SIAM J. MATH. ANAL. Vol. 54, No. 1, pp. 1131--1168

CAUCHY THEORY FOR GENERAL KINETIC VICSEK MODELS IN COLLECTIVE DYNAMICS AND MEAN-FIELD LIMIT APPROXIMATIONS[1]*

MARC BRIANT[2]†, ANTOINE DIEZ[3]‡, AND SARA MERINO-ACEITUNO[4]§

Abstract. In this paper we provide a local Cauchy theory both on the torus and in the whole space for general Vicsek dynamics at the kinetic level. We consider rather general interaction kernels, nonlinear viscosity, and nonlinear friction. Particularly, we include normalized kernels which display a singularity when the flux of particles vanishes. Thus, in terms of the Cauchy theory for the kinetic equation, we extend to more general interactions and complete the program initiated in [I. M. Gamba and M.-J. Kang, *Arch. Ration. Mech. Anal.*, 222 (2016), pp. 317--342] (where the authors assume that the singularity does not take place) and in [A. Figalli, M.-J. Kang, and J. Morales, *Arch. Ration. Mech. Anal.*, 227 (2018), pp. 869--896] (where the authors prove that the singularity does not happen in the spatially homogeneous case). Moreover, we derive an

[1] Received by the editors March 17, 2021; accepted for publication (in revised form) October 19, 2021; published electronically February 17, 2022.

https://doi.org/10.1137/21M1405885

Funding: This research was partially supported by the Austrian Science Fund (FWF) through the project F65. The work of the second author was supported by the EPSRC-Roth scholarship cofounded by the Engineering and Physical Sciences Research Council and the Department of Mathematics at Imperial College London. The work of the third author was supported by the Vienna Science and Technology Fund (WWTF) with a Vienna Research Groups for Young Investigators, grant VRG17-014.

[2] MAP5, Université de Paris, CNRS, F-75006 Paris, France (briant.maths@gmail.com).

[3] Department of Mathematics, Imperial College London, London, SW7 2AZ, UK (antoine.diez18@imperial.ac.uk).

[4] Faculty of Mathematics, University of Vienna, 1020 Vienna, Austria, and Department of Mathematics, University of Sussex, Brighton, BN1 9RH, UK (sara.merino@univie.ac.at).

explicit lower time of existence as well as a global existence criterion that is applicable, among other cases, to obtain a long time theory for nonrenormalized kernels and for the original Vicsek problem without any a priori assumptions. On the second part of the paper, we also establish the mean-field limit in the large particle limit for an approximated (regularized) system that coincides with the original one whenever the flux does not vanish. Based on the results proved for the limit kinetic equation, we prove that for short times, the probability that the dynamics of this approximated particle system coincides with the original singular dynamics tends to one in the many particle limit.

Key words. Vicsek model, Vicsek--Kolmogorov equation, collective dynamics, nonlinear Fokker--Planck equation on the sphere, normalized interaction kernels, mean-field limit, well-posedness

AMS subject classifications. 35A01, 92B05

DOI. 10.1137/21M1405885

1. Introduction.

1.1. Motivation. The emergence of collective motions among a group of selfpropelled agents is receiving a great deal of attention: flocks of birds [35], schools of fish [34], pedestrian dynamics [41], and microswimmers [10]. Many models have been proposed to explain these collective behaviors. Among them, the Vicsek model [23, 48] is one of the most studied. In this model all agents move at a constant speed while trying to adopt the same velocity as their neighbors, up to some noise.

In this introductory section we present the discrete dynamics for the Vicsek model as presented in [23] as well as its corresponding kinetic equation for the time evolution of the distribution of the particles. Up to now, there were no complete rigorous results on the existence of solutions and the derivation of the kinetic equation (mean-field

limit) for the Vicsek dynamics presented in [23]. The reason for this is the presence of a singularity in the dynamics that is reached when the local average velocity of the agents vanishes.

In contrast to [23], one can define a Vicsek model without a singularity [3, 15, 16, 32] for different modeling choices. However, these different modeling choices for the Vicsek model have profound mathematical implications (like appearance of phase transitions). For this reason, in this article we investigate a general form of the Vicsek model that includes the forms presented in [23] (with a singularity) and in [3, 32] (without singularity), as well as further extensions of the model like those in [4, 16, 25, 31]. We give conditions to have existence of solutions for the equations and investigate approximations to the mean-field limit for this general class of Vicsek models. All of these are detailed in the following sections.

1.2. Particular case: Vicsek model as in [23] and [3, 32]. In what follows we assume that the agents move in a domain Ω, with Ω being either the d-dimensional torus $\mathbb{T}^d$ or the whole space $\mathbb{R}^d$ for $d \geqslant 1$. The orientations of the agents are given elements on the sphere $\mathbb{S}^{d-1}$. Throughout the article, for a given vector $\omega \in \mathbb{R}^d$ we denote by $\mathbf{P}_{\omega^\bot}$ the orthogonal projection in $\mathbb{R}^d$ onto $(\mathbb{R}\omega)^\bot$. The time-continuous Vicsek model [23] considers N agents characterized by their positions $X^{i,N}(t) \in \Omega$, $i = 1,\dots,N$, and orientations $\omega^{i,N}(t) \in \mathbb{S}^{d-1}$ over time $t \geqslant 0$. The evolution of the system is given by the following Stratonovich stochastic differential equation, where $c, \nu, \sigma > 0$ are positive constants:

$$dX^{i,N} = c\,\omega^{i,N}\,dt, \tag{1.1}$$

$$d\omega^{i,N} = \nu \nabla_{\omega^{i,N}}(\omega^{i,N} \cdot \bar{\omega}_i)\,dt + \mathbf{P}_{(\omega^{i,N})^\bot} \circ \sqrt{2\sigma}\,dB_t^i. \tag{1.2}$$

Next, we explain this system of equations. Equation (1.1) describes the transport of the agents: agent i moves in the orientation $\omega^{i,N}$ at speed $c > 0$. Equation (1.2) gives the evolution of the orientations over time. It includes two competing forces. On one hand the first term in the form of a gradient represents organized motion: agents try to adopt the same orientation. The gradient ∇_ω is the gradient on the sphere, and the term $\bar{\omega}_i$ represents the average orientation of the neighbors around agent i. Therefore, the first term on the right-hand-side of (1.2) is a gradient flow which relaxes the value of the orientation $\omega^{i,N}$ towards the mean orientation of the neighboring particles $\bar{\omega}_i$. The constant $\nu > 0$ gives the speed of this relaxation. We will comment later on the different choices to compute the average orientation $\bar{\omega}_i$. These different choices give rise to the different models in [23] and [3, 32].

On the other hand B_t^i, $i = 1, \dots, N$, are N independent Brownian motions in $\mathbb{R}^d$ and they introduce noise in the dynamics, driving particles away from organized motion. The constant $\sigma > 0$ gives the intensity of the noise. The symbol "$\circ$" is used to specify that (1.2) has to be understood in the Stratonovich sense. This ensures that $\omega_i(t) \in \mathbb{S}^{d-1}$ for all times where the solution is defined (this will be proven later).

Now, formally at least, one can compute the time evolution for the distribution of agents $f = f(t,x,\omega)$ in space $x \in \Omega$ and orientations $\omega \in \mathbb{S}^{d-1}$ at time $t \geqslant 0$ as the number of particles $N \rightarrow \infty$ [23]. The dynamics for the distribution f is given by the following kinetic equation:

$$\partial_t f + c\,\omega \cdot \nabla_\omega f = \nabla_\omega \cdot \left[-\nu \mathbf{P}_{\omega^\bot} \bar{\omega}_f f + \sigma \nabla_\omega f \right] =: L(f), \tag{1.3}$$

where $\nabla_\omega \cdot$ denotes the divergence on the sphere $\mathbb{S}^{d-1}$ and where $\bar{\omega}_f(t,x)$ represents the average orientation of the particles around position x at time t. Notice that the projection term appears due to the fact that

$$\nabla_\omega (\omega \cdot \bar{\omega}) = \mathbf{P}_{\omega^\bot} \bar{\omega} \qquad \text{for } \bar{\omega} \in \mathbb{R}^d \text{ fixed.}$$

One can show that the operator L on the right-hand side of (1.3) can be rewritten as a nonlinear Fokker--Planck operator:

$$L(f) := \sigma \nabla_\omega \cdot \left[M_{\overline{\omega}_f} \nabla_\omega \left(\frac{f}{M_{\overline{\omega}_f}} \right) \right], \tag{1.4}$$

with

$$M_v(\omega) = \frac{\exp(\frac{\nu}{\sigma}(\omega\cdot v))}{Z}, \qquad Z = \int_{\mathbb{S}^{d-1}} \exp\left(\frac{\nu}{\sigma}(\omega\cdot v)\right) d\omega \tag{1.5}$$

for any $v \in \mathbb{S}^{d-1}$. The function M_v is a probability density on the sphere called the von Mises distribution. Notice that its mean is in the orientation given by v. When $\sigma \rightarrow \infty$ (large-noise limit) the von Mises distribution converges to the uniform distribution on the sphere.

We comment now on the different choices to define the average orientation around an agent i, denoted by $\overline{\omega}_i$ in (1.2). The choice considered has profound implications in the dynamics of the particles and in the derivation of the kinetic equation (1.3) (mean-field limit) and the derivation of macroscopic equations (equations for the particle density and mean orientation). In the mathematics literature, two options were first considered, given, in a compact way, by

$$\overline{\omega}_i = \frac{J_i^N}{\alpha + (1-\alpha)|J_i^N|}, \qquad J_i^N(t) = \sum_{j\neq i}^N K(|X_i(t) - X_j(t)|)\,\omega_j(t). \tag{1.6}$$

The two options correspond to the cases $\alpha = 0$ (as presented in [23]) or $\alpha = 1$ (as presented in [3, 32]). The kernel $K \geqslant 0$ is an interaction kernel that represents the weights given to the neighboring particles depending on their distance to particle i. A classical choice is to take K the indicator function of a ball of radius $R > 0$. The flux J_i^N is, indeed, an average of the orientations of the neighboring particles. In the kinetic equation these choices define the operator $\overline{\omega}_f$ as

$$\overline{\omega}_f(t,x) = \frac{\mathbf{J}_f(t,x)}{\alpha + (1-\alpha)|\mathbf{J}_f(t,x)|}, \qquad \mathbf{J}_f(t,x) = \int_{\Omega\times\mathbb{S}^{d-1}} K(|x-y|)\,\omega f(t,y,\omega)\,dy\,d\omega. \tag{1.7}$$

We compare next, the mathematical implications of these two choices:

Case $\boldsymbol{\alpha}$ = 1 (nonnormalization). If $\alpha = 1$, the average orientation $\overline{\omega}_i = J_i^N \notin \mathbb{S}^{d-1}$ is not a unit vector. This also holds in the kinetic equation (1.3) for $\overline{\omega}_f = \mathbf{J}_f \notin \mathbb{S}^{d-1}$ in (1.7). However, the case $\alpha = 1$ removes the singularity when $|\mathbf{J}_f| = 0$ or $|J_i^N| = 0$. In [32] the authors prove the well-posedness of the spatially homogeneous kinetic equation in any Sobolev space, and in [2] the authors prove the mean-field limit (here we will recover these results). As a counterpart, though, this choice for the average $\overline{\omega}_i$ makes the derivation of macroscopic equations for the density of the particles $\rho = \rho(t,x)$ and the mean orientation $\omega = \omega(t,x)$ of the agents more complex than in the case $\alpha = 0$. Specifically, two equilibria exist for the mean particle orientation ω depending on whether $\mathbf{J}_f \neq 0$ or $\mathbf{J}_f = 0$ [15, 16, 32]. In loose terms, if $\mathbf{J}_f \neq 0$, then the von Mises equilibria $M_{\mathbf{J}_f}$ is an equilibrium. The other equilibrium is given by the uniform distribution on the sphere, which corresponds to the case $\mathbf{J}_f = 0$ in $M_{\mathbf{J}_f}$. In different spatial regions (depending on the particle density ρ) one of these two equilibria is stable. Consequently, this gives rise to a bifurcation or phase transition: in some spatial regions the mean orientation of the agents is $\omega = 0$ (corresponding to disordered dynamics), and in other spatial regions the mean orientation is given by a unit vector $|\omega| = 1$ (corresponding to ordered dynamics or flocking). The interested reader can find all the details in [4, 15, 16, 32]. The presence of phase transitions enriches the dynamics in the sense that allows for a wider variety of patterns to arise. Understanding phase transition phenomena in the Vicsek model is also of great interest in the physics community [9, 48]. In particular, pattern formation has been studied through the simulation of the discrete dynamics for the Vicsek model and its variation. Band formation arises under some parameter conditions [6, 8, 9]. Phase transitions could be key to explaining the emergence of this band formation.

Case $\boldsymbol{\alpha} = 0$ (normalization). The first mathematical works on the Vicsek model correspond to this case [23]. The choice $\alpha = 0$ comes with the difficulty of dealing with singularities when the flux $J_i^N = 0$ (another example of collective dynamics with singularities, different from the Vicsek model, can be found in [45] for the Kuramoto model). However, assuming that the flux $\mathbf{J}_f$ in the kinetic equation does not vanish along the dynamics, then the macroscopic equations can be obtained more easily than in the case $\alpha = 1$ because there is a unique equilibria: the von Mises distribution in (1.5). Then the macroscopic equations for the particle density $\rho = \rho(t,x)$ and mean orientation $\bar{\omega} = \bar{\omega}(t,x) \in \mathbb{S}^{d-1}$ correspond to the self-organized hydrodynamics (SOH) given in [23]. This was the first formal derivation of the SOH dynamics. See also [38, 50] as well as [37] for later rigorous results. In this scenario a rigorous mean-field limit to derive the kinetic equation was missing as well as a Cauchy theory for the particle dynamics and the kinetic equation. Here we will not prove the mean-field limit starting from a particle system of the form (1.1)--(1.2), but from a modified system that we term the *approximated particle system*. This system does not have a singularity when the flux vanishes, and therefore it can be thought of as a regularization of the original particle dynamics. Proving the mean-field limit for particle systems with nonregular coefficients or with singularities is in general a difficult problem and, to the authors' knowledge, only very few results exist and focus on specific problems (see, e.g., [36]). Here, we will prove that with a probability which tends to one in the many-particle limit, for short times, solutions of the approximated particle system are also solutions of the Vicsek particle system. We investigate these questions in the present article.

Theorem 1.1 (Cauchy theory for general Vicsek models with normalization and mean field from approximated particle dynamics). *Suppose that* $\alpha = 0$ *(normalized case). Suppose that the kernel K is Lipschitz and bounded, and that* $f(0,x,\omega)$ *satisfies assumptions* (i) *and* (ii) *in Theorem* 2.4*. Then there exists a unique local-in-time solution to the kinetic equation* (1.7)*. Moreover, the kinetic equation* (1.7) *can be obtained as the mean-field limit of an "approximated" particle dynamics whose solutions are, with a probability which tends to one in the many-particle limit, also solutions to the particle Vicsek dynamics* (1.1)--(1.2)*.*

The precise definition of "solution" of the approximated particle system and what we mean by "a probability which tends to one in the many-particle limit" will be explained in Theorem 2.4 and in section 2.3, respectively. As mentioned before, the difficulty in proving the previous theorem is to show that there exists at least some time interval $t \in [0,T]$ such that $\mathbf{J}_f \neq 0$ and that for N large enough the probability that $|J_i^N| > 0$ goes to 1 as $N \rightarrow \infty$ for all $i = 1,\dots,N$ and $t \in [0,T]$. To prove a lower bound in the flux $\mathbf{J}_f \neq 0$ is, therefore, key. Indeed, in [33] the authors manage to prove well-posedness for the kinetic equation (1.3) assuming precisely that a priori all solutions satisfy $|\mathbf{J}_f|(t,x,\omega) \geqslant a > 0$ for all times, positions, and orientations. In [30] the authors prove well-posedness in the spatially homogeneous setting as long as initially $|\mathbf{J}_{f_0}| > 0$ (plus some other assumptions). This has been followed by further extensions in [39]. The spatially homogeneous case corresponds to spatially local kernels, i.e., $K(x-y) = \delta_0(x-y)$ (delta distribution).

To conclude this part, the choice on how to define the average orientation $\bar{\omega}_i$ has profound mathematical and modeling implications (like dealing with singularities or with phase transitions). For this reason, in this article we will consider a general class of Vicsek models to encompass different modeling choices—existing or yet to be developed.

1.3. General forms of the Vicsek model considered. As we saw in the previous section, modeling choices in the Vicsek model may produce important differences in the mathematical properties of the equations. For this reason, in this article we consider a wide class of Vicsek models based not only on the choice of the averaged orientation (1.2) but also on the particular shape of the friction ν, the viscosity σ, and the interaction kernel K. Specifically, we will allow ν and σ to be functions of f (the solution to the kinetic equation) at the kinetic level and functions of the empirical distribution (see (2.15)) at the particle level. The interaction kernel K can take the general form $K = K(t,x,x_\ast,\omega,\omega_\ast)$. The precise assumptions on ν, σ, and K can be found in section 2.1.

As a consequence, the results presented here apply to a wide breadth of Vicsektype models, like the ones in [15, 16, 25, 31]. In [15, 16] the authors consider functions $\nu = \nu(|\mathbf{J}_f|)$, $\sigma = \sigma(|\mathbf{J}_f|)$ that depend on the flux of the particles; in [31] the author considers a kernel K that is not isotropic; in [25] the authors consider a Vicsek-type model for alignment of discs. Beyond this, the approach presented here can be adapted to investigate other models like the ones in [4, 12, 13, 18, 17, 19, 20, 21, 22, 24, 42]. In [13] the authors couple the Vicsek model with a Kuramoto model. In [42] the author considers a Vicsek model with two species having different velocities. The models in [4, 18, 19, 20, 21] describe collective motion based on nematic alignment (where particles align by adopting the same direction of motion but not necessarily the same orientation). BGK versions of Vicsek-type models are considered in [12, 17]. In [22] the authors couple the Vicsek model with Stokes equations to model microswimmers. Finally, in [24] a model for the persistent turning walker with curvature "alignment" is presented.

1.4. Aims of the paper. In this paper, for a general form of the Vicsek model, we aim to

- (i) establish the well-posedness, at least locally in time, for the kinetic model (Theorem 2.4); existence of solutions will be proven in Lebesgue spaces;
- (ii) rigorously prove that the kinetic equation can be derived as the mean-field limit of some "approximated" agent-based dynamics in the limit $N \rightarrow +\infty$ (Theorems 2.10, 2.11, 2.12), and also show with a probability which tends to one in the many-particle limit realizations of the

"approximated" particle system that there are solutions of the original Vicsek system (Theorem 2.14);

(iii) derive a criterion that characterizes the global-in-time well-posedness of these systems (points (a) and (b) in Theorem 2.4) and apply it to several interesting cases (section 3.3.1); along the way we shall also open the discussion to longtime behavior and construct a free energy, but we shall not investigate further (section 3.3.2).

Our goal is to give constructive proofs rather than weak-compactness arguments and to work in Lebesgue spaces rather than more regular Sobolev spaces. Our strategy is thus to prove a fixed point contraction theory for the kinetic equation (2.4) and to prove the mean-field limit, using a coupling approach popularized by Sznitman [3, 5, 7, 47].

1.5. Structure of the paper. The paper is structured as follows. Section 2 describes our main results, first in the kinetic framework in section 2.2 and second for the particle dynamics and its mean-field limit in section 2.3. The proofs for the results in section 2.2 are given in section 3, and the proofs of the results in section 2.3 are given in section 4. In Appendix A the reader can find known results on stochastic differential equations and mean-field limits that are used in section 4 and that have been added here for the sake of completeness. The results in Appendix A are mainly from [5].

2. Main results and strategy.

2.1. Functional framework and notation. First let us give some notation for functional spaces. We shall work on Lebesgue spaces indexed by the variable into consideration: for any p in $[1,+\infty)$ we denote

$$\|f\|_{L^p_\omega} = \biggl(\int_{\mathbb{S}^{d-1}} |f(\omega)|^p\, d\omega\biggr)^{\frac{1}{p}} \quad \text{and} \quad \|f\|_{L^p_x} = \biggl(\int_{\Omega} |f(x)|^p\, dx\biggr)^{\frac{1}{p}}$$

and for a time-dependent function for any $T > 0$

$$\|f\|_{L^p_t} = \left(\int_{\mathbb{R}^+} |f(t)|^p\, dt\right)^{\frac{1}{p}} \quad \text{and} \quad \|f\|_{L^p_{[0,T)}} = \left(\int_0^T |f(t)|^p\, dt\right)^{\frac{1}{p}}.$$

Finally we denote several variable Lebesgue spaces for any q and r in $[1,+\infty)$,

$$\|f\|_{L^r_{[0,T)}L^q_x L^p_\omega} = \left(\int_0^T \left(\int_\Omega \|f(t,x,\cdot)\|^q_{L^p_\omega}\, dx\right)^{\frac{r}{q}} dt\right)^{\frac{1}{r}},$$

with direct modifications for p, q, or r being $+\infty$. Note that when two indexes are the same we shall use the shorthand notation $L^p_x L^p_\omega = L^p_{x,\omega}$.

For the mean-field limit results, we will consider the space $\mathcal{H} = (\Omega \times \mathbb{R}^d \times \mathcal{P}_2(\Omega \times \mathbb{R}^d))$, where $\mathcal{P}_2(\Omega \times \mathbb{R}^d)$ is the space of probability measures in $\Omega \times \mathbb{R}^d$ with finite second-order moment, i.e., $\mu \in \mathcal{P}_2(\Omega \times \mathbb{R}^d)$ if it fulfills

$$\int_{\Omega \times \mathbb{R}^d} |z|^2 \mu(dz) < \infty.$$

For $\mu, \mu' \in \mathcal{P}_2(\Omega \times \mathbb{R}^d)$, the 2-Wasserstein distance is given by

$$W_2(\mu,\mu') = \inf \Biggl\{ \left[\int_{(\Omega\times\mathbb{R}^d)^2} |z-z'|^2 \, \pi(dz,dz') \right]^{1/2};$$

$$\pi \in \mathcal{P}(\Omega \times \mathbb{R}^d)^2 \text{ with marginals } \mu \text{ and } \mu' \Biggr\}. \tag{2.1}$$

The distance W_2 induces the topology of weak convergence of measures and the convergence of all the moments of order up to 2 [5] (see also [49, p. 83]). The space $\mathcal{H}$ is a metric space with distance $d_{\mathcal{H}}$ given by

$$d_{\mathcal{H}}\Bigl((x,v,m),(y,w,p)\Bigr) = |x-y| + |v-w| + W_2(m,p).$$

The hypotheses we shall make on the nonlinearities are the following.

(H1) The interaction kernel $\mathbf{K}$ is Lipschitz, regular, and bounded in all the variables. More precisely, it is assumed that $\mathbf{K}(t,x,x_*,\omega,\omega_*) \in W^{1,\infty}_{t,x_*} W^{2,\infty}_{\omega_*} L^{\infty}_{x,\omega}$, and in the case $\Omega = \mathbb{R}^d$ we further assume that $\mathbf{K} \in L^{\infty}_{t,x,\omega} L^2_{x_*,\omega_*}$.

(H2) The viscosity satisfies that $\sigma : L^2_{[0,T],x,\omega} \longrightarrow L^{\infty}_{[0,T],x,\omega}$ is bounded from below and Lipschitz uniformly for any $T > 0$ in the sense there exists σ_0, σ_∞, and σ_{lip} such that for any $T > 0$,

$$\forall (f,g) \in \Bigl(L^2_{[0,T],x,\omega}\Bigr)^2, \quad 0 < \sigma_0 \leqslant \sigma(f) \leqslant \sigma_\infty,$$

$$\|\sigma(f) - \sigma(g)\|_{L^\infty_{[0,T],x,\omega}} \leqslant \sigma_{\mathrm{lip}} \|f-g\|_{L^2_{[0,T],x,\omega}}.$$

(H3) The friction is local in time and Lipschitz, that is, $\nu(f)(t,x,v) = \nu(f(t,\cdot,\cdot))(x,v)$, where $\nu : L^2_{x,\omega} \longrightarrow L^{\infty}_x W^{1,\infty}_\omega$:

$$\forall (f,g) \in \bigl(L^2_{x,\omega}\bigr)^2, \|\nu(f)\|_{L^\infty_x W^{1,\infty}_\omega} \leqslant \nu_\infty,$$

$$\|\nu(f) - \nu(g)\|_{L^\infty_{x,\omega}} \leqslant \nu_{\mathrm{lip}} \|f-g\|_{L^2_{x,\omega}}.$$

(H4) For the mean-field limit results we will assume further that σ, ν, and $\nabla_\omega \sigma$ are Lipschitz and bounded in W_2.

Remark 2.1.

- Note that (H2) and (H3) are satisfied when σ and ν do not depend on f, which is the case in most models so far. Ultimately, one would want $\sigma(f)$ to be local in time and Lipschitz, like $\nu(f)$. We think that this could be achieved with standard parabolic regularity methods for more regular initial data and $\sigma(f)$. The main issue for closing a fixed point argument is the speed of convergence to 0 of the gain of regularity of the solutions f that we did not manage to quantify; see Proposition 3.3 and Remark 3.4.
- In applications, typically the functions $\nu = \nu(t,x,\omega,f)$ and $\sigma = \sigma(t,x,\omega,f)$ are of the form

$$\nu(t,x,\omega,f) = \int_{\Omega\times\mathbb{S}^{d-1}} \widetilde{\nu}(t,x,\omega,x_*,\omega_*) f(t,x_*,\omega_*)\, dx_*\, d\omega_* \tag{2.2}$$

for some function $\widetilde{\nu}$, and analogously for σ. For example, in [23] the authors consider

$$\nu(t,x,\omega,f) = \omega\cdot\overline{\omega}_f(t,x),$$

with

$$\overline{\omega}_f(t,x) = \frac{\widetilde{\mathbf{J}}_f}{|\widetilde{\mathbf{J}}_f|}(t,x), \quad \widetilde{\mathbf{J}}_f(t,x) = \int_{\Omega\times\mathbb{S}^{d-1}} K(x-x_*,\omega_*) f(t,x_*,\omega_*)\, dx_*\, d\omega_*. \tag{2.3}$$

And in [16] the authors consider ν and σ of the form

$$\nu = \nu(t,x,f) = \nu(|\widetilde{\mathbf{J}}_f|(t,x)),$$

with $\widetilde{\mathbf{J}}_f$ given by (2.3).

2.2. Kinetic point of view: Strategy and results. We will consider the following kinetic equation for the distribution function $f = f(t,x,\omega)$ for $(t,x,\omega) \in \mathbb{R}_+ \times \Omega \times \mathbb{R}^d$ (with Ω either the d-dimensional torus $\mathbb{T}^d$ or the full space $\mathbb{R}^d$):

$$\partial_t f + c\,\omega\cdot\nabla_x f = \nabla_\omega\cdot(\sigma(f)\nabla_\omega f) + \nabla_\omega\cdot(\nu(f) f \mathbf{P}_{\omega^\perp}[\mathbf{\Psi}[f]]), \tag{2.4}$$

where ∇_ω, $\nabla_\omega\cdot$ are the gradient and divergence operators on the sphere; $c > 0$ is a positive constant and $\nu, \sigma, \mathbf{\Psi}$ are given functions.

Remark 2.2. Notice that we allow the function $\nu(f) < 0$. For example, in the kinetic equation for the Vicsek model in (1.3) this function corresponds to $-\nu$ for some constant ν strictly positive.

There exists an interaction kernel $\mathbf{K} : \mathbb{R}^+ \times \Omega^2 \times \bigl(\mathbb{S}^{d-1}\bigr)^2 \rightarrow \mathbb{R}^d$ that defines the flux $\boldsymbol{J}$ as

$$\int$$

$$\boldsymbol{J}[f](t,x,\omega) = \int_{\Omega\times\mathbb{S}} \mathbf{K}(t,x,x_*,\omega,\omega_*) f(t,x_*,\omega_*)\, dx_*\, d\omega_*. \tag{2.5}$$

From the flux $\boldsymbol{J}$ we define the functional $\boldsymbol{\Psi}$ as

$$\exists\, \alpha \in [0,1],\quad \boldsymbol{\Psi}[f](x,\omega) = \frac{\boldsymbol{J}[f](t,x,\omega)}{|\boldsymbol{J}[f](t,x,\omega)|_{\alpha}}, \tag{2.6}$$

where

$$|\boldsymbol{J}[f](t,x,\omega)|_{\alpha} = \alpha + (1-\alpha)|\boldsymbol{J}[f](t,x,\omega)|, \tag{2.7}$$

where $|\cdot|$ stands for the norm in $\mathbb{R}^d$. Note that when $\alpha = 0$ we talk about a normalized operator since $|\boldsymbol{\Psi}[f]| = 1$, whereas for $\alpha = 1$ we have a complete kernel operator.

The orthogonal kernel nonlinearity (2.4) we tackle is more general than the gradient-type interaction (2.8). Furthermore, it is also more physically relevant when one wants to derive the Kolmogorov–Vicsek type of kinetic equation from particle systems behavior [2, 3]. The main issue for the kinetic part is the possible degeneracy of $|\boldsymbol{\Psi}[f]|_{\alpha}$ as well as the nonlinearity of the dissipativity and viscosity that does not necessarily compensate for the aforementioned degeneracy, unlike existing works in the literature. The key part of our strategy is to provide first a quantification of the possible time of degeneracy of the nonlinearity combined to a study of the dependencies over σ and ν at a linear level.

Remark 2.3. In the literature, typically the operator K is of the form $\mathbf{K}(x,x_*,\omega,\omega_*) = \mathbf{K}(x-x_*,\omega_*)$. In this case it holds that

$$\mathbf{P}_{\omega^\bot}(\boldsymbol{\Psi}[f]) = \nabla_{\omega}\biggl(\int_{\Omega\times\mathbb{S}^{d-1}} \omega\cdot\mathbf{K}(x-x_*,\omega_*) f(x_*,\omega_*)\, dx_*\, d\omega_*\biggr)$$

because for a vector $\mathbf{X}$ in $\mathbb{R}^d$, $\nabla_{\omega}(\omega\cdot X) = \mathbf{P}_{\omega^\bot}(\mathbf{X})$. Therefore, in this case the Kolmogorov–Vicsek equation (2.4) can be rewritten as a gradient-type interaction

$$\partial_t f + c\,\omega\cdot\nabla_x f = \nabla_{\omega}\cdot(\sigma(f)\nabla_{\omega} f) + \nabla_{\omega}\cdot(\nu(f) f \nabla_{\omega}\psi[f]). \tag{2.8}$$

Theorem 2.4. *Let Ω be either $\mathbb{T}^d$ or $\mathbb{R}^d$, let α be in* $[0,1]$, *and let σ, ν, and $\mathbf{K}$ satisfy the hypotheses* (H1)-(H2)-(H3). *Let p belong to* $[2,+\infty]$ *and f_0 be such that*

(i) *f_0 is a nonnegative function in $L^1_{x,\omega}\cap L^p_{x,\omega}$ with mass*

$$\int_{\Omega\times\mathbb{S}^{d-1}} f_0(x,\omega)\, dx\, d\omega := M_0 > 0;$$

(ii) $\inf_{(x,\omega)\in\Omega\times\mathbb{S}^{d-1}} |\boldsymbol{J}[f_0](x,\omega)|_\alpha := J_0 > 0$, *where* $|\boldsymbol{J}[f_0]|_\alpha$ *was defined in* (2.7).

There exist a time $T_{\mathrm{max}} > 0$, *independent of p, and a unique weak solution f in*

$L^2\bigl([0,T_{\mathrm{max}}),L^1_{x,\omega}\cap L^2_{x,\omega}\bigr)$ *to* (2.4) *with* f_0 *as initial datum. Moreover, f is nonnegative on* $[0,T_{\mathrm{max}})$, *belongs to* $L^\infty\bigl([0,T_{\mathrm{max}}),L^p_{x,\omega}\bigr)\cap L^2\bigl([0,T_{\mathrm{max}}),L^2_xH^1_\omega\bigr)$, *and preserves the mass* M_0, *and one of the following holds:*

(a) $T_{\mathrm{max}} = +\infty$;

(b) $\lim_{t\to T_{\mathrm{max}}^-}\sup_{\Omega\times\mathbb{S}^{d-1}}\frac{\nu(f)}{|\boldsymbol{J}[f]|_\alpha} = +\infty$.

Remark 2.5. Of important note are the following consequences.

- Our Cauchy theory does not stand on any a priori assumption of the solutions and provides an explicit lower bound for T_{max} (see (3.18)–(3.20)). In particular, for $T < T_1$ (where $T_1 > 0$ is given in (3.20)), it holds that for all $t\in[0,T]$

$$|\boldsymbol{J}(t,x,\omega,f)| \geqslant J_0 - K_\infty M_0 T > 0, \tag{2.9}$$

where K_∞ is given by (3.17) and M_0 is as given in the theorem above. In what follows, we will define $c_\ast = c_\ast(T)$ as

$$c_\ast = J_0 - K_\infty M_0 T \qquad \text{for } T < T_1. \tag{2.10}$$

- The theorem above includes all the previous results made in L^2 or L^∞, in both the nonhomogeneous and the homogeneous cases (it suffices to consider $\mathbf{K} = k(x_\ast)\mathbf{K}(\omega,\omega_\ast)$ with $\int_\Omega k = 1$).
- The global existence criterion offers direct global existence for nonnormalized interactions $\alpha \neq 0$, but it also gives global existence for the original Vicsek equation with spatially homogeneous kernel $\mathbf{K}(t,x,x_\ast,\omega,\omega_\ast) = \omega_\ast$ and fully homogeneous viscosity and negative friction (i.e., only f and time dependent). Section 3.3.1 describes several general cases where global existence happens in the problematic and purely normalized case $\alpha = 0$.

Besides the issue of well-posedness, it is of great interest to understand the large time behavior of the solutions. We recall that section 3.3.1 proves global existence for different types of interactions, and it also exhibits a free energy which decreases along the flow. One cannot expect a general theory since it heavily depends on the shape of the kernel $\mathbf{K}$, but one can still wonder if there are equilibria and if they are attractive. For this purpose, kinetic equations with gradient nonlinearities (2.8) are often used because one can extract an explicit free energy functional decreasing along time, hence leading to the existence of equilibria and a hope for an asymptotic study of the solutions.

This has been tackled in [32, 14, 16], where the authors exhibited a decreasing energy functional

$$\mathcal{F}(t) = \int_{\mathbb{S}^{d-1}} f\ln(f)d\omega + \frac{1}{2}\int_0^{|\mathbf{J}_f|} \frac{\nu(s)}{\sigma(s)}ds \tag{2.11}$$

when \sigma and \nu are solely functions of | $\mathbf{J}_f$| . Associated to the latter is an energy dissipation allowing one to dig out explicitly the equilibria in the spatially homogeneous setting.

It appears that a free energy is also underlying in the general equation: (2.4) enjoys a free energy functional that decreases along the flow with an explicit energy dissipation. Namely,

$$\begin{aligned}\mathcal{F}[f](t) = &\int_{\Omega\times\mathbb{S}^{d-1}} f\ln(f)dxd\omega \\ &+ \int_0^t \int_{\Omega\times\mathbb{S}^{d-1}} f\left[\nabla_\omega\cdot(\nu(f)\mathbf{P}_{\omega^\perp}[\mathbf{\Psi}[f]]) - \frac{\nu(f)^2}{\sigma(f)}\left|\mathbf{P}_{\omega^\perp}[\mathbf{\Psi}[f]]\right|^2\right]dxd\omega ds\end{aligned} \tag{2.12}$$

and

$$\mathcal{D}[f](t) = \int_{\Omega\times\mathbb{S}^{d-1}} \sigma(f)f\left|\nabla_\omega\ln(f) + \frac{\nu(f)}{\sigma(f)}\mathbf{P}_{\omega^\perp}[\mathbf{\Psi}[f]]\right|^2 dxd\omega, \tag{2.13}$$

which will

be proven, in section 3.3, to satisfy along the flow

$$\frac{d}{dt}\mathcal{F}[f](t) = -\mathcal{D}[f](t).$$

Remark 2.6. Let us emphasize that, considering the spatially homogeneous case **K** = \omega \ast , (2.12) and (2.13) are the ones obtained in [14] with | $\mathbf{J}_f$| ^{2} instead of | $\mathbf{J}_f$| (in case of constant viscosity and friction we also recover the original Doi--Onsager free energy [43, 26]). For general gradient kernel interactions \psi [*f*] (2.8) that are symmetric in \omega and \omega \ast the free energy (2.12) becomes

$$\mathcal{F}(t) = \int_{\mathbb{S}^{d-1}} f\ln(f)d\omega + \frac{1}{2}\int_0^{\langle\psi[f],f\rangle_{L^2_{x,\omega}}} \frac{\nu(s)}{\sigma(s)}ds$$

for \nu and \sigma being solely functions of \langle \psi [*f*],*f*\rangle $_{L2x,\omega}$ (which equals | $\mathbf{J}_f$| ^{2}), hence offering a new view on the gradient structure where the natural dependencies are, in fact,

\langle \psi [*f*],*f*\rangle $_{L2x,\omega}$ = | $\mathbf{J}_f$| ^{2}.

One can immediately see that the energy dissipation vanishes on

$$\mathcal{E}_\infty = \left\{f_\infty \geqslant 0 \in L^\infty_t L^p_x H^2_\omega, \quad \nabla_\omega\left(\ln f_\infty\right)(t,x,\omega) = -\frac{\nu(f_\infty)}{\sigma(f_\infty)}\mathbf{P}_{\omega^\perp}[\mathbf{\Psi}[f_\infty]](t,x,\omega)\right\}, \tag{2.14}$$

and when it vanishes so does the right-hand side of the kinetic equation (2.4). The latter, with the decrease of \scrF [*f*](*t*)*,* offers a good chance that one could get a La Salle's invariance principle---in the spirit of [32]---when solutions are globally defined, namely, the solution draws closer to the set of local equilibria \scrE \infty .

Remark 2.7. If one establishes a La Salle's principle, then necessarily, for a global equilibrium *f*\infty , the quantity $\frac{\nu(f_\infty)}{\sigma(f_\infty)}\mathbf{P}_{\omega^\perp}[\mathbf{\Psi}[f_\infty]] = \nabla_\omega\left(\psi[f\right.$]) must be a tangential gradient---and then $f_\infty(t,x,\omega) = A(t,x)e^{-\psi[f](t,x,\omega)}$ and we recover generalized von

Mises equilibria as in previous works [32, 14, 16]. Hence, as mentioned earlier, a gradient structure pops out naturally but does not solely concern $\mathbf{P}_{\omega^\bot}[\mathbf{\Psi}[f]]$, as first considered in (2.8). However, this is beyond the scope of this article, and we did not investigate further.

2.3. Microscopic point of view and mean-field limit for the Vicsek kinetic equation: Strategy and results. We will prove that the kinetic equation can be obtained as the mean-field limit of some "approximated" (regularized) particle system. When there is normalization ($\alpha = 0$), we will show that realizations of the approximated dynamics are also solutions of the Vicsek particle system with a probability which tends to one in the many-particle limit (Theorem 2.14). Note, however, that for the case $\alpha = 0$, we will not prove the mean-field limit for the kinetic equation starting from particle systems of the form (1.1)--(1.2), nor its well-posedness.

We consider a system of N particles given by their positions $X_t^{i,N} \in \Omega$ and velocities $V_t^{i,N} \in \mathbb{R}^d$ over time $t \geqslant 0$. Notice that we shall consider that the equations are written for $v \in \mathbb{R}^d$ rather than in $\omega \in \mathbb{S}^{d-1}$, but we shall later prove that the velocities are restricted to the sphere with radius c, thus recovering the expected spherical dynamics. We recall the empirical distribution of the process $(X_t^{i,N}, V_t^{i,N})_{1\leqslant i\leqslant N}$:

$$\mu_t^N(x,v) := \frac{1}{N}\sum_{i=1}^N \delta_{(X_t^{i,N},V_t^{i,N})}(x,v) \quad (2.15)$$

As mentioned before, in this article we will prove results on two types of particle systems. The first one we will call the *general particle Vicsek* and the second one *approximated particle dynamics*. We define the general particle Vicsek as the particle system given by the following Stratonovich stochastic differential equation:

$$\begin{cases} dX_t^{i,N} = V_t^{i,N}dt, & (2.16a)\\ dV_t^{i,N} = \nu(\mu_t^N)\,\mathbf{P}_{(V_t^{i,N})^\perp}(\mathbf{\Psi}(X_t^{i,N},V_t^{i,N},\mu_t^N))dt \\ \qquad +\frac{1}{2}\mathbf{P}_{(V_t^{i,N})^\perp}[(\nabla_v\sigma(\mu_t^N))(X_t^{i,N},V_t^{i,N})]\,dt & (2.16b)\\ \qquad +\sqrt{2\sigma(\mu_t^N)}\,\mathbf{P}_{(V_t^{i,N})^\perp}\circ dB_t^i, & (2.16c)\\ X_t^{i,N}(t=0)=X_0^{i,N}, \quad V_t^{i,N}(t=0)=V_0^{i,N}, & (2.16d)\end{cases}$$

where $\mathbf{P}_{v^\bot}$ is the projection operator

$$\mathbf{P}_{v^\perp} = \mathrm{Id} - \frac{v\otimes v}{|v|^2},$$

where Id is the identity matrix, and $((B_t^i)_{t\geqslant})_{i=1,\dots,N}$ are independent Brownian motions in $\mathbb{R}^d$. The symbol "$\circ$" denotes that the stochastic differential equation (2.16) is in the Stratonovich convention.

The precise way in which the function $\mathbf{\Psi}$ is extended to $V_t \in \mathbb{R}^d$ is explained in section 4.2; see system (4.5).

The approximated particle system is given by similar equations, where the difference is that the functional $\mathbf{\Psi}$ is replaced by a functional τ_{ε_0} that will be made precise later:

$$
\begin{cases}
dX_t^{i,N} = V_t^{i,N} dt, & \text{(2.17a)}\\
dV_t^{i,N} = \nu(\mu_t^N)\,\mathbf{P}_{(V_t^{i,N})^\perp}(\tau_{\varepsilon_0}(X_t^{i,N}, V_t^{i,N}, \mu_t^N))dt & \\
\qquad + \frac{1}{2}\mathbf{P}_{(V_t^{i,N})^\perp}[(\nabla_v \sigma(\mu_t^N))(X_t^{i,N}, V_t^{i,N})]\,dt & \text{(2.17b)}\\
\qquad + \sqrt{2\sigma(\mu_t^N)}\,\mathbf{P}_{(V_t^{i,N})^\perp} \circ dB_t^i, & \text{(2.17c)}\\
X_t^{i,N}(t=0) = X_0^{i,N}, \quad V_t^{i,N}(t=0) = V_0^{i,N}. & \text{(2.17d)}
\end{cases}
$$

The function $\tau_{\varepsilon_0} = \tau_{\varepsilon_0}(x,v,m)$ is defined such that

$$\tau_{\varepsilon_0}(x,v,m) = \Psi(x,v,m) \text{ whenever } |\mathbf{J}(x,v,m)| \geqslant \varepsilon_0 \tag{2.18}$$

for some $\varepsilon_0 > 0$, and so that it is well-defined in $\mathscr{H}$ (i.e., without singularities). In particular, in the system above we will have that

$$\tau_{\varepsilon_0}(X_t^{i,N}, V_t^{i,N}, \mu_t^N) = \Psi(X_t^{i,N}, V_t^{i,N}, \mu_t^N) \text{ if } |\mathbf{J}(X_t^{i,N}, V_t^{i,N}, \mu_t^N)| \geqslant \varepsilon_0.$$

Because τ_{ε_0} does not present any singularities when $|\mathbf{J}| = 0$, the approximated particle system (2.17) can be thought of as a regularization of the Vicsek particle system (2.16).

Associated to the approximated particle system, we define the approximated kinetic equation given by

$$\partial_t f + c\omega \cdot \nabla_x f = \nabla_\omega \cdot (\sigma(f)\nabla_\omega f) + \nabla_\omega \cdot (\nu(f) f \mathbf{P}_{\omega^\bot}[\tau_{\varepsilon_0}[f]]). \tag{2.19}$$

Observe that when τ_{ε_0} is substituted by Ψ in the approximated kinetic equation (2.19), we obtain the Vicsek kinetic equation (2.4).

First, we will show the well-posedness for the approximated particle system (Theorem 2.10) and that in the mean-field limit gives the Vicsek kinetic equation (for short times) (Theorem 2.11 and Corollary 2.12).

Remark 2.8. Notice that the term (2.16b) in system (2.16) appears so that we obtain the kinetic equation (2.4). This is just a technicality: system (2.16) in Itô's convention corresponds to

$$
\begin{cases}
dX^{i,N} = V^{i,N} dt, & \text{(2.20a)}\\
dV^{i,N} = \nu(\mu^N)\, P_{(V^{i,N})^\perp}(\Psi(X^{i,N}, V^{i,N}, \mu^N))dt & \\
\qquad + P_{(\overline{V}^{i,N})^\perp}[(\nabla_v \sigma(\mu^N))(\overline{X}^{i,N}, \overline{V}^{i,N})]\,dt & \\
\qquad + \sqrt{2\sigma(\mu^N)}\, P_{(V^{i,N})^\perp} dB_t^i & \\
\qquad - 2(d-1)\sigma(\mu^N)\dfrac{V^{i,N}}{|V^{i,N}|^2}, & \text{(2.20b)}\\
X^{i,N}(t=0) = X_0^{i,N}, \quad V^{i,N}(t=0) = V_0^{i,N}. & \text{(2.20c)}
\end{cases}
$$

See also Appendix A.2 for more details. Without the extra term (2.16b) we would obtain a kinetic equation where the operator in ω is of the form (see (A.3))

$$\Delta_\omega(\sigma(f)f) - \frac{1}{2}\nabla_\omega \cdot (\nabla_\omega \sigma(f) f).$$

But with the extra term the operator ω in the kinetic equation is of the form

$$\Delta_\omega(\sigma(f)f) - \nabla_\omega \cdot (\nabla_\omega \sigma(f)f) = \nabla_\omega \cdot (\sigma(f)\nabla_\omega f),$$

which is the operator that we are dealing with in (2.4). Notice that if we wanted to obtain just a factor $\Delta_\omega(\sigma(f)f)$, then the constant in front of the extra term (2.16b) should be $-1/2$ rather than $1/2$. Notice also that if $\sigma = \sigma(t,x,f)$ does not depend on ω, then the extra term (2.16b) does not appear. This is the case, for example, when

$\sigma = \sigma(|\mathbf{J}\widetilde{f}|(t,x))$ for $\mathbf{J}\widetilde{f}$ defined in (2.3); see [16]. The extra term (2.16b) has the effect of relaxing $V^{i,N}$ towards the value taken by $\nabla_v \sigma$. It can be rewritten equivalently as a gradient

$$\frac{1}{2}\mathbf{P}_{(V_t^{i,N})^\perp}[(\nabla_v \sigma(\mu_t^N))(X_t^{i,N}, V_t^{i,N})]\,dt = \frac{1}{2}\nabla_{V^{i,N}}(V^{i,N} \cdot \nabla_v \sigma(\mu_t^N))(X_t^{i,N}, V_t^{i,N}),$$

where the gradient is on the sphere.

As announced before, the dynamics described by our agent-based model indeed force the velocities to have unit norm, as shown by the next lemma.

Lemma 2.9. *Suppose that* $(X_t^{i,N}, V_t^{i,N})$ *is a solution to the approximated particle dynamics* (2.17) *or the Vicsek particle dynamics* (2.16)*. Suppose that* $|V_0^{i,N}| = 1$ *for all* $i = 1,\dots,N$*; then it holds that*

$$|V_t^{i,N}| = 1 \qquad \forall i = 1,\dots,N,$$

for all times where the solution is defined.

Proof of Lemma 2.9*.* It is a direct check that

$$d|V_t^{i,N}|^2 = 0$$

using the Stratonovich chain rule (see, for example, [29, p. 122]) and the fact that $V \cdot \mathbf{P}_{V^\bot} = 0$. ∎

The potential degeneracy of $\mathbf{\Psi}$ breaks the Lipschitz regularity of the interactions and thus prevents standard agent-based well-posedness results from being applied. There exist results for non-Lipschitz interactions [2], but their singularities differ from the one at stake in the present article.

In the spirit of [3], we thus shall prove well-posedness of the system hand in hand with the mean-field limit with the aid of an auxiliary process which mixes microscopic dynamics with the mesoscopic distribution function f_t. The main difference with respect to [3] is that we need to deal with the fact that the noise coefficient σ is not constant. Let us make explicit the auxiliary process for the approximated particle system $(\bar{X}_t^{i,N}, \bar{V}_t^{i,N})_{t\geqslant 0}$, $i = 1,\dots,N$, solution of

$$\begin{cases} d\overline{X}_t^{i,N} = \overline{V}_t^{i,N}\,dt, & \text{(2.21a)}\\ d\overline{V}_t^{i,N} = \nu(f_t)\,\mathbf{P}_{(\overline{V}_t^{i,N})^\perp}(\tau_{\varepsilon_0}(\overline{X}_t^{i,N},\overline{V}_t^{i,N},f_t))dt \\ \qquad +\frac{1}{2}\mathbf{P}_{(\overline{V}_t^{i,N})^\perp}[(\nabla_\omega\sigma(f))(\overline{X}_t^{i,N},\overline{V}_t^{i,N})]\,dt \\ \qquad +\sqrt{2\sigma(f_t)}\,\mathbf{P}_{(\overline{V}_t^{i,N})^\perp}\circ dB_t^i, & \text{(2.21b)}\\ f_t = \mathrm{law}(\overline{X}_t^{i,N},\overline{V}_t^{i,N}), & \text{(2.21c)}\\ \overline{X}_t^{i,N}(t=0) = X_0^{i,N}, \quad \overline{V}_t^{i,N}(t=0) = V_0^{i,N}, & \text{(2.21d)}\end{cases}$$

where the initial data and the Brownian processes $(B_t^i)_{t\geqslant 0}$ are the same ones as in (2.17). Notice that all the processes $(\overline{X}_t^{i,N},\overline{V}_t^{i,N})$ are independent by construction, and they all have the same law f_t.

We will prove the following two results for the approximated particle system (2.17) and the approximated kinetic equation (2.19).

Theorem 2.10 (local-in-time existence and uniqueness of solutions for the approximated system). *Under the assumptions of Theorem* 2.4 *and assumption* (H4)*, let f_0 be a probability measure on $\mathbb{R}^d\times\mathbb{S}^{d-1}$ with finite second moment in $x\in\mathbb{R}^d$, and let $(X_0^{i,N},V_0^{i,N})$ for $i=1,\dots,N$ be N independent random variables with law f_0.*

The following hold:

- (i) *There exists a pathwise global unique solution to the SDE system* (2.17) *with initial data $(X_0^{i,N},V_0^{i,N})$ for $i=1,\dots,N$. Moreover, the solution is such that $|V_t^{i,N}|=1$.*
- (ii) *There exists a pathwise global unique solution to the auxiliary process* (2.21) *with initial data $(X_0^{i,N},V_0^{i,N})$ for $i=1,\dots,N$ and $|\overline{V}_t^{i,N}|=1$.*
- (iii) *There exists a global-in-time unique weak solution of the approximated kinetic equation* (2.19) *with initial datum f_0. The solution of the kinetic equation is the law of the process solution to the auxiliary system* (2.21)*, wherever the solution is defined.*

Along with this well-posedness result, we rigorously show its mean-field limit towards the approximated kinetic equation (2.19).

Theorem 2.11 (mean-field limit for the approximated system: propagation of chaos). *Under the assumptions of Theorem* 2.4 *and assumption* (H4)*, for the respective solutions $(X_t^{i,N},V_t^{i,N})_{t\geqslant 0}$ and $(\overline{X}_t^{i,N},\overline{V}_t^{i,N})_{t\geqslant 0}$ of* (2.17) *and* (2.21)*, for any $T>0$, it holds that*

$$\lim_{N\to\infty}\sup_{1\leqslant i\leqslant N}\mathbb{E}\Big[\sup_{0\leqslant t\leqslant T}\big(|X_t^{i,N}-\overline{X}_t^{i,N}|^2+|V_t^{i,N}-\overline{V}_t^{i,N}|^2\big)\Big]=0. \tag{2.22}$$

From this we deduce a mean-field limit result for the Vicsek kinetic equation.

Corollary 2.12 (mean-field limit for the Vicsek kinetic equation). *Suppose that the assumptions in Theorem* 2.4 *and assumption* (H4) *hold. Let f_t be the local-in-time solution of* (2.4) *given by Theorem* 2.4 *for $t\in[0,T)$. Then there exists an $\varepsilon_0>0$ such that the law of the auxiliary process* (2.21) *is precisely f_t for any $t\in[0,T)$.*

Consequently, Theorem 2.11 *holds for $(\overline{X}_t^i,\overline{V}_t^i)$ having law f_t solution to the kinetic Vicsek equation* (2.4)*.*

Proof. Take $T_0 \in [0,T)$ and define

$$\varepsilon_0 := \inf_{t\in[0,T_0]} \inf_{(x,v)} |\mathbf{J}(t,x,v,f_t)| > 0 .$$

Then, with this value of ε_0, the approximated kinetic equation (2.19) coincides with the Vicsek kinetic equation (2.4) for $t \in [0,T_0]$. Therefore, by Theorem 2.10, the solution of the Vicsek kinetic equation f_t is the law of the auxiliary process (2.21) for $t \in [0,T_0]$ and Theorem 2.11 applies. □

Corollary 2.12 implies the mean-field limit of the approximated dynamics (2.17) towards the Vicsek kinetic equation (2.4) for short times (in the case $\alpha \neq 0$, then $T = +\infty$ and the mean-field limit holds for all times).

Remark 2.13 (convergence of the measures). Expression (2.22) ensures the convergence as $N \rightarrow \infty$ of the law of the process $(X_t^{i,N}, V_t^{i,N})$ towards f_t for any i and $t \in [0,T]$. See the notes after Theorem 1.1 in [2] for more details. In particular, in [2] the following upper bound in 2-Wasserstein distance is shown:

$$W_2^2(f_t^{(1)}, f_t) \leqslant \mathbb{E}\left[|X_t^i - \overline{X}_t^i|^2 + |V_t^i - \overline{V}_t^i|^2\right] \leqslant \varepsilon(N),$$

where $\varepsilon(N) \rightarrow 0$ as $N \rightarrow \infty$ (by Theorem 2.11). The function $f^{(1)}$ denotes the first marginal of the N particle system. They also show that for any Lipschitz map φ

$$\mathbb{E}\left[\left|\frac{1}{N}\sum_{i=1}^{N}\varphi(X_t^i, V_t^i) - \int \varphi\, df_t\right|^2\right] \leqslant \varepsilon(N) + \frac{C}{N} \tag{2.23}$$

for some constant $C > 0$ independent of N.

Under more regularity assumptions, one can obtain explicit estimates on $\varepsilon(N)$. See Theorem 10 in [5] (section 1.3.4).

2.3.1. The Vicsek particle system when $\alpha = 0$. All the previous results for the mean-field limit correspond only to the approximated system (2.17), which is not singular when $\alpha = 0$. When the norm of the flux

$$|\mathbf{J}(t, X^{i,N}, V^{i,N}, \mu_t^N)| \geqslant \varepsilon_0 \tag{2.24}$$

for all $t \in [0,T]$ and all $i = 1,\dots,N$, the approximated particle dynamics (2.17) coincides with the Vicsek particle dynamics (2.16) for $t \in [0,T]$.

When $\alpha = 0$ (normalized case), every realization of the "approximated" particle system (2.17) such that $|\mathbf{J}(t,X_t^{i,N},V_t^{i,N},\mu_t^N)| > \varepsilon_0$ for all $t \in [0,T]$ and all $i = 1,\dots,N$ is also a solution to the Vicsek particle system. The next result shows that for short times (2.24) happens with a probability which tends to one in the manyparticle limit.

Theorem 2.14 (lower bound on the flux for the normalized case). *Suppose that we are under the assumptions of Theorem* 2.4 *and assumption* (H4). *Consider a time $T < T_1$ (where T_1 is defined in* (3.20)*) and $c_\ast = c_\ast(T) > 0$ given by* (2.10). *Then, for all $t \in [0,T]$ and all $\varepsilon_0 \in (0,c_\ast)$, it holds that*

$$\mathbb{P}\left(\inf_{t\in[0,T]} \inf_{(x,v)} |\mathbf{J}(t,x,v,\mu_t^N)| > \varepsilon_0\right) \geqslant 1 - \frac{1}{c_* - \varepsilon_0}\,\varepsilon(N), \tag{2.25}$$

where $\varepsilon(N) \rightarrow 0$ *as* $N \rightarrow \infty$.

Define $A^{(N)}_{T,\varepsilon_0}$ as the event such that

$$A^{(N)}_{T,\varepsilon_0} = \left\{ \text{for all } t \in [0,T] \text{ and all } i = 1,\dots,N,\ \left| \boldsymbol{J}(t, X^{i,N}_t, V^{i,N}_t, \mu^N_t) \right| > \varepsilon_0 \right\}.$$

Then, as a consequence of Theorem 2.14, it holds that for all $\varepsilon_0 \in (0, c_\ast)$

$$\mathbb{P}\left(A^{(N)}_{T,\varepsilon_0} \right) \geqslant 1 - \frac{1}{c_* - \varepsilon_0} \varepsilon(N), \tag{2.26}$$

and so with a probability which tends to one in the many-particle limit, in the sense of (2.26), realizations of the approximated particle dynamics (2.17) will have nonzero flux for short times, and they will also be solutions to the general Vicsek particle system (2.16).

Remark 2.15. Proving the well-posedness and the rigorous mean-field limit for the Vicsek particle system (2.16) in the normalized case ($\alpha = 0$) seems to at least require (using the current strategy) showing that $|\mathbf{J}(t,x,v,\mu^N_t)| > 0$ a.s. for all N large enough, for all times in some interval, and for all values of (x,v). It is not clear if this is even true. This would require some kind of generalization of a strong law of large numbers. Such a strong law of large numbers exists for infinite exchangeable sequences as a result of de Finetti's theorem (see [40]). However, to the best of our knowledge, there is not an analogous result for a hierarchy of finite exchangeable sequences.

3. The mesoscopic framework: Kinetic differential equation.

3.1. Linear equation: Dependence on coefficients and averaging positivity. The differential operator ∇_ω is the tangential gradient---also called Günter derivatives---on the sphere $\mathbb{S}^{d-1}$. It can be related to the standard gradient on $\mathbb{R}^d$ by

$$\nabla_\omega = \nabla_v - \langle \nabla_v, \omega \rangle_{\mathbb{R}^d} \omega = \mathbf{P}_{\omega^\bot} [\nabla_v],$$

where ∇_v is the Euclidean gradient for functions from $\mathbb{R}^d$ to $\mathbb{R}^d$. First of all let us emphasize that the tangential gradient ∇_ω displays very different behavior from the usual gradient. We give here three formulas that we shall use throughout the proofs. We refer the interested reader to [27, 1] for an introduction on tangential Günterderivatives and to [44, Appendix II] for the specific calculus of the following formulas (proven in dimension 3 but immediately extendable in dimension d). Integration by parts is allowed but generates an additional term in the direction of orientation ω:

$$\int_{\mathbb{S}^{d-1}} \nabla_\omega (f(\omega)) g(\omega) d\omega = - \int_{\mathbb{S}^{d-1}} f(\omega) \nabla_\omega (g(\omega)) d\omega + (d-1) \int_{\mathbb{S}^{d-1}} \omega f g \, d\omega. \tag{3.1}$$

The following explicit tangential derivatives for each coordinate ω_i of $\omega = (\omega_1, \dots, \omega_d)$ on the sphere also hold:

$$\nabla_\omega(\omega_i) = (\delta_{ij} - \omega_i \omega_j)_{1 \leqslant j \leqslant d} \quad \text{and} \quad \Delta_\omega(\omega_i) = -(d-1)\omega_i. \tag{3.2}$$

As mentioned in the introduction, (2.4) can be viewed as a nonlinear Fokker-Planck equation on the torus with local and nonlinear viscosity σ and drift ν. The purpose of this section is to study the associated linear equation when the nonlinearities are considered as given parameters. In other words we consider here the following differential problem:

$$\partial_t f + c\,\omega \cdot \nabla_x f = \nabla_\omega \cdot \bigl(\overline{\sigma}(t,x,\omega)\nabla_\omega f + f\mathbf{\Psi}(t,x,\omega)\bigr), \tag{3.3}$$

where $\overline{\sigma}$ and $\mathbf{\Psi}$ are given functions.

The issue of existence and uniqueness for (3.3) has been solved, for instance, in

[11, Appendix A] for velocities in $\mathbb{R}$ or $\mathbb{R}^2$ for constant $\overline{\sigma}$ with $\mathbf{\Psi}$ and $\nabla_\omega \cdot \mathbf{\Psi}$ being in $L^\infty_{x,\omega}$. However, their methods are directly applicable for velocities in $\mathbb{S}^{d-1}$ as shown in [33, Lemma 4.1] for constant $\overline{\sigma}$. The case of nonconstant $\overline{\sigma}$ is a straightforward adaptation of their proofs if $\overline{\sigma}$ belongs to $L^\infty_{t,x,\omega}$ and is uniformly bounded from below by a constant $\sigma^1 > 0$ (note that one could also adapt to tangential derivatives standard Galerkin-type methods for linear parabolic equations [28, Chapter 6], as done in [32] in the spatially homogeneous case). We thus omit the proof and state the following theorem.

Theorem 3.1. *Let $\overline{\sigma}$ be in $L^\infty_{t,x,\omega}$ uniformly bounded from below, $\overline{\sigma}(t,x,\omega) \geqslant \sigma_0 >$ Then for any f_0 in $L^2_{x,\omega}$ there exists a unique f in the space*

$$Y = \bigl\{u \in L^2_{t,x}H^1_\omega, \quad \partial_t u + \omega \cdot \nabla_x u \in L^2_{t,x}H^{-1}_\omega\bigr\},$$

solution to (3.3) *with initial datum f_0.*

Moreover, if f_0 is nonnegative, then f is nonnegative.

Section 3.1.1 is dedicated to L^p bounds and gain of regularity for solutions to (3.3), while section 3.1.2 studies the dependences on the viscosity and friction. To conclude, section 3.1.3 tackles the issue of an explicit lower bound on the vanishing time for velocity averaging densities.

3.1.1. L^p bounds and gain of regularity. The issue of $L^p_{x,\omega}$ estimates and gain of regularity had already been investigated in previously cited references for

[1] *, and let $\mathbf{\Psi}$ and $\nabla_\omega \cdot \mathbf{\Psi}$ be in $L^\infty_{t,x,\omega}$ with the orthogonal property $\mathbf{P}_{\omega^\bot}(\mathbf{\Psi}) = \mathbf{\Psi}$.*

particular σ and $\mathbf{\Psi}$, and we provide here an adapted version that fits our general setting.

Proposition 3.2. *Let p belong to $[2,+\infty]$ and let $f_0 \geqslant 0$ be in $L^1_{x,\omega} \cap L^p_{x,\omega}$ with mass $\| f_0\|_{L^1_{x,\omega}} = M_0$. Under the assumptions of Theorem 3.1 on σ and $\mathbf{\Psi}$, the solution f to (3.3) is in $L^p_{x,\omega}$ and preserves the mass M_0. More precisely it satisfies*

$$\forall t \geqslant 0, \| f(t)\|_{L^p_{x,\omega}} \leqslant e^{C_p(\overline{\sigma},\overline{\mathbf{\Psi}})t} \| f_0\|_{L^p_{x,\omega}},$$

and it also gains regularity in ω when $p < +\infty$ in the following sense:

$$\int_0^t e^{C_p(\overline{\sigma},\overline{\mathbf{\Psi}})(t-s)} \left\| f^{\frac{p-2}{2}} \nabla_\omega f(s)\right\|^2_{L^2_{x,\omega}} \|f(s)\|^{1-p}_{L^p_{x,\omega}}\, ds \leqslant \frac{2\|f_0\|_{L^p_{x,\omega}}}{(p-1)\overline{\sigma}_0} e^{C_p(\overline{\sigma},\overline{\mathbf{\Psi}})t},$$

where

$$C_p(\overline{\sigma},\overline{\mathbf{\Psi}}) = \left\|\nabla_\omega \cdot \overline{\mathbf{\Psi}}\right\|_{L^\infty_{t,x,\omega}} + \frac{\left\|\overline{\mathbf{\Psi}}\right\|^2_{L^\infty_{t,x,\omega}}}{(p-1)\overline{\sigma}_0}, \tag{3.4}$$

with $C_\infty = \lim_{p\rightarrow\infty} C_p$.

Proof of Proposition 3.2.

Step 1*: L^p bounds and gain of regularity.* Note that since ∇_ω and $\overline{\mathbf{\Psi}}$ are orthogonal to ω, we can perform on (3.3) integration by parts on $\mathbb{S}^{d-1}$ as if we were working with classical derivatives. Indeed, the orthogonality to ω exactly cancels the additional term in the tangential derivatives integration-by-parts formula (3.1). It yields,

for any $p \geqslant 2$,

$$\begin{aligned} I := \frac{1}{p}\frac{d}{dt}\|f(t)\|^p_{L^p_{x,\omega}} &= \int_{\Omega\times\mathbb{S}^{d-1}} \operatorname{sgn}(f(t,x,v))\,|f(t,x,v)|^{p-1}\,\partial_t f(t,x,\omega)\,dxd\omega \\ &= -\int_{\Omega\times\mathbb{S}^{d-1}} \operatorname{sgn}(f)\,|f|^{p-1}\,\nabla_x(c\omega f) - \int_{\Omega\times\mathbb{S}^{d-1}} \overline{\sigma}(t,x,\omega)\nabla_\omega\left(\operatorname{sgn}(f)\,|f|^{p-1}\right)\cdot\nabla_\omega f \\ &\quad + \int_{\Omega\times\mathbb{S}^{d-1}} \operatorname{sgn}(f)\,|f|^{p-1}\,\nabla_\omega\cdot\left(f\overline{\mathbf{\Psi}}\right)\,dxd\omega. \end{aligned}$$

We use that the first integrand on the right-hand side can be written as a divergence in x: $\nabla_x \cdot (c\omega\, |f|^p)$, whereas the second one can be written as $(p-1)\overline{\sigma}(t,x,\omega)|f|^{p-2}(\nabla_\omega f)^2$.

This leads to the following upper bound:

$$I \leqslant -(p-1)\overline{\sigma}_0 \left\|f^{\frac{p-2}{2}}\nabla_\omega f\right\|^2_{L^2_{x,\omega}} + \int_{\Omega\times\mathbb{S}^{d-1}} \left(|f|^p\,\nabla_\omega\cdot\overline{\mathbf{\Psi}} + \overline{\mathbf{\Psi}}\,|f|^{p-1}\,|\nabla_\omega f|\right) dxd\omega$$

$$
\begin{aligned}
&\quad - (p-1)\overline{\sigma}_0 \left\| f^{\frac{p-2}{2}} \nabla_\omega f \right\|^2_{L^2_{x,\omega}} + \left\| \nabla_\omega \cdot \overline{\Psi} \right\|_{L^\infty_{t,x,\omega}} \|f\|^p_{L^p_{x,\omega}} \\
&\leqslant \quad + \left\| \overline{\Psi} \right\|_{L^\infty_{t,x,\omega}} \|f\|^{\frac{p}{2}}_{L^p_{x,\omega}} \left\| f^{\frac{p-2}{2}} \nabla_\omega f \right\|_{L^2_{x,\omega}} \\
&\leqslant - \left((p-1)\overline{\sigma}_0 - \frac{\eta}{2} \left\| \overline{\Psi} \right\|^2_{L^\infty_{t,x,\omega}} \right) \left\| f^{\frac{p-2}{2}} \nabla_\omega f \right\|^2_{L^2_{x,\omega}} + \left(\left\| \nabla_\omega \cdot \overline{\Psi} \right\|_{L^\infty_{t,x,\omega}} + \eta^{-1} \right) \|f\|^p_{L^p_{x,\omega}}.
\end{aligned}
$$

Note that we used the Cauchy–Schwarz and Young inequalities for any $\eta > 0$. Choosing $\frac{p-1}{2}\overline{\sigma}_0 - \frac{\eta}{2} \left\| \overline{\Psi} \right\|^2_{L^\infty_{t,x,v}} = 0$ gives us

$$
\begin{aligned}
\frac{d}{dt} \|f(t)\|_{L^p_{x,\omega}} &= \frac{1}{p} \frac{d}{dt} \|f(t)\|^p_{L^p_{x,\omega}} \|f(t)\|^{1-p}_{L^p_{x,\omega}} \\
&\leqslant -\frac{(p-1)}{2}\overline{\sigma}_0 \left\| f^{\frac{p-2}{2}} \nabla_\omega f \right\|^2_{L^2_{x,\omega}} \|f\|^{1-p}_{L^p_{x,\omega}} + C_p \|f\|_{L^p_{x,\omega}}.
\end{aligned}
$$

From this we deduce, thanks to the Grönwall lemma, that for all $t \geqslant 0$

$$
\|f(t)\|_{L^p_{x,\omega}} + \frac{(p-1)}{2}\overline{\sigma}_0 \int_0^t e^{C_p(t-s)} \left\| f^{\frac{p-2}{2}}(s) \nabla_\omega f(s) \right\|^2_{L^2_{x,\omega}} \|f(s)\|^{1-p}_{L^p_{x,\omega}} \, ds \leqslant e^{C_p t} \|f_0\|_{L^p_{x,\omega}}. \tag{3.5}
$$

The latter is exactly the standard bounds in Proposition 3.2 for p in $[2, +\infty)$, obtained as an a priori estimate. The proof that f belongs to L^p follows standard methods (also used in [33]) where our kinetic equation is approximated by a linear iterative scheme for which each solution f_n is in L^p and (3.5) holds at each step. Then the uniqueness of the kinetic solution implies that $f_n \rightarrow f$ and taking the limit in (3.5) shows that f belongs to L^p and satisfies (3.5).

Step 2: Mass conservation and L^∞ bounds. Now let us suppose that f_0 is nonnegative, which implies that f is nonnegative at all time by Theorem 3.1. Hence, it holds that

$$
\begin{aligned}
\frac{d}{dt} \|f(t)\|_{L^1_{x,\omega}} &= \frac{d}{dt} \int_{\Omega \times \mathbb{S}^{d-1}} f(t,x,\omega) \, dx d\omega \\
&= - \int_{\Omega \times \mathbb{S}^{d-1}} \nabla_x \cdot (c\omega f) + \nabla_\omega \cdot \left(\overline{\sigma} \nabla_\omega f + f \overline{\Psi} \right) dx d\omega = 0
\end{aligned}
$$

because, again, the integration by parts (3.1) does not generate additional terms in the direction of ω. This concludes the preservation of the $L^1_{x,\omega}$-norm.

The $L^\infty_{x,\omega}$ estimates follows straight from the limit p goes to $+\infty$, since $f(t)$ belongs to $L^1_{x,\omega}$. □

3.1.2. Dependence on the coefficients. Our strategy to tackle the nonlinear equation is via a fixed point argument. We thus need to understand how solutions to (3.3) differ from one another when they are associated to different coefficients $\overline{\sigma}$ and $\overline{\Psi}$. The main issue in establishing an estimate on $\|f_1 - f_2\|_{L^p_{x,\omega}}$ of two different solutions relies on the fact that the gain of regularity proved in Proposition 3.2 is highly nonlinear as soon as $p > 2$. Moreover, we did not manage to quantify the way the gain of regularity vanishes at initial time, so we cannot close a direct L^∞_t fixed point method. We shall thus only study the dependence on the coefficient in $L^2_{t,x}$.

Proposition 3.3. *Let f_0 be a nonnegative function in $L^1_{x,\omega} \cap L^2_{x,\omega}$. Suppose that $\overline{\sigma}_1$, $\overline{\sigma}_2$ and $\overline{\boldsymbol{\Psi}}_1$, $\overline{\boldsymbol{\Psi}}_2$ satisfy the assumptions of Theorem* 3.1. *Let f_1 and f_2 be the solutions of* (3.3) *associated, respectively, to the coefficients $(\overline{\sigma}_1, \overline{\boldsymbol{\Psi}}_1)$ and $(\overline{\sigma}_2, \overline{\boldsymbol{\Psi}}_2)$ with initial datum f_0.*

Then the following holds for any $t \geqslant 0$:

$$\|(f_1 - f_2)(t)\|^2_{L^2_{x,\omega}} \leqslant \frac{4e^{3C_2 t}\|f_0\|^2_{L^2_{x,\omega}}}{\inf\{\overline{\sigma_1} + \overline{\sigma_2}\}} \int_0^t \left\|\left(\overline{\boldsymbol{\Psi}}_1 - \overline{\boldsymbol{\Psi}}_2\right)(s)\right\|^2_{L^\infty_{x,\omega}} ds$$
$$+ \frac{2}{\inf\{\overline{\sigma_1} + \overline{\sigma_2}\}} \left[\int_0^t e^{C_2(t-s)}\left(\|\nabla_\omega f_1\|^2_{L^2_{x,\omega}} + \|\nabla_\omega f_2\|^2_{L^2_{x,\omega}}\right) ds\right] \|\overline{\sigma}_1 - \overline{\sigma}_2\|^2_{L^\infty_{t,x,\omega}},$$

where $C_2 = C_2(\overline{\sigma}_1, \overline{\boldsymbol{\Psi}}_1) + C_2(\overline{\sigma}_2, \overline{\boldsymbol{\Psi}}_2)$ as defined in Proposition 3.2*.*

Remark 3.4. The main difficulty with dealing with local-in-time $\sigma(f)$ appears in the second term on the right-hand side above. Indeed, from Proposition 3.2 we see that $\|\nabla_\omega f_i\|^2_{L^2_{x,\omega}}$ is integrable on $[0,t]$, and thus $\int_0^t \|\nabla_\omega f_i\|^2_{L^2_{x,\omega}}$ vanishes when t goes to 0, thus almost managing to close a contraction fixed point argument. As we shall see later on, however, we need an explicit independence over σ (other than $\inf \sigma$ and $\sup \sigma$) to avoid any nonlinear breakdown of contraction theorem in short times. As a result an $L^2([0,T], L^2_{x,\omega})$ setting is required to get an extra integration in time. This explains hypothesis (H2), but one could also ask for more regularity for σ and f_0 to explicitly estimate the convergence to 0 of $\int_0^t \|\nabla_\omega f_i\|^2_{L^2_{x,\omega}}$ with parabolic regularity.

Proof of Proposition 3.3*.* To shorten notation we will use $\sigma^+ = \overline{\sigma}_1 + \overline{\sigma}_2$, $\sigma^- = \overline{\sigma}_1 - \overline{\sigma}_2$, $\boldsymbol{\Psi}^+ = \overline{\boldsymbol{\Psi}}_1 + \overline{\boldsymbol{\Psi}}_2$, and $\boldsymbol{\Psi}^- = \overline{\boldsymbol{\Psi}}_1 - \overline{\boldsymbol{\Psi}}_2$. Also we denote the constant constructed from (3.4) by $C_2 = \max\{C_2(\overline{\sigma}_1, \overline{\boldsymbol{\Psi}}_1), C_2(\overline{\sigma}_2, \overline{\boldsymbol{\Psi}}_2), C_2(\sigma^+, \boldsymbol{\Psi}^+), C_2(\sigma^-, \boldsymbol{\Psi}^-)\}$.

Using the algebraic identity $ab - cd = \frac{1}{2}[(a-c)(b+d) + (a+c)(b-d)]$ we find

$$\partial_t [f_1 - f_2] + c\omega\nabla_x \cdot [f_1 - f_2]$$
$$= \frac{1}{2}\nabla_\omega \cdot \left((\overline{\sigma}_1 + \overline{\sigma}_2)\nabla_\omega [f_1 - f_2]\right) + \frac{1}{2}\nabla_\omega \cdot \left([f_1 - f_2]\left(\overline{\boldsymbol{\Psi}}_1 + \overline{\boldsymbol{\Psi}}_2\right)\right)$$
$$+ \frac{1}{2}\nabla_\omega \cdot \left[(\overline{\sigma}_1 - \overline{\sigma}_2)\nabla_\omega (f_1 + f_2)\right] + \frac{1}{2}\nabla_\omega \cdot \left[(f_1 + f_2)\left(\overline{\boldsymbol{\Psi}}_1 - \overline{\boldsymbol{\Psi}}_2\right)\right]. \tag{3.6}$$

We start by bounding the terms in $f_1 - f_2$ as in Proposition 3.2 and obtain

$$\frac{1}{2}\frac{d}{dt}\|f_1 - f_2\|^2_{L^2_{x,\omega}} \leqslant -\frac{\sigma_0^+}{4}\|\nabla_\omega (f_1 - f_2)\|^2_{L^2_{x,\omega}} + C_2\|f_1 - f_2\|^2_{L^2_{x,\omega}}$$
$$- \frac{1}{2}\int_{\Omega \times \mathbb{S}^{d-1}} (f_1 - f_2)\nabla_\omega \cdot \left[\sigma^- \nabla_\omega (f_1 + f_2) + \boldsymbol{\Psi}^- (f_1 + f_2)\right] dx d\omega,$$

where $\overline{\sigma}_0^+ = \inf\{\overline{\sigma}_1 + \overline{\sigma}_2\}$. We bound the two terms inside the integral on the righthand side. First, thanks to an integration by parts (3.1),

$$\int_{\Omega\times\mathbb{S}^{d-1}} (f_1 - f_2)\nabla_\omega \cdot \bigl[\overline{\sigma}^- \nabla_\omega (f_1+f_2)\bigr]\, dx\,d\omega = -\int_{\Omega\times\mathbb{S}^{d-1}} \overline{\sigma}^- \nabla_\omega (f_1 - f_2)\cdot\nabla_\omega (f_1+f_2)\, dx\,d\omega.$$

The Cauchy--Schwarz and Young inequalities yield

$$\left|\int_{\Omega\times\mathbb{S}^{d-1}} (f_1 - f_2)\nabla_\omega \cdot \left[\overline{\sigma}^- \nabla_\omega (f_1+f_2)\right]\right| \leqslant \frac{\overline{\sigma}_0^+}{8}\left\|\nabla_\omega(f_1-f_2)\right\|^2_{L^2_{x,\omega}} + \frac{4\left\|\overline{\sigma}^-\right\|^2_{L^\infty_{x,\omega}}}{\overline{\sigma}_0^+}\left\|\nabla_\omega(f_1+f_2)\right\|^2_{L^2_{x,\omega}}. \tag{3.7}$$

The second integrand in (3.6) is dealt with in the same way, and we get

$$\left|\int_{\Omega\times\mathbb{S}^{d-1}} (f_1 - f_2)\nabla_\omega \cdot \left[\overline{\mathbf{\Psi}}^- (f_1+f_2)\right]\right| \leqslant \frac{\overline{\sigma}_0^+}{8}\left\|\nabla_\omega(f_1-f_2)\right\|^2_{L^2_{x,\omega}} + \frac{4\left\|\overline{\mathbf{\Psi}}^-\right\|^2_{L^\infty_{x,\omega}}}{\overline{\sigma}_0^+}\left\|(f_1+f_2)\right\|^2_{L^2_{x,\omega}}. \tag{3.8}$$

We then plug (3.7) and (3.8) into (3.6) and get

$$\frac{1}{2}\frac{d}{dt}\left\|f_1-f_2\right\|^2_{L^2_{x,\omega}} \leqslant C_2\left\|f_1-f_2\right\|^2_{L^2_{x,\omega}} + \frac{2\left\|\overline{\sigma}^-\right\|^2_{L^\infty_{x,\omega}}}{\overline{\sigma}_0^+}\left\|\nabla_\omega(f_1+f_2)\right\|^2_{L^2_{x,\omega}} + \frac{2\left\|\overline{\mathbf{\Psi}}^-\right\|^2_{L^\infty_{x,\omega}}}{\overline{\sigma}_0^+}\left\|(f_1+f_2)\right\|^2_{L^2_{x,\omega}},$$

on which we apply Grönwall's lemma to obtain

$$\left\|f_1-f_2\right\|^2_{L^2_{x,\omega}} \leqslant \frac{2}{\overline{\sigma}_0^+}\left\|\overline{\sigma}^-\right\|^2_{L^\infty_{t,x,\omega}}\int_0^t e^{C_2(t-s)}\left\|\nabla_\omega(f_1+f_2)\right\|^2_{L^2_{x,\omega}}\,ds + \frac{2}{\overline{\sigma}_0^+}\int_0^t e^{C_2(t-s)}\left\|(f_1+f_2)\right\|^2_{L^2_{x,\omega}}\left\|\overline{\mathbf{\Psi}}^-(s)\right\|^2_{L^\infty_{x,\omega}}\,ds. \tag{3.9}$$

To conclude we apply Proposition 3.2 to bound the second term on the right-hand side, which directly gives the expected result. ∎

3.1.3. Estimation of vanishing time for velocity averaging densities. Now we turn to an explicit estimation of the vanishing time of velocity averaging densities.

Proposition 3.5. *Let p belong to* $[2,+\infty]$ *and* $f_0 \geqslant 0$ *be in* $L^1_{x,\omega} \cap L^p_{x,\omega}$. *Let also* $\overline{\sigma}$ *and* $\overline{\mathbf{\Psi}}$ *satisfy the assumptions of Theorem* 3.1 *supplemented with* $\overline{\sigma}$ *Lipschitz. Let* f *be the solution to* (3.3) *associated to* f_0. *For any* $\mathbf{K}$ *in* $W^{1,\infty}_{t,x_*}W^{2,\infty}_{\omega_*}L^\infty_{x,\omega}$ *denote*

\int

$$\boldsymbol{J}(t,x,\omega) = (J_i(t,x,\omega))_{1\leqslant i\leqslant d} = \int_{\Omega\times\mathbb{S}^{d-1}} f(t,x_\ast,\omega_\ast)\mathbf{K}(t,x,x_\ast,\omega,\omega_\ast)\,dx_\ast\, d\omega_\ast.$$

Then the following holds:

$$\forall (t,x,\omega) \in \mathbb{R}^+\times \Omega \times \mathbb{S}^{d-1}, \quad |\boldsymbol{J}(t,x,\omega)| \geqslant |\mathbf{J}(0,x,\omega)| - K_\infty \|f_0\|_{L^1_{x,\omega}} t,$$

where we defined

$$K_\infty = \sqrt{\sum_{i=1}^d K_{i,\infty}^2} \tag{3.10}$$

with

$$K_{i,\infty} = \left(1+|c|+\|\overline{\sigma}\|_{L^\infty_{t,x}W^{1,\infty}_\omega} + \left\|\overline{\mathbf{\Psi}}\right\|_{L^\infty_{t,x,\omega}}\right)\|K_i\|_{W^{1,\infty}_{t,x_*}W^{2,\infty}_{\omega_*}L^\infty_{x,\omega}}. \tag{3.11}$$

Proof of Proposition 3.5. Since f_0 is nonnegative and belongs to $L^1_{x,\omega} \cap L^p_{x,\omega}$ we deduce from Theorem 3.1 and Proposition 3.2 that $f(t)$ is nonnegative and belongs to $L^1_{x,\omega} \cap L^p_{x,\omega}$ for all time. Hence, we can multiply (3.3) by $\mathbf{K}(x,x_\ast,\omega,\omega_\ast)$ and integrate by parts (3.1). Here again note that there are no additional terms along $\omega_\ast$ since $\overline{\mathbf{\Psi}}$ and $\nabla_{\omega_\ast}$ are both orthogonal to $\omega_\ast$. This yields

$$\begin{aligned}
\frac{1}{2}\frac{d}{dt}|J_i(t,x,\omega)|^2 &= J_i(t,x,\omega)\int_{\Omega\times\mathbb{S}^{d-1}} K_i(t,x,x_*,\omega,\omega_*)\frac{\partial f}{\partial t}(t,x_*,\omega_*)dx_*d\omega_* \\
&\quad + J_i(t,x,\omega)\int_{\Omega\times\mathbb{S}^{d-1}} f(t,x_*,\omega_*)\frac{\partial K_i}{\partial t}(t,x,x_*,\omega,\omega_*)dx_*d\omega_* \\
&= J_i(t,x,\omega)\int_{\Omega\times\mathbb{S}^{d-1}} f\left[\frac{\partial K_i}{\partial t} + c\omega\cdot\nabla_{x_*}K_i + \nabla_{\omega_*}\cdot(\overline{\sigma}\nabla_{\omega_*}K_i) - \overline{\mathbf{\Psi}}\cdot\nabla_{\omega_*}K_i\right]dx_*d\omega_* \\
&\geqslant -|J_i(t,x,\omega)|\left(1+|c|+\|\overline{\sigma}\|_{L^\infty_{t,x}W^{1,\infty}_\omega}+\left\|\overline{\mathbf{\Psi}}\right\|_{L^\infty_{t,x,\omega}}\right)\|K_i\|_{W^{1,\infty}_{t,x_*}W^{2,\infty}_{\omega_*}L^\infty_{x,\omega}}\|f(t)\|_{L^1_{x,\omega}} \\
&\geqslant -K_{i,\infty}\|f_0\|_{L^1_{x,\omega}}|J_i(t,x,\omega)|,
\end{aligned}$$

where we used conservation of mass along time. Summing over i we obtain

$$\begin{aligned}
\frac{1}{2}\frac{d}{dt}|\mathbf{J}(t,x,\omega)|^2 &\geqslant -\|f_0\|_{L^1_{x,\omega}}\sum_{i=1}^d K_{i,\infty}|J_i(t,x,\omega)| \\
&\geqslant -\|f_0\|_{L^1_{x,\omega}}\sqrt{\sum_{i=1}^d K_{i,\infty}^2}\,|\mathbf{J}|.
\end{aligned}$$

The latter implies that

$$|\mathbf{J}(t,x,\omega)| \geqslant |\mathbf{J}(0,x,\omega)| - \sqrt{\sum_{i=1}^{d} K_{i,\infty}^2}\, \|f_0\|_{L^1_{x,\omega}}\, t,$$

which concludes the proof. ∎

3.2. The local-in-time nonlinear Cauchy theory. The present section is devoted to the proof of Theorem 2.4 thanks to the linear study we presented in section 3.1. We shall prove existence, uniqueness, and the global existence criterion to

$$\partial_t f + c\,\omega\cdot\nabla_x f = \nabla_\omega\cdot(\sigma(f)\nabla_\omega f) + \nabla_\omega\cdot(\nu(f) f \mathbf{P}_{\omega^\bot}[\mathbf{\Psi}[f]]). \tag{3.12}$$

Proof of Theorem 2.4. We fix p in $[2,+\infty]$ and $f_0 \geqslant 0$ in $L^1_{x,\omega}\cap L^p_{x,\omega}$. We recall some assumptions of Theorem 2.4:

$$\int_{\Omega\times\mathbb{S}^{d-1}} f_0(x,\omega)\,dx\,d\omega = M_0 > 0, \quad \inf_{(x,\omega)\in\Omega\times\mathbb{S}^{d-1}} |\boldsymbol{J}[f_0](x,\omega)|_\alpha = J_0 > 0.$$

The strategy is to apply a contraction fix point argument so we start by defining a complete metric space on which we shall work. For any $M \geqslant M_0$ and $T > 0$ we define

$$\Gamma_T^M = \left\{ f \in L^2_{[0,T],x,\omega},\ \operatorname*{esssup}_{[0,T]} \|f(t)\|_{L^1_{x,\omega}} \leqslant M \ \text{ and } \operatorname*{esssup}_{[0,T]\times\Omega\times\mathbb{S}^{d-1}} \frac{|\nu(f(t))|}{|\boldsymbol{J}[f(t)](x,\omega)|_\alpha} \leqslant \frac{\nu_\infty}{\alpha+(1-\alpha)\frac{J_0}{2}} \right\}. \tag{3.13}$$

Note that since $J[f]$ is an integral over $(x_\ast,\omega_\ast)$ against $\mathbf{K}$ and because $\mathbf{K}$ is $L^2_{x_*,\omega_*}$ (by $(H1)$), a Cauchy sequence for the L^2-norm in Γ_T^M indeed converges in Γ_T^M.

For a function g in Γ_T^M we have that g belongs to $L^2_{[0,T],x,\omega}$, so by Fubini's theorem g belongs to $L^2([0,T],L^2_{x,\omega})$ and for almost every t in $[0,T]$ the function $g(t,\cdot,\cdot)$ belongs to $L^2_{x,\omega}$. Therefore $\sigma(g(t))$ and $\nu(g(t))$ are defined almost everywhere in $[0,T]$, and with hypotheses (H2) and (H3) one has

$$\sigma_0 \leqslant \sigma(g(t)) \leqslant \sigma_\infty \quad \text{and} \quad \|\nu(g(t))\|_{L^\infty_x W^{1,\infty}_\omega} \leqslant \nu_\infty \text{ almost everywhere in } [0,T].$$

Defining

Ψ

$$\left|\overline{\boldsymbol{\Psi}}(t,x,\omega)\right| \leqslant \nu_\infty \frac{\left|\int_{\Omega\times\mathbb{S}^{d-1}} g(t,x_*,\omega_*)\mathbf{K}(t,x,x_*,\omega,\omega_*)dx_*d\omega_*\right|}{|\boldsymbol{J}[g](t,x,\omega)|_\alpha}$$

$$(t,x,\omega) = \nu(g(t))(x,\omega)\mathbf{P}_{\omega^\bot}[\boldsymbol{\Psi}[g]],$$

hypothesis (H1) shows that for $t \in [0,T]$

$$\leqslant \nu_\infty \begin{cases} \|\mathbf{K}\|_{L^\infty_{t,x,x_*,\omega,\omega_*}} M & \text{if } \alpha = 1 \\ \frac{1}{1-\alpha}, & \text{if } \alpha \neq 1 \end{cases} := C_M^1, \tag{3.14}$$

and

$$\left|\nabla_\omega \cdot \overline{\boldsymbol{\Psi}}(t,x,\omega)\right| \leqslant \left|\nabla_\omega \nu(g)\cdot \mathbf{P}_{\omega^\perp}[\boldsymbol{\Psi}[g]]\right| + \left|\nu(g)\nabla_\omega \cdot \mathbf{P}_{\omega^\perp}[\boldsymbol{\Psi}[g]]\right|$$

$$\leqslant C_M^1 + \nu_\infty C_\alpha M \|\mathbf{K}\|_{L^\infty_{t,x,x_*,\omega_*} W^{1,\infty}_\omega} := C_M^2, \tag{3.15}$$

where $C_\alpha > 0$ only depends on α and equals 1 when $\alpha = 1$. Therefore thanks to Theorem 3.1 we are able to define

$$\begin{aligned} \Phi : \Gamma_T^M &\longrightarrow L^2_{[0,T],x,\omega} \\ g &\longmapsto \Phi(g) = f_g, \end{aligned}$$

where f_g is the solution on $[0,T] \times \Omega \times \mathbb{S}^{d-1}$ to the linear equation

$$\partial_t f + c\omega \cdot \nabla_x f = \nabla_\omega \cdot \left(\sigma(g)(t,x,\omega)\nabla_\omega f + f\nu(g)\mathbf{P}_{\omega^\bot}[\boldsymbol{\Psi}[g]]\right)$$

associated to the initial datum f_0.

We now show that for a specific T, Φ is in fact a contraction from Γ^M_T to Γ^M_T.

Due to Proposition 3.2 and since f_0 belongs to $L^2_{x,\omega}$ we see that $\Phi(g)$ belongs to $L^\infty([0,T])L^2_{x,\omega}) \subset L^2_{[0,T]\times\Omega\times\mathbb{S}^{d-1}}$ and also that for almost every $t \in [0,T]$

$$\|\Phi(g)(t)\|_{L^1_{x,\omega}} = M_0 < M.$$

Moreover, Proposition 3.5 gives that almost everywhere

$$|\boldsymbol{J}[\Phi(g)](t,x,\omega)| \geqslant J_0 - K_\infty M_0 T. \tag{3.16}$$

We defined K_∞ in (3.10) and recall it here:

$$K_\infty(f_0) = \left(1 + |c| + \sigma_\infty + \sigma_{lip} + C_M^1\right)\sqrt{\sum_{n=1}^d \|K_i\|^2_{W^{1,\infty}_{t,x_*}W^{2,\infty}_{\omega_*}L^\infty_{x,\omega}}}. \tag{3.17}$$

Therefore if we choose

$$T < T_0 := \frac{J_0}{2K_\infty M_0}, \tag{3.18}$$

the lower bound (3.16) implies

$$\forall t \in [0,T], \quad |\boldsymbol{J}[\Phi(g)](t,x,\omega)|_\alpha \geqslant \alpha + (1-\alpha)\frac{J_0}{2},$$

so that

$$\forall t \in [0,T], \quad \frac{\left|\,(\,(g))\right|}{\left|\boldsymbol{J}[\,(g)](t,x,\omega)\right|_{\,}} \leqslant \frac{\qquad \nu_\infty}{\ +(1 - \ )^{\frac{L_0}{2}}},$$

and we thus proved that for any $0 < T < T_0$, Φ maps Γ^{M}_{T} to itself.

It remains to prove that Φ is a contraction on Γ^{M}_{T} for T sufficiently small. By hypothesis (H2) we have

$$\|\sigma(g_1) - \sigma(g_2)\|_{L^\infty_{[0,T],x,\omega}} \leqslant \sigma_{\mathrm{lip}} \|g_1 - g_2\|_{L^2_{[0,T],x,\omega}}.$$

Also, thanks to the Lipschitz property (H3) of the friction ν and the kernel form of $\boldsymbol{\Psi}[g]$, one infers

$$\int_0^t \left\|\overline{\boldsymbol{\Psi}}(g_1) - \overline{\boldsymbol{\Psi}}(g_2)\right\|^2_{L^\infty_{x,\omega}} \leqslant \Psi_{\mathrm{lip}} \|g_1 - g_2\|^2_{L^2_{[0,t],x,\omega}},$$

where $\Psi_{\mathrm{lip}} > 0$ is a constant depending only on $\|\mathbf{K}\|_{L^\infty_{t,x,x_\ast,\omega,\omega_\ast}}$, M, α, Ψ_0, and ν_{lip}.

This inequality comes from the algebraic identity $ab - cd = \frac{(a-c)(b+d)}{2} + \frac{(a+c)(b-d)}{2}$.

We apply Proposition 3.3 to see that for any g_1 and g_2 in Γ^{M}_{T}

$$\|\Phi(g_1) - \Phi(g_2)\|^2_{L^2_{[0,T],x,\omega}} \leqslant \Lambda(T) \|g_1 - g_2\|^2_{L^2_{[0,T],x,\omega}},$$

where

$$\Lambda(T) = \frac{2\|f_0\|^2_{L^2_{x,\omega}}}{\sigma_0^2} T e^{3C_2T} \left(\Psi_{\mathrm{lip}}\sigma_0 + 2\sigma_{\mathrm{lip}}\right). \tag{3.19}$$

Note that we used Proposition 3.2 to obtain explicit constants. To conclude, it suffices to choose

$$T < \min\left\{T_0,\ \sup_{t>0} \Lambda(t) \leqslant \frac{1}{4}\right\} := T_1 \tag{3.20}$$

because then for any g_1 and g_2 in Γ^{M}_{T} we have proved

$$\|\Phi(g_1) - \Phi(g_2)\|_{L^2_{[0,T],x,\omega}} \leqslant \frac{1}{2}\|g_1 - g_2\|_{L^2_{[0,T],x,\omega}},$$

which implies that Φ is a contraction on Γ^{M}_{T}.

We thus proved the local existence and uniqueness of a solution f to the nonlinear kinetic equation (2.4) in $L^2_{[0,T_{\max}),x,\omega}$ with $T_{\mathrm{max}} \geqslant T_1$. The solution f belongs to $L^\infty\left([0,T_{\max}), L^p_{x,\omega}\right) \cap L^2\left([0,T_{\max}), L^2_x H^1_\omega\right)$ due to Proposition 3.2 since $\sigma(f)$ and $\nu(f)$ are well-defined and satisfy the required assumptions, as we saw above. At last, thanks to the global-in-time Cauchy theory for the linear equation given by Theorem 3.1 we see that if $\lim_{t\to T^-_{\max}} \sup_{\Omega\times\mathbb{S}^{d-1}} \frac{\nu(f)}{|\boldsymbol{J}[f]|_\alpha} < +\infty$, then we can apply our fixed point argument starting at T_{max}, and therefore T_{max} must be $+\infty$. This concludes the proof of Theorem 2.4. ∎

3.3. Global existence and decay of the free energy. The present section focuses on some important cases where one is ensured to have a globally defined solution to the fully nonlinear kinetic equation (2.4). Then we get a look into the free energy and its energy dissipation.

3.3.1. Important examples of global solutions. Theorem 2.4 offers a direct and easy way to check whether the solutions obtained are globally defined. Indeed, under the assumptions (H1)-(H2)-(H3), if one shows that $\inf_{\Omega\times\mathbb{S}^{d-1}} |\boldsymbol{J}[f]|_\alpha$ cannot vanish along the flow, then indeed $\sup_{\Omega\times\mathbb{S}^{d-1}} \frac{|\nu(f)|}{|\boldsymbol{J}[f]|_\alpha}$ cannot explode in finite time.

The nonnormalized case: $\boldsymbol{\alpha} \neq \mathbf{0}$. The discussion above implies that when $\alpha \neq 0$, then $|\boldsymbol{J}[f]|_\alpha \geqslant \frac{1}{\alpha} > 0$ for any function f so that $\boldsymbol{\Psi}[f]$ is well-defined. Therefore for any $\alpha \neq 0$ the solutions constructed in Theorem 2.4 are global in time. This is the case of the equations looked at in [43, 26, 32], among others.

The nonasymptotic normalized case: $\boldsymbol{\nu}/|\boldsymbol{J}[\boldsymbol{f}]|$ prevents any explosion. Here we are interested in the normalized framework $\alpha = 0$, but when the coefficient $\nu(f)$ compensates for a possible vanishing of $|\boldsymbol{J}[f]|$, namely, if we add the assumption that $f \mapsto \frac{\nu(f)}{|\boldsymbol{J}[f]|}$ is a bounded function when $|\boldsymbol{J}[f]|$ tends to zero.

Again, in this particular case the solutions constructed in Theorem 2.4 are global in time. This is the case of equations looked at in [14, 16] where $\nu(f) = \nu(|\mathbf{J}_f|)$ with $\frac{\nu(y)}{y}$ bounded and is generalized here, for instance, to any $\nu(f) = \nu(|\boldsymbol{J}[f]|)$ with $\frac{\nu(y)}{y}$ bounded near zero.

The normalized case of a kernel with one coordinate with a strict sign. In this paragraph we assume that the kernel $\mathbf{K}(t,x,x_\ast,\omega,\omega_\ast)$ is such that one of its coordinates K_i has a strict sign in the following sense: there exists a positive constant $k_i > 0$ such that

$$\inf_{\mathbb{R}^+\times\Omega\times\Omega\times\mathbb{S}^{d-1}\times\mathbb{S}^{d-1}} K_i \geqslant k_i \quad \text{or} \quad \sup_{\mathbb{R}^+\times\Omega\times\Omega\times\mathbb{S}^{d-1}\times\mathbb{S}^{d-1}} K_i \leqslant -k_i.$$

It is a direct verification since, thanks to the positivity of f, either

$$\int_{\Omega\times\mathbb{S}^{d-1}} \int f K_i \, dx_\ast \, d\omega_\ast \geqslant k_i M_0 \quad \text{or} \quad \int_{\Omega\times\mathbb{S}^{d-1}} f K_i \, dx_\ast \, d\omega_\ast \leqslant -k_i M_0.$$

Therefore, $J_i(t,x,\omega)$ cannot vanish in finite time and again, in this particular case, the solutions constructed in Theorem 2.4 are global in time.

The Vicsek kernel operator. We study the special case when $\mathbf{K}(t,x,x_\ast,\omega,\omega_\ast) = \omega_\ast$ and normalized $\alpha = 0$, as in [14, 16, 33, 30] for the spatially homogeneous equation. Note that here we also obtain the global-in-time solution for the nonspatially homogeneous setting. We also assume that ν and σ

do not depend on x or ω and that the friction ν is negative (standard in all previous studies).

Contrary to the setting above, the present kernel can degenerate in the sense that one nonzero coordinate of $\mathbf{K}$ can possibly vanish later on but gives rise to another nonzero coordinate. As a consequence, the full $|\mathbf{K}|$ will not vanish in finite time. First we explicitly write

$$J_i(t,x,\omega) = J_i(t) = \int_{\Omega\times\mathbb{S}^{d-1}} \omega_{\ast i} f(t,x_\ast,\omega_\ast)\, dx_\ast\, d\omega_\ast \quad \text{and} \quad \mathbf{\Psi} = \frac{\boldsymbol{J}}{|\boldsymbol{J}|}.$$

As before let us consider the computation below for $0 \leqslant t < T_{\mathrm{max}}$, which in the present setting yields

$$\frac{1}{2}\frac{d}{dt}|\boldsymbol{J}|^2 = \sum_{i=1}^{d} J_i(t)\int_{\Omega\times\mathbb{S}^{d-1}} f\left[\sigma(f)(t)\Delta_{\omega_*}(\omega_{*i}) - \nu(f)(t)\mathbf{P}_{\omega_*^\perp}[\mathbf{\Psi}]\cdot\nabla_{\omega_*}(\omega_{*i})\right].$$

Thanks to the properties of the tangential derivatives we can compute directly with (3.2)

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}|\boldsymbol{J}(t)|^2 &= -(d-1)\sigma(f)(t)\,|\boldsymbol{J}(t)|^2 - \nu(f)(t)\int_{\Omega\times\mathbb{S}^{d-1}}\frac{1}{|\boldsymbol{J}(t)|}\mathbf{P}_{\omega_*^\perp}[\boldsymbol{J}(t)]\cdot\nabla_{\omega_*}(\omega_*\cdot\boldsymbol{J}(t))\\ &= -(d-1)\sigma(f)(t)\,|\boldsymbol{J}(t)|^2 - \frac{\nu(f)(t)}{|\boldsymbol{J}(t)|}\int_{\Omega\times\mathbb{S}^{d-1}}|\nabla_{\omega_*}(\omega_*\cdot\boldsymbol{J}(t))|^2\\ &\geqslant -(d-1)\sigma_0|\boldsymbol{J}(t)|^2.\end{aligned}$$

Note that we use the fact that the friction is negative. Using Grönwall's inequality we conclude that $|\boldsymbol{J}(t)|^2 \geqslant |\boldsymbol{J}(0)|^2 e^{-2(d-1)\sigma_0 t}$.

These computations were already done in a homogeneous regularized setting in [30], and our local theory allows us to carry them out directly. Such a lower bound thus implies the nonvanishing of $|\boldsymbol{J}(t)|$ in finite time, thus leading again to a global-in-time solution in this setting.

3.3.2. Decay of the free energy. We recall our definitions for the free energy and the energy dissipation (2.12)--(2.13):

$$\begin{aligned}\mathcal{F}[f](t) &= \int_{\Omega\times\mathbb{S}^{d-1}} f\ln(f)\,dx d\omega\\ &\quad + \int_0^t\int_{\Omega\times\mathbb{S}^{d-1}} f\left[\nabla_\omega\cdot(\nu(f)\mathbf{P}_{\omega^\perp}[\mathbf{\Psi}[f]]) - \frac{\nu(f)^2}{\sigma(f)}\left|\mathbf{P}_{\omega^\perp}[\mathbf{\Psi}[f]]\right|^2\right] dx d\omega ds,\\ \mathcal{D}[f](t) &= \int_{\Omega\times\mathbb{S}^{d-1}}\sigma(f)f\left|\nabla_\omega\ln(f) + \frac{\nu(f)}{\sigma(f)}\mathbf{P}_{\omega^\perp}[\mathbf{\Psi}[f]]\right|^2 dx d\omega.\end{aligned}$$

Let f_0, σ, ν, and $\mathbf{K}$ be as in Theorem 2.4. We now establish the decay of $\mathcal{F}[f](t)$:

$$\forall t \geqslant 0, \quad \frac{d}{dt}\mathcal{F}[f](t) = -\mathcal{D}[f](t). \tag{3.21}$$

It comes from direct computations. Indeed since $\nabla_{x/\omega}\left[1+\ln f\right] = \frac{\nabla_{x/\omega} f}{f}$ it follows by integrating by parts formula (3.1),

$$\begin{aligned}
\frac{d}{dt}\int_{\Omega\times\mathbb{S}^{d-1}} f\ln f(t,x,\omega)dxd\omega &= \int_{\Omega\times\mathbb{R}^d}\left[1+\ln f\right]\partial_t f(t,x,\omega)dxd\omega \\
&= \int_{\Omega\times\mathbb{S}^{d-1}} c\omega\cdot\nabla_x f\, dxd\omega \\
&\quad + \int_{\Omega\times\mathbb{S}^{d-1}} \left[\sigma(f)\nabla_\omega f + \nu(f) f\mathbf{P}_{\omega^\perp}[\mathbf{\Psi}[f]]\right]\cdot\frac{\nabla_\omega f}{f} dxd\omega \\
&= -\int_{\Omega\times\mathbb{S}^{d-1}} \left[\sigma(f)\frac{\left|\nabla_\omega f\right|^2}{f} + \nu(f)\mathbf{P}_{\omega^\perp}[\mathbf{\Psi}[f]]\cdot\nabla_\omega f\right] dxd\omega.
\end{aligned}$$

Hence,

$$\begin{aligned}
&\frac{d}{dt}\mathcal{F}[f](t) \\
&= -\int_{\Omega\times\mathbb{S}^{d-1}}\left[\sigma(f)\frac{\left|\nabla_\omega f\right|^2}{f} + 2\nu(f)\mathbf{P}_{\omega^\perp}[\mathbf{\Psi}[f]]\cdot\nabla_\omega f + \frac{\nu(f)^2}{\sigma(f)}\left|\mathbf{P}_{\omega^\perp}[\mathbf{\Psi}[f]]\right|^2 f\right] dxd\omega \\
&= -\mathcal{D}[f](t),
\end{aligned}$$

as claimed above.

4. Mean-field limit: Proofs of the results in section 2.3.

Let us first describe the strategy we are about to implement. At the core, we will use the coupling approach popularized by Sznitman [47] (which differs from classical BBGKY approaches [36]). We will prove Theorems 2.10 and 2.11 and follow the same steps as in [3] and the results in [5].

4.1. Preliminaries: Functional framework and notation.

In this section we will work in the space $\mathscr{H}$ defined in section 2.1 and with the 2-Wasserstein distance also defined there. We will need some results on Wasserstein distances in what follows, so we summarize them next (these results and proofs can be found in [49]). In our setting, the Wasserstein distance of order 1 is given by

$$W_1(m,p) := \inf \Biggl\{ \Biggl[\int_{(\Omega\times\mathbb{R}^d)^2} |z-u|\,\pi(dz,du) \Biggr]; \ \pi\in\mathscr{P}(\Omega\times\mathbb{R}^d)^2 \text{ with marginals } m \text{ and } p \Biggr\}.$$

This distance can be expressed in duality form using the Kantorovich--Rubinstein distance

$$W_1(m,p) = \sup_{\|\varphi\|_{lip}\leqslant 1}\left\{\int_{\Omega\times\mathbb{R}^d}\varphi(z)dm - \int_{\Omega\times\mathbb{R}^d}\varphi(z)dp\right\}. \tag{4.1}$$

Also, by Hölder's inequality, it holds that

$$W_1 \leqslant W_2. \tag{4.2}$$

4.2. The nonsingular dynamics. To prove Theorems 2.10 and 2.11, we will consider first modified versions of systems (2.17) and (2.21) where we work with Lipschitz coefficients. Being more precise, first we construct $\gamma : \mathbb{R}^d \rightarrow \mathbb{R}^d$ Lipschitz and bounded with

$$\gamma(v) = v, \quad \text{when } |v| \leqslant 2. \tag{4.3}$$

We will use the function γ as a substitute of the variable $V(t)$ in some parts of the equations. We do this to be able to apply known results of existence and uniqueness of solutions for stochastic differential equations that require the coefficients to be Lipschitz and bounded (see Theorem A.2). We will prove in a second stage that along the dynamics it holds that $|V(t)| = 1$; therefore $\gamma(V(t)) = V(t)$ and we recover the terms in the original equation. Second, we define a functional τ_0 as

$$\tau_0(x,v,m) = \tau_{\varepsilon_0}(x,\gamma(v),m). \tag{4.4}$$

We will show that τ_0 is Lipschitz and bounded in the whole $\mathscr{H}$.

Lemma 4.1. *The functions $r = r(x,v,m) := |\mathbf{J}|(x,\gamma(v),m) : \mathscr{H} \rightarrow \mathbb{R}_+$ and $\tau_0 = \tau_0(x,v,m) : \mathscr{H} \rightarrow \mathbb{R}_+$ are Lipschitz and bounded.*

Proof of Lemma 4.1. We notice, first, that since γ is bounded, then $K(x,\gamma(v),m)$ is bounded in all the variables and, therefore, $r_f(z)$ is bounded in $\mathscr{H}$. Now, letting $f,g \in \mathscr{P}_2(\Omega \times \mathbb{R}^d)$ and fixing $z \in \Omega \times \mathbb{R}^d$, we show next that $r_f := r(z,f)$ is Lipschitz in f. Consider $z,z' \in \Omega \times \mathbb{R}^d$ and for $z = (x,v)$ denote $\widetilde{z} = (x,\gamma(v))$. Now, we consider (the integrals are in $\Omega \times \mathbb{R}^d$)

$$\begin{aligned}
|r_f(z) - r_g(z')| &= \left| \left| \int K(\widetilde{z},y) f(dy) \right| - \left| \int K(\widetilde{z}',y)\, g(dy) \right| \right| \\
&\leqslant \left| \int [K(\widetilde{z},y) f(dy) - K(\widetilde{z}',y) g(dy)] \right| \\
&\leqslant \|K\|_{Lip} |\widetilde{z} - \widetilde{z}'| + \left| \int K(\widetilde{z}',y)[f(dy) - g(dy)] \right| \\
&\leqslant \|K\|_{Lip}(\|\gamma\|_{Lip} + 1)|z - z'| + \|K\|_{Lip} \sup_{\|\varphi\|_{Lip} \leqslant 1} \bigg| \int \varphi(y) f(dy) - \int \varphi(y) g(dy) \bigg| \\
&= \|K\|_{Lip}(\|\gamma\|_{Lip} + 1)|z - z'| + \|K\|_{Lip} W_1(f,g) \\
&\leqslant \|K\|_{Lip}(\|\gamma\|_{Lip} + 1)|z - z'| + \|K\|_{Lip} W_2(f,g),
\end{aligned}$$

where in the second line we used the reverse triangle inequality; in the third line we have used that $\int_{\mathbb{R}^{2d}} f = 1$ and that K is Lipschitz; in the fourth line we used that γ is Lipschitz and chose an arbitrary $z_0 \in \mathbb{R}^{2d}$; the fifth line is given by (4.1); and the last inequality follows from (4.2). Therefore, we conclude that $r = r_f(z)$ is Lipschitz in $\mathcal{H}$.

Next, we also show that τ_0 is Lipschitz and bounded in $\mathcal{H}$. First, whenever $r \geqslant \varepsilon_0$ we have that $(\alpha + (1-\alpha)r)^{-1}$ is also Lipschitz and bounded for all $\alpha \in [0,1]$; therefore, τ_0 is the product of Lipschitz and bounded functions, and hence is Lipschitz and bounded. ∎

With the function τ_0 we define the *nonsingular particle dynamics* as

$$\left\{ \begin{aligned} dX_t^{i,N} &= V_t^{i,N}\,dt, && \text{(4.5a)} \\ dV_t^{i,N} &= \nu(\mu^N)\mathbf{P}_{(V^{i,N})^\perp}(\tau_0(X^{i,N},V^{i,N},\mu^N))\,dt \\ &\quad + \frac{1}{2}\mathbf{P}_{(\overline{V}^{i,N})^\perp}[(\nabla_v\sigma(\mu^N))(\overline{X}^{i,N},\overline{V}^{i,N})]\,dt \\ &\quad + \sqrt{2\sigma(\mu^N)}\,\mathbf{P}_{(V^{i,N})^\perp}\circ dB_t^i && \text{(4.5b)} \end{aligned} \right.$$

(this is exactly system (2.17) substituting τ_0 by the function τ_{ε_0}), and the *nonsingular auxiliary process* is given by

$$\left\{ \begin{aligned} d\overline{X}_t^{i,N} &= \overline{V}_t^{i,N}\,dt, && \text{(4.6a)} \\ d\overline{V}^{i,N} &= \nu(f)\,\mathbf{P}_{(\overline{V}^{i,N})^\perp}(\tau_0(\overline{X}^{i,N},\overline{V}^{i,N},f))\,dt \\ &\quad + \frac{1}{2}\mathbf{P}_{(\overline{V}^{i,N})^\perp}[(\nabla_v\sigma(f))(\overline{X}^{i,N},\overline{V}^{i,N})]\,dt \\ &\quad + \sqrt{2\sigma(f)}\,\mathbf{P}_{(\overline{V}^{i,N})^\perp}\circ dB_t^i, && \text{(4.6b)} \\ f_t &= \mathrm{law}(\overline{X}^{i,N},\overline{V}^{i,N}) && \text{(4.6c)} \end{aligned} \right.$$

(again this is exactly system (2.21) after substituting τ_0 by τ_{ε_0}). Analogously, we also consider the *nonsingular kinetic equation* in $\Omega \times \mathbb{S}^{d-1}$ given by

$$\partial_t f + \omega \cdot \nabla_x f = \nabla_\omega \cdot (\sigma(f)\nabla_\omega f) + \nabla_\omega \cdot (\nu(f) f \nabla_\omega \tau_0(f)). \tag{4.7}$$

Note that under the assumptions of Theorem 2.4, (4.7) has global existence and uniqueness of solutions in the spaces stated in Theorem 2.4.

Remark 4.2. Notice that we are assuming that ν, σ, and $\nabla_v \sigma$ are bounded and Lipschitz for all $v \in \mathbb{R}^d$ rather than $v \in \mathbb{S}^{d-1}$ (see hypothesis (H4)). However, we just need ν, σ, and $\nabla_v\sigma$ to be Lipschitz and bounded in a neighborhood of $|v| = 1$, as we can use regularizing arguments as the one done to define the functional τ_0.

Proposition 4.3. *Theorems* 2.10 *and* 2.11 *hold by replacing in the statements system* (2.17) *by system* (4.5)*, system* (2.21) *by system* (4.6)*, and the kinetic equation* (2.4) *by the kinetic equation* (4.7)*.*

We prove this Proposition in section 4.4. The proof of Theorems 2.10 and 2.11 is direct, assuming Proposition 4.3 holds true.

Proof of Theorems 2.10 *and* 2.11 *assuming Proposition* 4.3*.* The result is direct by Lemma 2.9, since it implies that the nonsingular dynamics and the approximated dynamics coincide. ∎

4.3. Proof of Theorem 2.14. Let f_t be a solution to the Vicsek kinetic equation (2.4). By the proof of Theorem 2.4 we know that for $T < T_1$ (with T_1 given in (3.20)) it holds that

$$\forall (t,x,\omega) \in [0,T]\times\Omega\times\mathbb{S}^{d-1}, \quad |\boldsymbol{J}(t,x,\omega,f_t)| > c_*. \tag{4.8}$$

Therefore, for $t \leqslant T$ the approximated kinetic equation (2.19) coincides with the Vicsek kinetic equation (2.4).

Suppose now that for all $a > 0$ it holds that

$$\mathbb{P}\left(\sup_{t\in[0,T]}\sup_{(x,v)}\big||\mathbf{J}(t,x,v,\mu_t^N)| - |\mathbf{J}(t,x,v,f_t)|\big| \geqslant a\right) \leqslant \frac{1}{a}\varepsilon(N), \tag{4.9}$$

where $\varepsilon(N) \rightarrow 0$ as $N \rightarrow \infty$. If (4.9) holds, then we have that

$$\begin{aligned}\mathbb{P}\biggl(\inf_{t\in[0,T]}\inf_{(x,v)}|\mathbf{J}(t,x,v,\mu_t^N)| > \varepsilon_0\biggr) &= 1 - \mathbb{P}\biggl(\inf_{t\in[0,T]}\inf_{(x,v)}|\mathbf{J}(t,x,v,\mu_t^N)| \leqslant \varepsilon_0\biggr)\\ &\geqslant 1 - \mathbb{P}\left(\inf_{t\in[0,T]}\inf_{(x,v)}\big||\mathbf{J}(t,x,v,f_t) - \mathbf{J}(t,x,v,\mu_t^N)| - |\mathbf{J}(t,x,v,f_t)|\big| \leqslant \varepsilon_0\right),\end{aligned} \tag{4.10}$$

where we used the inequality

$$\inf_{t\in[0,T]}\inf_{(x,v)}|\mathbf{J}(t,x,v,\mu_t^N)| \geqslant \inf_{t\in[0,T]}\inf_{(x,v)}\big||\mathbf{J}(t,x,v,\mu_t^N) - \mathbf{J}(t,x,v,f_t)| - |\mathbf{J}(t,x,v,f_t)|\big|,$$

which follows from the triangular inequality ($||a| - |b|| \leqslant |a - b|$ for all $a,b \in \mathbb{R}$).

The bound

$$\begin{aligned}&\inf_{t\in[0,T]}\inf_{(x,v)}\big||\mathbf{J}(t,x,v,f_t) - \mathbf{J}(t,x,v,\mu_t^N)| - |\mathbf{J}(t,x,v,f)|\big|\\ &\geqslant \inf_{t\in[0,T]}\inf_{(x,v)}|\mathbf{J}(t,x,v,f_t)| - |\mathbf{J}(t,x,v,f_t) - \mathbf{J}(t,x,v,\mu_t^N)|\\ &\geqslant \inf_{t\in[0,T]}\inf_{(x,v)}|\mathbf{J}(t,x,v,f_t)| - \sup_{(x,v)}|\mathbf{J}(t,x,v,f_t) - \mathbf{J}(t,x,v,\mu_t^N)|\\ &\geqslant c_\ast - \sup_{t\in[0,T]}\sup_{(x,v)}|\mathbf{J}(t,x,v,f_t) - \mathbf{J}(t,x,v,\mu_t^N)|\end{aligned}$$

holds by (4.8). Using this inequality in combination with (4.10), we deduce that

$$\mathbb{P}\left(\inf_{t\in[0,T]}\inf_{(x,v)}|\mathbf{J}(t,x,v,\mu_t^N)| > \varepsilon_0\right) \geqslant 1-\mathbb{P}\left(\sup_{t\in[0,T]}\sup_{(x,v)}\left|\mathbf{J}(t,x,v,\mu_t^N)-\mathbf{J}(t,x,v,f)\right| \geqslant c_* - \varepsilon_0\right).$$

Finally, applying (4.9) in this inequality, we obtain (2.25), from which we conclude the proof of the theorem.

We are left with checking that (4.9) holds. We have the following by applying Markov's inequality:

$$\mathbb{P}\left(\sup_{t\in[0,T]}\sup_{(x,v)}\left|\mathbf{J}(t,x,v,\mu_t^N)-\mathbf{J}(t,x,v,f_t)\right| \geqslant a\right) \leqslant \frac{\mathbb{E}\left[\sup_{t\in[0,T]}\sup_{(x,v)}\left|\mathbf{J}(t,x,v,\mu_t^N)-\mathbf{J}(t,x,v,f_t)\right|\right]}{a}. \quad (4.11)$$

Now, let $C(\mathbf{K})$ be the Lipschitz constant of $\mathbf{K}$ (which does not depend on (x,v)). By the Kantorovich characterization of the 1-Wasserstein distance given in (4.1) we have that

$$\left|\mathbf{J}(t,x,v,\mu_t^N)-\mathbf{J}(t,x,v,f_t)\right| \leqslant C(\mathbf{K})W_1(\mu_t^N,f).$$

We take first the supremum over (x,v) on the previous expression, but observe that the right-hand side is independent of (x,v), and then we take expectations. The result is that

$$(4.12)\quad \mathbb{E}\left[\sup_{t\in[0,T]}\sup_{(x,v)}\left|\mathbf{J}(t,x,v,\mu_t^N)-\mathbf{J}(t,x,v,f_t)\right|\right] \leqslant C(\mathbf{K})\mathbb{E}\left[\sup_{t\in[0,T]}W_1(\mu_t^N,f_t)\right].$$

By the triangle inequality, it holds that

$$W_1(\mu_t^N,f_t) \leqslant W_1(\mu_t^N,\overline{\mu}_t^N) + W_1(\overline{\mu}_t^N,f_t),$$

where $\overline{\mu}_t^N$ denotes the empirical measure associated to the nonlinear process (2.21). Now, applying (A.5), we have that

$$\begin{aligned}\mathbb{E}\sup_{t\in[0,T]}W_1(\mu_t^N,\overline{\mu}_t^N) &\leqslant \mathbb{E}\sup_{t\in[0,T]}W_2(\mu_t^N,\overline{\mu}_t^N)\\ &\leqslant \mathbb{E}\sup_{t\in[0,T]}\left(\frac{1}{N}\sum_{i=1}^N\left(|X_t^{i,N}-\overline{X}_t^{i,N}|^2+|V_t^{i,N}-\overline{V}_t^{i,N}|^2\right)\right)^{1/2}\\ &\leqslant \sup_{1\leqslant i\leqslant N}\mathbb{E}\sup_{t\in[0,T]}\left(|X_t^{i,N}-\overline{X}_t^{i,N}|^2+|V_t^{i,N}-\overline{V}_t^{i,N}|^2\right) := \widetilde{\varepsilon}(N),\end{aligned}$$

where in the second inequality we used the identity (A.5) and in the third inequality we used the Cauchy--Schwarz inequality to get rid of the square root and that $\sup(|b|+|c|) \leqslant \sup|b| + \sup|c|$ and $\sup|b|^2 = (\sup|b|)^2$ for $b,c \in \mathbb{R}$. We also recall that $W_1 \leqslant W_2$. It also holds that

$$\mathbb{E} \sup_{t\in[0,T]} W_1(\overline{\mu}_t^N, f_t) \leqslant \mathbb{E} \sup_{t\in[0,T]} W_2(\overline{\mu}_t^N, f_t) \leqslant \mathbb{E}\left[\sup_{t\in[0,T]} W_2(\overline{\mu}_t^N, f_t)^2\right]^{1/2}.$$

To bound this last quantity we will consider the path solutions to the auxiliary particle system (2.21), i.e., the path solutions $(\overline{X}_t^i, \overline{V}_t^i)_{t\in[0,T]}$ as $\mathcal{C} := C([0,T], \Omega \times \mathbb{S}^{d-1})$valued random variables. Denote by $f_{[0,T]} \in \mathcal{P}(\mathcal{C})$ their common law and its associated empirical measure

$$\overline{\mu}_{[0,T]}^N := \frac{1}{N}\sum_{i=1}^N \delta_{(\overline{X}_t^i, \overline{V}_t^i)_{t\in[0,T]}} \in \mathcal{P}(\mathcal{C}).$$

Since the space $\mathcal{C}$ with the norm $\| w\|_\infty := \sup_{t\in[0,T]} |w_t|$ is a Banach space, one can define the Wasserstein distance on $\mathcal{P}(\mathcal{C})$:

$$W_2^{(T)}(m_1, m_2) = \inf\left\{ \left[\int \|w_1 - w_2\|_\infty^2 \, m(dw_1, dw_2)\right]^{1/2} \biggr\} \; m \in \mathcal{P}_2(\mathcal{C}\times\mathcal{C}) \text{ with marginals } m_1 \text{ and } m_2.$$

One can check that for $m, m' \in \mathcal{P}(\mathcal{C})$ it holds that $W_2(m_t, m'_t) \leqslant W_2^{(t)}(m, m') \leqslant W_2^{(T)}(m, m')$ for $t \in [0,T]$. Therefore

$$\mathbb{E}\left[\sup_{t\in[0,T]} W_2(\overline{\mu}_t^N, f_t)^2\right] \leqslant \mathbb{E}\left[W_2^{(T)}(\overline{\mu}_{[0,T]}^N, f_{[0,T]})^2\right] \to 0 \text{ as } N\to\infty,$$

where the limit is a consequence of Lemma A.4 (notice that, following the proof in [5], this lemma can be applied to random variables in $\mathcal{C}$). As a consequence we have that

$$\mathbb{E}\sup_{t\in[0,T]} W_1(\overline{\mu}_t^N, f_t) \to 0 \text{ as } N\rightarrow\infty.$$

From these estimates, we have that

$$\mathbb{E}\Big[\sup_{t\in[0,T]} W_1(\mu_t^N, f)\Big] \to 0 \quad \text{as } N\rightarrow\infty, \tag{4.13}$$

since $\widetilde{\varepsilon}(N) \rightarrow 0$ as $N\rightarrow\infty$ by Theorem 2.11.

Finally, combining (4.11), (4.12), and (4.13), we conclude (4.9).

4.4. Proof of Proposition 4.3. The proof of Proposition 4.3 follows closely the methodologies in [3] and [5], which are based on the Sznitman approach [47]. The main difference is that the interaction rate ν and the noise coefficient σ are considered to be constant in [3]. In particular, this gives rise to an extra term in the equations (coming from (2.16b)).

Step 1. Regularized version of the nonsingular dynamics. The nonsingular particle dynamics (4.5) written in It\^o's convention corresponds to (see section A.2 for more details)

$$
\begin{cases}
dX^{i,N} = V^{i,N} dt, & (4.14a)\\
dV^{i,N} = \nu(\mu^N)\,\mathbf{P}_{(V^{i,N})^\perp}(\tau_0(X^{i,N}, V^{i,N}, \mu^N))dt \\
\qquad + \sqrt{2\sigma(\mu^N)}\,\mathbf{P}_{(V^{i,N})^\perp} dB^i_t \\
\qquad + (d-1)\sigma(\mu^N)\dfrac{V^{i,N}}{|V^{i,N}|^2} dt. & (4.14b)
\end{cases}
$$

And the It\^o formulation for the nonsingular auxiliary process (4.6) is given by

$$
\begin{cases}
d\overline{X}^{i,N} = \overline{V}^{i,N} dt, & (4.15a)\\
d\overline{V}^{i,N} = \nu(f)\,\mathbf{P}_{(\overline{V}^{i,N})^\perp}(\tau_0(\overline{X}^{i,N}, \overline{V}^{i,N}, f))dt \\
\qquad + \mathbf{P}_{(\overline{V}^{i,N})^\perp}[(\nabla_v \sigma(\mu^N))(\overline{X}^{i,N}, \overline{V}^{i,N})]\,dt \\
\qquad + \sqrt{2\sigma(f)}\,\mathbf{P}_{(\overline{V}^{i,N})^\perp} dB^i_t \\
\qquad + (d-1)\sigma(f)\dfrac{\overline{V}^{i,N}}{|\overline{V}^{i,N}|^2} dt, \\
f_t = \mathrm{law}(\overline{X}^{i,N}, \overline{V}^{i,N}). & (4.15b)
\end{cases}
$$

Remark 4.4. Notice that the solutions of (4.14) and (4.15) fulfill $|V_t|^2 = |V_0|^2$ in the velocities for all times where the solution is defined (this is shown as in Lemma 2.9).

Step 2. Existence and uniqueness for the regularized particle system: Proof of part (i) of Theorem 2.10 for (4.5). We consider now a regularized version of systems (4.14) and (4.15) using two functions τ_1 and τ_2, both Lipschitz and bounded and satisfying

$$\tau_1(v) = \mathbf{P}_{v^\perp} = \mathrm{Id} - \frac{v\otimes v}{|v|^2} \qquad \text{if } |v| \geqslant 1/2,$$

$$\tau_2(v) = \frac{v}{|v|^2} \qquad \text{if } |v| \geqslant 1/2 .$$

With these functions we defined the *regularized particle dynamics* as

$$
\begin{cases}
dX^{i,N} = V^{i,N} dt, & (4.16a)\\
dV^{i,N} = \nu(\mu^N)\,\tau_1(V^{i,N})(\tau_0)(X^{i,N}, V^{i,N}, \mu^N)dt \\
\qquad + \tau_1(V^{i,N})[(\nabla_v \sigma(\mu^N))(\overline{X}^{i,N}, \overline{V}^{i,N})]\,dt \\
\qquad + \sqrt{2\sigma(\mu^N)}\,\tau_1(V^{i,N})dB^i_t \\
\qquad + (d-1)\sigma(\mu^N)\tau_2(V^{i,N})dt. & (4.16b)
\end{cases}
$$

Remark 4.5. Notice that the functions τ_0, τ_1, and τ_2 are introduced to regularize the original system in the sense that we obtain a new system where all the coefficients are Lipschitz and bounded in $\mathcal{H}$. This regularity allows us to apply classical results of existence of solutions and convergence, as we will see next. At the same time, when

$|V| = 1$ and $|\mathbf{J}| \geqslant \varepsilon_0$ we recover the approximated particle equations.

First, we have by Lemma 4.1 that τ_0 is a Lipschitz bounded function and, moreover, all the coefficients are also Lipschitz bounded (using that the product of Lipschitz bounded functions is Lipschitz bounded; see also Remark 4.2). So we have existence and uniqueness of pathwise solutions for (4.16) (see [5, Theorem 1.2], which for completeness we have added in simplified form as Theorem A.2 in the appendix.).

Second, one can check as in Lemma 2.9 that $|V_t^{i,N}| = |V_0^{i,N}| = 1$. Therefore, the solution to the regularized system (4.16) is also a solution to the nonsingular system (4.5).

Step 3. Existence and uniqueness for an auxiliary regularized process of (2.21): Proof of part (ii) of Theorem 2.10 for (4.6). Similarly as before, we consider a regularized version of the nonsingular auxiliary process (4.6) given by

$$\begin{cases} d\overline{X}^{i,N} = \overline{V}^{i,N} dt, & (4.17a)\\ d\overline{V}^{i,N} = \nu(f)\, \tau_1(\overline{V}^{i,N})(\tau_0(\overline{X}^{i,N}, \overline{V}^{i,N}, f)dt \\ \qquad +\tau_1(\overline{V}^{i,N})[(\nabla_v \sigma(\mu^N))(\overline{X}^{i,N}, \overline{V}^{i,N})]\, dt & (4.17b)\\ \qquad +\sqrt{2\sigma(f)}\, \tau_1(\overline{V}^{i,N}) dB_t^i \\ \qquad +(d-1)\sigma(f)\tau_2(\overline{V}^{i,N})dt & (4.17c)\\ f_t = \mathrm{law}(\overline{X}^{i,N}, \overline{V}^{i,N}). & (4.17d) \end{cases}$$

Since (4.17) has bounded and Lipschitz coefficients in $\mathscr{H}$ (this can be seen as in the previous proof), it admits a pathwise unique global solution (this statement is shown in [5, Theorem 1.7], which is a generalization of Sznitman's strategy [47, Theorem 1.1]; see Theorem A.3 in the appendix). We can prove as in Lemma 2.9 that $d|\overline{V}_t|^2 = 0$, so $|\overline{V}_t| = 1$ for all times. Therefore, the solution to (4.17) is also the solution of the nonsingular auxiliary system (4.6).

Step 4. Existence and uniqueness for the nonsingular kinetic equation (4.7): Proof of part (iii) of Theorem 2.10 for the nonsingular case, (4.6) and (4.7). Next we show that part (iii) of Theorem 2.10 holds for the nonsingular auxiliary process (4.6) and the nonsingular kinetic equation (4.7).

We can apply Theorem 2.4 to the nonsingular kinetic equation (4.7). Moreover, since ν is bounded and τ_0 is also bounded we have global-in-time existence of solutions for (4.7).

Now, we also know that the kinetic equation (2.4) has existence and uniqueness of weak solutions in $L^\infty([0, T_{max}), L^1_{x,\omega} \cap L^p_{x,\omega})$ for p in $[2,\infty]$. Let f_t be the weak solution of the kinetic equation for $t \in [0,T_{max})$. We also now that for $\varepsilon_0 > 0$ small enough there is a time $T_{\varepsilon_0} \leqslant T_{max}$ such that $|\mathbf{J}_{f_t}| > \varepsilon_0$ for all $t \in [0,T_{\varepsilon_0})$. Therefore, for $t \in [0,T_{\varepsilon_0})$ the solution f_t of the kinetic equation is also a solution of the nonsingular kinetic equation (4.7).

The fact that the Fokker--Planck equation for the nonsingular auxiliary process $(\overline{X}_t,\overline{V}_t)$ (4.7) corresponds to the nonsingular kinetic equation (4.7) is proven in [3] (this is an application of Itô's formula; see, for example, [47, Remark 1.2], [5, Remark

1.8], and (A.3)). The proof that $(\overline{X}_t, \overline{V}_t)$ has law f_t (local-in-time) is carried out in the same way as in [3].

Step 5. Mean-field limit for the nonsingular dynamics: Proof of Theorem 2.11 for nonsingular dynamics (4.5) and (4.6). The proof of Theorem 2.11 is an application of the Sznitman approach, generalized in [5].

We can apply the result in [5, Theorem 1.10] (a condensed version can be found in Theorem A.5 in the appendix).

Remark 4.6. The control in the particle position is immediate from the one in the velocities:

$$|X_T^{i,N} - \overline{X}_T^{i,N}|^2 = \left| \int_0^T V_s^{i,N}\, ds - \int_0^T \overline{V}_s^{i,N} \right|^2 \leqslant T \sup_{0 \leqslant t \leqslant T} |V_t^{i,N} - \overline{V}_t^{i,N}|^2 .$$

Appendix A. Some known results on SDE. For the sake of completeness, we introduce in this section known results for stochastic differential equations that we used in the main part of the document.

A.1. Results on existence of solutions and large-particle limits. The results and proofs from this section can be found in a more general form in [5]. The original approach to proving the large-particle limit is due to Sznitman and can be found in [47].

General stochastic differential equations. Consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geqslant 0}, \mathbb{P})$ and an $(\mathcal{F}_t)_{t \geqslant 0}$-Brownian motion B taking values in $\mathbb{R}^m$. Consider the stochastic differential equation giving the evolution for $Z_t \in \mathbb{R}^p$ for $t \geqslant 0$:

$$dZ_t = b(Z_t)dt + a(Z_t)dB_t, \tag{A.1}$$

with initial data $Z(t=0) = Z_0 \in \mathbb{R}^p$, where the coefficients $a : \mathbb{R}^p \rightarrow \mathbb{R}^{p \times m}$ and $b : \mathbb{R}^p \rightarrow \mathbb{R}^p$ are Lipschitz and bounded. (This is the form taken by the SDE (4.16) with $p = 2d$ and $Z_t = (X_t, V_t)$.)

Definition A.1 (see Definition 1.1 in [5]). *An* $(\mathcal{F}_t)_{t \geqslant 0}$*-adapted continuous process* $(Z_t)_{0 \leqslant t \leqslant T}$ *is a solution to* (A.1) *if*

$$Z_t = Z_0 + \int_0^T b(Z_s)\, ds + \int_0^T a(Z_s)\, dB_s, \quad 0 \leqslant t \leqslant T.$$

Theorem A.2 (adapted from Theorem 1.2 in [5]). *Let us assume that $Z_0 \in L^2$ is independent of B and that the coefficients b and a are Lipschitz and bounded. Then there exists a unique solution of* (A.1)*. Moreover, $Z_t \in L^2$ for all $t < T$, with T finite.*

Results for nonlinear SDE. We consider a nonlinear SDE of the form

$$dZ_t = b(Z_t, \mathcal{L}(Z_t))dt + a(Z_t, \mathcal{L}(Z_t))dB_t, \tag{A.2}$$

where $\mathcal{L}(Z)$ denotes the distribution or law of the random element Z. Here we also assume that $b : \mathbb{R}^p \times \mathcal{P}_2(\mathbb{R}^p) \rightarrow \mathbb{R}^p$ and $a : \mathbb{R}^p \times \mathcal{P}_2(\mathbb{R}^p) \rightarrow \mathbb{R}^{p \times m}$ are bounded and

Lipschitz. Specifically, in this case to be Lipschitz means that for all $z,z' \in \mathbb{R}^d$ and all $\mu,\mu' \in \mathcal{P}(\mathbb{R}^d)$ it holds that

$$|(b,a)(z,\mu) - (b,a)(z',\mu')| \leqslant c(|z - z'| + W^{(2)}(\mu,\mu'))$$

for some constant c (see section 4.1 for more details). System (A.2) is the form taken by the auxiliary system (4.17).

Theorem A.3 (from Theorem 1.7 in [5]). *Let us assume that $Z_0 \in L^2$ is independent of B, and that the coefficients b and a are bounded and Lipschitz. Then there exists a unique solution to* (A.2)*.*

For $\varphi \in C_b^2(\mathbb{R}^p)$, Itô's formula applied to the process Z_t gives (see Remark 1.8 in [5])

$$\varphi(Z_t) = \varphi(Z_0) + \int_0^t \Big[\frac{1}{2}\operatorname{trace}[a(Z_s,\mathcal{L}(Z_s))^T a(Z_s,\mathcal{L}(Z_s))]D^2\varphi(Z_s) + b(Z_s,\mathcal{L}(Z_s))D\varphi(Z_s)\Big]ds + \int_0^t D\varphi(Z_s)\,a(Z_s,\mathcal{L}(Z_s))\,dB_s, \tag{A.3}$$

where the exponent T denotes the transpose and D^2 denotes the Hessian matrix.

Lemma A.4 (law of large numbers; Lemma 1.9 in [5]). *Let $\mu \in \mathcal{P}_2(\mathbb{R}^p)$, and let $(Z_i)_{i\in\mathbb{N}}$ be a sequence of independent and identically distributed random variables with common law μ. For each $N \geqslant 1$ denote by μ^N the empirical distribution associated to the first N elements of the sequence, i.e.,*

$$\mu^N(z) = \frac{1}{N}\sum_{i=1}^N \delta_{Z_i}(z).$$

Then it holds that

$$\lim_{N\rightarrow\infty} \mathbb{E}[W_2(\mu^N,\mu)^2] = 0. \tag{A.4}$$

Letting $z_i \in \mathbb{R}^p$ and $z_i' \in \mathbb{R}^p$ for $i = 1,\dots,N$, it holds (see [5, (1.24)]) that

$$W^{(2)}\left(\frac{1}{N}\sum_{i=1}^N \delta_{z_i},\ \frac{1}{N}\sum_{i=1}^N \delta_{z_i'}\right) \leqslant \left(\frac{1}{N}\sum_{i=1}^N |z_i - z_i'|^2\right)^{1/2}. \tag{A.5}$$

Particle approximations (from [5, section 1.3.4]*).*

Theorem A.5 (extracted and adapted from section 1.3.4 in [5]). *Consider*

$$Z_t^i = Z_0^i + \int_0^T b(Z_s^i,\mathcal{L}(Z_s^i)), ds + \int_0^T a(Z_s^i,\mathcal{L}(Z_s^i))\,dB_s^i \tag{A.6}$$ *and*

$$Z_t^{i,N} = Z_0^i + \int_0^T b(Z_s^{i,N}, \mu_s^N)\, ds + \int_0^T a(Z_s^{i,N}, \mu_s^N)\, dB_s^i, \tag{A.7}$$

where μ_s^N is the empirical distribution of the N particles. Assume that a, b are bounded and Lipschitz in $\mathbb{R}^p \times \mathcal{P}_2(\mathbb{R}^p)$. Then it holds that

$$\lim_{N\to\infty} \sup_{1\leqslant i\leqslant N} \mathbb{E}\left[\sup_{0\leqslant t\leqslant T} |Z_t^{i,N} - Z_t^i|^2\right] = 0. \tag{A.8}$$

A.2. Stratonovich to Itô's convention. We use the results presented in [29] as well as [46, V.30, Theorem (30.14)]. Particularly, if Z_t is a solution to the Stratonovich SDE

$$\begin{cases} dZ = b(Z,t)dt + a(Z,t) \circ dB_t & \text{(A.9a)}\\ Z(t=0) = Z_0, & \text{(A.9b)} \end{cases}$$

where B is an m-dimensional Brownian motion, and $b : \mathbb{R}^p \times [0,T] \rightarrow \mathbb{R}^p$ and $a : \mathbb{R}^p \times [0,T] \rightarrow M^{p\times m}$ (the space of real matrices of dimension $p \times m$) and are such that there is existence and uniqueness of solutions for the SDE (A.9), then Z_t is a solution to the Itô SDE:

$$\begin{cases} dZ = \left[b(Z,t) + \frac{1}{2}c(Z,t)\right] dt + a(Z,t)dB_t & \text{(A.10a)}\\ Z(t=0) = Z_0, & \text{(A.10b)} \end{cases}$$

where

$$c_i(z,t) = \sum_{k=1}^{m}\sum_{j=1}^{n} \frac{\partial a_{ik}}{\partial z_j}(z,t) a_{jk}(z,t).$$

In our case $a : \mathbb{R}^{2d} \rightarrow \mathbb{R}^{2d\times d}$ with

$$a(x,v) = \begin{pmatrix} 0_{d\times d} \\ \eta(x,v) \end{pmatrix},$$

where $0_{d\times d}$ is a $d \times d$ zero-matrix and

$$\eta(x,v) = \alpha(x,v) P_{v^\perp};\quad \alpha(x,v) = \sqrt{2\sigma(f)(x,v)}.$$

With this we have that $c : \mathbb{R}^{2d} \rightarrow \mathbb{R}^{2d}$, with $c_i = 0$ for $i = 1,\dots,d$ since $\gamma_{jk} = 0$ for all $j = 1,\dots,d$ and any $k = 1,\dots,d$. Now, for $i = d+1,\dots,2d$ we have that

$$c_i(x,v) = \sum_{k=1}^{d}\sum_{j=1}^{d} \frac{\partial \eta_{ik}}{\partial v_j}(x,v)\eta_{jk}(x,v).$$

Using that $\eta_{jk}(x,v) = \alpha(x,v)\left(\delta_{j=k} - \frac{v_j v_k}{|v|^2}\right)$, we compute the previous expression and obtain that

$$c(x,v) = \left(0_{1\times d},\ 2(d-1)\sigma(f)(x,v)\frac{v}{|v|^2} + P_{v^\perp}[(\nabla_v \sigma(f))(x,v)]\right).$$

Acknowledgment. The authors also thank Amic Frouvelle for useful discussions and to the referees for very useful comments.

REFERENCES

[1] A. Bendali and S. Tordeux, *Extension of the Gu\"nter Derivatives to Lipschitz Domains and Application to the Boundary Potentials of Elastic Waves*, preprint, https://hal.archivesouvertes.fr/hal-01395952, 2016.

[2] F. Bolley, J. A. Canizo, and J. A. Carrillo\~ , *Stochastic mean-field limit: Non-Lipschitz forces and swarming*, Math. Models Methods Appl. Sci., 21 (2011), pp. 2179--2210, https://doi.org/10.1142/S0218202511005702.

[3] F. Bolley, J. A. Canizo, and J. A. Carrillo\~ , *Mean-field limit for the stochastic Vicsek model*, Appl. Math. Lett., 25 (2012), pp. 339--343, https://doi.org/10.1016/j.aml.2011.09.
011, http://www.sciencedirect.com/science/article/pii/S0893965911004113. [4]
M. Briant, N. Meunier, and L. Navoret, work in progress.

[5] R. Carmona, *Lectures on BSDEs, Stochastic Control, and Stochastic Differential Games with Financial Applications*, SIAM, Philadelphia, 2016.

[6] J.-B. Caussin, A. Solon, A. Peshkov, H. Chate, T. Dauxois, J. Tailleur, V. Vitelli, and\'
D. Bartolo, *Emergent spatial structures in flocking models: A dynamical system insight*, Phys. Rev. Lett., 112 (2014), 148102.

[7] L.-P. Chaintron and A. Diez, *Propagation of Chaos: A Review of Models, Methods and Applications*, preprint, https://arxiv.org/abs/2106.14812, 2021.

[8] H. Chate, F. Ginelli, G. Gr\' egoire, F. Peruani, and F. Raynaud\' , *Modeling collective motion: Variations on the Vicsek model*, Eur. Phys. J. B, 64 (2008), pp. 451--456.

[9] H. Chate, F. Ginelli, G. Gr\' egoire, and F. Raynaud\' , *Collective motion of self-propelled particles interacting without cohesion*, Phys. Rev. E, 77 (2008), 046113.

[10] I. David, P. Kohnke, G. Lagriffoul, O. Praud, F. Plouarboue, P. Degond, and X. Dru-\' art, *Mass sperm motility is associated with fertility in sheep*, Animal Reproduction Sci., 161 (2015), pp. 75--81.

[11] P. Degond, *Global existence of smooth solutions for the Vlasov-Fokker-Planck equation in* 1 *and* 2 *space dimensions*, Ann. Sci. Ecole Norm. Sup. (4), 19 (1986), pp. 519--542,\' http://eudml.org/doc/82185.

[12] P. Degond, A. Diez, A. Frouvelle, and S. Merino-Aceituno, *Phase transitions and macroscopic limits in a BGK model of body-attitude coordination*, Int. J. Nonlinear Sci., 30 (2020), pp. 2671--2736.

[13] P. Degond, G. Dimarco, and T. B. N. Mac, *Hydrodynamics of the Kuramoto--Vicsek model of rotating self-propelled particles*, Math. Models Methods Appl. Sci., 24 (2014), pp. 277-325.

[14] P. Degond, A. Frouvelle, and J.-G. Liu, *A note on phase transitions for the Smoluchowski equation with dipolar potential*, in Hyperbolic Problems: Theory, Numerics, Applications (Padova, Italy, 2012), Appl. Math. 8, AIMS, 2012, pp. 179--192, https://hal.archivesouvertes.fr/hal-00765704.

[15] P. Degond, A. Frouvelle, and J.-G. Liu, *Macroscopic limits and phase transition in a system of self-propelled particles*, J. Nonlinear Sci., 23 (2013), pp. 427--456.

[16] P. Degond, A. Frouvelle, and J.-G. Liu, *Phase transitions, hysteresis, and hyperbolicity for self-organized alignment dynamics*, Arch. Ration. Mech. Anal., 216 (2015), pp. 63--115, https://doi.org/10.1007/s00205-014-0800-7

[17] P. Degond, A. Frouvelle, S. Merino-Aceituno, and A. Trescases, *Alignment of selfpropelled rigid bodies: From particle systems to macroscopic equations*, in International Workshop on Stochastic Dynamics out of Equilibrium, Springer, 2017, pp. 28--66.

[18] P. Degond, A. Frouvelle, S. Merino-Aceituno, and A. Trescases, *Quaternions in collective dynamics*, Multiscale Model. Simul., 16 (2018), pp. 28--77, https://doi.org/10.1137/ 17M1135207.

[19] P. Degond, A. Manhart, and H. Yu, *A continuum model for nematic alignment of selfpropelled particles*, Discrete Contin. Dyn. Syst. Ser. B, 22 (2017), pp. 1295--1327.

[20] P. Degond, A. Manhart, and H. Yu, *An age-structured continuum model for myxobacteria*, Math. Models Methods Appl. Sci., 28 (2018), pp. 1737--1770.

[21] P. Degond and S. Merino-Aceituno, *Nematic alignment of self-propelled particles: From particle to macroscopic dynamics*, Math. Models Methods Appl. Sci., 30 (2020), pp. 1935-1986.

[22] P. Degond, S. Merino-Aceituno, F. Vergnet, and H. Yu, *Coupled self-organized hydrodynamics and Stokes models for suspensions of active particles*, J. Math. Fluid Mech., 21 (2019), 6.

[23] P. Degond and S. Motsch, *Continuum limit of self-driven particles with orientation interaction*, Math. Models Methods Appl. Sci., 18 (2008), pp. 1193--1215.

[24] P. Degond and S. Motsch, *A macroscopic model for a system of swarming agents using curvature control*, J. Stat. Phys., 143 (2011), pp. 685--714.

[25] P. Degond and L. Navoret, *A multi-layer model for self-propelled disks interacting through alignment and volume exclusion*, Math. Models Methods Appl. Sci., 25 (2015), pp. 2439-2475.

[26] M. Doi, *Molecular dynamics and rheological properties of concentrated solutions of rodlike polymers in isotropic and liquid crystalline phases*, J. Polymer Sci. Polymer Phys. Ed., 19 (1981), pp. 229--243, https://doi.org/10.1002/pol.1981.180190205.
[27] R. Duduchava, *On Poincar\'e, Friedrichs and Korns Inequalities on Domains and Hypersurfaces*, preprint, https://arxiv.org/abs/1504.01677, 2015.
[28] L. C. Evans, *Partial Differential Equations*, 2nd ed., Grad. Stud. Math. 19, American Mathematical Society, Providence, RI, 2010, https://doi.org/10.1090/gsm/019.
[29] L. C. Evans, *An introduction to Stochastic Differential Equations*, American Mathematical Society, Providence, RI, 2012.
[30] A. Figalli, M.-J. Kang, and J. Morales, *Global well-posedness of the spatially homogeneous Kolmogorov--Vicsek model as a gradient flow*, Arch. Ration. Mech. Anal., 227 (2018), pp. 869--896, https://doi.org/10.1007/s00205-017-1176-2.
[31] A. Frouvelle, *A continuum model for alignment of self-propelled particles with anisotropy and density-dependent parameters*, Math. Models Methods Appl. Sci., 22 (2012), 1250011.
[32] A. Frouvelle and J.-G. Liu, *Dynamics in a kinetic model of oriented particles with phase transition*, SIAM J. Math. Anal., 44 (2012), pp. 791--826, https://doi.org/10.1137/ 110823912.
[33] I. M. Gamba and M.-J. Kang, *Global weak solutions for Kolmogorov--Vicsek type equations with orientational interactions*, Arch. Ration. Mech. Anal., 222 (2016), pp. 317--342, https://doi.org/10.1007/s00205-016-1002-2.
[34] C. K. Hemelrijk and H. Hildenbrandt, *Schools of fish and flocks of birds: Their shape and internal structure by self-organization*, Interface Focus, 2 (2012), pp. 726--737.
[35] H. Hildenbrandt, C. Carere, and C. Hemelrijk, *Self-organized aerial displays of thousands of starlings: A model*, Behavioral Ecology, 21 (2010), pp. 1349--1359, https://doi.org/10. 1093/beheco/arq149.
[36] P.-E. Jabin and Z. Wang, *Mean Field Limit for Stochastic Particle Systems*, in Active Particles, Vol. 1, Springer, 2017, pp. 379--402.

[37] N. Jiang, Y.-L. Luo, and T.-F. Zhang, *Coupled self-organized hydrodynamics and Navier-Stokes models: Local well-posedness and the limit from the self-organized kinetic-fluid models*, Arch. Ration. Mech. Anal., 236 (2020), pp. 329--387.
[38] N. Jiang, L. Xiong, and T.-F. Zhang, *Hydrodynamic limits of the kinetic self-organized models*, SIAM J. Math. Anal., 48 (2016), pp. 3383--3411, https://doi.org/10.1137/15M1035665.
[39] M.-J. Kang and J. Morales, *Dynamics of a Spatially Homogeneous Vicsek Model for Oriented Particles on the Plane*, preprint, https://arxiv.org/abs/1608.00185, 2016.
[40] J. F. Kingman, *Uses of exchangeability*, Ann. Probab., 6 (1978), pp. 183--197.
[41] B. Maury, A. Roudneff-Chupin, and F. Santambrogio, *A macroscopic crowd motion model of gradient flow type*, Math. Models Methods Appl. Sci., 20 (2010), pp. 1787--1821.
[42] L. Navoret, *A two-species hydrodynamic model of particles interacting through self-alignment*, Math. Models Methods Appl. Sci., 23 (2013), pp. 1067--1098.
[43] L. Onsager, *The effects of shape on the interaction of colloidal particles*, Ann. New York Acad. Sci., 51 (1949), pp. 627--659, https://doi.org/10.1111/j.1749-6632.1949.tb27296.x.
[44] F. Otto and A. E. Tzavaras, *Continuity of velocity gradients in suspensions of rod-like molecules*, Comm. Math. Phys., 277 (2008), pp. 729--758, https://doi.org/10.1007/s00220007-0373-5.
[45] D. Poyato, *Filippov Flows and Mean-field Limits in the Kinetic Singular Kuramoto Model*, preprint, https://arxiv.org/abs/1903.01305, 2019.
[46] L. C. G. Rogers and D. Williams, *Diffusions, Markov Processes and Martingales. Volume* 2*: Ito\^ Calculus*, Cambridge University Press, 2000.
[47] A.-S. Sznitman, *Topics in propagation of chaos*, in Ecole d'\'et\'e de probabilit\'es de Saint-Flour XIX--1989, Springer, 1991, pp. 165--251.
[48] T. Vicsek, A. Czirok, E. Ben-Jacob, I. Cohen, and O. Shochet\' , *Novel type of phase transition in a system of self-driven particles*, Phys. Rev. Lett., 75 (1995), pp. 1226--1229.
[49] C. Villani, *Optimal Transport: Old and New*, Grundlehren Math. Wiss. 338, Springer Science \& Business Media, 2008.
[50] T.-F. Zhang and N. Jiang, *A local existence of viscous self-organized hydrodynamic model*, Nonlinear Anal. Real World Appl., 34 (2017), pp. 495--506.