Microscopic, mesoscopic, and macroscopic descriptions of the Euler alignment system with adaptive communication strength

Roman Shvydkoy [1] and Trevor Teolis [1]

[1] *Department of Mathematics, Statistics and Computer Science, University of Illinois at Chicago, Chicago, IL 60607, USA*

Abstract. This a continuation of our previous joint work on s-model in [*Well-posedness and long time behavior of the Euler Alignment System with adaptive communication strength.* arXiv preprint arXiv:2310.00269, 2023]. The s-model, introduced by the first author in [*Environmental averaging.* arXiv preprint arXiv:2211.00117, 2022], is an alignment model with the property that the strength of the alignment force, s, is transported along an averaged velocity field. The transport of the strength is designed so that it admits an e-quantity, $e = \partial_x u + \mathrm{s}$, which controls regularity in 1D analogously to the classical Cucker-Smale case. The utility of the s-model is that it has the versatility to behave qualitatively like the Motsch-Tadmor model, for which global regularity theory is not known.

This paper aims to put the s-model on more firm physical grounds by formulating and justifying the microscopic and mesoscopic descriptions from which it arises. A distinctive feature of the microscopic system is that it is a discrete-continuous system: the position and velocity of the particles are discrete objects, while the strength is an active continuum scalar function. We establish a rigorous passage from the microscopic to the mesoscopic description via the Mean Field Limit and a passage from the mesoscopic to the macroscopic description in the monokinetic and Maxwellian limiting regimes. We present a survey of such results for the Cucker-Smale model and explain how to extend these arguments to the s-model, where the strength of the alignment force is transported. We also address the long-time behavior of the kinetic Fokker-Planck-Alignment equation by establishing the relaxation to the Maxwellian in 1D when the velocity averaging is the Favre averaging. As a supplement to the numerical results already presented in our previous work, we provide additional numerical evidence, via a particle simulation, that the s-model behaves qualitatively like Motsch-Tadmor.

Contents

E-mail address: shvydkoy@uic.edu, tteoli2@uic.edu.

Date: September 6, 2024.

1991 *Mathematics Subject Classification.* 35Q35, 35Q83, 35Q84 37A60, 92D25.

Key words and phrases. mean field limit, hydrodynamic limit, collective behavior, alignment, Cucker-Smale model, Motsch-Tadmor model, environmental averaging.

Email for correspondence: tteoli2@uic.edu.

Acknowledgment. This work was supported in part by NSF grant DMS-2405326.

1. Introduction

Collective patterns and behavior can emerge from simply interacting agents. Flocks of birds, schools of fish, and flashing of fireflies are examples of systems of simply interacting agents where there is a clear emergent behavior of the whole system. This has spurred much interest in recent years in the mathematical modeling of this large class of phenomena. The question is: how do we create a model that describes the behavior of a large class of communicating agents and is mathematically tractable in the sense of well-posedness analysis and provability of the emergent phenomena?

Many classical alignment models possess some of these features. The 3-zone model was developed in 1987 by Craig Reynolds, a computer animator, who was interested in developing realistic models of birds for cinema, see [14]. It is characterized by a repulsion force in the close range, alignment in the medium range, and attraction in the far range. The repulsion force makes it challenging to describe the long-time behavior due to the lack of a natural equilibrium state. The Viscek model proposed in [22] incorporates stochastic forces that facilitate phase transitions as the strength of noise varies – an important feature of emergent dynamics. It evolves in discrete time steps whereby agents get assigned the average and normalized velocities of nearby agents. Phase transitions were observed numerically, however the model at the moment lacks a satisfactory analysis, see [23]. In 2007 [4, 5] Cucker and Smale proposed another alignment model with the following key features: it weights the alignment force with a radial interaction kernel ϕ inversely proportional to the distance between agents. The model allows for an array of different dynamics and, most importantly, it was the first model that allowed a proof of alignment which depended only on the initial state and not on perpetual connectivity of the flock. Moreover, the long-time behavior holds under the large crowd limit $N \to \infty$, which allows to carry such results over to the kinetic and macroscopic descriptions, [3, 20]. The universality of description and mathematical tractability of the model made it the subject of a series of studies surveyed for example in

[1, 21, 15, 11]. Despite its success, however, the Cucker-Smale model does not respond well in certain modeling scenarios. Motsch and Tadmor argued in [12] that in heterogeneous formations when, say, a small subflock separates itself from a distant large flock, its internal forces become annihilated by the latter if subjected to the Cucker-Smale protocol [12] creating unrealistic behavior. A proposed renormalization of the averaging operation restored the balance of forces and the alignment results under long range communication were proved similar to those of Cucker and Smale at the microscopic level and in the large crowd limit, [13]. Such renormalization, however, destroys the symmetry of the alignment force and possible vacuum formation makes the system more singular, which is the reason for the lack of a coherent well-posedness theory at the moment.

The s-model has been designed to have an adaptive averaging protocol which allows for many similar regularity properties as those of the Cucker-Smale model, while maintaining the same level of universality as that of the Motsch-Tadmor model. One downside of the adaptive as opposed to prefixed protocol is the lack of control over kinematic properties of the alignment force, which makes the analysis of the long-time behavior a challenging but interesting problem, see [16, 19].

This paper aims to (1) introduce the microscopic and mesoscopic counterparts to the previously studied macroscopic description of the s-model in [19]; (2) to establish a rigorous passage between the levels of description; and (3) to establish the relaxation to a thermodynamic state in one dimension for the mesoscopic description when the velocity averaging is the Favre averaging.

Before we turn to the technical description of the results let us present a more detailed motivation of the s-model by giving a brief survey of the classical Cucker-Smale and Motsch-Tadmor alignment models as well as the more general environmental averaging models from which it originated.

1.1. **Cucker-Smale model.** The pressureless Euler alignment system based on the Cucker-Smale model on $\mathbb{T}^n$ or $\mathbb{R}^n$ is given by

$$\begin{cases} \partial_t \rho + \nabla \cdot (\mathbf{u}\rho) = 0 \\ \partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} = (\mathbf{u}\rho)_\phi - \mathbf{u}\rho_\phi \end{cases} \tag{1}$$

where ρ, $\mathbf{u}$ are the density and velocity of the flock and $\phi \geqslant 0$ is a smooth radially decreasing communication kernel. The notation $f_\phi = f * \phi$ denotes a convolution. Provided the kernel has a fat tail (i.e. is non-integrable),

$$\int_0^\infty \phi(r) dr = \infty \tag{2}$$

this model admits alignment and flocking: there exists $\delta > 0$ such that

$$\lim_{t\to\infty} \operatorname{diam} \operatorname{supp}(\rho) < \infty, \qquad \lim_{t\to\infty} \sup_{x,y\in \operatorname{supp}(\rho)} |\mathbf{u}(t,x) - \mathbf{u}(t,y)| \leqslant C e^{-\delta t}$$

It is a well-known result of Carillo, Choi, Tadmor, and Tan [2] that

$$e = \partial_x u + \rho_\phi, \qquad \partial_t e + \partial_x (ue) = 0$$

provides a threshold regularity criterion in 1D: the solution remains smooth iff $e_0 \geqslant 0$. In fact, it is not only useful for global regularity; it is also instrumental in proving 1D strong flocking (i.e. convergence to a limiting distribution) and estimating the limiting distribution, [10, 18].

1.2. **Motsch-Tadmor model.** It was observed in [12] that the Cucker-Smale model displays unrealistic behavior in heterogeneous formations. In particular, when there is a small mass flock and a faraway, large mass flock, the internal dynamics of the small mass flock are hijacked by the large mass flock. The Motsch-Tadmor model was introduced in order to resolve this behavior. It was proposed that the strength of the alignment force on a particle should be scaled by the total influence on that particle. With this modification, we get:

$$\begin{cases} \partial_t \rho + \nabla \cdot (\mathbf{u}\rho) = 0 \\ \partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} = \frac{1}{\rho_\phi}((\mathbf{u}\rho)_\phi - \mathbf{u}\rho_\phi) \end{cases} \tag{3}$$

We will explain what this fix does to restore the balance of forces in Section 1.6.1.

It was shown in [13] that (3) aligns under the same fat tail kernel condition (2). The cost is that it no longer possesses the e-quantity and, as a result, there is no known 1D threshold regularity criterion. The s-model aims to fix this lack of a threshold regularity criterion while also retaining the desired qualitative behavior of the Motsch-Tadmor model in heterogeneous formations. Before we introduce it, we will mention the environmental averaging models from which it originated.

1.3. **Environmental Averaging Models.** Many alignment models are characterized by an alignment force which pushes the velocity towards an averaged velocity field. The Vicsek, Cucker-Smale, and Motsch-Tadmor models are among many examples. For the Cucker-Smale model, we can rewrite the alignment force as

$$F = s_\rho([\mathbf{u}]_\rho - \mathbf{u}), \qquad \mathrm{s}_\rho = \rho_\phi, \;\; [\mathbf{u}]_\rho = \frac{(\mathbf{u}\rho)_\phi}{\rho_\phi}.$$

The so-called strength s_ρ of the alignment force is a fixed function which depends on the density; and $[\mathbf{u}]_\rho$ is an averaged velocity field. Written in this form, it is more clear that it is an alignment model– the velocity $\mathbf{u}$ is being pushed towards the averaged velocity $[\mathbf{u}]_\rho$. Alignment forces which take this more general form are called environmental averaging models and the macroscopic version is given by:

$$\begin{cases} \partial_t \rho + \nabla \cdot (\mathbf{u}\rho) = 0 \\ \partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} = \mathrm{s}_\rho([\mathbf{u}]_\rho - \mathbf{u}). \end{cases} \tag{4}$$

They generalize the Cucker-Smale model by treating s_ρ and $[\mathbf{u}]_\rho$ as arbitrary modeling components in a proper functional framework. The general theory of environmental averaging models has been developed extensively in [16].

An important class of models are those for which the velocity averaging $[\mathbf{u}]_\rho$ has an integral representation against a smooth kernel Φ_ρ:

$$[\mathbf{u}]_\rho = \int \Phi_\rho(x,y)\mathbf{u}(y)\rho(y)\,\mathrm{d}y, \qquad \Phi_\rho \text{ is smooth.} \tag{5}$$

Remark 1.1. Models whose velocity averaging has the form (5) will be an important sub-class for the as-yet defined s-model since the regularity of the model is tied to the regularity of the kernel $\Phi_\rho(x, y)$.

The representation (5) holds for many classical models: Cucker-Smale (M_{CS}), Motsch-Tadmor (M_{MT}), the overmollified version of (M_{MT}) introduced in [17],

$$[\mathbf{u}]_\rho = \left(\frac{(\mathbf{u}\rho)_\phi}{\rho_\phi}\right)_\phi, \qquad \mathrm{s}_\rho = 1, \tag{M_ϕ}$$

the segregation model based on a partition of unity $\sum_{l=1}^{L} g_l = 1$,

$$[\mathbf{u}]_\rho = \sum_{l=1}^{L} g_l(x) \frac{\int_\Omega u g_l \rho \, \mathrm{d}y}{\int_\Omega g_l \rho \, \mathrm{d}y}, \qquad \mathrm{s}_\rho = 1, \tag{M_{seg}}$$

and other multi-flock, multi-species variants of the above, see [16]. All of these models possess a representation kernel as shown in table 1.3.

	Strength s_ρ	Kernel $\Phi_\rho(x,y)$
M_{CS}	ρ_ϕ	$\frac{\phi(x-y)}{\rho_\phi(x)}$
M_{MT}	1	$\frac{\phi(x-y)}{\rho_\phi(x)}$
M_ϕ	1	$\int_\Omega \frac{\phi(x-z)\phi(y-z)}{\rho_\phi(z)} \, \mathrm{d}z$
M_{seg}	1	$\sum_{l=1}^{L} \frac{g_l(x) g_l(y)}{\int_{\mathbb{T}^n} g_l \rho \, \mathrm{d}x}$

We can see that many of the models have the density-dependent renormalization at its core, $\mathbf{u}_F := \frac{(\mathbf{u}\rho)_\phi}{\rho_\phi}$. This is known as the Favre filtration, which was introduced in the context of non-homogeneous turbulence in [6]. In our context, we will refer to it as the "Favre averaging" since it represents the averaged velocity.

Definition 1.2. The *Favre-averaged* velocity is given by

$$\mathbf{u}_F := \frac{(\mathbf{u}\rho)_\phi}{\rho_\phi}$$

1.4. **The s-model.** The s-model is a descendant of the environmental averaging models and was introduced in [16] at the macroscopic level. The only adendum to the environmental averaging models is that s is no longer a fixed function of the density, but instead evolves according to its own evolution equation along the average velocity field. This has a profound impact on the 1D regularity theory at the macroscopic level of description, and subsequently it lends to stronger long-time behavior results, all of which resemble the favorable behavior of the Cucker-Smale model in this more general setting.

It was observed in [16] that the existence of the e-quantity in the Cucker-Smale case is owed to the transport of the strength function, $\mathrm{s}_\rho = \rho_\phi$, along the Favre-averaged velocity field:

$$\partial_t \rho_\phi + \partial_x(\rho_\phi \mathbf{u}_F) = 0.$$

The new model therefore proposes that is more natural for the strength s to evolve according to its own transport equation along the averaged velocity field:

$$\partial_t \mathrm{s} + \partial_x(\mathrm{s}[\mathbf{u}]_\rho) = 0.$$

The s-model is given by

$$\begin{cases} \partial_t \rho + \nabla \cdot (\mathbf{u}\rho) = 0 \\ \partial_t \mathrm{s} + \nabla \cdot (\mathrm{s}[\mathbf{u}]_\rho) = 0, \qquad \mathrm{s} \geqslant 0 \\ \partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} = \mathrm{s}([\mathbf{u}]_\rho - \mathbf{u}) \end{cases} \tag{SM}$$

and, by design, it admits the e-quantity in 1D

$$e = \partial_x u + \mathrm{s}, \quad \partial_t e + \partial_x (ue) = 0.$$

The special case where the velocity averaging is the Favre averaging, i.e. $[\mathbf{u}]_\rho = \mathbf{u}_F$ (see Definition 1.2), bears a similar alignment force to the classical Cucker-Smale and Motsch-Tadmor alignment models. Indeed, we can cast it into a similar form by introducing a new variable w, which we call the "weight", defined by $\mathrm{s} = \mathrm{w}\rho_\phi$. With this change of variables, w satisfies a pure transport along the Favre-averaged velocity field

$$\partial_t \mathrm{w} + \mathbf{u}_F \cdot \nabla \mathrm{w} = 0$$

and the system becomes

$$\begin{cases} \partial_t \rho + \nabla \cdot (\mathbf{u}\rho) = 0 \\ \partial_t \mathrm{w} + \mathbf{u}_F \cdot \nabla \mathrm{w} = 0, \qquad \mathrm{w} \geqslant 0 \\ \partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} = \mathrm{w}((\mathbf{u}\rho)_\phi - \mathbf{u}\rho_\phi) \end{cases} \tag{WM}$$

We obtain a model that looks like a hybrid of Cucker-Smale and Motsch-Tadmor. We refer to this particular variant of the s-model as the w-model. Setting $\mathrm{w}_0 = 1$, we recover the Cucker-Smale model; and setting $\mathrm{w}_0 = 1/(\rho_0)_\phi$, we recover the Motsch-Tadmor model at the initial time (at later times, the transport of the weight w will cause it to deviate from the Motsch-Tadmor weight). Due to the existence of the e-quantity and the similar structure of the alignment force to the Cucker-Smale model, many classical results of the Cucker-Smale case were extended to the w-model in our joint work [19]. The following is a list of the results which were obtained:

(a) Local well-posedness and global well-posedness for small data in multi-D;
(b) Global well-posedness in 1D and for unidirectional flocks in multi-D under the threshold $e \geqslant 0$;
(c) L^∞-based alignment; conditional L^2-based alignment;
(d) Strong flocking in 1D (i.e. convergence of the density to a limiting distribution);
(e) Estimates on the limiting distribution of the flock in 1D.

The importance of the w-model is that it retains the regularity and alignment characteristics of the Cucker-Smale system while also having the versatility to behave qualitatively like the Motsch-Tadmor model (when $\mathrm{w}_0 = 1/(\rho_0)_\phi$). In other words, the w-model can be thought of as a more analytically tractable version of the Motsch-Tadmor model. Numerical evidence of the qualitative similarities between the w-model and the Motsch-Tadmor model has been presented at the macroscopic level in 1D in [19]. Here, we present, perhaps more clear, numerical evidence of this qualitative behavior by showcasing simulations of the microscopic system in Section 1.6. First, we introduce the microscopic and mesoscopic versions of the s-model, and hence also the w-model as a particular case.

1.5. **Microscopic and mesoscopic levels of description.** The s-model was introduced in [16] at the level of the hydrodynamic description, while the microscopic and mesoscopic levels have remained unattended. To fill this gap, we introduce and justify them here. Due to its ancestral relation to the Cucker-Smale model, the microscopic and mesoscopic systems for the s-model will have a similar structure to the Cucker-Smale versions, albeit with important differences. As a reference point for the reader, we will first present the microscopic and mesoscopic descriptions for the Cucker-Smale model followed by those for the s-model, highlighting the differences between the two.

Remark 1.3. We could have also juxtaposed the s-model to the environmental averaging models, which also have a well developed theory at all levels of description.

1.5.1. *Cucker-Smale: microscopic and mesoscopic descriptions.* The classical discrete Cucker-Smale model, originally introduced by Cucker and Smale in [4], is given by an ODE system of N interacting agents:

$$\begin{cases} \dot{x}_i = v_i \\ \dot{v}_i = \lambda \sum_{j=1}^{N} m_j \phi(x_i - x_j)(v_j - v_i). \end{cases} \tag{6}$$

where x_i, v_i are the velocity and positions of the agents and $\lambda > 0$ is a scalar that affects the strength of the alignment force. Of course, it aligns and flocks under the same fat tail condition on the kernel, (2). As the number of particles $N \to \infty$, it is more convenient to look at the evolution of the probability density, $f(t, x, v)$, of finding a particle at position x and velocity v at time t instead of tracking individual particle trajectories. The evolution equation for the probability density, $f(t, x, v)$, is given by a kinetic Vlasov equation derived according to the BBGKY hierarchy [9]:

$$\partial_t f + v \cdot \nabla_x f + \lambda \nabla_v \cdot [f F(f)] = 0, \tag{7}$$
$$F(f)(t, x, v) = \int_{\mathbb{R}^{2n}} \phi(|x - y|)(w - v) f(t, y, w) \, \mathrm{d}w \, \mathrm{d}y.$$

The passage from the (6) to (7) in the scaling regime where the range of the communication between particles remains independent of N, i.e. the *mean field limit*, was established in [8].

In practice, one cannot physically observe the probability density function. The physically observable (or macroscopic) variables are the density and the momentum, which are defined by

$$\rho(t, x) = \int_{\mathbb{R}^n} f(t, x, v) \, \mathrm{d}v, \qquad (\rho \mathbf{u})(t, x) = \int_{\mathbb{R}^n} v f(t, x, v) \, \mathrm{d}v. \tag{8}$$

To determine the evolution of macroscopic variables, one computes the v-moments of the kinetic Vlasov equation and obtains

$$\begin{cases} \partial_t \rho + \nabla \cdot (\mathbf{u} \rho) = 0 \\ \partial_t (\rho \mathbf{u}) + \nabla \cdot (\rho \mathbf{u} \otimes \mathbf{u}) + \nabla \cdot \mathcal{R} = \rho((\mathbf{u}\rho)_\phi - \mathbf{u} \rho_\phi), \end{cases} \tag{9}$$

where $\mathcal{R}$ is the Reynolds stress tensor

$$\mathcal{R}(t, x) = \int_{\mathbb{R}^n} (v - \mathbf{u}(t, x)) \otimes (v - \mathbf{u}(t, x)) f(t, x, v) \, \mathrm{d}v. \tag{10}$$

This results in the closure problem: the system (9) depends not only on the macroscopic variables ρ and $\mathbf{u}$, but also on the kinetic distribution f. The system can be closed by adding a local alignment force which pushes the distribution f to be concentrated around the macroscopic velocity and thus removes the dependence on f. Two such alignment forces have been considered in the literature corresponding to the monokinetic and Maxwellian limiting regimes. For background, we will survey the monokinetic and Maxwellian limits for the Cucker-Smale model in Section 3.2.

Remark 1.4. The macroscopic pressureless Euler alignment system (1) arises from the monokinetic limiting regime. The as-yet defined macroscopic isentropic Euler alignment system (34) (discussed in Section 3.2) arises from the Maxwellian limit.

1.5.2. s-*model: microscopic and mesoscopic descriptions.* The microscopic s-model stands uniquely from the other microscopic alignment models as a discrete-continuous system. The strength satisfies its transport equation along the averaged velocity field and it is therefore not a discrete object. Instead, it is an active continuum scalar function that is transported along the continuous field $[\mathbf{u}^N]_{\rho^N}(\cdot, x)$. As far as the alignment force is concerned, it only keeps track of s and $[\mathbf{u}^N]_{\rho^N}$ at the discrete points x_i.

The discrete density $\rho^N \in \mathcal{P}_M(\mathbb{T}^n)$ is given by $\rho^N = \sum_{j=1}^N m_j\delta_{x_j}$ and the discrete velocity is given by $u^N = \sum_{j=1}^N v_j \mathbb{1}_{x_j}$. The microscopic description of (SM) is then given by an ODE system describing the position and velocity of particles coupled with a PDE describing the transport of the strength function:

$$\begin{cases} \dot{x}_i = v_i \\ \dot{v}_i = \lambda \mathrm{s}(x_i)([\mathbf{u}^N]_{\rho^N}(x_i) - v_i), \qquad \rho^N = \sum_{j=1}^N m_j\delta_{x_j}, \;\; u^N = \sum_{j=1}^N v_j \mathbb{1}_{x_j}, \\ \partial_t \mathrm{s} + \nabla_x \cdot (\mathrm{s}[\mathbf{u}^N]_{\rho^N}) = 0. \end{cases} \tag{11}$$

Once again, $\lambda > 0$ is a scalar that affects the strength of the alignment force. Although $[\mathbf{u}^N]_{\rho^N}$ appears to be a discrete object, observe that when the velocity averaging satisfies the integral representation (5), the velocity averaging is given by

$$[\mathbf{u}^N]_{\rho^N}(t,x) = \sum_{j=1}^N m_j \Phi_{\rho^N}(x, x_j) v_j(t). \tag{12}$$

In this case, $[\mathbf{u}^N]_{\rho^N}$ is a smooth object whenever Φ_{ρ^N} is smooth. The well-posedness of this discrete-continuous system for smooth kernels Φ_{ρ^N} is proved in Section 4.

The corresponding mesoscopic description is given by a Vlasov-type equation coupled with a transport of the strength function:

$$\begin{cases} f_t + v \cdot \nabla_x f + \lambda \nabla_v(\mathrm{s}(v - [\mathbf{u}]_\rho) f) = 0 \\ \partial_t \mathrm{s} + \nabla_x \cdot (\mathrm{s}[\mathbf{u}]_\rho) = 0. \end{cases} \tag{13}$$

For the mesoscopic case, when the velocity averaging satisfies the integral representation (5), the velocity averaging is given by

$$[\mathbf{u}]_\rho(x) = \int_{\mathbb{T}^n} \Phi_\rho(x,y) \int_{\mathbb{R}^n} v f(t,y,v) \,\mathrm{d}v \,\mathrm{d}y.$$

From table 1.3, we see that the discrete Cucker-Smale model (6) can be recovered by setting $\mathrm{s} = \rho_\phi^N$ and $\Phi_\rho(x,y) = \phi(x-y)/\rho_\phi(x)$. Since ρ_ϕ automatically satisfies the transport along the averaged velocity, the strength equation drops out. For the kinetic system, (7) is recovered similarly by setting $\mathrm{s} = \rho_\phi$ and $\Phi_\rho(x,y) = \phi(x-y)/\rho_\phi(x)$.

For the s-model, the mean field passage from the microscopic description (11) to the mesoscopic description (13) necessitates some uniform regularity on s and $[\mathbf{u}]_\rho$. Let us make the following observations of the Cucker-Smale model, which will influence how we handle the mean field passage for the s-model:

- the regularity of the strength comes for free when the communication kernel ϕ is smooth, and
- the regularity of the velocity averaging is tied to the regularity of the kernel Φ_ρ .

For the s-model, the regularity of the strength will not come for free, but we will see later that it is also tied to the regularity of the kernel Φ_ρ. Therefore, instead of imposing regularity conditions on the solution, we will subordinate the uniform regularity assumptions to the kernel Φ_ρ, (Reg1) and (Reg2), which are sufficient for achieving the desired regularity on the velocity averaging and the strength, provided that the L^1 norm of the momentum is bounded. The mean field limit under these uniform regularity assumptions on the kernel Φ_ρ will be established in Section 5. This passage will simultaneously show existence and uniqueness of solutions to (13) and that the solutions arise as a limit, in some sense, of solutions to (11).

Taking the v-moments of the Vlasov equation (13), we obtain

$$\begin{cases} \partial_t \rho + \nabla \cdot (\mathbf{u}\rho) = 0 \\ \partial_t \mathrm{s} + \nabla \cdot (\mathrm{s}[\mathbf{u}]_\rho) = 0 \\ \partial_t(\rho\mathbf{u}) + \nabla \cdot (\rho\mathbf{u} \otimes \mathbf{u}) + \nabla \cdot \mathcal{R} = \rho \mathrm{s}([\mathbf{u}]_\rho - \mathbf{u}), \end{cases} \tag{14}$$

where the density ρ and momentum $\rho\mathbf{u}$ are defined as in (8) and $\mathcal{R}$ is the same Reynolds Stress tensor that appeared in the Cucker-Smale case, (10). This results in the same closure problem and it will be addressed in the same way by adding a local alignment force. The monokinetic and Maxwellian limits for the s-model will be proved in Section 6.

Remark 1.5. The pressureless macroscopic system (SM) arises from the monokinetic limiting regime. The as-yet defined isentropic macroscopic system (58) (introduced in Section 6) arises from the Maxwellian limit.

1.6. **Numerical evidence for similar qualitative behavior to Motsch-Tadmor.** To make the case for physical relevance of the s-model, we present numerical evidence, at the microscopic level, that the w-model with $\mathrm{w}_0 = 1/\rho_\phi$, and hence the s-model, displays similar qualitative behavior to that of Motsch-Tadmor in heterogeneous formations. We first clarify the qualitative behavior that we are seeking.

1.6.1. *Qualitative behavior of Motsch-Tadmor.* Here, we reproduce the motivation for the Motsch-Tadmor model given in [12]. Let m, M with $M \gg 1 \approx m$ be the masses of the small and large flock, respectively. Let $I = \{i : \text{agent } i \text{ is in the small flock}\}$ and define J similarly for the large flock. Suppose that the flocks are far enough away so that $M\sum_{j\in J}\phi(|x_i - x_j|) < < 1$ for all $i \in I$ and that the agents in the small flock are close enough so that $\phi(|x_i - x_{i'}|) \approx 1$ for all $i, i' \in I$. Then for any agent i, the alignment force of the Cucker-Smale system (6) can be written as

$$\begin{aligned} \dot{v}_i =& \frac{\lambda}{m+M}\Big(\sum_{j\in J} m_j \phi(|x_i - x_j|)(v_j - v_i) + \sum_{i'\in I} m_{i'}\phi(|x_i - x_{i'}|)(v_{i'} - v_i)\Big) \\ :=& A_J + A_I. \end{aligned}$$

Since the velocities are bounded and the flocks are far away, $A_J \ll 1$. On the other hand, since $M \gg m$, we also have $A_I \ll 1$. The result is that $\frac{d}{dt}v_i \ll 1$ for all $i \in I$ so the small flock proceeds with essentially no force on it. The small flock is, in this sense, "hijacked" by the large flock.

To see how the Motsch-Tadmor fixes this issue, let us introduce the discrete version of (3):

$$\begin{cases} \dot{x}_i = v_i \\ \dot{v}_i = \frac{\lambda}{\sum_{j=1}^N m_j \phi(|x_i - x_j|)} \sum_{j=1}^N m_j \phi(x_i - x_j)(v_j - v_i). \end{cases} \tag{15}$$

Under the same assumptions and notation as before, we estimate the alignment force. Since the large and small flock are far away, the strength of the alignment force any agent in the small flock, $i \in I$, is approximately

$$\frac{\lambda}{\sum_{i' \in I}^N m_{i'} \phi(|x_i - x_{i'}|) + \sum_{j \in J}^N m_j \phi(|x_i - x_j|)} \approx \frac{\lambda}{\sum_{i' \in I}^N m_{i'} \phi(|x_i - x_{i'}|)}.$$

The interactions with the agents from the large flock also drop out as before. So, the total alignment force on an agent $i \in I$ is approximately

$$\frac{d}{dt}v_i \approx \frac{\lambda}{\sum_{i' \in I}^N m_{i'} \phi(|x_i - x_{i'}|)} \sum_{i' \in I} m_{i'} \phi(|x_i - x_{i'}|)(v_{i'} - v_i)\Big)$$

The small flock then behaves according to the Cucker-Smale model, but independently of the large flock. In our w-model simulation, we are looking for this qualitative behavior. In particular, we aim to see:

- (Q_{cs}) For the Cucker-Smale model: The small flock proceeds linearly as if there were no force on it.
- (Q_w) For the w-model with $w_0 = 1/(\rho_0)_\phi$: The small flock behaves according to Cucker-Smale, but independently of the large flock. The velocities of the small flock will therefore align to the average velocity of the small flock.

1.6.2. *Discrete* w*-model.* Let us state the discrete w-model as a special case of (11). Setting $\Phi_{\rho^N}(x,y) = \phi(x-y)/(\rho^N)_\phi$ where $\rho^N \in \mathcal{P}_M(\mathbb{T}^n)$ is given by $\sum_{j=1}^N m_j \delta_{x_j(t)}$ in (12), we get

$$[\mathbf{u}^N]_{\rho^N} = \sum_{j=1}^N \phi(|x_i - x_j|)(v_j - v_i).$$

Recalling that $s = w\rho_\phi^N$, we obtain the discrete w-model:

$$\begin{cases} \dot{x}_i = v_i \\ \dot{v}_i = \lambda w(x_i) \sum_{j=1}^N m_j \phi(|x_i - x_j|)(v_j - v_i) \\ \partial_t w + [\mathbf{u}^N]_{\rho^N} \cdot \nabla_x w = 0. \end{cases} \tag{16}$$

1.6.3. *Numerical simulation.* The solutions are computed on the 2D unit torus, $\mathbb{T}^2$. We consider initial data consisting of a small mass and a faraway large mass flock. The aim is to illustrate that, in solutions to the Cucker-Smale system, the small flock proceeds linearly as if there were no force on it (Q_{cs}); and that, in solutions to the w-model (with $w_0 = 1/(\rho_0)_\phi$), the small flock aligns to the average velocity of the small flock, independently of the large flock (Q_w).

The parameters of the experiment are as follows.

- The scalar strength of the alignment force is $\lambda = 10$.
- $\phi(r) = \frac{1}{(1+r^2)^{80/2}}$.
- ρ_0^N is identical in the Cucker-Smale and the w-model simulation. It is shown in Figure 1 (and Figure 2) as the leftmost picture.

The kernel is periodized so that the distance r measures the distance on $\mathbb{T}^2$. The large mass flock is indicated by red particles and the small mass flock is indicated by green particles. Each red particle has 100 times the mass of a green particle.

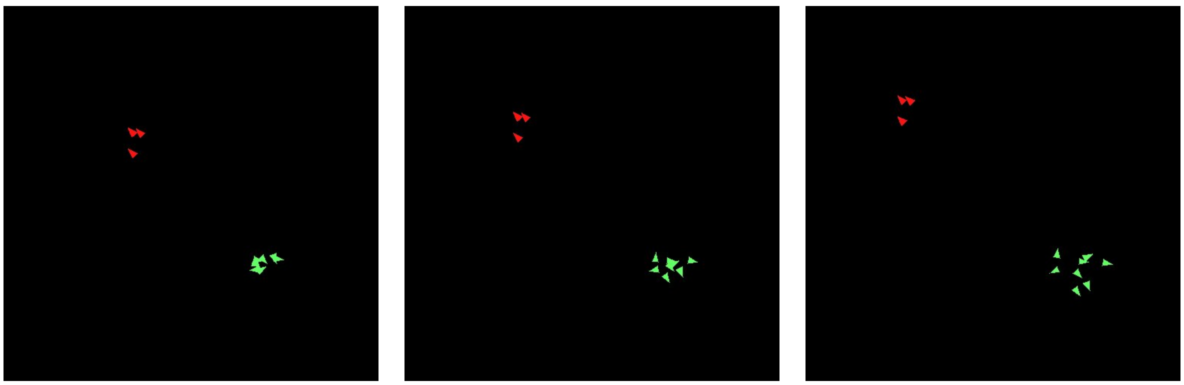

FIGURE 1. The computed solution of the **Cucker-Smale** model at three different time steps. The leftmost image is the initial configuration of the flock and time moves left to right.

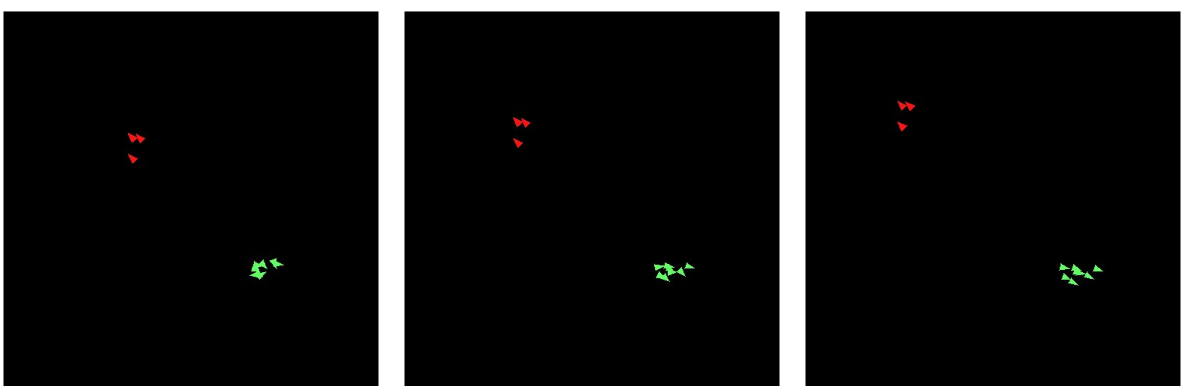

FIGURE 2. The computed solution of the w**-model with Motsch-Tadmor initial data**, i.e. $\mathrm{w}_0 = 1/\rho_\phi$, at three different time steps. The leftmost image is the initial configuration of the flock and time moves left to right.

We see that the small agents in the Cucker-Smale case proceed linearly while the small agents in w-model case align to the average velocity of the small flock. This is the desired qualitative behavior.

1.7. **Relevant quantities and Notation.** Let us introduce relevant quantities and notation. $\mathbb{T}^n$ is the n-dimensional Torus. $\mathcal{P}_M(\Omega)$ denotes the set of measures on Ω with total mass $M = \int_\Omega \rho(x)\,\mathrm{d}x$. $R > 0$ is a bound on the maximum velocity of the initial flock, i.e. $\operatorname{supp}\mu_0 \subset \mathbb{T}^n \times B_R$ where B_R is the ball of radius R in $\mathbb{R}^n$ centered around zero. C^k is the space of k continuously differentiable functions with the usual norm $\|f\|_{C^k} = \sum_{i=0}^k \|f\|_{C^i}$. As in the introduction, we will write $f_\phi := f * \phi$ to denote convolutions. Subscripts $-$ and $+$ will be used as a shorthand for infima and suprema. For instance, $f_- = \inf_{x\in\mathbb{T}^n} f(x)$, $f_+ = \sup_{x\in\mathbb{T}^n} f(x)$. We will use $(f_1, f_2) = \int_{\mathbb{T}^n} f_1 f_2\,\mathrm{d}x$ to denote the L^2 inner product and $(f_1, f_2)_h = \int_{\mathbb{T}^n} f_1 f_2 h\,\mathrm{d}x$ to denote the weighted L^2 inner product. As per the notation in [16], we will use the notation $\kappa_\rho = \mathrm{s}\rho$. For instance, $(\cdot,\cdot)_{\kappa_\rho}$ denotes the L^2 inner product with respect to the measure κ_ρ. When a constant C depends on a parameter α, for instance, we will write $C := C(\alpha)$.

Finally, let us define J, which will be a key quantity in the inheritance of regularity of the strength.

Definition 1.6. We will use J to denote the maximum L^1 norm of the velocity with respect to the density measure on the time interval $[0, T]$:

$$J := \sup_{t\in[0,T]} \|\mathbf{u}(t,\cdot)\|_{L^1(\rho)} = \sup_{t\in[0,T]} \int_\Omega (|\mathbf{u}|\rho)(t,x)\,\mathrm{d}x.$$

1.8. **Outline.** The remaining goal of this paper is to establish the passage between the microscopic, mesoscopic, and macroscopic levels of descriptions of the s-model and to address the relaxation to the Maxwellian in one dimension for the mesoscopic description.

A key observation we make in this paper is that, when $J < \infty$ (given in Definition 1.6) and $[\mathbf{u}]_\rho$ satisfies the integral representation (5), the regularity conditions needed on s can be subordinated to regularity conditions on the kernel $\Phi_\rho(x, y)$, (Reg1) and (Reg1). That is, s and $[\mathbf{u}]_\rho$ inherit regularity from the kernel Φ_ρ. This inherited regularity resembles the 'uniform regularity' conditions outlined in [16] for the environmental averaging models. The only result considered in the paper for which we do not have inheritance of regularity is the Maxwellian limit. There, we instead resort to a set of continuity conditions on the strength and the velocity averaging, (R1)-(R4), which are stated in Section 6.2. Notably, these continuity conditions hold a priori for the w-model when the communication kernel ϕ is bounded away from zero. The w-model therefore serves as an important example where all of our results hold.

The rest of the paper will be organized as follows. In Section 2, we will show that, when $J < \infty$, the velocity averaging and strength inherit the regularity of the kernel. To prepare our arguments for the passage from the microscopic to macroscopic system for the s-model, we will survey the results related to the passage for the environmental averaging models, but catered to the Cucker-Smale model, in Sections 3.1 and 3.2. Indeed, the mathematical tools and arguments used in the passage for the s-model follow a similar outline to those used for the environmental averaging models, but adapted to accommodate for the transport of the strength. In preparation for the mean field limit, the well-posedness of the microscopic s-model (11) is established in Section 4. The mean field limit is proved in Section 5. The hydrodynamic limits are proved in Section 6. Finally, in Section 7, we establish the relaxation to the Maxwellian in 1D for the mesoscopic w-model (66) provided the variation of the weight is small.

1.9. **Assumptions.** Our results will be stated for the torus $\mathbb{T}^n$. Letting $\rho, \rho', \rho'' \in \mathcal{P}_M(\mathbb{T}^n)$, we assume throughout the paper, unless stated otherwise, that $[\mathbf{u}]_\rho$ has the integral representation (5) and that its reproducing kernel $\Phi_\rho(x,y) \geqslant 0$ satisfies the following uniform regularity assumptions:

$$\|\partial^k_{x,y}\Phi_\rho\|_\infty \leqslant C(k,M) \tag{Reg1}$$

$$\|\partial^k_{x,y}(\Phi_{\rho'} - \Phi_{\rho''})\|_\infty \leqslant C(k,M)W_1(\rho',\rho''), \tag{Reg2}$$

where W_1 is the Wasserstein-1 distance. In order to guarantee that $[\mathbf{u}]_\rho$ is an averaging operator, we also assume that Φ_ρ is right stochastic:

$$\int_{\mathbb{T}^n} \Phi_\rho(x,y)\rho(y)\,\mathrm{d}y = 1. \tag{17}$$

Remark 1.7. Table 1.3 lists models for which $[\mathbf{u}]_\rho$ satisfies the integral representation (5). All of these models– namely, M_{CS}, M_{MT}, M_ϕ, and M_{seg}– have kernels $\Phi_\rho(x,y)$ which satisfy (Reg1) and (Reg2) when $\phi \geqslant c > 0$ and is smooth.

Remark 1.8. The Maxwellian limit is the only result for we which these assumptions do not guarentee that the strength inherits the regularity of the kernel. There, we instead introduce a set of continuity assumptions, (R1)-(R4), which are verified for the w-model when $\phi \geqslant c > 0$.

2. Inherited Regularity from the Kernel

Our main assumptions (stated in Section 1.9) together with $J < \infty$ (given in Definition 1.6) imply that the velocity averaging and the strength inherit regularity from kernel. The inherited regularity is recorded in Proposition 2.1 and Proposition 2.3 below and will be the key ingredient to establishing the mean field limit and the monokinetic limit.

Proposition 2.1. *(Inherited regularity of the velocity averaging) Suppose that the velocity averaging has the integral representation* (5) *and that the kernel satisfies the uniform regularity assumptions* (Reg1) *and* (Reg2)*. Let $\rho_0, \rho_0', \rho_0'' \in \mathcal{P}_M(\mathbb{T}^n)$. If $J < \infty$, then for all $k \geqslant 0$:*

$$\|\partial^k[\boldsymbol{u}]_\rho\|_\infty \leqslant C_1 \tag{uReg1}$$

$$\|\partial^k([\boldsymbol{u}']_{\rho'} - [\boldsymbol{u}'']_{\rho''})\|_\infty \leqslant C_1W_1(\rho',\rho'') + C_2W_1(\boldsymbol{u}'\rho', \boldsymbol{u}''\rho''), \tag{uReg2}$$

where $C_1 := C_1(k,M,J)$ and $C_2 := C_2(k,M)$.

Proof. For (**u**Reg1), place all of the derivatives on the kernel and use For (**u**Reg2), we write

$$\begin{aligned}
[\mathbf{u}']_{\rho'} - [\mathbf{u}'']_{\rho''} &= \int_{\mathbb{T}^n} \Phi_{\rho'}(x,y)\mathbf{u}'(y)\rho'(y)\,\mathrm{d}y - \int_{\mathbb{T}^n} \Phi_{\rho''}(x,y)\mathbf{u}''(y)\rho''(y)\,\mathrm{d}y \\
&= \int_{\mathbb{T}^n} (\Phi_{\rho'}(x,y) - \Phi_{\rho''}(x,y))(\mathbf{u}'\rho')(y)\,\mathrm{d}y + \int_{\mathbb{T}^n} \Phi_{\rho''}(x,y)((\mathbf{u}'\rho')(y) - (\mathbf{u}''\rho'')(y))\,\mathrm{d}y.
\end{aligned}$$

Once again, placing the derivatives on the kernel and using $J < \infty$ on the first term, we arrive at (**u**Reg2). □

Remark 2.2. $J < \infty$ holds a priori for the microscopic and mesoscopic s-model, (11) and (13), due to the maximum principle on the velocity. The maximum principle also holds for the mesoscopic s-model with the strong local alignment force for the monokinetic limiting regime, (59) (the justification is provided in Section 6.1). However, for the mesoscopic s-model with strong Fokker-Planck penalization force, (62), there is no control on J and, therefore, there is no inheritance of regularity from the kernel.

The strength subsequently inherits regularity from the velocity averaging.

Proposition 2.3. *(Inherited regularity of the strength) Suppose that the uniform regularity conditions on the velocity averaging* (**u**Reg1) *and* (**u**Reg2) *hold. Let* $\rho_0', \rho_0'' \in \mathcal{P}_M(\mathbb{T}^n)$*. If* $\mathrm{s}', \mathrm{s}'' \in C([0,T]; C^{k+1}(\mathbb{T}^n))$ *with* $\mathrm{s}_0' = \mathrm{s}_0''$ *where* s' *solves* $\partial_t \mathrm{s}' + \nabla_x \cdot (\mathrm{s}'[\boldsymbol{u}']_{\rho'}) = 0$ *and* s'' *solves* $\partial_t \mathrm{s}'' + \nabla_x \cdot (\mathrm{s}''[\boldsymbol{u}'']_{\rho''}) = 0$*, then for all* $k \geqslant 0$*:*

$$\|\partial^k \mathrm{s}'\|_\infty \leqslant C \tag{sReg1}$$

$$\|\partial^k (\mathrm{s}' - \mathrm{s}'')\|_\infty \leqslant C\|\partial^k (s_0' - s_0'')\|_\infty + C \sup_{t\in[0,T]} \big(W_1(\rho', \rho'') + W_1(\boldsymbol{u}'\rho', \boldsymbol{u}''\rho'')\big)(t), \tag{sReg2}$$

where $C := C(k, M, J, T)$*.*

Proof. We only prove (sReg2) since (sReg1) follows a similar computation. Taking the k^{th} partial derivative of the strength equation, we get

$$\partial_t \partial^k (\mathrm{s}' - \mathrm{s}'') + \nabla \cdot \partial^k ([\mathbf{u}']_{\rho'} \mathrm{s}' - [\mathbf{u}'']_{\rho''} \mathrm{s}'') = 0.$$

This can be written as

$$\partial_t \partial^k (\mathrm{s}' - \mathrm{s}'') + \nabla \cdot \partial^k (([\mathbf{u}']_{\rho'} - [\mathbf{u}'']_{\rho''})\mathrm{s}' + [\mathbf{u}'']_{\rho''}(\mathrm{s}' - \mathrm{s}'')) = 0.$$

Evaluating at a point of maximum of $\partial^k(\mathrm{s}' - \mathrm{s}'')$, the term $[\mathbf{u}'']_{\rho''}\partial^{k+1}(\mathrm{s}' - \mathrm{s}'')$ drops out. We obtain for some constant A depending on k:

$$\partial_t \|\partial^k (\mathrm{s}' - \mathrm{s}'')\|_\infty \leqslant A(k)\big(\|[\mathbf{u}']_{\rho'} - [\mathbf{u}'']_{\rho''}\|_{C^{k+1}} \|\mathrm{s}'\|_{C^{k+1}} + \|[\mathbf{u}'']_{\rho''}\|_{C^{k+1}} \|\mathrm{s}' - \mathrm{s}''\|_{C^k}\big).$$

Summing over k, we obtain for some constant A' depending on k:

$$\partial_t \|\mathrm{s}' - \mathrm{s}''\|_{C^k} \leqslant A'(k)\big(\|[\mathbf{u}']_{\rho'} - [\mathbf{u}'']_{\rho''}\|_{C^{k+1}} \|\mathrm{s}'\|_{C^{k+1}} + \|[\mathbf{u}'']_{\rho''}\|_{C^{k+1}} \|\mathrm{s}' - \mathrm{s}''\|_{C^k}\big).$$

Applying (**u**Reg1), (**u**Reg2), and (sReg1) along with Gronwall's inequality yields constants $C_1'(k, M, J)$, $C_2'(k, M, J, T)$ such that

$$\|\mathrm{s}' - \mathrm{s}''\|_{C^k} \leqslant e^{C_1' T} \|s_0' - s_0''\|_{C^k} + C_2' e^{C_1' T} \sup_{t\in[0,T]} \big(W_1(\rho', \rho'') + W_1(\mathbf{u}'\rho', \mathbf{u}''\rho'')\big)(t, \cdot)$$

Setting $C = \max\{C_2' e^{C_1' T}, e^{C_1' T}\}$ gives (sReg2). □

3. Passage from microscopic to macroscopic for Cucker-Smale

It will be helpful to survey the mean field and hydrodynamic arguments for the Cucker-Smale model in order to understand the extension of our argument to the s-model. Indeed, these limiting arguments for the s-model fit into the same framework as the arguments used in the Cucker-Smale case. Each section will outline the argument for the Cucker-Smale case and conclude with the statement of the corresponding theorem for the s-model. The mean field and hyrdoynamic limits for the s-model are stated in Theorems 3.2, 3.6, and 3.8. The proofs are in Section 5 and 6.

3.1. **Mean field limit (Cucker-Smale survey).** We will include details for pieces of the argument that do not depend (in a significant way) on s and leave the pieces which depend on s for Section 5. To rigorously pass from the discrete system (6) to the kinetic one (7), we would like to show that solutions to (7) are unique and that the solution arises as a limit, in some sense, of solutions to the discrete equation (6). To make sense of discrete solutions converging to the kinetic one, the solutions to (6) and (7) must lie in the same space. We therefore consider measure-valued solutions to both systems. For (6), a measure-valued solution is given by the empirical measure

$$\mu_t^N(x,v) = \sum_{i=1}^N m_i \delta_{x_i(t)} \otimes \delta_{v_i(t)} \tag{18}$$

where $(x_i(t), v_i(t))_{i=1}^N$ solve (6). To make sense of measure valued solutions for (7), we define a weak solution as follows.

Definition 3.1. Fix a time $T > 0$ and an integer $k \geqslant 0$. We say $\mu \in C_{w^*}([0,T]; \mathcal{P}_M(\mathbb{T}^n \times \mathbb{R}^n))$ is a *weak* solution to (7) if for all $g(t,x,v) \in C_0^\infty([0,T] \times \mathbb{T}^n \times \mathbb{R}^n)$ and for all $0 < t < T$,

$$\begin{aligned}\int_{\mathbb{T}^n\times\mathbb{R}^n} g(t,x,v)d\mu_t(x,v) &= \int_{\mathbb{T}^n\times\mathbb{R}^n} g(0,x,v)d\mu_0(x,v) \\ &\quad + \int_0^t \int_{\mathbb{T}^n\times\mathbb{R}^n} (\partial_\tau g + v\cdot\nabla_x g + \lambda F(\mu_\tau)\cdot\nabla_v g d\mu_\tau(x,v).\end{aligned} \tag{19}$$

Observe that the empirical measure (18) is a solution to (19) if and only if $(x_i(t), v_i(t))$ solve the discrete system (6). We will always assume that $\operatorname{supp}\mu_0 \subset \mathbb{T}^n \times B_R$ for some fixed $R > 0$. We now aim to show the existence and uniqueness of weak solutions and that they arise as weak limits of these empirical measures (18), where the weak limit can be topologized by the Wasserstein-1 metric:

$$W_1(\mu,\nu) = \sup_{Lip(g)\leqslant 1} \Big| \int_{\mathbb{T}^n\times\mathbb{R}^n} g(\omega)(d\mu(\omega) - d\nu(\omega))\Big|.$$

The Wasserstein-1 distance metrizes the weak topology as long as the measures lie on some common compact set. Owing to the maximum principle on the velocity equation, this is the case here: $\operatorname{supp}\mu_t \subset \mathbb{T}^n \times B_R$ for all $t \in [0,T)$.

The goal will be to establish a stability estimate in the Wasserstein-1 metric: for $\mu_t', \mu_t'' \in \mathcal{P}_M(\mathbb{T}^n \times B_R)$, there exists a constant $C(M,R,T)$ such that

$$W_1(\mu_t', \mu_t'') \leqslant C(M,R,T) W_1(\mu_0', \mu_0''). \tag{20}$$

For then, a Cauchy sequence μ_0^N with $W_1(\mu_0^N, \mu_0) \to 0$ yields a Cauchy sequence μ_t^N with $W_1(\mu_t^N, \mu_t) \to 0$ for some $\mu_t \in C_{w^*}([0,T]; \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n))$. That μ is a solution to (19) will follow from taking limits.

Stability in the Wasserstein-1 metric is tied to the stability of the characteristic flow of (19).

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t} X(t,s,x,v) = V(t,s,x,v), & X(s,s,x,v) = x, \\ \frac{\mathrm{d}}{\mathrm{d}t} V(t,s,x,v) = \lambda F(\mu_t)(X,V), & V(s,s,x,v) = v. \end{cases} \tag{21}$$

Indeed, let $X(t,x,v) := X(t,0,x,v)$ and $V(t,0,x,v) = V(t,x,v)$ and $(x,v) = \omega$. Given $h \in C_0^\infty(\mathbb{R}^{2n})$, define the test function $g(s,\omega) = h(X(t,s,\omega), V(t,s,\omega))$. Then

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

Plugging g into (19), we see that μ_t is the push-forward measure of μ_0 along the characteristic flow (X,V):

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} h(X(t,\omega), V(t,\omega))\, \mathrm{d}\mu_0(\omega) = \int_{\mathbb{T}^n \times \mathbb{R}^n} h(\omega)\, \mathrm{d}\mu_t(\omega).$$

Letting $X' := X'(t,\omega)$, $V' := V'(t,\omega)$ and similarly for X'', V'', we have for any $h \in Lip(\mathbb{R}^n)$ with $Lip(h) \leqslant 1$,

$$\begin{aligned}
\int_{\mathbb{T}^n \times \mathbb{R}^n} h(\omega)(d\mu_t' - d\mu_t'') &= \int_{\mathbb{T}^n \times \mathbb{R}^n} h(X',V') d\mu_0' - \int_{\mathbb{T}^n \times \mathbb{R}^n} h(X'',V'') d\mu_0'') \\
&= \int_{\mathbb{T}^n \times \mathbb{R}^n} h(X',V')(d\mu_0' - d\mu_0'') \\
&+ \int_{\mathbb{T}^n \times \mathbb{R}^n} (h(X',V') - h(X'',V'')) d\mu_0'' \\
&\leqslant (\|\nabla X'\|_\infty + \|\nabla V'\|_\infty) W_1(\mu_0', \mu_0'') + \|X' - X''\|_\infty + \|V' - V''\|_\infty.
\end{aligned}$$

Therefore,

$$W_1(\mu_t', \mu_t'') \leqslant (\|\nabla X'\|_\infty + \|\nabla V'\|_\infty) W_1(\mu_0', \mu_0'') + \|X' - X''\|_\infty + \|V' - V''\|_\infty. \tag{22}$$

The stability estimate (20) reduces to establishing the following stability estimates on the characteristic flow:

$$\|\nabla X\|_\infty + \|\nabla V\|_\infty \leqslant C(M,R,T) \tag{23}$$

$$\|X' - X''\|_\infty + \|V' - V''\|_\infty \leqslant C(M,R,T) W_1(\mu_0', \mu_0''). \tag{24}$$

This in turn implies the Wasserstein-1 stability (20). Since these estimates will depend on the regularity of the strength, we will address these details for the s-model in Lemmas 5.2 and 5.3.

Wasserstein-1 stability implies $\lim_{N\to\infty} W_1(\mu_t^N, \mu_t) = 0$ for each $t \in [0,T]$ and for some $\mu_t \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$. The last piece is to show $\mu \in C_{w^*}([0,T]; \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n))$ is a weak solution to (19). The weak convergence $W_1(\mu_t^N, \mu_t) \to 0$ immediately implies that the linear terms in (19) converge. For the non-linear term, the strength enters so we will address this for the s-model later in Lemma 5.5.

We record the corresponding mean field limit theorem for the s model, which will be proven in Section 5.

Theorem 3.2. *(s-model: Mean Field Limit) Suppose that the velocity averaging $[\boldsymbol{u}]_\rho$ has the integral representation (5) with the kernel satisfying (Reg1) and (Reg2). Given $T > 0$, $R > 0$, $k \geqslant 0$, $\mu_0 \in \mathcal{P}_M(\mathbb{T}^n \times B_R)$, and $\mathrm{s}_0 \in C^\infty(\mathbb{T}^n)$, then there exists a unique weak solution $(\mu, \mathrm{s}) \in C_{w^*}([0,T]; \mathcal{P}_M(\mathbb{T}^n \times B_R)) \times C([0,T]; C^k(\mathbb{T}^n))$ to (13). Moreover, the solution can be obtained as a limit of empirical measures (18), μ_t^N, with corresponding strength functions s_t^N: If $\mathrm{s}_0^N = \mathrm{s}_0$ and the empirical measures μ_0^N are constructed from agents $(x_i^0, v_i^0) \in \mathbb{T}^n \times \mathbb{R}^n$ with total mass $M = \sum_{i=1}^N m_i = \int_{\mathbb{T}^n \times \mathbb{R}^n} d\mu_0^N(x,v)$ in such a way that $W_1(\mu_0^N, \mu_0) \to 0$, then*

(a) $\sup_{t\in[0,T]} W_1(\mu_t^N, \mu_t) \to 0$, and

(b) $\sup_{t\in[0,T]} \|\partial^k(\mathrm{s}_t^N - \mathrm{s}_t)\|_\infty \to 0$ *for any* $k \geqslant 0$ *where* s *solves the transport equation in* (13)*,* $\partial_t \mathrm{s} + \nabla\cdot(\mathrm{s}[\boldsymbol{u}]_\rho) = 0$*.*

Remark 3.3. The strength s enters in the stability estimates on the characteristic flow and in showing the limits (a) and (b). The inherited regularity from the kernel will play a crucial role in each of these.

3.2. **Hydrodynamic limits (Cucker-Smale survey).** Once again, we will present the core arguments of the hydrodynamic limits for the Cucker-Smale case which do not depend (in a significant way) on s and leave the pieces which depend on s for Section 6.

To resolve the closure problem for (9), a strong local alignment force, F_{la}, is added to the kinetic Vlasov equation,

$$\partial_t f^\epsilon + v\cdot\nabla_x f^\epsilon + \nabla_v\cdot[f^\epsilon F(f^\epsilon)] + \frac{1}{\epsilon}F_{\mathrm{la}}(f^\epsilon) = 0. \tag{25}$$

Under the right local alignment force F_{la}, the corresponding macroscopic system formed from taking the v-moments will lose its dependence on the kinetic solution f^ϵ as $\epsilon \to 0$. We will consider two such local alignment forces, F_{la}, corresponding to the monokinetic regime, where the distribution is forced to be concentrated around the macroscopic velocity $\mathbf{u}$; and maxwellian regime, where the distribution is forced to be a Gaussian distribution centered around the macroscopic velocity $\mathbf{u}$. We will denote the macroscopic density and momentum corresponding to the distribution f^ϵ by ρ^ϵ and $\mathbf{u}^\epsilon\rho^\epsilon$:

$$\rho^\epsilon(t,x) = \int_{\mathbb{R}^n} f^\epsilon(t,x,v)\,\mathrm{d}v, \qquad (\rho^\epsilon\mathbf{u}^\epsilon)(t,x) = \int_{\mathbb{R}^n} v f^\epsilon(t,x,v)\,\mathrm{d}v.$$

3.2.1. *Monokinetic regime.* Under the monokinetic local alignment force,

$$F_{\mathrm{la}} = F_{\mathrm{mono}} = \nabla_v[f^\epsilon(\mathbf{u}^\epsilon - v)] = 0, \tag{26}$$

the probability density f^ϵ is forced to the monokinetic ansatz

$$f(t,x,v) = \rho(t,x)\delta(v - \mathbf{u}(t,x)) \tag{27}$$

where $(\rho, \mathbf{u})$ solves:

$$\begin{cases} \partial_t\rho + \nabla\cdot(\mathbf{u}\rho) = 0 \\ \partial_t(\rho\mathbf{u}) + \nabla\cdot(\rho\mathbf{u}\otimes\mathbf{u}) = \rho((\mathbf{u}\rho)_\phi - \mathbf{u}\rho_\phi). \end{cases} \tag{28}$$

If we know a priori that the solution remains non-vacuous, i.e. $\rho > 0$, then we can divide the momentum equation by ρ in order to rewrite it as an equation on the velocity $\mathbf{u}$. We arrive at the pressureless Euler alignment system from the introduction (1). The system (28) has the advantage that the velocity equation obeys the maximum principle (as does (SM)): $\|\mathbf{u}\|_{L^\infty} \leqslant \|\mathbf{u}_0\|_{L^\infty}$, which lends it to alignment analysis. For this reason, it has received a lot of attention in the literature.

This monokinetic limit was first proved by Figalli and Kang for non-vacuous solutions $\rho > 0$ on the torus $\mathbb{T}^n$ in [7]. It was later improved to allow vacuous solutions on the open space $\mathbb{R}^n$ in [15] and extended to environmental averaging models in [16]. There, the local aligment is modified to force the system to a special averaged velocity field instead of the rough velocity field $\mathbf{u}^\epsilon$.

$$F_{\mathrm{la}} = F_{\mathrm{mono\ reg}} = \nabla_v[f^\epsilon(\mathbf{u}^\epsilon_\delta - v)] = 0 \tag{29}$$

where $\mathbf{u}_\delta^\epsilon$ is a special mollification given by

$$\mathbf{u}_\delta = \left(\frac{(\mathbf{u}\rho)_{\Psi_\delta}}{\rho_{\Psi_\delta}}\right)_{\Psi_\delta} \tag{30}$$

for some smooth mollifier $\Psi_\delta(x) = \frac{1}{\delta^n}\Psi(x/\delta)$. This special mollification has the key approximation property that $\mathbf{u}_\delta$ is close to $\mathbf{u}$ for small δ with a bound independent of ρ. The following approximation lemma can be found in [16, Lemma 9.1].

Lemma 3.4. *For any $\boldsymbol{u} \in Lip$ and for any $1 \leqslant p < \infty$, one has*

$$\|\boldsymbol{u}_\delta - \boldsymbol{u}\|_{L^p(\rho)} \leqslant C\delta\|\boldsymbol{u}\|_{Lip}$$

where $C > 0$ depends only on Ψ and p.

We will be interested in extending this argument with the local alignment force $F_{\text{mono reg}}$ from the environmental averaging models to the s-model, see Theorem 3.6. The convergence to the monokinetic ansatz is measured in the Wasserstein-2 metric, which is given by

$$W_2^2(f_1, f_2) = \inf_{\gamma \in \prod(f_1, f_2)} \int_{\mathbb{T}^{2n}\times\mathbb{R}^{2n}} |w_1 - w_2|^2 \mathrm{d}\gamma(w_1, w_2).$$

The modulated kinetic energy, $e(f^\epsilon|\mathbf{u})$, which will be crucial in controlling $W_2^2(f_t^\epsilon, f_t)$, is defined by:

$$e(f^\epsilon|\mathbf{u}) = \int_{\mathbb{T}^n\times\mathbb{R}^n} |v - \mathbf{u}(x)|^2 f^\epsilon(x, v)\,\mathrm{d}v\,\mathrm{d}x. \tag{31}$$

Let us state the theorem for the Cucker-Smale model.

Theorem 3.5. *(Cucker-Smale: Monokinetic Limit) Let $(\rho, \boldsymbol{u})$ be a classical solution to (1) on the time interval $[0, T]$ and let f be the monokinetic ansatz (27). Suppose that $f_0^\epsilon \in C_0^k(\mathbb{T}^n \times \mathbb{R}^n)$ is a family of initial conditions satisfying:*

(i) $\operatorname{supp} f_0^\epsilon \subset \mathbb{T}^n \times B_R$ *for any fixed $R > 0$, and*
(ii) $W_2(f_0^\epsilon, f_0) < \epsilon$

Then there exists a constant $C(M, R, T)$ such that for all $t \in [0, T]$ one has

$$W_2(f_t^\epsilon, f_t) \leqslant C\sqrt{\epsilon + \frac{\delta}{\epsilon}}.$$

In order to control $W_2(f_t^\epsilon, f_t)$, it suffices to control

$$\int_{\mathbb{T}^{2n}\times\mathbb{R}^{2n}} |w_1 - w_2|^2 \mathrm{d}\gamma(w_1, w_2)$$

for a particular $\gamma_t \in \prod(f_t^\epsilon, f_t)$. The natural choice for γ_t is given by the flow

$$\partial_t\gamma + v_1 \cdot \nabla_{x_1}\gamma + v_2 \cdot \nabla_{x_2}\gamma + \nabla_{v_1}[\gamma \mathrm{s}^\epsilon(v_1 - \mathbf{u}_F^\epsilon) + \frac{1}{\epsilon}(v_1 - \mathbf{u}_\delta^\epsilon)] + \nabla_{v_2}[\gamma \mathrm{s}(v_2 - \mathbf{u}_F)] = 0,$$

where $\mathbf{u}_F^\epsilon = (\mathbf{u}^\epsilon\rho^\epsilon)_\phi/\rho_\phi^\epsilon$. Since $\gamma_t \in \prod(f_t^\epsilon, f_t)$,

$$W := \int_{\mathbb{T}^{2n}\times\mathbb{R}^{2n}} |w_1 - w_2|^2 \mathrm{d}\gamma_t(w_1, w_2) \geqslant W_2^2(f_t^\epsilon, f_t)$$

So, we aim to control W. We split it into the potential and kinetic components

$$W = \int_{\mathbb{T}^{2n}\times\mathbb{R}^{2n}} |v_1 - v_2|^2 \mathrm{d}\gamma_t + \int_{\mathbb{T}^{2n}\times\mathbb{R}^{2n}} |x_1 - x_2|^2 \mathrm{d}\gamma_t := W_v + W_x$$

For W_x, we have

$$\frac{d}{dt}W_x \leqslant W_x + W_v.$$

For W_v, we have

$$\begin{aligned} W_v &\leqslant \int_{\mathbb{T}^{2n}\times\mathbb{R}^{2n}} |v_1 - \mathbf{u}(x_1)|^2 \mathrm{d}\gamma_t + \int_{\mathbb{T}^{2n}\times\mathbb{R}^{2n}} |\mathbf{u}(x_1) - \mathbf{u}(x_2)|^2 \mathrm{d}\gamma_t + \int_{\mathbb{T}^{2n}\times\mathbb{R}^{2n}} |\mathbf{u}(x_2) - v_2|^2 \mathrm{d}\gamma_t \\ &\leqslant \int_{\mathbb{T}^n\times\mathbb{R}^n} |v - \mathbf{u}(x)|^2 f^\epsilon(x,v)\,\mathrm{d}v\,\mathrm{d}x + C\int_{\mathbb{T}^{2n}\times\mathbb{R}^{2n}} |x_1 - x_2|^2 \mathrm{d}\gamma_t \\ &= e(f^\epsilon|\mathbf{u}) + CW_x. \end{aligned}$$

We obtain

$$\begin{cases} \frac{d}{dt}W_x \lesssim e(f^\epsilon|\mathbf{u}) + W_x \\ W_v \lesssim e(f^\epsilon|\mathbf{u}) + W_x. \end{cases}$$

Expanding $e(f^\epsilon|\mathbf{u})$, we get

$$e(f^\epsilon|\mathbf{u}) = \mathcal{E}_\epsilon - \int_{\mathbb{R}^n} \rho^\epsilon \mathbf{u}^\epsilon \cdot \mathbf{u} + \frac{1}{2}\int_{\mathbb{R}^n} \rho^\epsilon |\mathbf{u}|^2\,\mathrm{d}x,$$

where $\mathcal{E}_\epsilon$ is the kinetic energy:

$$\mathcal{E}_\epsilon = \frac{1}{2}\int |v|^2 f^\epsilon(x,v)\,\mathrm{d}x\,\mathrm{d}v.$$

Under no assumptions on the strength, it is shown in [16, Theorem 9.2] that

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} e(f^\epsilon|\mathbf{u}) \lesssim e(f^\epsilon|\mathbf{u}) &+ \frac{1}{\epsilon}\int_{\mathbb{T}^n} \rho^\epsilon(\mathbf{u}^\epsilon_\delta - \mathbf{u}^\epsilon)\cdot\mathbf{u}\,\mathrm{d}x \\ &+ (\mathbf{u}^\epsilon - \mathbf{u}, [\mathbf{u}^\epsilon]_{\rho^\epsilon} - \mathbf{u}^\epsilon)_{\kappa_{\rho^\epsilon}} + \int_{\mathbb{T}^n} \rho^\epsilon(\mathbf{u} - \mathbf{u}^\epsilon)\cdot \mathrm{s}([\mathbf{u}]_\rho - \mathbf{u})\,\mathrm{d}x. \end{aligned}$$

The local alignment term is controlled with the help of Lemma 3.4, see [16, Theorem 9.2]:

$$\frac{1}{\epsilon}\int_{\mathbb{T}^n} \rho^\epsilon(\mathbf{u}^\epsilon_\delta - \mathbf{u}^\epsilon)\cdot\mathbf{u}\,\mathrm{d}x \lesssim \frac{\delta}{\epsilon}.$$

So, it remains to estimate the alignment terms

$$A := (\mathbf{u}^\epsilon - \mathbf{u}, [\mathbf{u}^\epsilon]_{\rho^\epsilon} - \mathbf{u}^\epsilon)_{\kappa_{\rho^\epsilon}} + \int_{\mathbb{T}^n} \rho^\epsilon(\mathbf{u} - \mathbf{u}^\epsilon)\cdot \mathrm{s}([\mathbf{u}]_\rho - \mathbf{u})\,\mathrm{d}x.$$

The estimate on A does depend on the regularity of the strength and it is addressed in Theorem 3.6, which is proved in Section 6.1.

Theorem 3.6. *(s-model: Monokinetic Limit) Suppose the velocity averaging has the integral representation* (5) *with a kernel* Φ *satisfying the regularity conditions* (Reg1) *and* (Reg2)*. Let* $(\rho, \mathrm{s}, \boldsymbol{u})$ *be a smooth solution to* (SM) *on the time interval* $[0, T]$ *with mass* M*. Let* $f = \rho(t,x)\delta(v - \boldsymbol{u}(t,x))$*. Suppose* $\mathrm{s}_0^\epsilon \in C^\infty(\mathbb{T}^n)$*,* $f_0^\epsilon \in C_0^k(\mathbb{T}^n \times \mathbb{R}^n)$ *is a family of initial conditions satisfying:*

(i) $\operatorname{supp} f_0^\epsilon \subset \mathbb{T}^n \times B_R$*, for any fixed* $R > 0$*.*
(ii) $W_2(f_0^\epsilon, f_0) < \epsilon$
(iii) $\mathrm{s}_0^\epsilon = \mathrm{s}_0 \in C^\infty(\mathbb{T}^n)$.

Then there exists a constant $C(M,R,T)$ such that for all $t \in [0,T]$:

$$W_2(f_t^\epsilon, f_t) \leqslant C\sqrt{\epsilon + \frac{\delta}{\epsilon}}.$$

As a consequence of this and Proposition 2.3, $\sup_{t\in[0,T]} \|(\mathrm{s}^\epsilon - \mathrm{s})(t,\cdot)\|_\infty \to 0$.

3.2.2. *Maxwellian regime.* Let us rewrite the equation for the macroscopic density and momentum (9) as

$$\begin{cases} \partial_t \rho + \nabla\cdot(\mathbf{u}\rho) = 0 \\ \partial_t(\rho\mathbf{u}) + \nabla\cdot(\rho\mathbf{u}\otimes\mathbf{u}) + \nabla\rho + \nabla\cdot\mathcal{R} = \rho((\mathbf{u}\rho)_\phi - \mathbf{u}\rho_\phi) \end{cases} \tag{32}$$

where

$$\mathcal{R}(t,x) = \int_{\mathbb{T}^n} ((v-\mathbf{u})\otimes(v-\mathbf{u}) - \mathbb{I}) f \,\mathrm{d}v. \tag{33}$$

Under the local alignment force,

$$F_{\mathrm{la}} = F_{\mathrm{max\ reg}} = \Delta_v f^\epsilon + \nabla_v \cdot ((v - \mathbf{u}_\delta^\epsilon) f^\epsilon),$$

where $\mathbf{u}_\delta$ is given in (30), the probability density f^ϵ is forced to the Maxwellian $\mu(t,x)$,

$$\mu(t,x) = \frac{\rho(t,x)}{(2\pi)^{n/2}} e^{\frac{|v-\mathbf{u}(t,x)|^2}{2}},$$

where $(\rho, \mathbf{u})$ solve the isentropic Euler alignment system:

$$\begin{cases} \partial_t \rho + \nabla\cdot(\mathbf{u}\rho) = 0 \\ \partial_t(\rho\mathbf{u}) + \nabla\cdot(\rho\mathbf{u}\otimes\mathbf{u}) + \nabla\rho = \rho((\mathbf{u}\rho)_\phi - \mathbf{u}\rho_\phi). \end{cases} \tag{34}$$

As in the monokinetic limit, we are interested in extending this argument with the local alignment force $F_{\mathrm{max\ reg}}$ from the environmental averaging models to the s-model, see Theorem 3.8. The distance between f^ϵ and the Maxwellian μ is controlled by the relative entropy $\mathcal{H}(f^\epsilon|\mu)$, which is defined as:

$$\mathcal{H}(f^\epsilon|\mu) = \int_{\mathbb{T}^n\times\mathbb{R}^n} f^\epsilon \log\frac{f^\epsilon}{\mu} \,\mathrm{d}v\,\mathrm{d}x. \tag{35}$$

Due to the classical Csiszár-Kullback inequality, for some constant $c > 0$,

$$c\|f^\epsilon - \mu\|_{L^1(\mathbb{T}^n\times\mathbb{R}^n)} \leqslant \mathcal{H}(f^\epsilon|\mu),$$

it suffices to show that $\mathcal{H}(f^\epsilon|\mu) \to 0$. We state the theorem for the Cucker-Smale model.

Theorem 3.7. *(Cucker-Smale: Maxwellian limit) Let $(\rho, \mathbf{u})$ be a smooth, non-vacuous solution to (34) on the torus $\mathbb{T}^n$ and on the time interval $[0,T]$. Suppose that $f_0^\epsilon \in C_0^k(\mathbb{T}^n\times\mathbb{R}^n)$ is a family of initial conditions satisfying*

$$\lim_{\epsilon\to 0} \mathcal{H}(f_0^\epsilon|\mu_0) = 0.$$

Then for $\delta = o(\epsilon)$,

$$\sup_{t\in[0,T]} \mathcal{H}(f^\epsilon|\mu) \to 0.$$

Breaking the relative entropy into the kinetic and macroscopic parts, we have

$$\mathcal{H}(f^\epsilon|\mu) = \mathcal{H}_\epsilon + \mathcal{G}_\epsilon, \tag{36}$$

$$\mathcal{H}_\epsilon = \int_{\mathbb{T}^n\times\mathbb{R}^n} \big(f^\epsilon \log f^\epsilon + \frac{1}{2}|v|^2 f^\epsilon\big)\,\mathrm{d}v\,\mathrm{d}x + \frac{n}{2}\log(2\pi), \tag{37}$$

$$\mathcal{G}_\epsilon = \int_{\mathbb{T}^n} \big(\frac{1}{2}\rho^\epsilon|\mathbf{u}|^2 - \rho^\epsilon\mathbf{u}^\epsilon\cdot\mathbf{u} - \rho^\epsilon\log\rho\big)\,\mathrm{d}x. \tag{38}$$

We then seek to estimate $\mathcal{H}_\epsilon$ and $\mathcal{G}_\epsilon$. Let us define the Fisher information, $\mathcal{I}_\epsilon$, which will be relevant for these estimates:

$$\mathcal{I}_\epsilon = \int_{\mathbb{T}^n\times\mathbb{R}^n} \frac{|\nabla_v f^\epsilon + (1+\epsilon \mathrm{s}^\epsilon/2)(v-\mathbf{u}^\epsilon)f^\epsilon|^2}{f^\epsilon}\,\mathrm{d}v\,\mathrm{d}x. \tag{39}$$

$$\tag{40}$$

The following identities are from [16]:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}_\epsilon = &-\frac{1}{\epsilon}\int_{\mathbb{T}^n\times\mathbb{R}^n}\Big[\frac{|\nabla_v f^\epsilon|^2}{f^\epsilon} + 2\nabla_v f^\epsilon\cdot(v-\mathbf{u}^\epsilon_\delta) + |v-\mathbf{u}^\epsilon_\delta|^2 f^\epsilon\Big]\,\mathrm{d}v\,\mathrm{d}x \\ &-\frac{1}{\epsilon}[(\mathbf{u}^\epsilon_\delta,\mathbf{u}^\epsilon)_{\rho^\epsilon} - (\mathbf{u}^\epsilon_\delta,\mathbf{u}^\epsilon_\delta)] \\ &-\int_{\mathbb{T}^n\times\mathbb{R}^n} \mathrm{s}^\epsilon[\nabla_v f^\epsilon\cdot(v-[\mathbf{u}^\epsilon]_{\rho^\epsilon}) + v\cdot(v-[\mathbf{u}^\epsilon]_{\rho^\epsilon})f^\epsilon]\,\mathrm{d}v\,\mathrm{d}x\end{aligned} \tag{41}$$

and

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{G}_\epsilon = &\int_{\mathbb{T}^n}[\nabla\mathbf{u}:\mathcal{R}^\epsilon - \rho^\epsilon(\mathbf{u}^\epsilon-\mathbf{u})\cdot\nabla\mathbf{u}\cdot(\mathbf{u}^\epsilon-\mathbf{u})]\,\mathrm{d}x + \frac{1}{\epsilon}\int_{\mathbb{T}^n}\rho^\epsilon(\mathbf{u}^\epsilon_\delta-\mathbf{u}^\epsilon)\cdot\mathbf{u}\,\mathrm{d}x \\ &+\|\mathbf{u}^\epsilon\|^2_{L^2(\kappa^\epsilon)} - (\mathbf{u}^\epsilon,[\mathbf{u}^\epsilon]_{\rho^\epsilon})_{\kappa^\epsilon} + \delta A.\end{aligned} \tag{42}$$

where δA is the same alignment term used in the proof of the monokinetic case. From (41), dropping the first two terms, we obtain

$$\begin{aligned}\frac{d}{dt}\mathcal{H}_\epsilon &\leqslant -\int_{\mathbb{T}^n\times\mathbb{R}^n}\mathrm{s}^\epsilon[\nabla_v f^\epsilon\cdot(v-[\mathbf{u}^\epsilon]_{\rho^\epsilon}) + v\cdot(v-[\mathbf{u}^\epsilon]_{\rho^\epsilon})f^\epsilon]\,\mathrm{d}v\,\mathrm{d}x \\ &= \int_{\mathbb{T}^n\times\mathbb{R}^n}\mathrm{s}^\epsilon f^\epsilon\,\mathrm{d}v\,\mathrm{d}x - \int_{\mathbb{T}^n\times\mathbb{R}^n}\mathrm{s}^\epsilon|v|^2 f^\epsilon\,\mathrm{d}v\,\mathrm{d}x + (\mathbf{u}^\epsilon,[\mathbf{u}^\epsilon]_{\rho^\epsilon})_{\kappa^\epsilon}.\end{aligned}$$

This estimate will be used to show that $\mathcal{H}_\epsilon$ is bounded which will in turn be used to control the Reynolds stress term $\mathcal{R}^\epsilon$ in the equation for $\mathcal{G}_\epsilon$. All that is needed to complete this estimate is the boundedness of s^ϵ, which will indeed hold due to the inherited regularity of the kernel. We will nonetheless delay this detail until the proof of Theorem 3.8 (in order to isolate the pieces of the argument which depend on the regularity of the strength). A stronger estimate on $\mathcal{H}_\epsilon$ is achieved by retaining the first dissipative term in (41). From [16], the first term and the last term in (41) combine to control the Fisher information and we obtain

$$\frac{d}{dt}\mathcal{H}_\epsilon \leqslant -\frac{1}{\epsilon}\mathcal{I}_\epsilon + \frac{\epsilon}{4}\int_{\mathbb{T}^n\times\mathbb{R}^n}\mathrm{s}^\epsilon|v-\mathbf{u}^\epsilon|^2 f^\epsilon\,\mathrm{d}v\,\mathrm{d}x - \|\mathbf{u}^\epsilon\|_{L^2(\kappa_\epsilon)^2} + (\mathbf{u}^\epsilon,[\mathbf{u}^\epsilon]_{\rho^\epsilon})_{\kappa^\epsilon}. \tag{43}$$

Turning to the terms in (42) which do not depend on the strength, we have due to the assumed smoothness of $\mathbf{u}$,

$$\Big| \int_{\mathbb{T}^n} \rho^\epsilon (\mathbf{u}^\epsilon - \mathbf{u}) \cdot \nabla \mathbf{u} \cdot (\mathbf{u}^\epsilon - \mathbf{u}) \,\mathrm{d}x \Big| \lesssim \int_{\mathbb{T}^n} \rho^\epsilon |\mathbf{u}^\epsilon - \mathbf{u}|^2 \,\mathrm{d}x \leqslant \mathcal{H}(f^\epsilon | \mathbf{u}).$$

The local alignment term is the same as in the monokinetic case and is handled the same way with the help of Lemma 3.4, see [16, Theorem 9.2]:

$$\frac{1}{\epsilon} \int_{\mathbb{T}^n} \rho^\epsilon (\mathbf{u}^\epsilon_\delta - \mathbf{u}^\epsilon) \cdot \mathbf{u} \,\mathrm{d}x \lesssim \frac{\delta}{\epsilon}.$$

The remaining estimates depend on the regularity of the strength and will be addressed in the proof of Theorem 3.8. The Fokker-Planck penalization force destroys any uniform control on the momentum uniformly in ϵ so the strength will not inherit the regularity of the kernel in this case. Instead, we identify a set of continuity conditions (R1)-(R4) which are sufficient for proving the Maxwellian Limit. Notably, these conditions hold for the w-model, which we show (along with the proof of Theorem 3.8) in Section 6.2.

Theorem 3.8. *(s-model: Maxwellian Limit) Suppose the continuity conditions (R1)-(R4) hold. Let $(\rho, \mathrm{s}, \boldsymbol{u})$ be a smooth, non-vacuous solution to (58) on the torus $\mathbb{T}^n$ and on the time interval $[0,T]$ with mass M. Suppose also that $\mathrm{s}^\epsilon_0 \in C^\infty(\mathbb{T}^n)$ and $f^\epsilon_0 \in C^k_0(\mathbb{T}^n \times \mathbb{R}^n)$ is a family of initial conditions satisfying*

(i) $\mathcal{H}(f^\epsilon_0 | \mu_0) \to 0$ *as* $\epsilon \to 0$,
(ii) $\mathrm{s}^\epsilon_0 = \mathrm{s}_0 \in C^\infty(\mathbb{T}^n)$.

Then for $\delta = o(\epsilon)$,

$$\sup_{t \in [0,T]} \mathcal{H}(f^\epsilon | \mu) \to 0.$$

As a consequence of this and Proposition 2.3, $\sup_{t\in[0,T]} \|(\mathrm{s}^\epsilon - \mathrm{s})(t,\cdot)\|_\infty \to 0$.

4. Well-posedness of the discrete-continuous s-model

The well-posedness of the discrete-continuous system must be established before addressing the mean field passage to the kinetic system. This section is devoted to proving global existence and uniqueness of solutions to (11) in class

$$(\{x_i(t)\}_{i=1}^N, \{v_i(t)\}_{i=1}^N, \mathrm{s}(t)) \in C([0,T]; \mathbb{R}^{nN} \times \mathbb{R}^{nN} \times H^k(\mathbb{T}^n)), \quad k > n/2 + 2.$$

The choice of $k > n/2 + 2$ guarantees that $\|\partial^2 \mathrm{s}\|_\infty \lesssim \|\mathrm{s}\|_{H^k}$ for an arbitrary second order partial derivative ∂^2 by the Sobolev Embedding Theorem.

Let us recall our main assumptions which will be relevant here: the velocity averaging satisfies the integral representation (5) and the kernel Φ_{ρ^N} is smooth. Written in the terms of the discrete empirical data $\rho^N, \mathbf{u}^N$, the velocity averaging becomes:

$$[\mathbf{u}^N]_{\rho^N}(t,x) = \sum_{j=1}^N m_j \Phi_{(\rho^N)}(x, x_j) v_j(t), \quad (\rho^N)(t,x) = \sum_{j=1}^N m_j \delta_{x_j(t)}.$$

From Proposition 2.1, $[\mathbf{u}^N]_{\rho^N}$ inherits the regularity of the kernel and, in particular, satisfies ($\mathbf{u}$Reg1). For the remainder of this section, we will use $C := C(k, M)$ to denote a constant depending on k and M. For instance, the inherited regularity of velocity averaging ($\mathbf{u}$Reg1) can be written: $\|\partial^k[\mathbf{u}^N]_{\rho^N}\|_\infty \leqslant C$. The constant C may change line by line. Consider the following regularized version of (11):

$$\begin{cases} \dot{x}_i = v_i \\ \dot{v}_i = \lambda \mathrm{s}(x_i)([\mathbf{u}^N]_{\rho^N}(x_i) - v_i) \\ \partial_t \mathrm{s} + \nabla_x \cdot (\mathrm{s}[\mathbf{u}^N]_{\rho^N}) = \epsilon\Delta \mathrm{s}. \end{cases} \tag{44}$$

For the moment, we avoid writing the explicit dependence of x_i, v_i, s on ϵ for the sake of brevity. They are not to be confused with solutions to the unregularized system (11). Let $X = C([0,T];\mathbb{R}^{2nN} \times H^k(\mathbb{T}^n))$ and define $Z(t) := (\{x_i(t)\}_{i=1}^N, \{v_i(t)\}_{i=1}^N, \mathrm{s}(t))$. The norm $\|\cdot\|_X$ is given by:

$$\|Z\|_X = \sup_{t\in[0,T]} \max_{i=1\dots N} \|x_i(t)\| + \|v_i(t)\| + \|\mathrm{s}(t,\cdot)\|_{H^k(\mathbb{R}^n)}.$$

Define the map $\mathcal{F}$ by

$$\mathcal{F}(Z(t)) = \begin{bmatrix} 1 \\ 1 \\ e^{\epsilon t\Delta} \end{bmatrix} Z_0 + \int_0^t \begin{bmatrix} 1 \\ 1 \\ e^{\epsilon(t-\tau)\Delta} \end{bmatrix} A(\tau) d\tau, \tag{45}$$

where A represents all of the non-laplacian terms. Existence and uniqueness of solutions to (44) amounts to showing that $\mathcal{F} : B_1(Z_0) \mapsto B_1(Z_0)$ and that it is a contraction mapping, where $B_1(Z_0)$ is the ball of radius 1 centered at Z_0 in X. Contraction will follow similarly from invariance. For invariance, we aim to show that:

$$\|\mathcal{F}(Z(t)) - Z_0\|_X \leqslant \left\| e^{\epsilon t\Delta} Z_0 - Z_0 \right\|_X + \left\| \int_0^t e^{\epsilon(t-\tau)\Delta} A(Z(\tau)) d\tau \right\|_X \leqslant 1.$$

The first term is small for small T due to the continuity of the heat semigroup. For second term, we will treat each component individually. The x_i-component gives

$$\|x_i(t) - x_i(0)\| \leqslant T \max_i \|v_i(0)\|.$$

For the v_i-component, we have

$$\|v_i(t) - v_i(0)\| \leqslant T\|\mathrm{s}\|_\infty (\|[\mathbf{u}^N]_{\rho^N}\|_\infty + \max_i \|v_i(0)\|) \leqslant CT(\|Z_0\| + 1).$$

For the s-component, we will use the analyticity property of the heat semigroup:

$$\|\nabla e^{\epsilon t\Delta} f\|_{L^2} \leqslant \frac{1}{\sqrt{\epsilon t}} \|f\|_{L^2}$$

along with the product estimate to get

$$\begin{aligned} \left\| \int_0^t e^{\epsilon(t-\tau)\Delta} \nabla_x \cdot (\mathrm{s}[\mathbf{u}^N]_{\rho^N}) d\tau \right\|_{H^k} &\leqslant \frac{2T^{1/2}}{\epsilon^{1/2}} \sup_{t\in[0,T]} \|\mathrm{s}[\mathbf{u}^N]_{\rho^N}\|_{H^k} \\ &\leqslant \frac{2T^{1/2}}{\epsilon^{1/2}} \sup_{t\in[0,T]} (\|\mathrm{s}\|_\infty \|[\mathbf{u}^N]_{\rho^N}\|_{H^k} + \|\mathrm{s}\|_{H^k} \|[\mathbf{u}^N]_{\rho^N}\|_\infty) \\ &\leqslant \frac{2T^{1/2}}{\epsilon^{1/2}} (\|Z_0\|_X + 1). \end{aligned}$$

For small enough T, we have invariance. Contractivity of $\mathcal{F}$ follows from similar estimates. This time interval of existence T could depend on ϵ. To establish that the there is a common time interval of existence independent of ϵ, we establish an ϵ-independent energy estimate.

For the x_i and v_i components, the estimates follow easily from the invariance estimates. We record them here.

$$\|x_i(t)\| \leqslant \|x_i(0)\| + t \max_i \|v_i(0)\|,$$

and

$$\|v_i(t)\| \leqslant \|v_i(0)\| + t\|\mathrm{s}\|_\infty(\|[\mathbf{u}^N]_{\rho^N}\|_\infty + \max_i \|v_i(0)\|) \leqslant C(1 + t\|\mathrm{s}\|_{H^k}).$$

For the strength, we multiply (44) by $\partial^{2j}\mathrm{s}$ and integrate by parts. We obtain for any $0 \leqslant j \leqslant k$,

$$\frac{d}{dt}\|\mathrm{s}\|_{\dot{H}^j} = \int \nabla_x \cdot [\mathbf{u}^N]_{\rho^N} |\partial^j \mathrm{s}|^2 \,\mathrm{d}x - \int (\partial^j([\mathbf{u}^N]_{\rho^N} \cdot \nabla_x \mathrm{s}) - [\mathbf{u}^N]_{\rho^N} \cdot \nabla_x \partial^j \mathrm{s})\partial^j \mathrm{s} \,\mathrm{d}x \qquad (46)$$
$$- \int \partial^j((\nabla_x \cdot [\mathbf{u}^N]_{\rho^N})\mathrm{s})\partial^j \mathrm{s} \,\mathrm{d}x - \epsilon \int |\partial^j \nabla \mathrm{s}|^2 \,\mathrm{d}x.$$

Dropping the ϵ term coming from the Laplacian, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\mathrm{s}\|^2_{\dot{H}^j} \leqslant C\big(\|\mathrm{s}\|^2_{\dot{H}^j} + \|\mathrm{s}\|_{\dot{H}^j}\|\nabla \mathrm{s}\|_\infty + \|\mathrm{s}\|_{\dot{H}^j}\|\mathrm{s}\|_\infty\big), \quad \text{for all } 0 \leqslant j \leqslant k.$$

Since $k > n/2 + 2$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\mathrm{s}\|^2_{H^k} \leqslant C\|\mathrm{s}\|^2_{H^k}. \qquad (47)$$

We will now denote the explicit dependencies on ϵ and take $\epsilon \to 0$. From Grownwall, we conclude that $Z^\epsilon(t)$ exists on a common time interval independent of ϵ. Writing the equation for $(\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{s}^\epsilon)^2$, we have

$$\Big\|\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{s}^\epsilon\Big\|^2_{L^2} \leqslant C\|\mathrm{s}^\epsilon\|_{H^1} + \epsilon\|\mathrm{s}^\epsilon\|_{H^2} \leqslant C\|\mathrm{s}^\epsilon\|_{H^k}.$$

By local well-posedness $Z^\epsilon \in C([0,T]; R^{2nN} \times H^k(\mathbb{T}^n))$ so that $\frac{d}{dt}Z^\epsilon \in L^2([0,T]; \mathbb{R}^{2nN} \times L^2(\mathbb{T}^n))$. By the Aubin-Lions lemma, we obtain a subsequence, which we denote again by Z^ϵ such that $Z^\epsilon \to Z^0$ in $C([0,T]; R^{2nN} \times H^{k-1}(\mathbb{T}^n))$. Since $H^{k-1}(\mathbb{T}^n)$ is dense in $H^k(\mathbb{T}^n)$, we have $Z^0 \in C_w([0,T]; \mathbb{R}^{2nN} \times H^k(\mathbb{T}^n))$. Finally, since $k > n/2 + 2$, the terms A^ϵ converge pointwise to A. Taking $\epsilon \to 0$ in the Duhamel formula (45), we get

$$Z^0(t) = Z^0_0 + \int_0^t A(\tau) d\tau.$$

That is, $Z^0 \in C_w([0,T]; \mathbb{R}^{2nN} \times H^k(\mathbb{T}^n))$ solves (11). Finally, we note that due to the ϵ-independent energy estimate (47), $\|Z^0\|_X$ remains bounded for any finite time and thus exists for all time. This concludes existence and uniqueness of solutions to (11) on the global time interval $[0, \infty)$.

5. Mean Field Limit

We establish the passage from from the discrete system (11) to the kinetic system (13) in the mean field limit for the s-model, Theorem 3.2. The outline for this argument is discussed in Section 3.1 for the Cucker-Smale model. We will skip the pieces of the argument which are covered there (in particular, we will skip the pieces of the argument in which s has an insignificant role).

The empirical measures are given by:

$$\mu_t^N(x,v) = \sum_{i=1}^N m_i \delta_{x_i(t)} \otimes \delta_{v_i(t)}, \qquad (x_i(t), v_i(t), \mathrm{s}(t,x))_{i=1}^N \quad \text{solve (11)}. \tag{48}$$

To make sense of measure valued solutions, we consider a weak version of (13).

Definition 5.1. Fix a time $T > 0$ and an integer $k \geqslant 0$. We say the pair (μ, s) with $\mu \in C_{w^*}([0,T]; \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n))$ and $\mathrm{s} \in C([0,T]; C^k(\mathbb{T}^n))$ is a *weak* solution to (13) if for all $g(t,x,v) \in C_0^\infty([0,T] \times \mathbb{T}^n \times \mathbb{T}^n)$ and for all $0 < t < T$,

$$\begin{cases} \int_{\mathbb{T}^n \times \mathbb{R}^n} g(t,x,v) d\mu_t(x,v) & = \int_{\mathbb{T}^n \times \mathbb{R}^n} g(0,x,v) d\mu_0(x,v) \\ & \quad + \int_0^t \int_{\mathbb{T}^n \times \mathbb{R}^n} (\partial_\tau g + v \cdot \nabla_x g + \mathrm{s}([\mathbf{u}]_\rho - v) \cdot \nabla_v g) d\mu_\tau(x,v) \\ \partial_t \mathrm{s} + \nabla_x \cdot (\mathrm{s}[\mathbf{u}]_\rho) = 0. \end{cases} \tag{49}$$

In other words, μ solves the first equation weakly and the strength, s, solves the second equation strongly.

As in the Cucker-Smale case, the empirical measure (48) is a solution to (49) if and only if $(x_i(t), v_i(t), \mathrm{s}(t,x))$ solve the discrete system (11). The well-posedness of (49) for empirical measure-valued solutions is therefore equivalent to the well-posedness of the discrete-continuous system (11) established in Section 4. To show existence of general weak solutions to (49), we will show that the weak solution arises a weak limit of the empirical measures (48). The goal is to establish the Wasserstein-1 stability estimate (20).

Stability in the Wasserstein-1 metric amounts to showing the stability estimates (23) and (24) on the characteristic flow:

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t} X(t,s,x,v) = V(t,s,x,v), & X(s,s,x,v) = x \\ \frac{\mathrm{d}}{\mathrm{d}t} V(t,s,x,v) = \mathrm{s}(X)([\mathbf{u}]_\rho(X) - V), & V(s,s,x,v) = v, \end{cases} \tag{50}$$

which will be established in the following Lemmas. Let us first note that the total momentum $\int_{\mathbb{T}^n} \mathbf{u}\rho \,\mathrm{d}x = \int_{\mathbb{T}^n \times \mathbb{R}^n} v d\mu_t(x,v)$ in (49) is conserved (by plugging in $g = v$). In particular, $J < \infty$ and we can apply Propositions 2.1 and 2.3 and use the inherited regularity from the kernel. By a similar argument to the Cucker-Smale case, we also have that μ_t is the push-forward of μ_0 along the characteristic flow:

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} h(X(t,\omega), V(t,\omega)) \,\mathrm{d}\mu_0(\omega) = \int_{\mathbb{T}^n \times \mathbb{R}^n} h(\omega) \,\mathrm{d}\mu_t(\omega), \qquad \omega = (x,v) \tag{51}$$

Lemma 5.2. *(Deformation Tensor Estimates) Suppose that the velocity averaging has the integral representation* (5) *and that the kernel satisfies the uniform regularity assumptions* (Reg1) *and* (Reg2)*. Let* (μ, s) *be a weak solution to* (49) *on* $[0,T]$ *with characterstics* X, V *given in* (50)*. Then*

$$\|\nabla X\|_\infty + \|\nabla V\|_\infty \leqslant C(M,R,J,T).$$

Proof. Differentiating (21),

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t} \nabla X = \nabla V \\ \frac{\mathrm{d}}{\mathrm{d}t} \nabla V = \nabla X^T \nabla \mathrm{s}(X)([\mathbf{u}]_\rho(X) - V) + \mathrm{s}(X) \nabla X^T \nabla [\mathbf{u}]_\rho(X) - \mathrm{s}(X) \nabla V. \end{cases}$$

By the maximum principle on the velocity and the inherited regularity (**u**Reg1), (**u**Reg2), and (sReg1)

$$\frac{\mathrm{d}}{\mathrm{d}t}(\|\nabla V\|_\infty + \|\nabla X\|_\infty) \leqslant C'(M,R,J,T)(\|\nabla X\|_\infty + \|\nabla V\|_\infty).$$

We conclude by Gronwall. □

Lemma 5.3. *(Continuity Estimates) Suppose that the velocity averaging has the integral representation* (5) *and that the kernel satisfies the uniform regularity assumptions* (Reg1) *and* (Reg2)*. Let* $(\mu',\mathrm{s}'), (\mu'',\mathrm{s}'')$ *be weak solutions to* (49) *on* $[0,T]$*; and let* X', X'', V', V'' *be the corresponding characteristics given by* (50)*. Then*

$$\|X'-X''\|_\infty + \|V'-V''\|_\infty \leqslant C(M,R,J,T)W_1(\mu_0',\mu_0'').$$

Proof. We have

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}(X'-V') = V'-V'' \\ \frac{\mathrm{d}}{\mathrm{d}t}(V'-V'') = \mathrm{s}'(X')([\mathbf{u}']_{\rho'}(X')-V') - \mathrm{s}''(X'')([\mathbf{u}'']_{\rho''}(X'')-V''). \end{cases}$$

By the maximum principle on the velocity and inherited regularity conditions (**u**Reg1), (**u**Reg2), (sReg1), (sReg2),

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\|V'-V''\|_\infty &\leqslant \|\mathrm{s}'(X')-\mathrm{s}'(X'')\|_\infty\|[\mathbf{u}']_{\rho'}(X')-V'\|_\infty + \|\mathrm{s}'(X'')-\mathrm{s}''(X'')\|_\infty\|[\mathbf{u}']_{\rho'}(X')-V'\|_\infty \\ &\quad + \|\mathrm{s}''(X'')\|_\infty\|\|[\mathbf{u}'']_{\rho''}(X')-V'-[\mathbf{u}'']_{\rho''}(X'')+V''\|_\infty \\ &\leqslant C(M,R,J,T)\Big(\|V'-V''\|_\infty + \sup_{t\in[0,T]}(W_1(\rho',\rho'')+W_1(\mathbf{u}'\rho',\mathbf{u}''\rho''))(t)\Big). \end{aligned}$$

Combining with $\frac{\mathrm{d}}{\mathrm{d}t}\|X'-X''\|_\infty \leqslant \|V'-V''\|_\infty$, we get

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t}\big(\|X'-X''\|_\infty + \|V'-V''\|_\infty\big) \qquad (52) \\ &\quad\leqslant C(M,R,J,T)\Big(\|X'-X''\|_\infty + \|V'-V''\|_\infty + \sup_{t\in[0,T]}(W_1(\rho',\rho'')+W_1(\mathbf{u}'\rho',\mathbf{u}''\rho''))(t)\Big). \end{aligned}$$

To estimate $W_1(\rho',\rho'')$ and $W_1(\mathbf{u}'\rho',\mathbf{u}''\rho'')$, we use the fact that μ_t is the push-forward of μ_0, (51). Fix $\|g\|_{Lip} \leqslant 1$. Since $M'=M$,

$$\begin{aligned} \int_{\mathbb{T}^n} g(x)(d\rho_t'-d\rho_t'') &= \int_{\mathbb{T}^n\times\mathbb{R}^n} g(x)(d\mu_t'-d\mu_t'') = \int_{\mathbb{T}^n\times\mathbb{R}^n} g(X')d\mu_0' - \int_{\mathbb{T}^n\times\mathbb{R}^n} g(X'')d\mu_0'' \\ &= \int_{\mathbb{T}^n\times\mathbb{R}^n} (g(X')-g(X''))d\mu_0' + \int_{\mathbb{T}^n\times\mathbb{R}^n} g(X'')(d\mu_0'-d\mu_0'') \\ &\leqslant M\|X'-X''\|_\infty + \|\nabla X''\|_\infty W_1(\mu_0',\mu_0''). \end{aligned}$$

For $W_1(\mathbf{u}'\rho', \mathbf{u}''\rho'')$, we have

$$\begin{aligned}\int_{\mathbb{T}^n} g(x)(d(\mathbf{u}'\rho'_t) - d(\mathbf{u}'\rho''_t)) &= \int_{\mathbb{T}^n\times\mathbb{R}^n} vg(x)(d\mu'_t - d\mu''_t)\\ &= \int_{\mathbb{T}^n\times\mathbb{R}^n} V'g(X')d\mu'_0 - \int_{\mathbb{T}^n\times\mathbb{R}^n} V''g(X'')d\mu''_0\\ &= \int_{\mathbb{T}^n\times\mathbb{R}^n} (V'g(X') - V''g(X''))d\mu'_0 + \int_{\mathbb{T}^n\times\mathbb{R}^n} V''g(X'')(d\mu'_0 - d\mu''_0)\\ &\leqslant M\|g\|_\infty\|V' - V''\|_\infty + M\|X' - X''\|_\infty R\\ &+ \big(\|g\|_\infty\|\nabla V''\|_\infty + R\|\nabla X''\|_\infty\big)W_1(\mu'_0, \mu''_0).\end{aligned}$$

These estimates hold uniformly in time. Plugging these into (52) and using Lemma 5.2, we conclude by Gronwall. □

An immediate Corollary is that the regularity conditions (**u**Reg2) and (sReg2) can be restated in terms of the distance $W_1(\mu'_0, \mu''_0)$.

Corollary 5.4. *Suppose that the velocity averaging has the integral representation* (5) *and that the kernel satisfies the uniform regularity assumptions* (Reg1) *and* (Reg2)*. Let* $(\mu', \mathrm{s}'), (\mu'', \mathrm{s}'')$ *be weak solutions to* (49) *on* $[0, T]$*. Then*

$$\sup_{t\in[0,T]} \big(W_1(\rho', \rho'') + W_1(\boldsymbol{u}'\rho', \boldsymbol{u}''\rho'')\big)(t) \leqslant C_1(M, R, J, T)W_1(\mu'_0, \mu''_0),$$

and

$$\sup_{t\in[0,T]} \|\partial^k([\boldsymbol{u}']_{\rho'} - [\boldsymbol{u}'']_{\rho''})(t, \cdot)\|_\infty \leqslant C_2(k, M, R, J, T)W_1(\mu'_0, \mu''_0), \tag{$\mathbf{u}$Lip2'}$$

$$\sup_{t\in[0,T]} \|\partial^k(\mathrm{s}' - \mathrm{s}'')(t, \cdot)\|_\infty \leqslant C_3(k, M, R, J, T)\Big(\|\partial^k(\mathrm{s}'_0 - \mathrm{s}''_0)\|_\infty + W_1(\mu'_0, \mu''_0)\Big). \tag{sLip2'}$$

Lemmas 5.2, 5.3 and inequality (22) imply the desired Wasserstein-1 stability (20). We conclude that the empirical measures μ^N_t converge in the Wasserstein-1 metric to some $\mu \in C_{w^*}([0, T]; \mathcal{P}(\mathbb{T}^n \times B_R))$ uniformly on $[0, T]$ (the weak* continuity of μ owes to the weak* continuity of the empirical measures and the uniform convergence on $[0, T]$). In addition, by Corollary 5.4, for any $k \geqslant 0$, s^N_t converges in C^k to some $\mathrm{s}_t \in C^k(\mathbb{T}^n)$ uniformly on $[0, T]$. It remains to show that the limiting pair (μ, s) is in fact a weak solution to (13) in the sense of Definition 5.1.

Lemma 5.5. *Suppose a sequence of solutions* $\mu^N \in C_{w^*}([0, T]; \mathcal{P}(\mathbb{T}^n \times B_R))$ *converge weakly pointwise, i.e.* $\mu^N_t \to \mu_t$*, uniformly for all* $t \in [0, T]$*; and suppose that for any* $k \geqslant 0$, s^N_t *converges in* $C^k(\mathbb{T}^n)$ *to* s_t*, i.e.* $\|\partial^k(\mathrm{s}^N_t - \mathrm{s}_t)\|_\infty \to 0$*, uniformly for all* $t \in [0, T]$*. Then* $(\mu, \mathrm{s}) \in C_{w^*}([0, T]; \mathcal{P}(\mathbb{T}^n \times B_R)) \times C([0, T], C^k(\mathbb{T}^n))$ *is a weak solution to* (13)*.*

Proof. In the following proof, $C := C(M, R, J, T)$ and $C' := C(k, M, R, J, T)$. We have already observed that, due to Corollary (5.4), for any k, s^N_t converges in C^k to some $\mathrm{s}_t \in C^k(\mathbb{T}^n)$ uniformly on $[0, T]$. In addition $\partial_t \mathrm{s}^N = -\nabla \cdot (\mathrm{s}^N[\mathbf{u}^N]_{\rho^N})$ is uniformly bounded on $[0, T]$ due to (**u**Reg1) and (sReg1). So, $\mathrm{s} \in C([0, T]; C^k(\mathbb{T}^n))$. To show that s solves the transport equation $\partial_t \mathrm{s} + \nabla \cdot (\mathrm{s}[\mathbf{u}]_\rho) = 0$, let $\tilde{X}$ be the characteristic flow,

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{X}(t, \alpha) = [\mathbf{u}]_\rho(t, \tilde{X}(t, \alpha)), \quad \tilde{X}(0, \alpha) = \alpha.$$

Similarly, let $\tilde{X}^N$ be the characteristic flow for the empirical strength s_t^N which solves $\partial_t \mathrm{s}^N + \nabla\cdot(\mathrm{s}^N[\mathbf{u}^N]_{\rho^N}) = 0$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{X}^N(t,\alpha) = [\mathbf{u}^N]_{\rho^N}(t,\tilde{X}^N(t,\alpha)), \qquad \tilde{X}^N(0,\alpha)=\alpha.$$

Abbreviating $\tilde{X}(t,\alpha), \tilde{X}^N(t,\alpha)$ by $\tilde{X}, \tilde{X}^N$ and using (**u**Reg1) and (**u**Lip2'), we have for all $k \geqslant 0$

$$\begin{aligned} &\|\partial^k([\mathbf{u}^N]_{\rho^N}(\tau,\tilde{X}^N) - [\mathbf{u}]_\rho(\tau,\tilde{X}))\|_\infty \\ &\quad\leqslant \|\partial^k([\mathbf{u}^N]_{\rho^N}(\tau,\tilde{X}) - [\mathbf{u}]_\rho(\tau,\tilde{X}))\|_\infty + \|\partial^k([\mathbf{u}^N]_{\rho^N}(\tau,\tilde{X}^N) - [\mathbf{u}^N]_{\rho^N}(\tau,\tilde{X}))\|_\infty \\ &\quad\leqslant \|\partial^k([\mathbf{u}^N]_{\rho^N}(\tau,\tilde{X}) - [\mathbf{u}]_\rho(\tau,\tilde{X}))\|_\infty + \|\partial^{k+1}[\mathbf{u}^N]_{\rho^N}\|_\infty\|\tilde{X}^N - \tilde{X}\|_\infty \\ &\quad\leqslant C'\big(W_1(\mu_0^N,\mu_0) + \|\tilde{X}^N - \tilde{X}\|_\infty\big). \end{aligned} \tag{53}$$

As a result,

$$\frac{\mathrm{d}}{\mathrm{d}t}\|(\tilde{X}^N - \tilde{X})(t,\cdot)\|_\infty \leqslant C\big(W_1(\mu_0', \mu_0'') + \|(\tilde{X}^N - \tilde{X})(t,\cdot)\|_\infty\big), \quad \text{for all } t\in[0,T).$$

By Gronwall and $\tilde{X}(0,\alpha) = \tilde{X}^N(0,\alpha) = \alpha$, we obtain

$$\|(\tilde{X}^N - \tilde{X})(t,\cdot)\|_\infty \leqslant CW_1(\mu_0^N,\mu_0), \quad \text{for all } t\in[0,T). \tag{54}$$

Now solving along the characteristic $\tilde{X}^N$,

$$\mathrm{s}^N(t,\tilde{X}^N(t,\alpha)) = \mathrm{s}_0^N(\alpha)\exp\Big\{\int_0^t \nabla\cdot[\mathbf{u}^N]_{\rho^N}(\tau,\tilde{X}^N(s,\alpha))d\tau\Big\}.$$

With (53) and (54) in hand, we obtain uniform in time convergence to

$$\mathrm{s}(t,\tilde{X}(t,\alpha)) = \mathrm{s}_0(\alpha)\exp\Big\{\int_0^t \nabla\cdot[\mathbf{u}]_\rho(\tau,\tilde{X}(s,\alpha))d\tau\Big\}.$$

So, s_t is a solution to $\partial_t\mathrm{s} + \nabla(\mathrm{s}[\mathbf{u}]_\rho) = 0$. Turning to the convergence of the empirical measures, the weak convergence $W_1(\mu_t^N,\mu_t)\to 0$ immediately implies that the linear terms in (49) converge. Let us address the nonlinear term.

$$\begin{aligned} &\int_0^t\int_{\mathbb{R}^{2n}} \nabla_v g\mathrm{s}^N([\mathbf{u}^N]_{\rho^N} - v))d\mu_\tau^N(x,v) - \int_0^t\int_{\mathbb{R}^{2n}} \nabla_v g\mathrm{s}([\mathbf{u}]_\rho - v))d\mu_\tau(x,v) \\ &\qquad\leqslant \|\nabla_v g\|_\infty \int_0^T\int_{\mathbb{R}^{2n}} \|\mathrm{s}^N([\mathbf{u}^N]_{\rho^N} - v)) - \mathrm{s}([\mathbf{u}]_\rho - v))\|_\infty d\mu_\tau^N(x,v) \\ &\qquad+ \|\nabla_v g\|_\infty \int_0^T\int_{\mathbb{R}^{2n}} \|\mathrm{s}([\mathbf{u}]_\rho - v))\|_\infty d(\mu_\tau^N(x,v) - \mu_\tau(x,v)). \end{aligned}$$

The second term goes to zero by weak convergence. For the first term, we simply use the regularity conditions (**u**Reg1), (**u**Lip2'), (sReg1), (sLip2') to get

$$\begin{aligned} \|\mathrm{s}^N([\mathbf{u}^N]_{\rho^N} - v)) - \mathrm{s}([\mathbf{u}]_\rho - v))\|_\infty &\leqslant \|\mathrm{s}^N - \mathrm{s}\|_\infty\|[\mathbf{u}^N]_{\rho^N}\|_\infty + \|\mathrm{s}\|_\infty\|[\mathbf{u}^N]_{\rho^N} - [\mathbf{u}]_\rho\|_\infty \\ &\leqslant CW_1(\mu_0^N,\mu_0). \end{aligned}$$

This gives $\|\mathrm{s}^N([\mathbf{u}^N]_{\rho^N} - v)) - \mathrm{s}([\mathbf{u}]_\rho - v))\|_\infty \to 0$ uniformly on $[0,T]$. □

6. Hydrodynamic Limits

The goal of this section is to establish a passage from the kinetic description (13) to the corresponding macroscopic description in the monokinetic and Maxwellian limiting regimes, Theorems 3.6 and 3.8. The arguments presented here contain similar estimates to the hydrodynamic limiting arguments presented for environmental averaging models. We will therefore often refer to Section 3.2 where relevant quantities and estimates are presented in the context Cucker-Smale model (i.e. when the regularity of the strength is known).

From the introduction, the v-moments of the mesoscopic system (13) yields (14) with the Reynolds Stress tensor $\mathcal{R}$ given by (10). The system is closed by adding a local alignment force $F(f^\epsilon)$ to the kinetic equation (13):

$$\begin{cases} \partial_t f^\epsilon + v\cdot\partial_x f^\epsilon = \nabla_v\cdot(\mathrm{s}^\epsilon(v-[\mathbf{u}^\epsilon]_{\rho^\epsilon})f^\epsilon) + \frac{1}{\epsilon}F(f^\epsilon) \\ \partial_t \mathrm{s}^\epsilon + \nabla\cdot(\mathrm{s}[\mathbf{u}^\epsilon]_{\rho^\epsilon}) = 0. \end{cases} \tag{55}$$

We will consider two such local alignment forces corresponding to the monokinetic and Maxwellian regimes. In the Monokinetic regime, $F(f^\epsilon)$ is given by (26). This monokinetic local alignment force pushes the kinetic solution towards the monokinetic ansatz

$$f(t,x,v) = \rho(t,x)\delta(v-\mathbf{u}(t,x)),$$

where $(\rho,\mathrm{s},\mathbf{u})$ solve (SM) (Theorem 3.6). The Reynolds stress tensor becomes zero and the macroscopic description becomes

$$\begin{cases} \partial_t\rho + \nabla\cdot(\mathbf{u}\rho) = 0 \\ \partial_t\mathrm{s} + \nabla\cdot(\mathrm{s}[\mathbf{u}]_\rho) = 0 \\ \partial_t(\rho\mathbf{u}) + \nabla\cdot(\rho\mathbf{u}\otimes\mathbf{u}) = \rho\mathrm{s}([\mathbf{u}]_\rho - \mathbf{u}). \end{cases} \tag{56}$$

In the non-vacuous case, we can divide the $\mathbf{u}$-equation by ρ and this becomes (SM). The inherited regularity of the strength from the kernel will once again play a crucial role in the proof of the monokinetic limit.

For the Maxwellian regime, we will rewrite the v-moments system (14) as:

$$\begin{cases} \partial_t\rho + \nabla\cdot(\mathbf{u}\rho) = 0 \\ \partial_t\mathrm{s} + \nabla\cdot(\mathrm{s}[\mathbf{u}]_\rho) = 0 \\ \partial_t(\rho\mathbf{u}) + \nabla\cdot(\rho\mathbf{u}\otimes\mathbf{u}) + \nabla\rho + \nabla\cdot\mathcal{R} = \rho\mathrm{s}([\mathbf{u}]_\rho - \mathbf{u}), \end{cases} \tag{57}$$

where $\mathbf{u}_\delta$ is given in (30) and

$$\mathcal{R}(t,x) = \int_{\mathbb{T}^n}((v-\mathbf{u})\otimes(v-\mathbf{u}) - \mathbb{I})\varphi\,\mathrm{d}v.$$

To close this system, we consider a strong Fokker-Planck force $F(f^\epsilon)$ given by (29), which pushes the kinetic solution towards the Maxwellian,

$$\mu(t,x) = \frac{\rho(t,x)}{(2\pi)^{n/2}}e^{\frac{-|v-\mathbf{u}(t,x)|^2}{2}},$$

where $(\rho, \mathrm{s}, \mathbf{u})$ solve (58) (Theorem 3.8). The Reynolds stress term vanishes and the system becomes

$$\begin{cases} \partial_t \rho + \nabla \cdot (\mathbf{u}\rho) = 0 \\ \partial_t \mathrm{s} + \nabla \cdot ([\mathbf{u}]_\rho \mathrm{s}) = 0 \\ \partial_t(\rho \mathbf{u}) + \nabla \cdot (\rho \mathbf{u} \otimes \mathbf{u}) + \nabla \rho = \rho \mathrm{s}([\mathbf{u}]_\rho - \mathbf{u}). \end{cases} \tag{58}$$

Unfortunately, the total momentum of solutions to (55) with the strong Fokker-Planck force $F(f^\epsilon)$ (29) is not necessarily bounded uniformly in ϵ, so Propsitions 2.1 and 2.3 do not apply and s does not inherit regularity from the kernel. We've instead identified a set of sufficient continuity conditions for the solution (R1)-(R4) given in Section 6.2. An important class of models for which these conditions can be verified is the w-model with the interaction kernel ϕ bounded from below away from zero.

6.1. **Monokinetic Regime.** We consider (13) under the local alignment force (26):

$$\begin{cases} \partial_t f^\epsilon + v \cdot \partial_x f^\epsilon = \nabla_v \cdot (\mathrm{s}^\epsilon (v - [\mathbf{u}^\epsilon]_{\rho^\epsilon}) f^\epsilon) + \frac{1}{\epsilon} \nabla_v \cdot ((v - \mathbf{u}_\delta^\epsilon) f^\epsilon) \\ \partial_t \mathrm{s}^\epsilon + \nabla \cdot (\mathrm{s}^\epsilon [\mathbf{u}^\epsilon]_{\rho^\epsilon}) = 0, \end{cases} \tag{59}$$

where $\mathbf{u}_\delta$ is the special mollification given in (30).

We will first verify the maximum principle on the velocity so that we can use the inherited regularity of the strength. The characteristic equations of (59) are given by:

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t} X(t,s,x,v) = V(t,s,x,v), & X(s,s,x,v) = x \\ \frac{\mathrm{d}}{\mathrm{d}t} V(t,s,x,v) = \mathrm{s}(X)([\mathbf{u}^\epsilon]_{\rho^\epsilon}(X) - V) + \frac{1}{\epsilon}(\mathbf{u}_\delta^\epsilon - V), & V(s,s,x,v) = v. \end{cases} \tag{60}$$

As in the Cucker-Smale and s-model, the measure $f_t(x,v)\,\mathrm{d}x\,\mathrm{d}v$ is the push-forward of $f_0(x,v)\,\mathrm{d}x\,\mathrm{d}v$ along the characteristic flow. Letting $\omega := (x,v)$, $X' := X(t,\omega')$, $V' := V(t,\omega')$, and using the right stochasticity of Φ_ρ, (17), we compute:

$$\begin{aligned} \mathrm{s}(X)([\mathbf{u}^\epsilon]_{\rho^\epsilon}(X) - V) &= \int_{\mathbb{T}^n \times \mathbb{R}^n} \mathrm{s}(X) \Phi_{\rho^\epsilon}(X,x) v f_t^\epsilon(x,v)\,\mathrm{d}x\,\mathrm{d}v - V \\ &= \int_{\mathbb{T}^n \times \mathbb{R}^n} \mathrm{s}(X) \Phi_{\rho^\epsilon}(X,x)(v - V) f_t^\epsilon(x,v)\,\mathrm{d}x\,\mathrm{d}v \\ &= \int_{\mathbb{T}^n \times \mathbb{R}^n} \mathrm{s}(X) \Phi_{\rho^\epsilon}(X,X')(V' - V) f_0^\epsilon(\omega')\,\mathrm{d}\omega'. \end{aligned}$$

Considering compactly supported initial data, $\operatorname{supp} f_0 \subset \mathbb{T}^n \times B_R$, and evaluating at a point of maximum of V, we have

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} \mathrm{s}(X) \Phi_{\rho^\epsilon}(X,X')(V' - V_+) f_0^\epsilon(x,v)\,\mathrm{d}\omega' \leqslant 0, \qquad V_+(t) = \max_{(x,v) \in \mathbb{T}^n \times B_R} |V(t,0,x,v)|.$$

For the local alignment term, recall that $\mathbf{u}_\delta^\epsilon$ is the averaging for the M_ϕ model with kernel given in Table (1.3) with $\phi = \Psi_\delta$ for a smooth mollifier Ψ_δ. Therefore, it is handled similarly. Denoting the kernel by $\Phi_{\rho^\epsilon,\delta}$, we have:

$$\frac{1}{\epsilon} \int_{\mathbb{T}^n \times \mathbb{R}^n} \Phi_{\rho^\epsilon,\delta}(X,X')(V' - V_+) f_0(x,v)\,\mathrm{d}\omega' \leqslant 0.$$

Using the classical Rademacher Lemma, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\|V\|_\infty \leqslant 0.$$

Of course, this implies boundedness of the macroscopic velocity:

$$|\mathbf{u}^\epsilon| = \frac{\left|\int_{B_R} v f^\epsilon dv\right|}{\int_{B_R} f^\epsilon dv} \leqslant \frac{R\int_{B_R} f^\epsilon dv}{\int_{B_R} f^\epsilon dv} = R.$$

With the maximum principle in hand, we can readily apply the inheritance of regularity. Furthermore, since $\operatorname{supp} f^\epsilon, \operatorname{supp} f \subset \mathbb{T}^n \times B_R$, we have $W_1(\mathbf{u}^\epsilon\rho^\epsilon, \mathbf{u}\rho) \leqslant CW_1(f^\epsilon, f) \leqslant CW_2(f^\epsilon, f)$. Similarly, $W_1(\rho^\epsilon, \rho) \leqslant CW_2(f^\epsilon, f)$. We can then modify (**u**Reg2) and (sReg2) to get

$$\|[\mathbf{u}^\epsilon]_{\rho^\epsilon} - [\mathbf{u}]_\rho\|_{L^\infty} \leqslant C(M,J)W_2(f^\epsilon, f) \tag{uReg2'}$$

$$\|\mathrm{s}^\epsilon - \mathrm{s}\|_{L^\infty} \leqslant C(M,J,T)\sup_{t\in[0,T]} W_2(f^\epsilon, f)(t) \quad \text{provided } \mathrm{s}_0^\epsilon = \mathrm{s}_0. \tag{sReg2'}$$

Since we are working on $\mathbb{T}^n$, the L^∞ norm can be interchanged with L^2 norm (with the only difference being the constant C, which is inconsequential for our arguments).

Proof. (proof of Theorem 3.6) When an identity or inequality does not depend on the regularity of the strength and can be found in the proof for the Cucker-Smale case, we refer the reader to Section 3.2.1. In order to control $W_2(f_t^\epsilon, f_t)$, we consider the flow γ_t whose marginals are f_t^ϵ and f_t (i.e. $\gamma_t \in \prod(f_t^\epsilon, f_t)$)

$$\partial_t\gamma + v_1\cdot\nabla_{x_1}\gamma + v_2\cdot\nabla_{x_2}\gamma + \nabla_{v_1}[\gamma\mathrm{s}^\epsilon(v_1 - [\mathbf{u}^\epsilon]_{\rho^\epsilon}) + \frac{1}{\epsilon}(v_1 - \mathbf{u}_\delta^\epsilon)] + \nabla_{v_2}[\gamma\mathrm{s}(v_2 - [\mathbf{u}]_\rho)] = 0.$$

Since $\gamma_t \in \prod(f_t^\epsilon, f_t)$,

$$W := \int_{\mathbb{T}^{2n}\times\mathbb{R}^{2n}} |w_1 - w_2|^2 \delta\mu_t(w_1, w_2) \geqslant W_2^2(f_t^\epsilon, f_t).$$

So, we aim to control W. Splitting it into the potential and kinetic components, we get:

$$W = \int_{\mathbb{T}^{2n}\times\mathbb{R}^{2n}} |v_1 - v_2|^2 \mathrm{d}\gamma_t + \int_{\mathbb{T}^{2n}\times\mathbb{R}^{2n}} |x_1 - x_2|^2 \mathrm{d}\gamma_t := W_v + W_x.$$

We will from here on use the notation $A \lesssim B$ as a shorthand for $A \leqslant C(M,J,T)B$. As in the Cucker-Smale case, we get

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}W_x \lesssim e(f^\epsilon|\mathbf{u}) + W_x \\ W_v \lesssim e(f^\epsilon|\mathbf{u}) + W_x, \end{cases} \tag{61}$$

where the modulated kinetic energy $e(f^\epsilon|\mathbf{u})$ is given in (31). The modulated kinetic energy can also be estimated as in the Cucker-Smale case:

$$\frac{\mathrm{d}}{\mathrm{d}t}e(f^\epsilon|\mathbf{u}) \lesssim e(f^\epsilon|\mathbf{u}) + \frac{\delta}{\epsilon} + (\mathbf{u}^\epsilon - \mathbf{u}, [\mathbf{u}^\epsilon]_{\rho^\epsilon} - \mathbf{u}^\epsilon)_{\kappa_{\rho^\epsilon}} + \int_{\mathbb{T}^n} \rho^\epsilon(\mathbf{u} - \mathbf{u}^\epsilon)\cdot\mathrm{s}([\mathbf{u}]_\rho - \mathbf{u})\,\mathrm{d}x.$$

The alignment terms can be written as

$$\begin{aligned}\delta A &= (\mathbf{u}^\epsilon - \mathbf{u}, [\mathbf{u}^\epsilon]_{\rho^\epsilon} - \mathbf{u}^\epsilon)_{\kappa_{\rho^\epsilon}} + \int_{\mathbb{T}^n} \rho^\epsilon(\mathbf{u} - \mathbf{u}^\epsilon) \cdot \mathrm{s}([\mathbf{u}]_\rho - \mathbf{u}) \,\mathrm{d}x \\ &= \int_{\mathbb{T}^n} \rho^\epsilon(\mathbf{u} - \mathbf{u}^\epsilon) \cdot \left(\mathrm{s}^\epsilon[\mathbf{u}^\epsilon]_{\rho^\epsilon} - \mathrm{s}[\mathbf{u}]_\rho + \mathrm{s}\mathbf{u} - \mathrm{s}^\epsilon \mathbf{u}^\epsilon\right) \mathrm{d}x \\ &= (\mathbf{u}^\epsilon - \mathbf{u}, [\mathbf{u}^\epsilon - \mathbf{u}]_{\rho^\epsilon})_{\kappa_{\rho^\epsilon}} + \int_{\mathbb{T}^n} \rho^\epsilon(\mathbf{u} - \mathbf{u}^\epsilon) \cdot (\mathrm{s}^\epsilon[\mathbf{u}]_{\rho^\epsilon} - \mathrm{s}[\mathbf{u}]_\rho) \,\mathrm{d}x \\ &\quad - \int_{\mathbb{T}^n} \rho^\epsilon|\mathbf{u} - \mathbf{u}^\epsilon|^2 \mathrm{s}^\epsilon \,\mathrm{d}x + \int_{\mathbb{T}^n} \rho^\epsilon(\mathbf{u} - \mathbf{u}^\epsilon) \cdot \mathbf{u}(\mathrm{s}^\epsilon - \mathrm{s}) \,\mathrm{d}x := I + II + III + IV.\end{aligned}$$

By (**u**Reg1) and (sReg1),

$$I \lesssim \int_{\mathbb{T}^n} \rho^\epsilon |\mathbf{u}^\epsilon - \mathbf{u}|^2 \,\mathrm{d}x.$$

From (**u**Reg2') and (sReg2') and the assumed smoothness of s, we get

$$II \lesssim e(f^\epsilon|\mathbf{u}) + \sup_{t\in[0,T]} W_2^2(f^\epsilon, f).$$

The term III is negative so we drop it. Finally, from (sReg2') and the assumed smoothness of $\mathbf{u}$,

$$IV \lesssim e(f^\epsilon|\mathbf{u}) + \sup_{t\in[0,T]} W_2^2(f^\epsilon, f).$$

Using (61), the estimate for the modulated kinetic energy becomes

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} e(f^\epsilon|\mathbf{u}) &\lesssim e(f^\epsilon|\mathbf{u}) + \frac{\delta}{\epsilon} + \sup_{t\in[0,T]} W_2^2(f^\epsilon, f) \\ &\leqslant e(f^\epsilon|\mathbf{u}) + \frac{\delta}{\epsilon} + \sup_{t\in[0,T]} W_x + \sup_{t\in[0,T]} W_v \\ &\lesssim \frac{\delta}{\epsilon} + \sup_{t\in[0,T]} (W_x + e(f^\epsilon|\mathbf{u})).\end{aligned}$$

All together,

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t} W_x \lesssim e(f^\epsilon|\mathbf{u}) + W_x \\ \frac{\mathrm{d}}{\mathrm{d}t} e(f^\epsilon|\mathbf{u}) \lesssim \frac{\delta}{\epsilon} + \sup_{t\in[0,T]}(W_x + e(f^\epsilon|\mathbf{u})). \end{cases}$$

Since the initial value of $e(f^\epsilon|\mathbf{u}) + W_x$ is smaller than ϵ, Gronwall implies

$$e(f^\epsilon|\mathbf{u}) + W_x \lesssim \epsilon + \frac{\delta}{\epsilon},$$

and by (61), $W_v \lesssim \epsilon + \frac{\delta}{\epsilon}$. □

6.2. **Maxwellian Limit.** We consider (13) under a strong Fokker-Planck penalization force (29):

$$\begin{cases} \partial_t f^\epsilon + v \cdot \partial_x f^\epsilon = \nabla_v \cdot (\mathrm{s}^\epsilon (v - [\mathbf{u}^\epsilon]_{\rho^\epsilon}) f^\epsilon) + \frac{1}{\epsilon}\left[\Delta_v f^\epsilon + \nabla_v \cdot (v - \mathbf{u}^\epsilon_\delta f^\epsilon)\right] \\ \partial_t \mathrm{s}^\epsilon + \nabla \cdot (\mathrm{s}^\epsilon [\mathbf{u}^\epsilon]_{\rho^\epsilon}) = 0 \end{cases} \tag{62}$$

To obtain the equation for the macroscopic quantities ρ^ϵ and $\rho^\epsilon \mathbf{u}^\epsilon$, we take the v-moments to get:

$$\begin{cases} \partial_t \rho^\epsilon + \nabla \cdot (\mathbf{u}^\epsilon \rho^\epsilon) = 0 \\ \partial_t \mathrm{s}^\epsilon + \nabla \cdot (\mathrm{s}^\epsilon [\mathbf{u}^\epsilon]_{\rho^\epsilon}) = 0 \\ \partial_t (\rho \mathbf{u}^\epsilon) + \nabla \cdot (\rho \mathbf{u}^\epsilon \otimes \mathbf{u}^\epsilon) + \nabla \rho^\epsilon + \nabla \cdot \mathcal{R}^\epsilon = \rho^\epsilon \mathrm{s}^\epsilon ([\mathbf{u}^\epsilon]_{\rho^\epsilon} - \mathbf{u}^\epsilon) + \frac{1}{\epsilon} \rho^\epsilon (\mathbf{u}^\epsilon_\delta - \mathbf{u}^\epsilon), \end{cases} \tag{63}$$

where $\mathbf{u}^\epsilon_\delta$ is given in (30) and

$$\mathcal{R}^\epsilon(t,x) = \int_{\mathbb{T}^n} ((v - \mathbf{u}^\epsilon) \otimes (v - \mathbf{u}^\epsilon) - \mathbb{I}) f^\epsilon \, \mathrm{d}v.$$

To justify the Maxwellian limit, we will assume the following continuity conditions on solutions to (63) and on the averaging $[\cdot]_\rho$:

$$\mathrm{s}^\epsilon \text{ is bounded uniformly in } \epsilon : \sup_{t \in [0,T]} \|\mathrm{s}^\epsilon(t,\cdot)\|_\infty \leqslant C \tag{R1}$$

$$\|\mathrm{s}^\epsilon [\mathbf{u}^\epsilon]_{\rho^\epsilon} - \mathrm{s}[\mathbf{u}]_\rho\|_{L^2} \lesssim \sup_{t \in [0,T]} (W_1(\rho^\epsilon, \rho) + W_1(\mathbf{u}^\epsilon \rho^\epsilon, \mathbf{u}\rho))(t,\cdot) \tag{R2}$$

$$\|\mathrm{s}^\epsilon - \mathrm{s}\|_{L^2} \lesssim \sup_{t \in [0,T]} (W_1(\rho^\epsilon, \rho) + W_1(\mathbf{u}^\epsilon \rho^\epsilon, \mathbf{u}\rho))(t,\cdot) \quad \text{provided } \mathrm{s}^\epsilon_0 = \mathrm{s}_0 \tag{R3}$$

$$[\cdot]_\rho : L^2(\rho) \mapsto L^2(\rho) \text{ is bounded uniformly over } \rho \in \mathcal{P}_M(\Omega) \text{ with equal mass} \tag{R4}$$

Note that (R4) clearly holds with our usual assumption on the velocity averaging and kernel. These conditions hold for the w-model, which we will show at the end of this section.

Define the Maxwellians associated to solutions $(\rho, \mathrm{s}, \mathbf{u})$ to (58) and solutions $(\rho^\epsilon, \mathrm{s}^\epsilon, \mathbf{u}^\epsilon)$ to (63) by

$$\mu(t,x) = \frac{\rho(t,x)}{(2\pi)^{n/2}} e^{\frac{|v - \mathbf{u}(t,x)|^2}{2}}, \quad \mu^\epsilon(t,x) = \frac{\rho^\epsilon(t,x)}{(2\pi)^{n/2}} e^{\frac{|v - \mathbf{u}^\epsilon(t,x)|^2}{2}}.$$

and the relative entropy by

$$\begin{aligned} \mathcal{H}(f^\epsilon|\mu) &= \mathcal{H}(f^\epsilon|\mu^\epsilon) + \mathcal{H}(\mu^\epsilon|\mu) \\ \mathcal{H}(\mu^\epsilon|\mu) &= \frac{1}{2} \int_{\mathbb{T}^n} \rho^\epsilon |\mathbf{u}^\epsilon - \mathbf{u}|^2 \, \mathrm{d}x + \int_{\mathbb{T}^n} \rho^\epsilon \log(\rho^\epsilon/\rho) \, \mathrm{d}x. \end{aligned}$$

Due to the Csiszár-Kullback inequality, $\mathcal{H}(f^\epsilon|\mu) \to 0$ implies

$$\begin{aligned} \rho^\epsilon &\to \rho \\ \mathbf{u}^\epsilon \rho^\epsilon &\to \mathbf{u}\rho \end{aligned}$$

in $L^1(\mathbb{T}^n)$. In particular, $\mathcal{H}(f^\epsilon|\mu) \to 0$ implies that

$$\begin{aligned} W_1(\rho^\epsilon, \rho) &\to 0, \\ W_1(\mathbf{u}^\epsilon \rho^\epsilon, \mathbf{u}\rho) &\to 0. \end{aligned}$$

Proof. (Proof of Theorem 3.8) When an identity or inequality does not depend on the regularity of the strength and can be found in the proof for the Cucker-Smale case, we refer the reader to Section 3.2.2. As in the Cucker-Smale case, we break the relative entropy into the kinetic and macroscopic parts, $\mathcal{H}(f^\epsilon|\mu) = \mathcal{H}_\epsilon + \mathcal{G}_\epsilon$ with $\mathcal{H}_\epsilon$, $\mathcal{G}_\epsilon$ given (37) and (38). By (41),

$$\begin{aligned}\frac{d}{dt}\mathcal{H}_\epsilon =& -\frac{1}{\epsilon}\int_{\mathbb{T}^n\times\mathbb{R}^n}\Big[\frac{|\nabla_v f^\epsilon|^2}{f^\epsilon} + 2\nabla_v f^\epsilon\cdot(v-\mathbf{u}^\epsilon_\delta + |v-\mathbf{u}^\epsilon_\delta|^2 f^\epsilon)\Big]\\ &-\frac{1}{\epsilon}[(\mathbf{u}^\epsilon_\delta,\mathbf{u}^\epsilon)_{\rho^\epsilon} - (\mathbf{u}^\epsilon_\delta,\mathbf{u}^\epsilon_\delta)_{\rho^\epsilon}]\\ &-\int_{\mathbb{T}^n\times\mathbb{R}^n} \mathrm{s}^\epsilon[\nabla_v f^\epsilon\cdot(v-[\mathbf{u}^\epsilon]_{\rho^\epsilon}) + v\cdot(v-[\mathbf{u}^\epsilon]_{\rho^\epsilon})f^\epsilon]\,\mathrm{d}v\,\mathrm{d}x\\ &:= A_1 + A_2 + A_3.\end{aligned}$$

A_1 is strictly negative so it can be dropped; according to [16] (due to ball-positivity of the overmollified M_ϕ model), A_2 is also strictly negative. And

$$A_3 = n\int_{\mathbb{T}^n\times\mathbb{R}^n} \mathrm{s}^\epsilon f^\epsilon\,\mathrm{d}v\,\mathrm{d}x - \int_{\mathbb{T}^n\times\mathbb{R}^n}\mathrm{s}^\epsilon|v|^2 f^\epsilon\,\mathrm{d}v\,\mathrm{d}x + (\mathbf{u}^\epsilon, [\mathbf{u}^\epsilon]_{\rho^\epsilon})_{\kappa^\epsilon}.$$

Due to boundedness of s^ϵ (R1),

$$A_3 \leqslant c_1 + c_1\|\mathbf{u}^\epsilon\|^2_{L^2(\rho^\epsilon)} + (\mathbf{u}^\epsilon, [\mathbf{u}^\epsilon]_{\rho^\epsilon})_{\kappa^\epsilon}.$$

Recalling that $\mathcal{E}_\epsilon = \int |v|^2 f^\epsilon(x,v)\,\mathrm{d}v\,\mathrm{d}x = \|\mathbf{u}^\epsilon\|^2_{L^2(\rho^\epsilon)} \leqslant \mathcal{H}_\epsilon$ and using $L^2(\rho)$ boundedness of $[\cdot]_\rho$, we have

$$A_3 \leqslant c_1 + c_2\mathcal{E}_\epsilon \leqslant c_3 + c_4\mathcal{H}_\epsilon,$$

where the last inequality can be found in [16, Section 7.3]. Thus,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}_\epsilon \leqslant c_3 + c_4\mathcal{H}_\epsilon,$$

and by Gronwall $\mathcal{H}_\epsilon$ is bounded uniformly in ϵ. Uniform boundedness of $\mathcal{H}_\epsilon$ in ϵ will be used to control the Reynolds stress term. By (43) and boundedness of s^ϵ (R1),

$$\frac{d}{dt}\mathcal{H}_\epsilon \leqslant -\frac{1}{\epsilon}\mathcal{I}_\epsilon + c\epsilon e(f^\epsilon|\mathbf{u}^\epsilon) - \|\mathbf{u}^\epsilon\|_{L^2(\kappa_\epsilon)^2} + (\mathbf{u}^\epsilon, [\mathbf{u}^\epsilon]_{\rho^\epsilon})_{\kappa^\epsilon}.$$

Recall that $\mathcal{I}_\epsilon$ is the Fisher information and is given by (39). Turning to $\mathcal{G}_\epsilon$, the evolution equation is given by (42). The term δA is the same alignment term used in the proof of the monokinetic case. As in the monokinetic case, the inherited regularity is precisely what is needed to control this term. The difference here from the monokinetic case is that we use (R2) and (R3) instead of (**u**Lip2'), (sLip2'); and we retain the term $\int_{\mathbb{T}^n}\rho^\epsilon|\mathbf{u}^\epsilon - \mathbf{u}|^2\,\mathrm{d}x$ and bound it by $\mathcal{H}_\epsilon$ instead of the modulated kinetic energy. We get

$$\delta A \lesssim \int_{\mathbb{T}^n}\rho^\epsilon|\mathbf{u}^\epsilon-\mathbf{u}|^2\,\mathrm{d}x + \sup_{t\in[0,T]}(W_1^2(\rho^\epsilon,\rho) + W_1^2(\mathbf{u}^\epsilon\rho^\epsilon,\mathbf{u}\rho))(t,\cdot) \tag{64}$$

and

$$\begin{aligned}W_1^2(\mathbf{u}^\epsilon\rho^\epsilon,\mathbf{u}\rho) &\leqslant \|\mathbf{u}^\epsilon\rho^\epsilon - \mathbf{u}\rho\|^2_{L^1}\\ &\leqslant \|\rho^\epsilon(\mathbf{u}^\epsilon-\mathbf{u})\|^2_{L^1} + \|\mathbf{u}\|^2_\infty\|\rho^\epsilon-\rho\|^2_{L^1}.\end{aligned}$$

By Hölder inequality,

$$\|\rho^\epsilon(\mathbf{u}^\epsilon-\mathbf{u})\|^2_{L^1} \lesssim \|\mathbf{u}^\epsilon-\mathbf{u}\|^2_{L^2(\rho_\epsilon)} \leqslant \mathcal{H}(\mu^\epsilon|\mu) \leqslant \mathcal{H}(f^\epsilon|\mu).$$

By Csiszár-Kullback inequality

$$\|\rho^\epsilon - \rho\|_{L^1}^2 \lesssim \mathcal{H}(\rho^\epsilon|\rho) \leqslant \mathcal{H}(\mu^\epsilon|\mu) \leqslant \mathcal{H}(f^\epsilon|\mu).$$

So $W_1^2(\mathbf{u}^\epsilon\rho^\epsilon, \mathbf{u}\rho) \lesssim \mathcal{H}(f^\epsilon|\mu)$ and therefore by (64)

$$\delta A \lesssim \mathcal{H}(f^\epsilon|\mu).$$

The Reynolds stress term can be written

$$\mathcal{R}^\epsilon = \int_{\mathbb{R}^n} [2\nabla_v\sqrt{f^\epsilon} + (1+\epsilon \mathrm{s}^\epsilon/2)(v-\mathbf{u}^\epsilon)\sqrt{f^\epsilon}] \otimes [(v-\mathbf{u}^\epsilon)\sqrt{f^\epsilon}]\,\mathrm{d}v - \epsilon \mathrm{s}^\epsilon/2 \int_{\mathbb{R}^n} (v-\mathbf{u}^\epsilon)\otimes(v-\mathbf{u}^\epsilon)\,\mathrm{d}v.$$

Note that

$$e(f^\epsilon|\mathbf{u}^\epsilon) = \mathcal{E}_\epsilon - \frac{1}{2}\int_{\mathbb{T}^n} \rho^\epsilon|\mathbf{u}^\epsilon|^2\,\mathrm{d}x \leqslant \mathcal{E}_\epsilon \leqslant \mathcal{H}_\epsilon.$$

Due to boundednes of s^ϵ (R1) and the boundedness of $\mathcal{H}_\epsilon$, we get

$$\mathcal{R}^\epsilon \lesssim \sqrt{e(f^\epsilon|\mathbf{u}^\epsilon)\mathcal{I}_\epsilon} + \epsilon e(f^\epsilon|\mathbf{u}^\epsilon) \lesssim \sqrt{\mathcal{I}_\epsilon} + \epsilon.$$

The remaining terms do not depend on s^ϵ are estimated in Section 3.2.1. In total, we get

$$\frac{d}{dt}\mathcal{G}_\epsilon \lesssim \mathcal{H}(f^\epsilon|\mu) + \sqrt{\mathcal{I}_\epsilon} + \epsilon + \frac{\delta}{\epsilon} + \|\mathbf{u}^\epsilon\|_{L^2(\kappa^\epsilon)} - (\mathbf{u}^\epsilon, [\mathbf{u}^\epsilon]_{\rho^\epsilon})_{\kappa^\epsilon}.$$

Combining the estimates on $\mathcal{H}_\epsilon$ and $\mathcal{G}_\epsilon$ and using $\sqrt{\mathcal{I}_\epsilon} \leqslant \frac{1}{2\epsilon}\mathcal{I}_\epsilon + 2\epsilon$ we arrive at

$$\frac{d}{dt}\mathcal{H}(f^\epsilon|\mu) \lesssim \mathcal{H}(f^\epsilon|\mu) - \frac{1}{\epsilon}\mathcal{I}_\epsilon + \epsilon + \frac{\delta}{\epsilon} + \sqrt{\mathcal{I}_\epsilon} \lesssim \mathcal{H}(f^\epsilon|\mu) + \epsilon + \frac{\delta}{\epsilon}.$$

Gronwall concludes the proof. □

We conclude this section by verifying (R1)-(R4) for the w-model for $\phi \geqslant c > 0$. Recall that $[\mathbf{u}]_\rho = \mathbf{u}_F = (\mathbf{u}\rho)_\phi/\rho_\phi$ and $\mathrm{s} = \mathrm{w}\rho_\phi$ where w solves the pure transport along $\mathbf{u}_F$ given in (WM). Let $(\rho, \mathrm{s}, \mathbf{u})$ be a smooth solution to (58) and $(\rho^\epsilon, \mathrm{s}^\epsilon, \mathbf{u}^\epsilon)$ be a solution to (63); and let $(\rho, \mathrm{w}, \mathbf{u})$ and $(\rho^\epsilon, \mathrm{w}^\epsilon, \mathbf{u}^\epsilon)$ be the corresponding change of variables. For (R1), w^ϵ is transported so s^ϵ is clearly bounded. For (R2), we have

$$\begin{aligned}\|\mathrm{s}^\epsilon[\mathbf{u}^\epsilon]_{\rho^\epsilon} - \mathrm{s}[\mathbf{u}]_\rho\|_{L^2} &= \|\mathrm{w}^\epsilon(\mathbf{u}^\epsilon\rho^\epsilon)_\phi - \mathrm{w}(\mathbf{u}\rho)_\phi\|_{L^2}\\ &\leqslant \|\mathrm{w}^\epsilon - \mathrm{w}\|_{L^2}\|(\mathbf{u}\rho)_\phi\|_\infty + \|(\mathbf{u}^\epsilon\rho^\epsilon)_\phi - (\mathbf{u}\rho)_\phi\|_{L^2}\|\mathrm{w}^\epsilon\|_{L^\infty}.\end{aligned}$$

To control $\|\mathrm{w}^\epsilon - \mathrm{w}\|_{L^2}$, we turn to control the velocities $\|\mathbf{u}_F^\epsilon - \mathbf{u}_F\|_\infty$. Since $\phi \geqslant c > 0$,

$$\|\mathbf{u}_F^\epsilon - \mathbf{u}_F\|_\infty = \left\|\frac{\rho_\phi(\mathbf{u}^\epsilon\rho^\epsilon)_\phi - \rho_\phi^\epsilon(\mathbf{u}\rho_\phi)}{\rho_\phi\rho_\phi^\epsilon}\right\|_\infty \lesssim \Big(W_1(\rho^\epsilon,\rho) + W_1(\mathbf{u}^\epsilon\rho^\epsilon, \mathbf{u}\rho)\Big).$$

Integrating $\partial_t(\mathrm{w}^\epsilon - \mathrm{w})$ along characteristics, we get:

$$\|\mathrm{w}^\epsilon - \mathrm{w}\|_{L^2} \lesssim \sup_{t\in[0,T]} (W_1(\rho^\epsilon,\rho) + W_1(\mathbf{u}^\epsilon\rho^\epsilon,\mathbf{u}\rho))(t,\cdot).$$

Moreover,

$$\|\mathrm{w}^\epsilon(\mathbf{u}^\epsilon\rho^\epsilon)_\phi - \mathrm{w}(\mathbf{u}\rho)_\phi\|_{L^2} \lesssim \sup_{t\in[0,T]} (W_1(\rho^\epsilon,\rho) + W_1(\mathbf{u}^\epsilon\rho^\epsilon,\mathbf{u}\rho))(t,\cdot).$$

(R3) follows similarly. For (R4), the Favre averaging is clearly a bounded operator on L^2 when the ϕ is bounded from below away from zero.

7. Relaxation to Global Maxwellian in 1D

Adding noise with the strength coefficient to the velocity equation leads to the Fokker-Planck-Alignment system given by

$$\begin{cases} \partial_t f + v \cdot \nabla_x f + \nabla_v \cdot (\mathrm{s}(v - [\mathbf{u}]_\rho) f) = \sigma \mathrm{s} \Delta_v f \\ \partial_t \mathrm{s} + \nabla_x \cdot (\mathrm{s}[\mathbf{u}]_\rho) = 0. \end{cases} \tag{65}$$

We study the case where the velocity averaging is given by the Favre averaging: $\mathbf{u}_F := (\mathbf{u}\rho)_\phi / \rho_\phi$ (i.e. the w-model). Recall that the weight w is defined by $\mathrm{s} = \mathrm{w}\rho_\phi$ and the system becomes

$$\begin{cases} \partial_t f + v \cdot \nabla_x f = \mathrm{w}\rho_\phi \nabla_v \cdot (\sigma \nabla_v f - v f) + \mathrm{w}(\mathbf{u}\rho)_\phi \cdot \nabla_v f \\ \partial_t \mathrm{w} + \mathbf{u}_F \cdot \nabla_x \mathrm{w} = 0. \end{cases} \tag{66}$$

Let us discuss well-posedness of (66). General well-posedness theory for kinetic alignment equations based on a predefined strength s_ρ has been developed in [16]. The new system, under the uniform regularity assumptions (Reg1) and (Reg2), which for the Favre filtration (with $\Phi_\rho(x, y) = \phi(x - y)/\rho_\phi(x)$) means all-to-all interactions

$$\inf_\Omega \phi = c_0 > 0, \tag{67}$$

falls directly into the same framework of [16]. In fact, in this case every flock is automatically "thick", meaning that

$$\inf_\Omega \rho_\phi = c_0 M > 0.$$

We also have the a priori bounded energy

$$\dot{\mathcal{E}} \leqslant c_1 \mathcal{E} + c_2, \quad \mathcal{E}(t) = \frac{1}{2} \int_{\Omega \times \mathbb{R}^n} |v|^2 f_t(x, v) \, \mathrm{d}x \, \mathrm{d}v$$

on any finite time interval. Indeed, multiplying (66) by $\frac{1}{2}|v|^2$ and integrating by parts, we have:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \frac{1}{2} \int_{\Omega \times \mathbb{R}^n} |v|^2 f \, \mathrm{d}x \, \mathrm{d}v &= \int_{\Omega \times \mathbb{R}^n} \mathrm{w}\rho_\phi v \cdot (v f - \sigma \nabla_v f) \, \mathrm{d}v \, \mathrm{d}x - \int_{\Omega \times \mathbb{R}^n} \mathrm{w}(\mathbf{u}\rho)_\phi \cdot v f \, \mathrm{d}v \, \mathrm{d}x \\ &= \int_{\Omega \times \mathbb{R}^n} \mathrm{w}\rho_\phi \cdot |v|^2 f \, \mathrm{d}v \, \mathrm{d}x + n\sigma \int_{\Omega \times \mathbb{R}^n} \mathrm{w}\rho_\phi f \, \mathrm{d}v \, \mathrm{d}x - \int_{\Omega \times \mathbb{R}^n} \mathrm{w}(\mathbf{u}\rho)_\phi \cdot v f \, \mathrm{d}v \, \mathrm{d}x \\ &= \int_\Omega \mathrm{w}\rho_\phi |\mathbf{u}|^2 \rho \, \mathrm{d}x + n\sigma \int_\Omega \mathrm{w}\rho_\phi \rho \, \mathrm{d}x - \int_\Omega \mathrm{w}(\mathbf{u}\rho)_\phi \cdot (\mathbf{u}\rho) \, \mathrm{d}v \, \mathrm{d}x. \end{aligned}$$

Since w remains bounded, the first term bounds the energy and the second term is bounded. For the last term, we estimate,

$$\int_\Omega \mathrm{w}(\mathbf{u}\rho)_\phi \cdot (\mathbf{u}\rho) \, \mathrm{d}v \, \mathrm{d}x \leqslant \mathrm{w}_+ \phi_+ \|\mathbf{u}\|^2_{L^1(\rho)} \lesssim \mathrm{w}_+ \phi_+ \|\mathbf{u}\|^2_{L^2(\rho)} = \mathrm{w}_+ \phi_+ \mathcal{E},$$

which yields the desired energy inequality. Since $\|\mathbf{u}\|_{L^1(\rho)} \lesssim \|\mathbf{u}\|_{L^2(\rho)} = \sqrt{\mathcal{E}}$, we have the inheritance of regularity:

$$\|\partial_x^k \mathbf{u}_F\|_\infty < C(k, M, J, T),$$

for any $k \in \mathbb{N}$. Consequently, according to (a trivial adaptation of) [16, Theorem 7.1] for any data in the classical weighted Sobolev spaces $f_0 \in H^m_q(\Omega \times \mathbb{R}^n)$ and $\mathrm{w}_0 \in C^m$, $\inf_\Omega \mathrm{w}_0 > 0$, where

$$H^m_q(\Omega \times \mathbb{R}^n) = \left\{ f : \sum_{k+l \leqslant m} \sum_{|\mathbf{k}|=k, |\mathbf{l}|=l} \int_{\Omega \times \mathbb{R}^n} |\langle v \rangle^{q-k-l} \partial_x^{\mathbf{k}} \partial_v^{\mathbf{l}} f|^2 \, \mathrm{d}v \, \mathrm{d}x < \infty \right\}, \quad q > m, \tag{68}$$

there exists a unique global in time classical solution in the same data space. Let us note that the transport of w preserves the lower bound on w uniformly in time, and by the automatic thickness condition we have the ellipticity coefficient $\mathrm{w}\rho_\phi$ uniformly bounded away from zero as well.

Let us now recall a set of conditions on a given global solution f that are sufficient to imply global relaxation of f to the Maxwellian

$$\mu_{\sigma, \bar{\mathbf{u}}} = \frac{1}{|\mathbb{T}^n|(2\pi\sigma)^{n/2}} e^{\frac{|v-\bar{\mathbf{u}}|^2}{2\sigma}},$$

where $\bar{\mathbf{u}} = \frac{1}{M} \int_\Omega \mathbf{u}\rho \, \mathrm{d}x$. We note that this total momentum $\bar{\mathbf{u}}$ is not preserved in time generally, but rather satisfies the equation

$$\partial_t \bar{\mathbf{u}} = \int_\Omega (\mathbf{u}_F - \mathbf{u}) \, \mathrm{d}\kappa_\rho, \qquad \mathrm{d}\kappa_\rho = \mathrm{s}\rho \, \mathrm{d}x. \tag{69}$$

We can't determine whether $\bar{\mathbf{u}}$ eventually settles to a fixed vector, as it does for all the classical alignment models including those that do not preserve the momentum.

So, according to [16, Proposition 8.1] a given solution f converges to the Maxwellian in the sense defined in (72) provided there exists a set of fixed constants $c_0, \ldots > 0$ such that

(i) $c_0 \leqslant \mathrm{s} \leqslant c_1$ and $\|\nabla \mathrm{s}\|_\infty \leqslant c_2$ for all ρ.

(ii)

$$\sup \left\{ (\mathbf{u}, [\mathbf{u}]_\rho)_{\kappa_\rho} : \mathbf{u} \in L^2(\kappa_\rho), \bar{\mathbf{u}} = 0, \|\mathbf{u}\|_{L^2(\kappa_\rho)} = 1 \right\} \leqslant 1 - \epsilon_0 \tag{70}$$

(iii) $\|\mathrm{s}[\cdot]_\rho\|_{L^2(\rho) \to L^2(\rho)} + \|\nabla_x(\mathrm{s}[\cdot]_\rho)\|_{L^2(\rho) \to L^2(\rho)} \leqslant c_3$.

We will be able to show (i)-(iii) in one dimensional case. Let us discuss these conditions starting from (i). The key condition here is $\|\partial_x \mathrm{w}\|_\infty \leqslant c_2$. We can establish such uniform control in 1D only. Indeed, since

$$\partial_t \partial_x \mathrm{w} + \partial_x(u_F \partial_x \mathrm{w}) = 0,$$

we can see that $\partial_x \mathrm{w}$ satisfies the same continuity equation as ρ_ϕ (this only holds in 1D). Thus,

$$\partial_t \frac{\partial_x \mathrm{w}}{\rho_\phi} + u_F \partial_x \frac{\partial_x \mathrm{w}}{\rho_\phi} = 0,$$

Since initially $\frac{|\partial_x \mathrm{w}|}{\rho_\phi} \leqslant C$ for some large $C > 0$ due to the all-to-all interaction assumption, this bound will persist in time. Hence, $\|\nabla \mathrm{w}\|_\infty \leqslant C\|\rho_\phi\|_\infty \leqslant c_2$.

Condition (iii) follows from (i). Indeed, since w, ∇w remain uniformly bounded, it reduces to

$$\int_\Omega |(u\rho)_\phi|^2 \rho \, \mathrm{d}x + \int_\Omega |(u\rho)_{\nabla\phi}|^2 \rho \, \mathrm{d}x \leqslant c_3 \|u\|^2_{L^2(\rho)}.$$

This follows by the Hölder inequality.

Finally, let us get to (ii). We reduce the computation of the spectral gap to the low-energy method and the classical Cucker-Smale model. To this end, we assume that $\phi = \psi * \psi$ where $\psi > 0$ is another positive kernel. In other words, ϕ is Bochner-positive. In order to establish (70) it suffices to show that

$$(u,u)_{\kappa_\rho} - (u,[u]_\rho)_{\kappa_\rho} \geqslant \epsilon_0 (u,u)_{\kappa_\rho}.$$

Let us start by symmetrizing and using cancellation in the second obtained integral:

$$\begin{aligned}(u,u)_{\kappa_\rho} - (u,[u]_\rho)_{\kappa_\rho} &= \int_{\Omega\times\Omega} u(x)\cdot(u(x)-u(y))\rho(x)\rho(y)\phi(x-y)\mathrm{w}(x)\,\mathrm{d}y\,\mathrm{d}x \\ &= \frac{1}{2}\int_{\Omega\times\Omega} |u(x)-u(y)|^2\rho(x)\rho(y)\phi(x-y)\mathrm{w}(x)\,\mathrm{d}y\,\mathrm{d}x \\ &- \frac{1}{2}\int_{\Omega\times\Omega} |u(y)|^2\rho(x)\rho(y)(\mathrm{w}(x)-\mathrm{w}(y))\phi(x-y)\,\mathrm{d}y\,\mathrm{d}x = I + II.\end{aligned}$$

Notice that

$$I \geqslant \mathrm{w}_- \int_{\Omega\times\Omega} u(x)\cdot(u(x)-u(y))\rho(x)\rho(y)\phi(x-y)\,\mathrm{d}y\,\mathrm{d}x = \mathrm{w}_-[(u,u)_{\rho\phi\rho} - (u,[u]_\rho)_{\rho\phi\rho}].$$

The difference of the inner productions inside the bracket represents exactly the spectral gap of the Cucker-Smale model computed in [16, Proposition 4.16]. From that computation it follows that

$$(u,u)_{\rho\phi\rho} - (u,[u]_\rho)_{\rho\phi\rho} \geqslant cM^3 (u,u)_{\rho\phi\rho} \geqslant c\mathrm{w}_- M^3 (u,u)_{\kappa_\rho},$$

where c depends only on the kernel ψ.

Let us now estimate II:

$$II \leqslant (\mathrm{w}_+ - \mathrm{w}_-)(u,u)_{\rho\phi\rho} \leqslant \frac{\mathrm{w}_+ - \mathrm{w}_-}{\mathrm{w}_-}(u,u)_{\kappa_\rho}.$$

We can see that provided

$$\frac{\mathrm{w}_+ - \mathrm{w}_-}{\mathrm{w}_-} \leqslant \frac{1}{2} c\mathrm{w}_- M^3,$$

we obtain

$$(u,u)_{\kappa_\rho} - (u,[u]_\rho)_{\kappa_\rho} \geqslant \epsilon_0 (u,u)_{\kappa_\rho}.$$

which is the needed result.

Let us collect now all the assumptions we have made and state the main result.

Theorem 7.1. *Suppose $n = 1$ and the kernel is Bochner-positive, $\phi = \psi * \psi$, with $\inf \psi > 0$. Then any initial distribution $f_0 \in H^m_q(\Omega \times \mathbb{R}^n)$ and strength $\mathrm{w}_0 \in C^m$ satisfying the following small variation assumption*

$$\frac{\sup(\mathrm{w}_0) - \inf(\mathrm{w}_0)}{\inf(\mathrm{w}_0)^2} \leqslant cM^3, \tag{71}$$

for some absolute $c > 0$, gives rise to a global classical solution f, w which relaxes to the Maxwellian exponentially fast

$$\|f(t) - \mu_{\sigma,\bar{u}}\|_{L^1(\mathbb{T}^n\times\mathbb{R}^n)} \leqslant c_1 \sigma^{-1/2} e^{-c_2\sigma t}. \tag{72}$$

We note once again that the solution relaxes to a moving Maxwellian centered around a time-dependent momentum $\bar{u}$. There two conceivable mechanisms of stabilizing $\bar{u}$. One is sufficiently fast alignment:

$$\int_0^\infty \sup_{x,y} |u(x,t) - u(y,t)| \,\mathrm{d}t < \infty, \tag{73}$$

which our relaxation seems to be not strong enough to imply. And another is stabilization of the density to a uniform distribution $\rho \to \frac{1}{|\Omega|}$, which we do have from (72) and in the case when $\mathrm{w} \equiv \mathrm{const}$ it does imply exponential stabilization of the momentum for Favre-based models, see [16]. If w varies, even if $\rho = \frac{1}{|\Omega|}$, from (69) we have,

$$\partial_t \bar{u} = \int_\Omega (u_F - u) \,\mathrm{d}\kappa_\rho = \frac{1}{|\Omega|} \int_\Omega (u_\phi - u\|\phi\|_1) \mathrm{w} \,\mathrm{d}x.$$

We can see that unless (73) holds, the persistent variations of w may keep this term large leaving a possibility forever moving momentum $\bar{u}$.

References

[1] G. Albi, N. Bellomo, L. Fermo, S.-Y. Ha, J. Kim, L. Pareschi, D. Poyato, and J. Soler. Vehicular traffic, crowds, and swarms: From kinetic theory and multiscale methods to applications and research perspectives. *Math. Models Methods Appl. Sci.*, 29(10):1901–2005, 2019.

[2] J. A. Carrillo, Y.-P. Choi, E. Tadmor, and C. Tan. Critical thresholds in 1D Euler equations with non-local forces. *Math. Models Methods Appl. Sci.*, 26(1):185–206, 2016.

[3] J. A. Carrillo, M. Fornasier, J. Rosado, and G. Toscani. Asymptotic flocking dynamics for the kinetic Cucker-Smale model. *SIAM J. Math. Anal.*, 42(1):218–236, 2010.

[4] F. Cucker and S. Smale. Emergent behavior in flocks. *IEEE Trans. Automat. Control*, 52(5):852–862, 2007.

[5] F. Cucker and S. Smale. On the mathematics of emergence. *Jpn. J. Math.*, 2(1):197–227, 2007.

[6] A. Favre. Turbulence: Space-time statistical properties and behavior in supersonic flows. *The Physics of Fluids*, 26(10):2851–2863, 1983.

[7] A. Figalli and M.-J. Kang. A rigorous derivation from the kinetic Cucker-Smale model to the pressureless Euler system with nonlocal alignment. *Anal. PDE*, 12(3):843–866, 2019.

[8] S.-Y. Ha and J.-G. Liu. A simple proof of the Cucker-Smale flocking dynamics and mean-field limit. *Commun. Math. Sci.*, 7(2):297–325, 2009.

[9] S.-Y. Ha and E. Tadmor. From particle to kinetic and hydrodynamic descriptions of flocking. *Kinet. Relat. Models*, 1(3):415–435, 2008.

[10] T. M. Leslie and R. Shvydkoy. On the structure of limiting flocks in hydrodynamic Euler Alignment models. *Math. Models Methods Appl. Sci.*, 29(13):2419–2431, 2019.

[11] P. Minakowski, P. B. Mucha, J. Peszek, and E. Zatorska. Singular Cucker-Smale dynamics. In *Active particles, Vol. 2*, Model. Simul. Sci. Eng. Technol., pages 201–243. Birkhäuser/Springer, Cham, 2019.

[12] S. Motsch and E. Tadmor. A new model for self-organized dynamics and its flocking behavior. *J. Stat. Phys.*, 144(5):923–947, 2011.

[13] S. Motsch and E. Tadmor. Heterophilious dynamics enhances consensus. *SIAM Rev.*, 56(4):577–621, 2014.

[14] C. W. Reynolds. Flocks, herds and schools: A distributed behavioral model. *ACM SIGGRAPH Computer Graphics*, 21:25–34, 1987.

[15] Roman Shvydkoy. *Dynamics and analysis of alignment models of collective behavior*. Nečas Center Series. Birkhäuser/Springer, Cham, [2021] ©2021.

[16] Roman Shvydkoy. Environmental averaging. *arXiv preprint arXiv:2211.00117*, 2022.

[17] Roman Shvydkoy. Global hypocoercivity of kinetic Fokker-Planck-Alignment equations. *Kinet. Relat. Models*, 15(2):213–237, 2022.

[18] R. Shvydkoy and E. Tadmor. Eulerian dynamics with a commutator forcing II: Flocking. *Discrete Contin. Dyn. Syst.*, 37(11):5503–5520, 2017.

[19] Roman Shvydkoy and Trevor Teolis. Well-posedness and long time behavior of the Euler Alignment System with adaptive communication strength. *arXiv preprint arXiv:2310.00269*, 2023

[20] E. Tadmor and C. Tan. Critical thresholds in flocking hydrodynamics with non-local alignment. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 372(2028):20130401, 22, 2014.

[21] Eitan Tadmor. On the mathematics of swarming: emergent behavior in alignment dynamics. *Notices Amer. Math. Soc.*, 68(4):493–503, 2021.

[22] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet. Novel type of phase transition in a system of self-driven particles. *Physical Review Letters*, 75(6):1226–1229, 1995.

[23] T. Vicsek and A. Zefeiris. Collective motion. *Physics Reprints*, 517:71–140, 2012.