{x x}|_{\infty} \leq c \|\rho\|_{H^{1/2+\sigma/2}}$, we have

$$\leq c_5 \|\rho\|_{H^{1/2+\sigma/2}}^2 + c_6 \|\rho\|_{H^{1/2+\sigma/2}}^2 \leq c_7 + \varepsilon \|\rho\|_{H^{1/2+\sigma/2}}^2.$$  (402)
In the next we simply use Hölder inequality:

\[\left| \int_T q(\rho) \rho_{xx} \rho_x^2 \mathcal{L}^s \rho \, dx \right| \leq |\rho_{xx}| |\rho_x|^2 \|\mathcal{L}^s \rho\|_{H^s} \leq \|\rho\|_{\mathcal{H}^{s+2/2}}^{\frac{2}{3+s}} \|\rho_x\|_{\mathcal{H}^{s+2/2}}^{\frac{2}{3+s}} \|\mathcal{L}^s \rho\|_{\mathcal{H}^{s+2/2}}^{\frac{2s-2}{3+s}} \]

\[= \|\rho\|_{\mathcal{H}^{s+2/2}}^{\frac{2s+2}{3+s}} \leq c_7 + \varepsilon \|\rho\|_{\mathcal{H}^{s+2/2}}^{2}.
\]

The same strategy applies for the last one:

\[\left| \int_T q(\rho) \rho_{xx} \rho_x^2 \mathcal{L}^s \rho_x \, dx \right| \leq |\rho_{xx}| |\rho_x|^2 \|\mathcal{L}^s \rho_x\|_{H^{s+1}} \leq \|\rho\|_{\mathcal{H}^{s+2/2}}^{\frac{2}{3+s}} \|\rho\|_{\mathcal{H}^{s+2/2}}^{\frac{1}{3+s}} \|\mathcal{L}^s \rho_x\|_{\mathcal{H}^{s+2/2}}^{\frac{2s-2}{3+s}} \]

\[= \|\rho\|_{\mathcal{H}^{s+2/2}}^{\frac{2s+2}{2}} \leq c_7 + \varepsilon \|\rho\|_{\mathcal{H}^{s+2/2}}^{2}.
\]

Let us now get back to (392). The last two terms are obviously of same local type. The two terms in the middle are of the same type too. There we replace \( u_{xx} \)

\[(398)\quad u_{xx} = \rho^2 Z + \rho^{-1} \rho_x u_x + c_{loc}(\rho^{-1} \rho_x \mathcal{L}^s \rho - \mathcal{L}^s \rho_x) + c_{nl}(q_1 \rho_x^3 + q_2 \rho_x \rho_{xx} - q_3 \rho_{xxx})
\]

where \( q \)'s are some functions of \( \rho \), and most importantly, \( q_3 = \mu(\rho)/\rho^2 > 0 \). This results into \( \rho_{xx} \rho_x Z \), already estimated in (399); and the series of terms

\[\rho_{xx} \rho_{xx}^2 u_x, \; \rho_{xx} \rho_{xx}^2 \mathcal{L}^s \rho, \; \rho_{xx} \rho_x \mathcal{L}^s \rho_x, \; \rho_{xx} \rho_x^2 \mathcal{L}^s \rho_x, \; \rho_{xx} \rho_x^4, \; \rho_{xx}^3,
\]

all of which have been protocoted above. Finally, in the first and main term in (392) we use (398) to obtain

\[g \rho_{xxx} u_{xx} = \rho^{-1} h'(\rho) \rho_{xxx} Z + g \rho_{xxx} (\rho_x u_x + \rho_x \mathcal{L}^s \rho - \mathcal{L}^s \rho_x) + g \rho_{xxx} (q_1 \rho_x^3 + q_2 \rho_x \rho_{xx} - q_3 \rho_{xxx}),
\]

The first term is precisely the one that cancels with (391). In the rest, we integrate by parts to relieve one derivative from \( \rho_{xxx} \) in all but the final term above, which is strictly dissipative. In the local terms, integrate by parts to relieve one derivative from \( \rho_{xxx} \) and letting \( g_i(r) := g(r)q_i(r) \), we obtain

\[-g \rho_{xxx} \rho_x u_x = g \rho_{xx} \rho_x^2 u_x + g \rho_{xx} \rho_x \rho_{xx} u_{xx} + g \rho_{xx} \rho_x \rho_{xx} u_{xx}
\]

\[-g_1 \rho_{xxx} \rho_x^2 = g_1 \rho_{xx} \rho_x^4 + 3g_1 \rho_{xx} \rho_x^2,
\]

\[-g_2 \rho_{xxx} \rho_x \rho_{xx} = \frac{1}{2} g_2 \rho_{xx}^3 + \frac{1}{2} g_2 \rho_{xx}^3
\]

upon integration. We obtain (up to the opposite sign)

\[(399)\quad g' \rho_{xx} \rho_x^2 u_x + g \rho_{xx} \rho_x^2 u_x + g \rho_{xx} \rho_x u_{xx} + g \rho_{xx} \rho_x^2 \mathcal{L}^s \rho + g \rho_{xx} \mathcal{L}^s \rho + g \rho_{xx} \rho_x \mathcal{L}^s \rho_x
\]

\[-g \rho_{xx} \rho_x \mathcal{L}^s \rho_x - g \rho_{xx} \rho_x \mathcal{L}^s \rho_x - g \rho_{xx} \rho_x \mathcal{L}^s \rho_{xx}
\]

\[+ g \rho_{xx} \rho_x u_{xx} + g_1 \rho_{xx} \rho_x^4 + \left( 3g_1 + \frac{1}{2} g_2 \right) \rho_{xx} \rho_x^2 + \frac{1}{2} g_2 \rho_{xx}^3 - g_3 \rho_{xx}^2.
\]

All of the terms have been included in the lists, except \( g \rho_{xx} \rho \mathcal{L}^s \rho_{xx} \) and \( g q_3 \rho_{xx}^2 \) which are dissipative. This is obvious for the local term:

\[(400)\quad \int_T g_3(\rho) \rho_{xx}^2 \, dx \geq c_1 \|\rho\|_{\mathcal{H}^3}^2.
\]
As to the non-local term, we perform the same argument – by symmetrization, and noting that variation of \( \rho g(\rho) \) results in a variation of \( \rho \), we find

\[
\int_T \rho g(\rho) \rho_{xx} \mathcal{L}^s \rho_{xx} \, dx \leq -c_1 \|\rho\|_{\dot{H}^{2+s/2}}^2 + \|\rho_{xx}\|_{\dot{H}^{1/2}} \|\rho\|_{\dot{H}^{2+s/2}}
\]
since \( \dot{H}^{s/2} \) norm is uniformly bounded,

\[
\leq -c_1 \|\rho\|_{\dot{H}^{2+s/2}}^2 + \|\rho\|_{\dot{H}^{2+s/2}}^{\frac{2}{\alpha} + 1} \leq -c_8 \|\rho\|_{\dot{H}^{2+s/2}}^2 + c_9.
\]

Collecting the obtained estimates we arrive at

\[
\frac{d}{dt} \mathcal{H}_2 \leq -c' \|\rho\|_{\dot{H}^{2+s/2}}^2 + c'' \mathcal{H}_2 + c'''.
\]

This proves uniform boundedness of \( \mathcal{H}_2 \) and \( \rho \in L^2 \dot{H}^{2+s/2} \). Also from the corrector we obtain \( \rho \in L^\infty \dot{H}^2 \). Since \( Z \in L^2 \) this translates into \( u_{xx} + \partial_x \mathcal{L}^s \rho + \rho_{xxx} \in L^2 \) uniformly. This puts \( u \in L^\infty \dot{H}^{3-s} \cap L^2 \dot{H}^2 \). Collectively we obtain

\[
\begin{align*}
(402) & \quad u \in L^\infty(0,T;H^{3-s}) \cap L^2(0,T;H^2), \\
(403) & \quad \rho \in L^\infty(0,T;H^2) \cap L^2(0,T;H^{s/2+2}).
\end{align*}
\]

8.4.3. \( \mathcal{H}_n \)-entropy: closing the argument. Let us note that in the previous calculations the requirements on \( s \) relax to just \( s > 1 \). It is now clear that can construct a hierarchy of higher order entropies in the form

\[
\mathcal{H}_n = \frac{1}{2} \int_T \rho Z_n^2 \, dx + \int_T \pi_n \, dx
\]

(404)

\[
\pi_n = \frac{1}{2} h'(\rho) \left( \frac{\partial^{n}_x \rho}{\rho}\right)^2,
\]

where at the core of \( Z_n \) is the term \( \partial^n_x u + c_{n0} \partial^{n-1}_x \mathcal{L}^s \rho + c_{n00} q(\rho) \partial^{n+1}_x \rho \). The argument above extends easily with identical steps to this general case. As a result we obtain uniform boundedness of the \( n \)-th entropy and corresponding \( L^2 \)-integrability of the enstrophy. This puts our solution into the classes

\[
\begin{align*}
(405) & \quad u \in L^\infty(0,T;H^{n+1-s}) \cap L^2(0,T;H^n), \\
(406) & \quad \rho \in L^\infty(0,T;H^n) \cap L^2(0,T;H^{s/2+n}).
\end{align*}
\]

We thus have shown that

\[
\sup_{T \in [0,T^*]} \|\rho\|_{L^\infty(0,T;H^n)} + \sup_{T \in [0,T^*]} \|\rho\|_{L^2(0,T;H^{s/2+n})} + \sup_{T \in [0,T^*]} \|u\|_{L^\infty(0,T;H^{n+1-s})} + \sup_{T \in [0,T^*]} \|u\|_{L^2(0,T;H^n)}
\]

\[
\leq F(\|\rho_0, u_0\|_{H^n(T) \times H^{n+1-s}(T)}, \|f\|_{L^\infty(0,T^*;C^n)}, \frac{1}{L}, T^*) < \infty
\]

for \( n \geq 2 \). Appealing to local existence, established by Prop. 8.6, the solution can be extended past \( T^* \).
References


University of Illinois at Chicago, 60607
Email address: shvydkoy@uic.edu