Vehicular traffic, crowds, and swarms: From kinetic theory and multiscale methods to applications and research perspectives

G. Albi

Department of Computer Science, University of Verona, Italy
giacomo.albi@univr.it

N. Bellomo*

Politecnico of Torino and Collegio Carlo Alberto, Torino, Italy
nicola.bellomo@polito.it

L. Fermo

Department of Mathematics and Computer Science,
University of Cagliari, Italy
fermo@unica.it

S.-Y. Ha

Department of Mathematical Sciences and Research Institute of Mathematics,
Seoul National University, Korea

Korea Institute for Advanced Study, Korea
syha@snu.ac.kr

J. Kim

Department of Mathematical Sciences, Seoul National University, Korea
jhkim206@snu.ac.kr

L. Pareschi

Department of Mathematics and Computer Science,
University of Ferrara, Italy
lorenzo.pareschi@unife.it

D. Poyato† and J. Soler‡

Departamento de Matemática Aplicada
and Research Unit "Modeling Nature" (MNat),
University of Granada, 18071-Granada, Spain
†*davidpoyato@ugr.es*
‡*jsoler@ugr.es*

*Corresponding author

This paper presents a review and critical analysis on the modeling of the dynamics of vehicular traffic, human crowds and swarms seen as living and, hence, complex systems. It contains a survey of the kinetic models developed in the last 10 years on the aforementioned topics so that overlapping with previous reviews can be avoided. Although the main focus of this paper lies on the mesoscopic models for collective dynamics, we provide a brief overview on the corresponding micro and macroscopic models, and discuss intermediate role of mesoscopic model between them. Moreover, we provide a number of selected challenging research perspectives for readers' attention.

Keywords: Active particles; kinetic equations; vehicular traffic; human crowds; swarms; multiscale vision.

AMS Subject Classification: 34A34, 90B20, 74A25, 76N10, 92D25, 35B40, 35L60, 35Q31, 90B20, 70F99, 74A25, 76N10, 92B20

1. Aims and Plan of the Paper

The modeling of self-propelled particles, as it is known, can be developed at the mesoscopic (kinetic) scale by a suitable generalization of the tools of the mathematical kinetic theory, classically referred to the celebrated Boltzmann equation.[68] The mesoscopic scale acts as a bridge between the microscopic scale corresponding to individual-based models and the macroscopic one which generates hydrodynamic type models.

The important difference of the kinetic theory of classical particles, with respect to the theory applied to self-propelled particles, is that in the former case interactions are reversible and satisfy conservation of mass, momentum and energy, while in the latter case interactions are generated by individual and/or collective strategies which pursue a certain idea of individual and/or collective well-being. As a consequence, interactions are nonlinearly additive, non-reversible, and generally also non-local. These features are properly enlightened in Refs. 192 and 18, while nowadays the term *active particles* is used to refer to particles having these specific features.[23]

Scientists who are involved in the modeling of self-propelled particles, more in general of active particles, occasionally dispute on the choice of the most appropriate scale to be used. It is often argued that only the microscopic approach is the most appropriate for systems having a finite number of degrees, while the macroscopic representation requires unrealistic assumptions on the continuity of the matter and consequently it kills some heterogeneity features of the individual behaviors.

The kinetic theory approach can be regarded as an approximation of discrete states by which the overall state of the system is defined by a probability distribution over the microscopic states. This approach needs the assumption of continuity of the said probability density function over position and velocity of the particles (or other

representative variables such as oscillation phase, frequency, aggregation capacity, etc.) which is reasonable only when their number is sufficiently large in some sense to be properly specified. Therefore, kinetic type models can provide a reasonable description of the collective dynamics just when individual-based models have to face the technical complexity of multiple interactions and excessively large number of equations.

Our bias suggests that modeling presents, at each scale, a number of advantages, both analytic and computational, but at the same time, a number of restrictions appears. Hence, rather than disputing on the selection of the most appropriate scale, we suggest that a comparison can be made only once the modeling approach has been developed at each scale, possibly focusing on well-defined case studies.

It is worth stressing that the derivation of models cannot be developed independently at each scale. In more detail, the derivation by kinetic theory methods requires a well-defined description of interactions at the microscopic scale, while the derivation of hydrodynamic models can be achieved by asymptotic methods applied to kinetic type models, where the distance between particles is made tending to zero.[50] Indeed, we are aware that the consistency of the modeling approach at each scale is a highly challenging perspective that will be discussed in the last part of our paper.

Our survey is developed taking these considerations into account. We will focus on the models of vehicular traffic, human crowds and swarms developed and presented in the last decade and, concerning with the first two, we also take advantage of the review.[24] The presentation accounts for the specific common features of the class of systems under consideration which should be taken into account in the modeling approach.

The paper is structured into seven sections and in some of them the authors draw the attention of the reader to possible research perspectives. In details:

Section 2 defines some basic requirements, shared by the three classes of behavioral systems, which represent a common target of all possible models required to be able to depict them. Indeed, the various models which are proposed in the existing literature do not yet exhaustively account for the features that have been listed above. However, an important objective of the modeling approach consists in the achievement of a descriptive ability of models toward the aforementioned features. In general, collective behaviors cannot be straightforwardly related to those of a few entities. This is a specific feature of all living, hence complex, systems due to their ability to develop a self-organizing intelligence. In addition, learning ability[53] progressively modifies the rules of the interactions.

Section 3 presents a review and critical analysis on the modeling of vehicular traffic and human crowds, where the first step of the modeling approach is the derivation of a mathematical structure suitable to capture the main feature of the systems specifically treated in the section. This structure is deemed to offer the conceptual framework for the derivation of models by implementing into it the mathematical description of interactions at the microscopic scale. A review is

presented on the mathematical models proposed in the literature referring them to the said structure and a mathematical statement of problems, typically initial-boundary value problems. Finally, some research perspectives are brought to the attention of the reader.

Section 4 introduces the mechanical and thermodynamical Cucker–Smale (CS) flocking models and their corresponding kinetic/hydrodynamic models. For the introduced models, we present analytical results such as global well-posedness, asymptotic emergent dynamics, uniform stability and uniform-in-time mean-field limit from the particle models to the corresponding kinetic models. We also discuss some open problems for future works and some possible applications of flocking algorithm in practical field of study.

Section 5 deals with the transition from kinetic descriptions to macroscopic equations that model collective behavior of species (swarming, flocking, schooling, synchronization, etc.). To understand the difficulties of obtaining macroscopic models, we will accordingly distinguish two main types of microscopic descriptions: First and second-order microscopic models, which give rise to a large family of variants: smooth (Lipschitz) or singular interaction kernels, anisotropies, heterogeneities, white noise or inertial effects, among others. In particular, we will recall why singular forces are strongly relevant in order to describe finite-time clustering of a population into subgroups. Although we shall present the general picture, we will just focus on two prototypical models of first and second-order: the Kuramoto model and the Cucker–Smale model. Obtaining macroscopic models for non-smooth interaction forces will require dealing with special tools like measure-valued solutions and solutions to the characteristic system in the Filippov sense together with an accurate control of the dissipation of kinetic energy due to alignment interactions.

Section 6 is devoted to computational methods and control problems applied to kinetic models. In the first part, we discuss a class of fast algorithms for mean-field swarming models based on Monte Carlo methods. The core idea of these methods is based on a binary interaction approximation of the dynamic, whose consistency is proved using the so-called grazing limit of the corresponding Boltzmann-type model. Numerical results show the efficiency of these methods compared to the quadratic cost required by a direct techniques for the non-local kinetic models. The second part of the section is devoted to optimal control of kinetic equations. Different type of control problems, and their numerical treatment are discussed. In particular we focus on: *selective control*, where the control action is localized in the phase-space domain; *leaders–followers mechanisms*: microscopic leaders and continuum followers are modeled by a coupled ODE-PDE system. Numerical experiments illustrate different applications in the context of swarming and pedestrian dynamics.

Section 7 looks ahead to research perspectives toward possible developments of the modeling approach with the aim of improving the descriptive ability of models. The main focus of this final section is on a multiscale vision of the modeling of

vehicular traffic, crowds and swarms starting from the main features to be captured by the modeling approach.

2. Complexity Features of Traffic, Crowds and Swarms

Large systems of driver-vehicles, humans in crowds, and living entities in swarms should be viewed as living systems which exhibit various complexity features typically of living entities and hence essentially different from those of the inert matter. These features can have an important impact on the overall collective dynamics. Indeed, unlike inert matter, the behavioral ability of living entities to develop specific strategies and to adapt them to the context makes observable effects which arise from causes that often do not appear evident. As a consequence, the overall dynamics of these systems cannot simply rely on deterministic causality principles.

The aforementioned behavioral strategy is inspired by interactions. It can be rational or even irrational,[4] although it is motivated by a well-defined goal. In the case when the strategy is rational, it may not even be the best possible one, and emergent collective irrational behaviors can be generated under certain circumstances. In some extreme cases, for instance in crowd dynamics, the presence of stress situations caused by incidents or overcrowding, can generate results rather distant from any predictable outcome.

A very first requirement for the modeling approach consists in looking for models which should have the ability to depict the behavioral features that have been outlined above. Therefore, following the conceptual approach proposed in Ref. 18, we list below some key features that are typical of the living systems under consideration which are different from those of the inert matter, see also Ref. 29.

(1) Ability to express a strategy. Living entities have the ability to develop specific strategies related to their organization ability. These strategies are not defined once and for all, but depend on the state of the entities in their surrounding environment including its geometrical shape and quality.

(2) Heterogeneity. The said strategy is heterogeneously distributed and it can also include different targets and groups, for instance, leaders who aim at driving all other entities to their own strategy. All types of heterogeneity can induce various stochastic features in the interactions. In particular, irrational behaviors of a few entities can generate large deviations from the standard dynamics observed in rationality situations.

(3) Nonlinear interactions. Interactions are nonlinearly additive and nonlocal as they could involve not only immediate neighbors but also distant entities. In some cases, the topological distribution of a fixed number of neighbors rather than all the entities in the visibility domain can play a prominent role in the interactions. In addition, quality of environment, namely weather conditions,

geometry of the venue, abrupt changes of directions, luminosity conditions, and many others can modify the dynamics of interactions.

(4) Learning dynamics. Living entities receive inputs from the external environment and have the ability to learn from past experience. Therefore, their strategic ability and the rules of interactions can evolve in time and space.

(5) Multiscale aspects. The mathematical approach always needs multiscale methods. Indeed, a single observation and representation scale are not generally sufficient to describe the overall collective dynamics of living systems. For instance, the dynamics at the microscopic scale defines the conceptual basis toward the derivation of models at the mesoscopic scale. Models at the higher scale, corresponding to observable macroscopic quantities can be obtained from kinetic models by letting the distance between individuals to zero.

The selection, we have proposed, has extracted the features which appear to us the most relevant ones. However, our choice does not claim to be exhaustive and possible additional ones can be included. In principle, one has to look for mathematical structures flexible enough to include the said additional ones, if consistent with the specific physical situation which is object of modeling. The main difficulty in the search of these structure appears to the attempt to capture the most relevant complexity features of living systems in general. Subsequently, one might specialize the approach to each specific class of systems by a detailed interpretation of interactions.

3. On the Modeling of Vehicular Traffic and Human Crowds

The idea of modeling vehicular traffic by means of the kinetic theory approach has been presented for the first time in the paper of Prigogine and Hermann.[202] This paper has the great merit of having introduced the use of mathematical tools of the kinetic theory to model the dynamics of self-propelled particles. This scientific initiative has been revisited in Ref. [193], where the heterogeneous behavior of the micro-scale system "driver-vehicle" has been proposed as a key issue of the model.

Both pioneer papers have been followed by a rich stream of scientific contributions, for example Refs. [82], [91], [164] and [186] which have already reviewed in Ref. [24] and hence their review is not repeated here. We simply remark that the novelty of Refs. [82] and [91] has been the development of a discrete velocity framework to take into account the granular nature of the traffic. Indeed, as observed also in Ref. [86], the speed of vehicles does not span the whole range of the admissible speeds to justify the continuity assumption of the probability distribution function over the velocity variable. This criticism has not been neglected by applied mathematicians who have attempted to account for it in different ways. For instance, discrete velocity models have been proposed both for vehicular traffic and human crowds as a possible constructive reply to the criticisms risen in Ref. [86].

Referring to the literature reviewed in Ref. [24] and specifically to discrete velocity models, let us stress the difference between Refs. [91] and [82]. In Ref. [91], the authors

have used a fixed velocity grid $\{v_i\}$ in which each velocity v_i, interpreted as velocity class, is constant with respect to both time and space. On the contrary in Ref. 82, the authors have introduced a grid in which each velocity v_i depends on the local traffic conditions by means of the macroscopic density.

Further developments have been proposed in the last decade, most of them as possible developments of Ref. 91, while the ideas of Ref. 82 have not been extensively exploited despite the fact that using a grid depending on the local density is a clever assumption which deserves attention. In fact, it contributes, at least partially, to the aforementioned inconsistency of the continuity assumption. The idea of using discrete velocity models appeared in various papers. For instance in Ref. 143, where a simple two velocity model with Enskog-type interactions has been developed and carefully investigated.

The kinetic theory approach has, in the case of crowd dynamics, a shorter story compared with that of vehicular traffic. The literature in the field mainly developed after the survey,[24] where only a few hints where given starting from a literature already well settled for the approach at the microscopic and macroscopic scales. Here, we simply mention that the known literature on kinetic models of crowds includes both discrete velocity models, for example Ref. 19, and models continuous over the velocity variable, for example Ref. 25.

The two classes of models treated in this section present some common features, but also specific features which differ from vehicular traffic to human crowds. These are mainly related to the specific venues where the dynamics occurs, namely roads, which can be confined by tollgates in highways or nodes in networks of road, in the case of vehicles and complex venues, which can include obstacles and walls, in the case of crowds. Both systems can be subdivided into different populations, technically called *functional subsystems*, which correspond to groups pursuing different strategies or even the same strategy, but expressed by different manners.

As in the classical kinetic theory, we refer to the *microscopic state* to denote the state of each individual, generally identified by position, velocity, as well as by an additional variable corresponding to the ability of the driver and/or the quality of the vehicle, in the case of vehicular traffic, and corresponding to the emotional state and/or walking ability in the case of human crowds. Moreover, we will consider the time and the position as independent variables and a probability distribution function over the microscopic state as a dependent variable.

Dimensionless variables are used by dividing each of them by a characteristic quantity of the system. Technical details are given in the next two subsections, while we simply mention here that the use of dimensionless quantities is an essential feature of modeling approach and, in addition, it contributes to a rational use of computational tools.

A description of the contents of the next subsections follows this general introduction. Subsections 3.1 and 3.2 deal with modeling, respectively, vehicular traffic and human crowds. The presentation follows analogous guidelines as it focuses firstly on the derivation of the mathematical structures which are appropriate to

capture the main phenomenological features of the two systems, and then focuses on a survey and a critical analysis of the research achievements known in the literature. Subsection 3.3 shows some sample simulations accounting also for the problem of validation of models. Subsection 3.4 proposes some research perspectives which are selected according to the authors' experience and bias.

3.1. *Kinetic models of vehicular traffic*

3.1.1. *Mathematical structures toward modeling*

Let us consider a one-dimensional flow of vehicles along a road of length ℓ and let us introduce the following dimensionless independent variables useful to describe the state of the system:

- Position x obtained by dividing the real space by the length ℓ;
- Velocity v scaled with respect to v_ℓ which denotes the maximum velocity which cannot be passed, simply for mechanical reasons by vehicles or for speed limits or by environmental conditions.
- Time t which is scaled with respect to the critical time t_c obtained by the ratio between ℓ and v_ℓ, where t_c corresponds to the time that the fast vehicles needs to travel on the whole length ℓ of the road.

In the sequel, we will also consider an additional variable u which, in addition to position and velocity, describes the microscopic state (x, v, u) of the driver-vehicle subsystem. Such variable, named activity, belongs to the interval $[0, 1]$ and identifies the quality of the driver-vehicle subsystem, where the value $u = 0$ is related to lowest quality, while the value $u = 1$ corresponds to the best ones.

In the following, we also take the quality of the environment into account by means of a parameter $\alpha \in [0, 1]$. Analogously to u, $\alpha = 0$ stands for the worst environmental conditions and $\alpha = 1$ for the best ones. Generally, α can also depend on the position x in order to consider, for instance, the presence of curves, local restrictions, and so on.

Once the microscopic state of the subsystem driver-vehicle is introduced, the state of the whole system can be described by a kinetic distribution function

$$f = f(t, x, v, u): \quad \mathbb{R}_+ \times [0, 1] \times [0, 1] \times [0, 1] \to \mathbb{R}_+ \tag{3.1}$$

such that $f(t, x, v, u)\, dx\, dv\, du$ is the infinitesimal number of vehicles that at time t are located in $[x, x + dx]$, travel with a speed belonging to $[v, v + dv]$ and have an activity in $[u, u + du]$.

Then, the macroscopic quantities like the car density can be computed as the zeroth-order moment of said probability distribution

$$\rho(t, x) = \int_0^1 \int_0^1 f(t, x, v, u) dv\, du, \tag{3.2}$$

while a further integration over the space

$$N(t) = \int_0^1 \int_0^1 \int_0^1 f(t, x, v, u) dx\, dv\, du \tag{3.3}$$

turns out the total number of vehicles at time t.

Let us point out that the density is referred to the maximum number of cars ρ_M corresponding to bumper-to-bumper traffic jam.

Analogously, we can also recover the local dimensionless mean velocity as the first-order moment

$$\xi(t, x) = \frac{1}{\rho(t, x)} \int_0^1 \int_0^1 v f(t, x, v, u) dv\, du, \tag{3.4}$$

which allow us to compute the flux

$$q(t, x) = \xi(t, x)\rho(t, x),$$

as well as higher order momenta that is the average kinetic energy E

$$E(t, x) = \frac{1}{2} \int_0^1 \int_0^1 v^2 f(t, x, v, u) dv\, du \tag{3.5}$$

and the variance of the velocity σ

$$\sigma(t, x) = \frac{1}{\rho(t, x)} \int_0^1 \int_0^1 [v - \xi(t, x)]^2 f(t, x, v, u) dv\, du. \tag{3.6}$$

Let us now derive the mathematical structure by stating an evolution equation in time and space for the distribution function f. To this end, we need to describe the interactions among the vehicles which occur in the visibility zone of the vehicle $\Omega \equiv \Omega(x) = [x,\ x + R]$ where R is the visibility length which depends on the quality of the environment.

The interactions are conservative in the sense that they preserve the total number of vehicles and are binary in the sense that they involve not more that two vehicles. Specifically, the interactions we are going to consider involve the *candidate vehicle* with state $w_* = (t, x, v_*, u_*)$ which is the "candidate" to change its state and the *field vehicle* with state $w^* = (t, x, v^*, u^*)$ which is the cause of such a change. In addition to these two vehicles, the so-called *test-vehicle* is also involved in the description of such interactions even if it is not involved directly. Its usefulness in identifying an ideal vehicle of the system whose microscopic state $w = (t, x, v, u)$ is targeted by a hypothetical observer.

Such interactions are described quantitatively by the following two quantities:

- The encounter rate η which models the frequency of the interactions among candidate and field vehicles;
- The transition probability density $\mathcal{A}$ which gives the probabilities that candidate vehicles get the test state after interacting with field vehicles.

Let us stress that the modeling approach is based on the assumption that these quantities depend not only on the microscopic state of the interacting particles, but also on the distribution function f which induces a nonlinearity in the models evidenced by square brackets. In addition, these interaction terms are allowed to depend also on the quality of the road modeled, as mentioned, by a parameter $\alpha \in [0, 1]$. Let us also remark that $\mathcal{A}$ is required to satisfy the *probability density* condition

$$\mathcal{A}[f;\alpha] \geq 0, \quad \int_{[0,1]} \mathcal{A}[f;\alpha](v_* \to v|w_*, w^*, \alpha)dv = 1, \tag{3.7}$$

for all possible inputs w_*, w^*, and parameter α.

The evolution equation reads as

$$\partial_t f(t,x,v,u) + v\,\partial_x f(t,x,v,u) = J[f](t,x,v,u), \tag{3.8}$$

where J is an operator acting on the distribution function f describing the interactions and their effects on the states of the vehicles. It is split into the difference of a gain operator $G[f]$ and a loss operator $L[f]$ which gives the amount of vehicles per unit time that get and lose the state (x, v, u), respectively. Specifically, such operators take the form

$$\begin{aligned} G[f](t,x,v,u) = \int_{[0,1]^3} &\eta[f](v_*, v^*, u_*, u^*, \alpha)\mathcal{A}[f;\alpha](v_* \to v|v_*, v^*, u_*, u^*, \alpha) \\ &\times f(t,x,v_*,u)\, f(t,x,v^*,u^*)dv_*\, dv^*\, du^*, \end{aligned}$$

and

$$L[f](t,x,v,u) = f(t,x,v,u) \int_{[0,1]^2} \eta(f; v_*, v^*, u_*, u^*, \alpha)\, f(t,x,v^*,u^*)dv^*\, du^*.$$

As announced at the beginning of Sec. 2, discrete microscopic states of the vehicles have been developed and extensively investigated in order to take into account the intrinsic granularity of the distribution of the vehicles.

With reference to the continuous model described above, in Ref. 32, the authors relax not only the continuous variable of the velocity v as done in Refs. 82 and 91 but also the activity variable u. Then they introduce two discrete grids with a fixed number of points $I_v = \{v_i\}_{i=1}^I$ and $I_u = \{u_j\}_{j=1}^J$ and identify the microscopic state of such a vehicle at time t by (t, x, v_i, u_j)

Hence, the physical system is described by $I \times J$ distribution functions

$$f_{ij} = f_{ij}(t,x) : \mathbb{R}_+ \times [0,\, 1] \to \mathbb{R}_+$$

such that $f_{ij}(t,x) = f(t,x,v_i,u_j)$ or in the distributional sense

$$f(t,x,v,u) = \sum_{i=1}^{I}\sum_{j=1}^{J} f_{ij}(t,x)\delta(v - v_i)\delta(u - u_j). \tag{3.9}$$

Thus, classical macroscopic average quantities can be easily derived. In details, the vehicle density ρ, the flux q and the average speed ξ are given by

$$\rho(t,x)=\sum_{i=1}^{I}\sum_{j=1}^{J}f_{ij}(t,x),\quad q(t,x)=\sum_{i=1}^{I}\sum_{j=1}^{J}v_j f_{ij}(t,x),\quad \xi(t,x)=\frac{q(t,x)}{\rho(t,x)}.$$

In this way, the general mathematical structure is the following:

$$\partial_t f_{ij}(t,x)+v_i\,\partial_x f_{ij}(t,x)=J_{ij}[f](t,x), \tag{3.10}$$

where the interaction operator reads as

$$J_{ij}[f](t,x)=G_{ij}[f]-L_{ij}[f]$$

with

$$G_{ij}[f]=\sum_{h,p=1}^{I}\sum_{k,q=1}^{J}\int_{\Omega_v}\eta[\rho(t,x^*),x]\mathcal{A}^{ij}_{hk,pq}(v_h\to v_i,u_k\to u_j|v_h,v_p,u_k,u_q,\rho(t,x^*))$$
$$\times\, f_{hk}(t,x)f_{pq}(t,x^*)dx^*,$$

and

$$L_{ij}[f]=f_{ij}(t,x)\int_{\Omega_v}\sum_{p=1}^{I}\sum_{q=1}^{J}\eta[\rho(t,x^*),x]f_{pq}(t,x^*)dx^*.$$

Here, the interaction rate η as well as the table of games $\mathcal{A}$ is assumed to depend on the density ρ. For details the interested reader can consult Ref. 32.

Some technical developments can be rapidly indicated looking ahead to research perspectives. In more details, the following are brought to the attention of the reader:

(1) Vehicles can be subdivided into a number n of functional subsystems (FS), corresponding to different types of vehicles, from slow trucks to cars with different speed ability, while the activity variable is left heterogeneous. Accordingly each FS is represented by a probability distribution function f_k with $k=1,\ldots,n$, while interactions involve vehicles belonging to all FSs.
(2) An additional subdivision can be introduced for the variable u which can be subdivided into tracts corresponding to discrete values of the driving ability.
(3) Multi-lane flows can be accounted by modeling the transition of vehicles from one lane to the other depending on the local density conditions in contiguous lanes.

In the next subsection, we discuss the discrete microscopic states models that have been developed after the paper.[32] For the sake of clarity, and in order to treat all the recent studies, we will focus on both the spatially homogeneous and inhomogeneous problem.

3.1.2. *A survey of mathematical models*

Let us now provide a survey of recent modeling approaches to vehicular traffic by kinetic theory methods. As already mentioned, different scientists developed the original idea of Ref. [91] toward new modeling issues and applications.

In details, a model which takes into account the driving ability of the vehicle-driver subsystem has been developed in Ref. [32], where a qualitative analysis (existence of solution) has been also presented for a road with periodic boundary conditions. Simulations have shown some interesting features of the predictive ability of the model. In particular numerical experimentations, focused on the interaction of a cluster of fast vehicles with a cluster of slow vehicles, have shown that interactions generate four, out of the initial two, clusters, namely a small groups of fast vehicles which rapidly passed the slow ones, a group of fast vehicles following the previous ones, a group of slow vehicles, and a small group of slow vehicles which were rapidly passed by fast vehicles, and consequently has been left behind.

Some simulations are also presented to show how the model reproduces empirical data. Indeed, the velocity diagram, where the mean velocity is represented as a function of the macroscopic density, is given as well as the fundamental diagram, describing as the macroscopic flux varies with the density, is presented. The test also depicts how the quality of the road modifies the shape of the diagram.

The role of the ability of drivers in real dynamics is not, as observed in Ref. [52], simply described by a parameter, but it should be depicted by a variable which might change in time and space due to interactions. This topic appears to be even more important in crowd dynamics as we shall see in the next subsection.

One of the possible criticisms to continuous modeling of vehicular traffic is that in the kinetic theory approach the number of vehicles is not sufficiently large to justify a continuous description of the probability distribution function over the microscopic state. The use of discrete velocities contributes to a realistic description referring to the velocity variable, but not to space. Accordingly a discrete model with a double discretization, space and velocity, has been proposed and applied in Ref. [103]. This model has been further developed toward the challenging objective of modeling traffic flows on networks,[105] thus developing a useful mathematical tool for applications.

Let us also underline that the application of models to real flow conditions has generated some interesting analytic problems. We have already mentioned the qualitative analysis of solutions and we can add the derivation of macroscopic, namely hydrodynamical, models from the underlying description at the micro-scale as delivered by the kinetic theory approach, as an example.[21]

In Ref. [174], the authors apply both Chapman–Enskog expansion and Grad's moment method in order to construct a second-order continuum traffic model which is very similar to the Navier–Stokes model for viscous fluids. Specifically, the model, which is able to simulate the standard traffic operations in real-life traffic, presents a traffic viscosity coefficient which is not introduced in an ad hoc way,

but comes into play through the derivation of a constitutive relation to the traffic pressure.

The kinetic theory of traffic proposed by Prigogine and Herman in which the Boltzmann equation is adapted to vehicular traffic is reviewed.[149] In particular, the paper contains a novel distribution of desired velocities that is more suitable for describing real traffic conditions and an analysis of the stationary velocity distribution at the transition between individual and collective flow patterns.

A kinetic model for vehicular traffic with a new structure which accounts for the heterogeneous composition of traffic flow is presented in Ref. 203. Here, the model is again built by assuming a discrete space of microscopic speeds and by expressing vehicle interactions in terms of transition probabilities among the admissible speed classes, but the approach differs from standard kinetic models in that the authors consider two distribution functions describing two classes of vehicles with different physical features, namely the typical length of a vehicle and its maximum speed.

In Ref. 204, following the classical Boltzmann-like setting of binary interactions, the authors study a kinetic model based on a continuous velocity space and analyze the space homogeneous case to study the asymptotic behavior of the distribution function together with the resulting flux-density diagrams. Later,[205] such a model has been extended to the case of more than one class of vehicles, characterized by a few parameters accounting for the microscopic differences which allow one to distinguish two (or more) type of vehicles.

Focusing on control problems, a mathematical approach to road risk mitigation issues has been developed in Ref. 223 based on the idea that the difference in the speeds of the vehicles is one of the major risk factors. Hence, two different possible control strategies for the reduction of the speed variance in the stream of vehicles have been proposed.

In Ref. 144, a two-dimensional kinetic traffic model, which takes into account speed changes both when vehicles interact along the road lanes and when they change lane, has been developed by means of suitable numerical methods, precisely structure preserving and direct Monte Carlo schemes. The authors use the model to compute theoretical speed density diagrams of traffic both along and across the lanes, including estimates of the data dispersion, and validate them against real data.

3.2. *Kinetic models of human crowds*

3.2.1. *Mathematical structures toward modeling*

The contents of this subsection is presented along the same guidelines followed in the review of the vehicular traffic. We refer both to human crowds in unbounded domain in $\mathbb{R}^2$ with walkers who move along one or more prefixed directions and to crowds confined in a two-dimensional domain Σ with one or more exit doors. The domain can include internal obstacles and the whole set of walls is denoted by $\partial\Sigma$.

The assessment of dimensionless parameters and variables simply requires a few technical modifications with respect to the case of vehicular traffic. In more detail:

(i) ℓ is selected either by the diameter of the circular domain containing Σ or, in the case of unbounded domain, by the diameter of the circular domain containing the domain Σ_0 initially occupied by the crowd.
(ii) Dimensionless position and velocity variables are dimensional vectors denoted, respectively, by $\mathbf{x}$ and $\mathbf{v}$, where the linear components of $\mathbf{x}$ are referred to ℓ, while the speed v is referred to v_ℓ, respectively, where v_ℓ is a limit velocity defined, as in Subsec. 3.1, by the limit speed of fast walkers in optimal quality areas.
(iii) Often it is useful using polar coordinates $\mathbf{v} = \{v\cos\theta, v\sin\theta\}$, with obvious meaning of notations.
(iv) Dimensionless time, denoted by t, is obtained by dividing the real time by $t_c = \ell/v_\ell$.
(v) The variable $u \in [0,1]$ denotes a specific emotional state which can be defined within the framework of each specific case study under consideration.
(vi) The parameter $\alpha \in [0,1]$ denotes the quality of the venue, where the dynamics occur.
(vii) Walkers have a visibility zone $\Omega_v = [\theta - \Theta, \theta + \Theta] \times [0, R]$, where Θ and R define, respectively, the visual angle and distance.

The overall state of the system is described by the distribution function over the state at the microscopic scale $f = f(t, \mathbf{x}, \mathbf{v}, u)$, or $f = f(t, \mathbf{x}, v, \theta, u)$ when polar coordinates are used, namely $\mathbf{v} = \{v\cos\theta, v\sin\theta\}$, where $\theta \in [0, 2\pi)$ denotes the direction of the velocity. The distribution function f is made to refer to ρ_M so that, if f is locally integrable, $f(t, \mathbf{x}, \mathbf{v}, u)\, d\mathbf{x}\, d\mathbf{v}\, du$ denotes the dimensionless density of vehicles which, at time t, are in the phase elementary domain $[x, x + dx] \times [v, v + dv] \times [u, u + du]$.

Macroscopic quantities, for example local density, total number of walkers, local dimensionless mean velocity and flow, can be obtained precisely as in the case of vehicular traffic simply accounting for the fact the now integration over space and velocity is in two dimensions. Calculations are not repeated, but we simply report the recovering of the local density when polar coordinates are used. In details, the local density, also referred to ρ_M, is given by

$$\begin{aligned}\rho(t, \mathbf{x}) &= \int_{D_\mathbf{v}} \int_0^1 f(t, \mathbf{x}, \mathbf{v}, u) d\mathbf{v}\, du \\ &= \int_0^1 \int_0^{2\pi} \int_0^1 f(t, \mathbf{x}, v, \theta, u) v\, dv\, d\theta\, du \end{aligned} \tag{3.11}$$

while the total number of pedestrians at time t is computed by an additional integration over space. Mean speed and flow can be computed by calculations analogous to those of Subsec. 3.2.1 simply by transferring them in a two-dimensional setting.

The derivation of a mathematical structure, deemed to provide a general framework for the derivation of models, can be developed referring to Ref. [25] along the same line followed for vehicular traffic. Hence interactions involve *candidate*, *field*, and the *test* particle which is representative of the whole system.

In general, the modeling of interactions can be achieved by means of the following terms:

— *The encounter rate* $\eta(f; \mathbf{v}_*, \mathbf{v}^*, u_*, u^*\alpha)$ which refers to the speed of interaction.
— *The transition probability density* $\mathcal{A}(\mathbf{v}_* \to \mathbf{v}|f, \alpha, \mathbf{v}_*, \mathbf{v}^*, u_*, u^*)$ whose integral over all outputs is equal to one, as in Eq. (3.7), for all input variables and parameters.

Both Ω_v and v_ℓ should be related to the shape and quality of the venue where the dynamics occurs. The geometrical shape, for instance the presence of walls, can reduce both visibility angle and distance.

Mathematical structures, suitable to provide conceptual basis for derivation of models, can be obtained by equating the transport of f to the net flow in the elementary volume of the space of microscopic states as it is induced by interactions. A quite general structure is as follows:

$$\begin{aligned}
&\partial_t f(t, \mathbf{x}, \mathbf{v}, u) + \mathbf{v} \cdot \nabla_{\mathbf{x}} f(t, \mathbf{x}, \mathbf{v}, u) \\
&\quad = \int_{D_\mathbf{v}} \int_{D_\mathbf{v}} \int_0^1 \int_0^1 \eta(f; \mathbf{v}_*, \mathbf{v}^*, u_*, u^*, \alpha) \mathcal{A}(\mathbf{v}_* \to \mathbf{v}|f, \alpha, \mathbf{v}_*, \mathbf{v}^*, u_*, u^*) \\
&\qquad \times f(t, \mathbf{x}, \mathbf{v}_*, u)\, f(t, \mathbf{x}, \mathbf{v}^*, u^*) d\mathbf{v}_*\, d\mathbf{v}^*\, du_* du^* \\
&\qquad - f(t, \mathbf{x}, \mathbf{v}, u) \int_{D_\mathbf{v}} \int_0^1 \eta(f; \mathbf{v}_*, \mathbf{v}^*, u_*, u^*, \alpha) f(t, \mathbf{x}, \mathbf{v}^*, u^*) d\mathbf{v}^*\, du^*. \qquad (3.12)
\end{aligned}$$

Some technical developments can be rapidly indicated looking ahead to research perspectives. In more details, the following are brought to the attention of the reader:

(1) Particles can be subdivided into a number n of functional subsystems corresponding to different types of walking directions or exits in case of evacuation.
(2) Semi-discrete (hybrid) models have been proposed where the velocity directions can attain a number of finite values, while the speed is continuous along each direction.
(3) Quantities which play a role in the dynamics of interactions are estimated by a weighted average within the visibility domain.

It is worth reporting the mathematical structures used for the two classes of models indicated above, as both of them have been used by models proposed in the literature. Let us first consider the case where the whole crowd is subdivided into

different groups, viewed as functional subsystems, while the state of each of them is delivered by the probability distribution function f_i with $i = 1, \ldots, m$. A technical generalization of Eq. (3.12) is as follows:

$$\begin{aligned}
&\partial_t f_i(t, \mathbf{x}, \mathbf{v}, u) + \mathbf{v} \cdot \nabla_{\mathbf{x}} f_i(t, \mathbf{x}, \mathbf{v}, u) \\
&\quad = \sum_{k=1}^{n} \int_{D_{\mathbf{v}}} \int_{D_{\mathbf{v}}} \int_0^1 \int_0^1 \eta_{ik}(\mathbf{f}; \mathbf{v}_*, \mathbf{v}^*, u_*, u^*, \alpha) \mathcal{A}_{ik}(\mathbf{v}_* \to \mathbf{v} | \mathbf{f}, \alpha, \mathbf{v}_*, \mathbf{v}^*, u_*, u^*) \\
&\qquad \times f_i(t, \mathbf{x}, \mathbf{v}_*, u)\, f_k(t, \mathbf{x}, \mathbf{v}^*, u^*) d\mathbf{v}_*\, d\mathbf{v}^*\, du_* du^* \\
&\qquad - f_i(t, \mathbf{x}, \mathbf{v}, u) \sum_{k=1}^{n} \int_{D_{\mathbf{v}}} \int_0^1 \eta_{ik}(\mathbf{f}; \mathbf{v}_*, \mathbf{v}^*, u_*, u^*, \alpha) f_k(t, \mathbf{x}, \mathbf{v}^*, u^*) d\mathbf{v}^*\, du^*,
\end{aligned} \tag{3.13}$$

where the encounter rate η_{ik} and the transition probability density $\mathcal{A}_{ik}$ account for interactions with the same FS and across FSs. Notice that if non-local interactions are modeled, the η, $\mathcal{A}$, and f depend on $\mathbf{x}$ and $\mathbf{x}^*$ and an additional integration in space over Ω_v is necessary in (3.12) and (3.13).

As mentioned, a variety of models have made use of discrete velocity models somehow inspired on the same principles presented for vehicular traffic. Here, the basic assumption is that the velocity can attain only a finite number of directions: $\{\theta^1 = 0, \ldots, \theta^i, \ldots, \theta^I = 2\,\pi/(n-1)\}$.

Mathematical structures can be derived referring to both to Eqs. (3.12) and (3.13). We report simply the former case leaving the latter to the interested reader. A simple technical generalization is as follows:

$$\begin{aligned}
&\partial_t f^i(t, \mathbf{x}, \mathbf{v}, u) + \mathbf{v} \cdot \nabla_{\mathbf{x}} f^i(t, \mathbf{x}, \mathbf{v}, u) \\
&\quad = \sum_{j=1}^{I} \int_0^1 \int_0^1 \eta^{ij}(\mathbf{f}; \mathbf{v}_*^i, \mathbf{v}^{*j}, u_*, u^*, \alpha) \mathcal{A}^{ij}(\mathbf{v}_*^i \to \mathbf{v}^j | \mathbf{f}, \alpha, \mathbf{v}_*^i, \mathbf{v}^{*j}, u_*, u^*) \\
&\qquad \times f^i(t, \mathbf{x}, \mathbf{v}_*^i, u)\, f^j(t, \mathbf{x}, \mathbf{v}^{*j}, u^*) du_* du^* \\
&\qquad - f^i(t, \mathbf{x}, \mathbf{v}^i, u) \int_0^1 \eta^{ij}(\mathbf{f}; \mathbf{v}_*^i, \mathbf{v}^{*j}, u_*, u^* \alpha) f^j(t, \mathbf{x}, \mathbf{v}^{*j}, u^*)\, du^*,
\end{aligned} \tag{3.14}$$

where f^i corresponds to the probability distribution function over the velocity $\mathbf{v}^i$, $\mathbf{f}$ is the set of all discrete velocities, η and $\mathcal{A}^{ij}$ are, respectively the encounter rate and the transition probability density related to each pair of interacting discrete velocities.

In addition, one might rapidly consider an additional subdivision into FSs, for instance FSs corresponding to different types of walking ability, models where the emotional state can attain different values and various others. However, we limit our presentation to the above structures, namely to those which have been effectively used in the literature.

3.2.2. *A survey of mathematical models*

Mathematical models of crowd dynamics have been proposed referring initially to the mathematical structure corresponding to discrete velocity directions delivered by Eq. (3.14). This modeling approach was introduced in Ref. 19 under the assumption that the speed was shared by all walkers, while the velocity direction was supposed to be distributed over a fixed number of velocity directions. A detailed modeling of the dynamics of interactions has been proposed by assuming that the candidate walker changes velocity direction by a weighted selection of three stimuli, namely the desire of following a prescribed local direction, avoiding overcrowded areas and attraction by the mean stream (following what all the others do).

The selection is weighted by the local density which enhances the desire of avoiding overcrowded areas, with respect to the prescribed velocity direction, as the density increases to the limit value $\rho = 1$. After this preliminary decision, a parameter modeling the level of stress is used as a weight to account for the attraction toward the main stream. Namely this attraction increases when the level of stress increases.

A technical development has been proposed in Ref. 3 by a hybrid model, where the velocity direction is assumed to attain a finite number of velocity directions, while the speed is assumed to be continuous in the domain $[0, 1]$ being modeled by the so-called velocity diagrams delivered by empirical data[212,214] which suggest that the dimensionless speed decays from one to zero when the density increases to the maximal value.

Discrete velocity models have been applied in Ref. 98 to compute the evacuation time from a venue that includes internal obstacles. The authors show how the qualitative analysis, proposed in Ref. 19, for the solutions to the initial vale problem in unbounded domains can be technically generalized to include interactions with boundaries. Computational methods based on splitting methods have been developed in Ref. 160 accounting also for nonlinear transport terms.

An analytic study has been developed in Ref. 17 to derive models at the macroscopic scale from the underlying description delivered by kinetic models. This approach refers the derivation of macroscopic models to the dynamics at the low, individual based, scale rather than by conservation equations closed by heuristic models of material behaviors.

The approach to modeling by continuous velocity variable has been developed in Refs. 25 and 26, where Ref. 25 is devoted to modeling behavioral interactions between walkers and between walkers and walls which generates nonlocal boundary conditions, while Ref. 26 is focused on the validation of models. In more details, validation has been interpreted as the ability of models to reproduce quantitatively data measured in steady uniform flows and qualitatively emerging behaviors that are observed corresponding to well-defined physical conditions.

The validation problem has been treated in Ref. 26 according to the idea that models should reproduce quantitatively empirical data which are generally obtained

in steady uniform flow and that, in addition, models should reproduce qualitatively collective emerging behaviors which are repetitively observed in appropriate physical conditions. An important aspect of the process is that both dynamics should not be artificially inserted into models, while they should be delivered by solution of problems.

The main source of quantitative empirical data is the so-called *velocity diagram* which reports the mean velocity versus density. This topic has been studied by various authors. We mention again that the interesting results are delivered in Refs. 212 and 214, where also dispersion of data is reported. Models should provide also the dependence of the velocity diagram on the quality of the venue where walkers mode, namely on α. Additional bibliography enlightened by sharp reasonings on the use of empirical data toward design and toning of models is reported in Ref. 210. Crowd behaviors related to interactions with signaling are reported in Ref. 161, while the bibliography on the interactions between crowd behaviors and specific features of the environment, e.g. smoke, is treated in Ref. 208.

This topic has been studied in Ref. 26, based on a model derived as a technical variation of Ref. 25. It has been shown that a simple model of interaction at the microscopic scale leads to a diagram parameterized with respect to α. This idea is analogous to that proposed in Ref. 104 for a simple discrete velocity model of vehicular traffic.

An additional analysis has been developed in Ref. 26 focusing on emerging behaviors in segregation problems which arise in counter-flows in corridors, where it has been observed that walkers which initially move randomly organize themselves into lanes. This is precisely what has been described by the model which has shown also how the segregation depends on the density of the crowd and on the emotional state of the walkers.

All models either based on discrete velocity assumptions or on the use of a continuous velocity require a detailed analysis of interactions at the microscopic scale, namely, on the analysis of the various strategies expressed by walkers to select their trajectories, choose, segregated paths, and react to the presence of other walkers. Interesting contributions have been delivered in Refs. 182 and 183, where the authors take advantage of empirical data delivered by experiments specifically organized to achieve the aforementioned interpretations.

Granular behaviors and contact problems have been studied in Refs. 102 and 175, while the implementation of the dynamics at the microscopic scale into possible hierarchies of models has been developed in Refs. 87 and 88. Additional research activity of the collection and interpretation of empirical data on crowd behaviors is reported in Refs. 80, 81 and 148.

All papers which have been reviewed above account for the emotional state of the walkers spans from the lowest value $u = 0$, which corresponds to the absence of walking desire, to the highest value $u = 1$ which corresponds to the highest level of stress. This state, as mentioned, increases the speed, but also the attraction toward the main stream. However, a very important problem has been left open, namely the

propagation in space of the emotional states by contagion/communication, in the crowd. Indeed, this propagation can play an important role in the case of incidents or of evacuation indications as an excess of stress, may be to perception of danger, might generate overcrowding. This topic has been initiated in Refs. 35 and 228 in the simple case of one-dimensional motion. The contagion dynamics is modeled by a consensus interaction somehow analogous to the BGK model of the Boltzmann equation.[68]

However, the modeling of contagion should account for communications by vocal or visual signs of walkers who transfer emotional state across the crowd. It is a problem of collective learning[53] which can induce significant modification in the overall self-organization, and hence on the collective dynamics, of the crowd. A modeling approach has been developed in Ref. 27.

3.3. *Sample simulations*

Two specific case studies have been selected here to show, also with tutorial aims, how the output of models can be interpreted.

Let us firstly consider the so-called *segregation dynamics in counter-flows* which has been treated in Ref. 25. The mathematical model specifically refers to the structure Eq. (3.13). The initial state of the flow is supposed to be delivered by walkers that move toward two opposite directions in a corridor, the flow is initially randomly distributed corresponding to two groups of walkers moving in a corridor toward two opposite directions, while gradually in time walkers segregate into four streams, namely two main streams moving to left and right segregated, respectively to an upper and lower band, and two small lanes close to the walls moving toward directions opposite to the main streams. Segregation appears only if the number density of walkers is sufficiently high, otherwise for low densities walkers move randomly in the corridor. Patterns with the four lanes are shown in Fig. 1.

These simulations have been obtained in Ref. 25, where no artificial dynamics has been additionally inserted to induce the aforementioned segregation, and where the following band-index has been proposed in the said paper to account

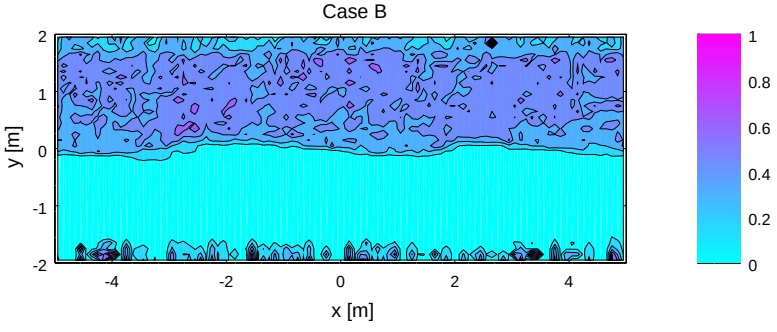


Fig. 1. Segregation into four lanes.

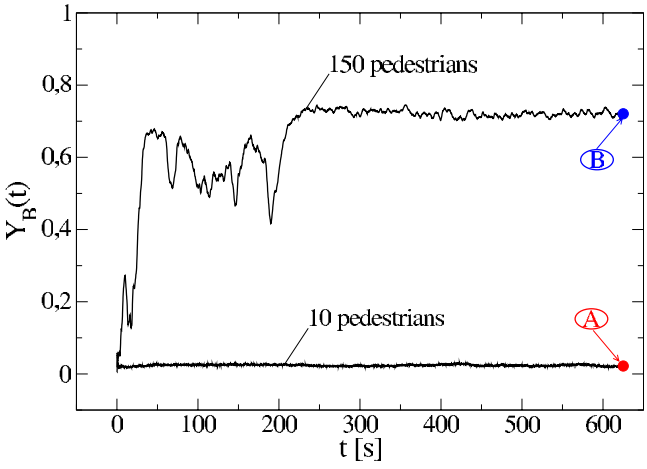


Fig. 2. Band Index for two walkers densities.

quantitatively assess the band index:

$$Y_{\text{B}}(t) = \frac{1}{L_x L_y} \int_0^{L_y} \left| \int_0^{L_x} \frac{\rho_1(t,\mathbf{x}) - \rho_2(t,\mathbf{x})}{\rho_1(t,\mathbf{x}) + \rho_2(t,\mathbf{x})} dx \right| dy. \tag{3.15}$$

Figure 2 shows how the Band Index reaches an asymptotic value after a transient time. In the case of 150 walkers the value is high and anticipates the segregation of lines, while in the case of 10 walkers the value of the Band Index keeps the initial value which corresponds to chaotic motion. These simulations have been used also to show the ability of the model to depict observed flow patterns, which are systematically observed, to be interpreted as emerging behavior. Indeed, this is one of the necessary steps toward the validation of models.[26]

The second case study refers to the evacuation dynamics from the platform of a metro-station, where individuals are, at initial time, uniformly distributed over the platform as shown in Fig. 3. Subsequently, due to an evacuation signal, people start moving towards the two exits as described by the model. Figure 3 shows the flow patterns at two subsequent times, say t_1 and $t_2 = 3\,t_1$. Simulations enlighten the densities of the flow patterns which indicate the onset and expansion of high density areas. Indeed, the ability of models to depict high density patterns can contribute to managing crisis situations.

3.4. *Critical analysis toward perspectives*

A review of various achievements on the modeling of vehicular traffic and human crowds has been presented in the previous subsections. The survey indicates that, despite some interesting achievements, a broad variety of problems appears to be still open to be viewed as possible research perspectives. However, rather than

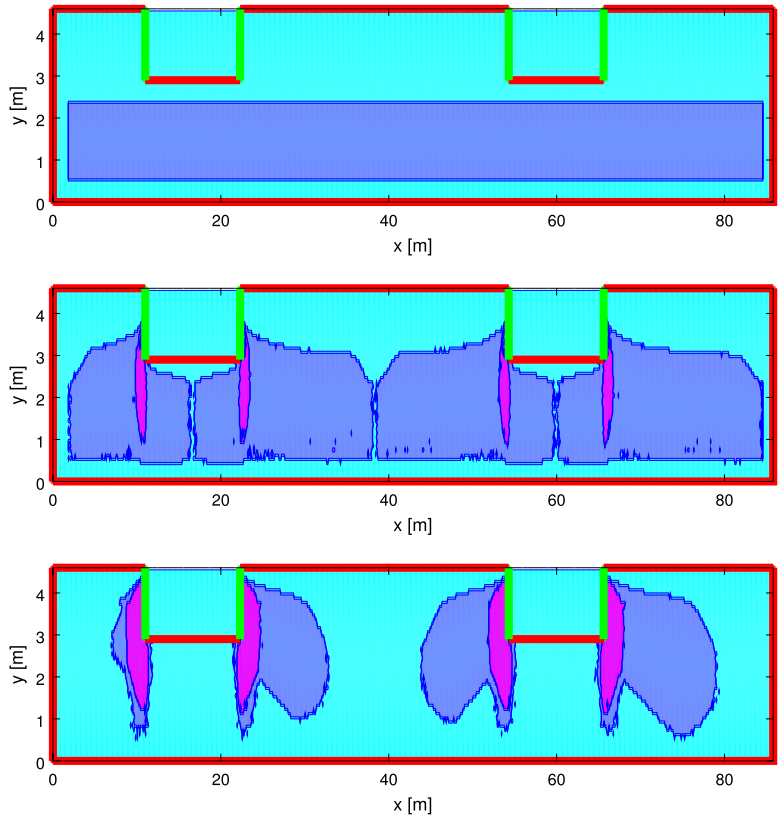


Fig. 3. From the top to the bottom: the initial distribution over the platform; the flow patterns at $t = t_1$ and the flow patterns at $t = 3t_1$.

presenting a long list of open problems we have selected a specific research perspective which can be addressed both to vehicular traffic and crowd dynamics.

An interesting open problem consists in showing how the general tools, reviewed in Sec. 3, can be developed into a *systems approach* and this more general tool can be used to support safety problems, where mathematical and computational tools can be addressed to support crisis managers and designers of venues according to safety requirements.[140,141,206,207,211,229,230]

A systems approach should account for the modeling hallmarks listed below, where the presentation enlightens the technical differences related to the two possible applications, namely traffic and crowds.

(1) The overall venue is constituted by a network of areas each of them labeled by a well-defined value of the parameter α corresponding to the quality of the streets or (venue). Moving across a network of streets (or of venues) implies moving through different values of the said parameter by junctions to be properly modeled.[45,78]

(2) Tuning of the parameters α should account that the quality of each area can be time dependent due to variable quality of the environment. For instance the presence of fog or smoke reduces the quality and hence the speed.
(3) The overall geometry of networks can be technically modified to account for different modifications induced by incidents, for instance the closure of stress or venues can be induced by incidents.
(4) Different qualities of drivers or walkers can be accounted for. In the case of crowds it might include the presence of individuals with limited mobility and trained leaders who are supposed to contribute to the evacuation process.

A systems approach, followed by simulations, can contribute to managing networks. As an example, in the case of crowds, the optimization of the design of geometry of venues by testing how different designs of the venue, however consistent with the overall design of buildings, induce safe evacuation conditions. Hence the approach appears necessary to support crisis managing by means of platforms where the process of selecting the most appropriate actions toward safety can be developed by predictive engines which refer the real flow dynamics to a database, where a huge number of simulations are stored. Some perspective ideas to achieve this important result have been proposed in Ref. [22].

In the case of vehicular traffic, the systems approach contributes to the modeling of the dynamics over networks and the subsequent optimization problems can be focused on the objective of reducing the time spent on roads with implication also to reduction of pollution of the environment.

The approach, in both cases vehicles and crowds, can take advantage of a multiscale vision which consists in developing a bottom-up derivation of models, where a detailed modeling of individual based interactions, namely at the microscopic scale, are used firstly to derive individual based, namely at the microscopic scale and, subsequently, to model the interaction terms η and $\mathcal{A}$ to derive kinetic type models, namely at the mesoscopic scale.[12] Subsequently, asymptotic methods can be developed to obtain hydrodynamical models from the underlying description at the lower scale. It is worth mentioning that an interesting research path has been initiated in Ref. [44], where a dynamics with interactions between vehicles and walkers has been studied referring to flow networks. A multiscale vision can definitely contribute to develop this interesting idea.

The micro-macro derivation looks ahead to a unified approach to physical sciences as inspired by the sixth Hilbert problem.[146] In fluid dynamics, this problem has been interpreted as the derivation of hydrodynamical models from the description delivered by the Boltzmann equation, the literature on the application of the approach to large systems of active particles can be recovered in Refs. [50] and [51]. This perspective refers both models of vehicular traffic and crowd dynamics, where some very preliminary results have been recently achieved,[17,21] however, limited to models in unbounded domains. This result has been achieved as a direct development of an approach initiated, as reviewed in Ref. [20], referring to multicellular

systems by considering a stochastic perturbation of the spatially homogeneous dynamics. Then, an expansion of the solution to the system is generated and properly truncated by ad hoc techniques, for instance moment closure or a detailed analysis of trend to equilibrium.

However, dealing with the same problem in domains with boundaries cannot be straightforwardly obtained by the same approach. In fact, the presence of obstacles and walls generates large deviations from equilibrium which cannot be treated in a general framework of small perturbation. Hence, it ought to be regarded as an open problems which requires mathematical tools to be properly treated. In addition to this approach, various other examples of limits based on averaging methods can be mentioned.[48,90,142]

4. Modeling on Swarming and Flocking of Agents

Collective behaviors of self-propelled agent (particle) systems often appear in our natural biological, chemical and physical systems, e.g. flocking of birds, schooling of fish, and herding of sheep are prototype examples of such collective phenomena. In these examples, self-propelled agents are willing to align with their neighboring agents. In this sequel, we use a jargon "*flocking*" to denote a phenomenon in which agent's velocities tend to common velocity asymptotically. From the pioneering work of Vicsek,[225] there have been many mechanical models describing flocking behavior of self-propelled systems. Among many flocking models in literature, our main interest in this section lies on the well-known Cucker–Smale (CS) model[84] which is one of the most successful model in applied mathematics. The CS model describes dynamic evolution of position and velocity variables of agents in d-dimensional Euclidean space. Let $(\mathbf{x}_i, \mathbf{v}_i) \in \mathbb{R}^{2d}$ be the position and velocity of the ith CS particle. Then, their dynamics is governed by the Cauchy problem to the CS model:

$$\begin{aligned} &\frac{d\mathbf{x}_i}{dt} = \mathbf{v}_i, \quad t > 0, \quad i = 1, \dots, N, \\ &\frac{d\mathbf{v}_i}{dt} = \frac{\kappa}{N} \sum_{j=1}^{N} \psi(\|\mathbf{x}_i - \mathbf{x}_j\|)(\mathbf{v}_j - \mathbf{v}_i), \\ &(\mathbf{x}_i(0), \mathbf{v}_i(0)) = (\mathbf{x}_{i0}, \mathbf{v}_{i0}). \end{aligned} \tag{4.1}$$

Here, $\|\cdot\|$ is the standard ℓ^2-norm in $\mathbb{R}^d$, the non-negative constant κ is called the coupling strength (gain) which characterizes the overall intensity of communication, and the function $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ is called the communication weight (or communication kernel). This function is in general nonincreasing, bounded and Lipschitz continuous function. The typical example of communication weight function is the CS type communication weight, depending on the parameter $\beta \in [0, \infty)$:

$$\psi_\beta(r) := \frac{1}{(1 + r^2)^{\beta/2}}. \tag{4.2}$$

The non-increasing property of communication weight implies that the intensity of communication becomes weaker as the distance between agents getting farther. The velocity flocking algorithm is encoded in the second equation of (4.1). At any time t, agents look around all other neighboring agents and adjust its velocity according to the weighted sum of a relative velocities. Thus, all agents can adjust their velocities toward the common value.

The kinetic theory for the flocking of CS particles was first introduced in Ref. 134, where a Vlasov type kinetic equation was derived by using standard BBGKY hierarchy argument based on the *molecular chaos* assumption which enable to close hierarchical system at the level of the one-particle distribution function. Let $f = f(t, \mathbf{x}, \mathbf{v})$ be the one-particle distribution function of the CS ensemble with velocity $\mathbf{v}$ at position $\mathbf{x}$, at time t. Then the kinetic CS equation reads as follows.

$$\begin{aligned} &\partial_t f + \mathbf{v} \cdot \nabla_{\mathbf{x}} f + \nabla_{\mathbf{v}} \cdot [F[f]f] = 0, \quad (\mathbf{x}, \mathbf{v}) \in \mathbb{R}^{2d}, \quad t > 0, \\ &F[f](t, \mathbf{x}, \mathbf{v}) := \kappa \int_{\mathbb{R}^{2d}} \psi(\|\mathbf{x} - \mathbf{x}^*\|)(\mathbf{v}^* - \mathbf{v}) f(t, \mathbf{x}^*, \mathbf{v}^*) d\mathbf{x}^* \, d\mathbf{v}^*. \end{aligned} \tag{4.3}$$

Although the CS flocking model has been quite successful for the description of flocking behavior, obviously, it has several defects, to name a few, it can only describe a mechanical movement of agents, ignoring the characteristics inherent in each agent. To resolve this deficiency, a new invariant of "Thermomechanical Cucker–Smale (TCS) model" was proposed in Ref. 133 motivated from the gas mixture of fluids. The TCS model incorporates the additional variable, called a *temperature* other than the mechanical variables $\mathbf{x}$ and $\mathbf{v}$. The exact TCS model can be written as

$$\begin{aligned} &\frac{d\mathbf{x}_i}{dt} = \mathbf{v}_i, \quad (\mathbf{x}_i, \mathbf{v}_i) \in \mathbb{R}^{2d}, \quad t > 0, \\ &\frac{d\mathbf{v}_i}{dt} = \frac{\kappa_1}{N} \sum_{j=1}^{N} \phi(\|\mathbf{x}_i - \mathbf{x}_j\|) \left(\frac{\mathbf{v}_j}{\theta_j} - \frac{\mathbf{v}_i}{\theta_i} \right), \\ &\frac{d\theta_i}{dt} = \frac{\kappa_2}{N} \sum_{j=1}^{N} \zeta(\|\mathbf{x}_i - \mathbf{x}_j\|) \left(\frac{1}{\theta_i} - \frac{1}{\theta_j} \right), \\ &(\mathbf{x}_i(0), \mathbf{v}_i(0), \theta_i(0)) = (\mathbf{x}_{i0}, \mathbf{v}_{i0}, \theta_{i0}). \end{aligned} \tag{4.4}$$

Here, $\theta_i = \theta_i(t) \in \mathbb{R}_+$ represents the temperature of the ith TCS particle. The coupling strengths κ_1 and κ_2 and communication weights ϕ and ζ are also defined similar to those of the CS model. Note that the TCS model is consistent with the physical principles such as Galilean invariance and entropy principle, and is reduced to the CS model (4.1) for the constant common temperature $\theta_i(t) = \theta^\infty$.

A BBGKY hierarchy can also be applied to the TCS model to yield a kinetic equation for one-particle distribution function $f = f(t, \mathbf{x}, \mathbf{v}, \theta)$ with thermomechanical observable $(\mathbf{x}, \mathbf{v}, \theta)$:

$$\partial_t f + \mathbf{v} \cdot \nabla_{\mathbf{x}} f + \nabla_{\mathbf{v}} \cdot [F[f]f] + \partial_\theta [G[f]f] = 0, \quad t > 0, \ (\mathbf{x}, \mathbf{v}, \theta) \in \mathbb{R}^{2d} \times \mathbb{R}_+,$$

$$F[f](t,\mathbf{x},\mathbf{v},\theta) := \kappa_1 \int_{\mathbb{R}^{2d}\times\mathbb{R}_+} \phi(\|\mathbf{x}-\mathbf{x}^*\|)\left(\frac{\mathbf{v}^*}{\theta^*}-\frac{\mathbf{v}}{\theta}\right) f(t,\mathbf{x}^*,\mathbf{v}^*,\theta^*)d\mathbf{x}^*d\mathbf{v}^*d\theta^*,$$

$$G[f](t,\mathbf{x},\theta) := \kappa_2 \int_{\mathbb{R}^{2d}\times\mathbb{R}_+} \zeta(\|\mathbf{x}-\mathbf{x}^*\|)\left(\frac{1}{\theta}-\frac{1}{\theta^*}\right) f(t,\mathbf{x}^*,\mathbf{v}^*,\theta^*)d\mathbf{x}^*d\mathbf{v}^*d\theta^*. \tag{4.5}$$

4.1. *Emergent dynamics of CS ensemble*

In this subsection, we will briefly survey the emergent dynamics of the CS flocking models, particle model (4.1), kinetic model (4.3) and hydrodynamic model (4.13).

4.1.1. *The particle CS model*

In this part, we present some selective results on the emergent dynamics for the particle CS model (4.1), including flocking estimates and uniform stability which are connected to the kinetic CS equation. First, we recall the concept of the (global (mono-cluster)) flocking for (4.1) as follows.

Definition 4.1. Let $\mathcal{P} := \{(\mathbf{x}_i,\mathbf{v}_i)\}_{i=1}^N$ be a solution to the CS model (4.1). Then, the ensemble $\mathcal{P}$ exhibits a (global (mono-cluster)) flocking, if and only if the following two relations hold.

(1) The relative positions are uniformly bounded:

$$\sup_{0\le t<\infty}\max_{1\le i\ne j\le N}\|\mathbf{x}_i(t)-\mathbf{x}_j(t)\|<\infty.$$

(2) The relative velocities decay to zero:

$$\lim_{t\to\infty}\max_{1\le i\ne j\le N}\|\mathbf{v}_i(t)-\mathbf{v}_j(t)\|=0.$$

From the CS model (4.1), it is easy to see that the system is Galilean invariant. Hence, without loss of generality, we may assume that the center of mass is fixed at the origin.

$$\mathbf{x}_c := \frac{1}{N}\sum_{i=1}^N \mathbf{x}_i = 0, \quad \mathbf{v}_c := \frac{1}{N}\sum_{i=1}^N \mathbf{v}_i = 0.$$

Under these assumptions, the flocking conditions in Definition 4.1 are equivalent to

$$\sup_{0\le t<\infty}\|\mathbf{x}_i(t)\|<\infty, \quad \lim_{t\to\infty}\|\mathbf{v}_i(t)\|=0, \quad 1\le i\le N.$$

A result on the emergence of flocking for the CS model was first reported in the paper of Cucker and Smale[84] for the communication weight of type (4.2), and this result was further generalized in Refs. 131 and 134. In particular, in Ref. 131, authors introduced a Lyapunov functional approach for a general communication

weight, and we briefly discuss the Lyapunov functional approach below. For the CS ensemble $\{(\mathbf{x}_i, \mathbf{v}_i)\}_{i=1}^N$, we set

$$\mathbb{X} := (\mathbf{x}_1, \dots, \mathbf{x}_N), \quad \mathbb{V} := (\mathbf{v}_1, \dots, \mathbf{v}_N),$$

$$\|\mathbb{X}\|^2 := \sum_{i=1}^N \|\mathbf{x}_i\|^2, \quad \|\mathbb{V}\| := \sum_{i=1}^N \|\mathbf{v}_i\|^2.$$

Then, these total ℓ^2-norms satisfy the following system of dissipative differential inequality[131]:

$$\left|\frac{d}{dt}\|\mathbb{X}\|\right| \le \|\mathbb{V}\|, \quad \frac{d}{dt}\|\mathbb{V}\| \le -\kappa\psi(\|\sqrt{2}\mathbb{X}\|)\|\mathbb{V}\|, \quad t > 0. \tag{4.6}$$

Now, we introduce a Lyapunov functional $\mathcal{L}_\pm$:

$$\mathcal{L}_\pm(t) := \|\mathbb{V}(t)\| \pm \kappa \int_0^{\|\mathbb{X}(t)\|} \psi(\sqrt{2}s)ds, \quad t \ge 0. \tag{4.7}$$

Then, we differentiate (4.11) with respect to (4.11) using (4.10) to get

$$\frac{d}{dt}\mathcal{L}_\pm(t) = -\psi(\sqrt{2}\|\mathbb{X}(t)\|)\left(\left|\frac{d}{dt}\|\mathbb{X}(t)\|\right| - \|\mathbb{V}(t)\|\right) \le 0,$$

which results in the stability estimate:

$$\|\mathbb{V}(t)\| + \kappa\left|\int_{\|\mathbb{X}(0)\|}^{\|\mathbb{X}(t)\|} \psi(\sqrt{2}s)ds\right| \le \|\mathbb{V}(0)\|, \quad t \ge 0.$$

Theorem 4.1. (Ref. 131) *Let $\{(\mathbf{x}_i, \mathbf{v}_i)\}_{i=1}^N$ be a solution to the CS model (4.1) with initial data $\{(\mathbf{x}_{i0}, \mathbf{v}_{i0})\}_{i=1}^N$ satisfying*

$$\|\mathbb{V}(0)\| < \kappa \int_{\|\mathbb{X}(0)\|}^\infty \psi(\sqrt{2}s)ds.$$

Then, there exists a $\mathbf{x}_M \ge 0$ such that

$$\|\mathbb{V}(0)\| = \kappa \int_{\|\mathbb{X}(0)\|}^{\mathbf{x}_M} \psi(\sqrt{2}s)ds, \quad \|\mathbb{X}(t)\| \le \mathbf{x}_M, \quad \|\mathbb{V}(t)\| \le \|\mathbb{V}(0)\|e^{-\kappa\psi(\sqrt{2}\mathbf{x}_M)t} t \ge 0.$$

Remark 4.1. Note that the result in Theorem 4.1 requires the condition on total ℓ^2-norm of initial data, namely $|\mathbf{x}|$ and $|\mathbf{v}|$, which depends on the number of particles N. However, to attain a uniform-in-time mean-field limit which will be presented in the next subsection, we need the flocking estimates with the condition on the initial data that is regardless of the number of particles. There have been several results of this kind, but we provide the most recent result[124] here. We present the flocking estimate in terms of the diameter of ensemble $D_\mathbf{x}$ and $D_\mathbf{v}$ defined as

$$D_\mathbf{x}(t) := \max_{1\le i\ne j\le N} \|\mathbf{x}_i - \mathbf{x}_j\|, \quad D_\mathbf{v}(t) := \max_{1\le i\ne j\le N} \|\mathbf{v}_i - \mathbf{v}_j\|.$$

Then, if initial data satisfy

$$D_\mathbf{v}(0) < \kappa \int_{D_\mathbf{x}(0)}^\infty \psi(s)ds,$$

there exists a positive constant $\mathbf{x}^\infty$ such that

$$\sup_{0\le t<\infty} D_{\mathbf{x}}(t) \le \mathbf{x}^\infty, \quad D_{\mathbf{v}}(t) \le D_{\mathbf{v}}(0)e^{-\kappa\psi(\mathbf{x}^\infty)t}, \quad t>0.$$

Finally, we close this subsection by presenting the uniform stability result. This uniform stability result is necessary to derive uniform-in-time mean-field limit from the particle CS model to the kinetic CS equation.

Theorem 4.2. (Ref. 124) *Suppose that initial data* $\{(\mathbf{x}_{i0},\mathbf{v}_{i0})\}_{i=1}^N$ *and* $\{(\tilde{\mathbf{x}}_{i0},\tilde{\mathbf{v}}_{i0})\}_{i=1}^N$ *satisfy*

$$\sum_{i=1}^N \mathbf{v}_{i0} = \sum_{i=1}^N \tilde{\mathbf{v}}_{i0} = 0, \quad \kappa > \max\left\{\frac{D_{\mathbf{v}}(0)}{\int_{D_{\mathbf{x}}(0)}^\infty \psi(s)ds}, \frac{D_{\tilde{\mathbf{v}}}(0)}{\int_{D_{\tilde{\mathbf{x}}(0)}}^\infty \psi(s)ds}\right\},$$

and let $\{(\mathbf{x}_i,\mathbf{v}_i\})_{i=1}^N$ *and* $\{(\tilde{\mathbf{x}}_i,\tilde{\mathbf{v}}_i\})_{i=1}^N$ *be two solutions to the CS model with the initial data* $\{(\mathbf{x}_{i0},\mathbf{v}_{i0})\}_{i=1}^N$ *and* $\{(\tilde{\mathbf{x}}_{i0},\tilde{\mathbf{v}}_{i0})\}_{i=1}^N$, *respectively. Then, there exists a generic constant* G *independent of* t *such that*

$$\begin{aligned}&\left(\sum_{i=1}^N \|\mathbf{x}_i-\tilde{\mathbf{x}}_i\|_p^q\right)^{1/q} + \left(\sum_{i=1}^N \|\mathbf{v}_i-\tilde{\mathbf{v}}_i\|_p^q\right)^{1/q}\\ &\qquad \le G\left[\left(\sum_{i=1}^N \|\mathbf{x}_{i0}-\tilde{\mathbf{x}}_{i0}\|_p^q\right)^{1/q} + \left(\sum_{i=1}^N \|\mathbf{v}_{i0}-\tilde{\mathbf{v}}_{i0}\|_p^q\right)^{1/q}\right],\end{aligned}$$

where $\|\cdot\|_p$ *denotes* ℓ^p*-norm of vector in* $\mathbb{R}^d$.

Remark 4.2. Note that for any given autonomous system $\dot{\mathbf{x}} = f(\mathbf{x})$ with smooth bounded function f, the difference between solutions $\mathbf{x}(t)$ and $\tilde{\mathbf{x}}(t)$ originated from $\mathbf{x}_0$ and $\tilde{\mathbf{x}}_0$ has stability up to any finite time:

$$\|\mathbf{x}(t)-\tilde{\mathbf{x}}(t)\| \le C\|\mathbf{x}_0-\tilde{\mathbf{x}}_0\|, \quad 0\le t\le T, \tag{4.8}$$

where C may depend on the final time T. The uniform stability result in Theorem 4.2 implies that stability estimate (4.8) holds for $T=\infty$ and the constant C does not depend on time anymore. This uniform stability estimate uses the flocking estimate in Theorem 4.1. See Ref. 124 for detailed derivation.

4.1.2. *The kinetic CS model*

In this part, we briefly sketch the kinetic theory for the CS ensemble. As mentioned before, the formal derivation of kinetic equation (4.3) from particle CS flocking model (4.1) can be done by applying BBGKY hierarchy to the Liouville equation for N-particle distribution function. We review the derivation of the kinetic CS equation introduced in Ref. 134.

Let $f^N = f^N(t,\mathbf{x}_1,\mathbf{v}_1,\ldots,\mathbf{x}_N,\mathbf{v}_N)$ be the N-particle probability density function on the state space $\mathbb{R}^{2Nd}$. Since all particles cannot be distinguished from each

other, we may assume that the density function f^N is symmetric in the exchange transform $(i \leftrightarrow j)$:

$$f^N(t, \dots, \mathbf{x}_i, \mathbf{v}_i, \dots, \mathbf{x}_j, \mathbf{v}_j, \dots) = f^N(t, \dots, \mathbf{x}_j, \mathbf{v}_j, \dots, \mathbf{x}_i, \mathbf{v}_i, \dots). \tag{4.9}$$

Then, N-particle distribution function f^N satisfies the Liouville equation:

$$\partial_t f^N + \sum_{i=1}^N \mathbf{v}_i \cdot \nabla_{\mathbf{x}_i} f^N + \frac{\kappa}{N} \sum_{i=1}^N \nabla_{\mathbf{v}_i} \cdot \left(\sum_{i=1}^N \psi(\|\mathbf{x}_i - \mathbf{x}_j\|)(\mathbf{v}_j - \mathbf{v}_i) f^N \right) = 0. \tag{4.10}$$

Now, we consider the one-particle density function, i.e. marginal density function $f^N(t, \mathbf{x}_1, \mathbf{v}_1)$ defined as

$$f^N(t, \mathbf{x}_1, \mathbf{v}_1) := \int_{\mathbb{R}^{2(N-1)d}} f^N(t, \mathbf{x}_1, \mathbf{v}_1, \mathbf{x}_-, \mathbf{v}_-) d\mathbf{x}_- \, d\mathbf{v}_-,$$

where

$$(\mathbf{x}_-, \mathbf{v}_-) := (\mathbf{x}_2, \mathbf{v}_2, \dots, \mathbf{x}_N, \mathbf{v}_N).$$

If we integrate (4.10) over $(\mathbf{x}_-, \mathbf{v}_-)$, we can obtain the PDE for the marginal density function. More precisely, the transportation part becomes

$$\int_{\mathbb{R}^{2(N-1)d}} \sum_{i=1}^N \mathbf{v}_i \cdot \nabla_{\mathbf{x}_i} f^N \, d\mathbf{x}_- \, d\mathbf{v}_- = \mathbf{v}_1 \cdot \nabla_{\mathbf{x}_1} f^N(t, \mathbf{x}_1, \mathbf{v}_1),$$

and the alignment force term can be calculated in a similar manner as

$$\begin{aligned} &\frac{\kappa}{N} \sum_{i=1}^N \int_{\mathbb{R}^{2(N-1)d}} \sum_{j=1}^N \nabla_{\mathbf{v}_i} \cdot (\psi(\|\mathbf{x}_i - \mathbf{x}_j\|)(\mathbf{v}_j - \mathbf{v}_i) f^N) d\mathbf{x}_- \, d\mathbf{v}_- \\ &\quad = \frac{\kappa}{N} \int_{\mathbb{R}^{2(N-1)d}} \sum_{j=2}^N \nabla_{\mathbf{v}_1} \cdot (\psi(\|\mathbf{x}_1 - \mathbf{x}_j\|)(\mathbf{v}_j - \mathbf{v}_1) f^N) d\mathbf{x}_- \, d\mathbf{v}_- \\ &\quad = \frac{\kappa(N-1)}{N} \int_{\mathbb{R}^{2(N-1)d}} \nabla_{\mathbf{v}_1} \cdot (\psi(\|\mathbf{x}_1 - \mathbf{x}_2\|)(\mathbf{v}_2 - \mathbf{v}_1) f^N) d\mathbf{x}_- \, d\mathbf{v}_-, \end{aligned}$$

where the last relation is due to (4.9). Now, we define the two-particle density function as

$$g^N(\mathbf{x}_1, \mathbf{v}_1, \mathbf{x}_2, \mathbf{v}_2, t) := \int_{\mathbb{R}^{2d(N-2)}} f^N \, d\mathbf{x}_3 \, d\mathbf{v}_3 \, \dots d\mathbf{x}_N \, d\mathbf{v}_N.$$

Then, the kinetic equation describing the dynamics of one-particle density function becomes

$$\begin{aligned} &\partial_t f^N + \mathbf{v}_1 \cdot \nabla_{\mathbf{x}_1} f^N + \kappa \left(1 - \frac{1}{N}\right) \nabla_{\mathbf{v}_1} \\ &\qquad \cdot \int_{\mathbb{R}^{2d}} \psi(\|\mathbf{x}_1 - \mathbf{x}_2\|)(\mathbf{v}_2 - \mathbf{v}_1) g^N \, d\mathbf{x}_2 \, d\mathbf{v}_2 = 0. \end{aligned} \tag{4.11}$$

Now, we take a mean-field limit ($N \to \infty$) and assume that there exist limiting functions $f(\mathbf{x}_1, \mathbf{v}_1, t)$ and $g(\mathbf{x}_1, \mathbf{v}_1, \mathbf{x}_2, \mathbf{v}_2, t)$ satisfying

$$f(\mathbf{x}_1, \mathbf{v}_1, t) := \lim_{N\to\infty} f^N(\mathbf{x}_1, \mathbf{v}_1, t),$$
$$g(\mathbf{x}_1, \mathbf{v}_1, \mathbf{x}_2, \mathbf{v}_2, t) := \lim_{N\to\infty} g^N(\mathbf{x}_1, \mathbf{v}_1, \mathbf{x}_2, \mathbf{v}_2, t).$$

Then, we take a formal limit $N \to \infty$ on (4.11) to obtain

$$\partial_t f + \mathbf{v}_1 \cdot \nabla_{\mathbf{x}_1} f + \kappa \nabla_{\mathbf{v}_1} \cdot \int_{\mathbb{R}^{2d}} \psi(\|\mathbf{x}_1 - \mathbf{x}_2\|)(\mathbf{v}_2 - \mathbf{v}_1) g \, d\mathbf{x}_2 \, d\mathbf{v}_2 = 0. \quad (4.12)$$

Since the kinetic equation (4.12) for $f(\mathbf{x}_1, \mathbf{v}_1, t)$ is not closed, we need further assumption to close equation (4.12) in an appropriate manner. For this closure condition, we use the *molecular chaos* assumption which means

$$g(\mathbf{x}_1, \mathbf{v}_1, \mathbf{x}_2, \mathbf{v}_2, t) = f(\mathbf{x}_1, \mathbf{v}_1, t) f(\mathbf{x}_2, \mathbf{v}_2, t).$$

Now, after relabeling variables $(\mathbf{x}_1, \mathbf{v}_1)$ and $(\mathbf{x}_2, \mathbf{v}_2)$ to $(\mathbf{x}, \mathbf{v})$ and $(\mathbf{x}^*, \mathbf{v}^*)$, respectively, we can close kinetic equation (4.12) as

$$\partial_t f + \mathbf{v} \cdot \nabla_{\mathbf{x}} f + \nabla_{\mathbf{v}} \cdot [F[f]f] = 0, \quad (\mathbf{x}, \mathbf{v}) \in \mathbb{R}^{2d}, \quad t > 0,$$
$$F[f](t, \mathbf{x}, \mathbf{v}) := \kappa \int_{\mathbb{R}^{2d}} \psi(\|\mathbf{x} - \mathbf{x}^*\|)(\mathbf{v}^* - \mathbf{v}) f(t, \mathbf{x}^*, \mathbf{v}^*) d\mathbf{x}^* \, d\mathbf{v}^*,$$

which is exactly the same as the previously introduced kinetic CS equation (4.3).

There have been numerous results about the kinetic CS equation. Among them, the first result about the kinetic CS equation is the finite-in-time stability and the rigorous proof of finite-time mean-field limit.[131] To describe the result in a precise way, we introduce several notations. We denote the set of all positive Radon measure by $\mathcal{M}(\mathbb{R}^{2d})$, and denote by d_{BL} the bounded Lipschitz distance between measures in $\mathcal{M}$:

$$d_{\mathrm{BL}}(\mu_1, \mu_2) := \sup_{g\in\mathcal{V}} \left| \int_{\mathbb{R}^{2d}} g d\mu_1 - \int_{\mathbb{R}^{2d}} g d\mu_2 \right|,$$

where

$$\mathcal{V} := \left\{ g : \mathbb{R}^{2d} \to \mathbb{R} : \|g\|_{L^\infty} \le 1 \text{ and } \mathrm{Lip}(g) := \sup_{\mathbf{x}\neq\mathbf{x}^*\in\mathbb{R}^{2d}} \frac{|g(\mathbf{x}) - g(\mathbf{x}^*)|}{\|\mathbf{x} - \mathbf{x}^*\|} \le 1 \right\}.$$

Theorem 4.3. (Ref. 134) *Suppose that initial measure $\mu_0 \in \mathcal{M}(\mathbb{R}^{2d})$ is compactly supported in $\mathbb{R}^{2d}$. Then, the following assertions hold.*

(1) *There exists a unique measure-valued solution $\mu = \mu(t)$ satisfying the kinetic CS equation (4.3) in the distributional sense.*

(2) *Suppose we choose a sequence of empirical measure* μ_0^N *of the form*

$$\mu_0^N = \frac{1}{N}\sum_{i=1}^N \delta_{(\mathbf{x}_i(0),\mathbf{v}_i(0))}$$

such that

$$\lim_{N\to\infty} d_{\mathrm{BL}}(\mu_0^N, \mu_0) = 0.$$

Consider the empirical measure $\mu^N(t)$ *generated from initial empirical measure* μ_0^N *and the particle CS dynamics*:

$$\mu^N(t) := \frac{1}{N}\sum_{i=1}^N \delta_{(\mathbf{x}_i(t),\mathbf{v}_i(t))}.$$

Then,

$$\lim_{N\to\infty} d_{\mathrm{BL}}(\mu^N(t), \mu(t)) = 0, \quad t \geq 0.$$

After this finite-time mean-field limit in Theorem 4.3, recently uniform-in-time mean-field limit was established in Ref. 124. Before we state the result on the uniform-in-time mean-field limit, we clarify the difference between the new result and finite time mean-field result in Theorem 4.3. The mean-field limit result in Theorem 4.3(2) holds only for fixed $t \geq 0$. Although this limit holds for arbitrary large time t, the decay depends on the final time t. In the following, we use the additional flocking estimates and asymptotic behavior to guarantee that, with a well-prepared initial data, the mean-field limit procedure can be made *uniform-in-time*, independent of the final time t. This is possible due to the uniform-in-time stability result in Theorem 4.2.

Theorem 4.4. (Ref. 124) *Suppose that initial measure* $\mu_0 \in \mathcal{M}(\mathbb{R}^{2d})$ *is compactly supported in* $\mathbb{R}^{2d}$, *and the coupling strength and communication weight satisfy*

$$\kappa > \frac{D_{\mathbf{v}}(0)}{\int_{D_{\mathbf{x}}(0)}^{\infty} \psi(s)ds},$$

where $D_{\mathbf{x}}$ *and* $D_{\mathbf{v}}$ *are defined parallel to those of the particle system*:

$$D_{\mathbf{x}}(t) := \sup_{\mathbf{x},\mathbf{x}^*\in \mathrm{supp}_{\mathbf{x}}\mu(t)} \|\mathbf{x}-\mathbf{x}^*\|, \quad D_{\mathbf{v}}(t) := \sup_{\mathbf{v},\mathbf{v}^*\in \mathrm{supp}_{\mathbf{v}}\mu(t)} \|\mathbf{v}-\mathbf{v}^*\|.$$

Then there exists a unique measure-valued solution $\mu = \mu(t)$ *to* (4.3) *with initial data* μ_0. *Moreover*, $\mu(t)$ *is approximated by the empirical measure* $\mu^N(t)$ *uniformly in time*:

$$\lim_{N\to\infty} \sup_{0\leq t<\infty} d_{\mathrm{BL}}(\mu^N(t), \mu(t)) = 0.$$

Remark 4.3. Note that the uniform-in-time mean-field limit holds for some class of initial data which guarantees the emergence of exponential flocking. In contrast, the finite-in-time mean-fields limit described in Theorem 4.3 holds for any initial data.

As in the particle CS model, the flocking behavior in kinetic level has been studied using a functional approach. In Refs. 64 and 134, the flocking behavior of the kinetic CS equation with communication weight (4.2) was studied and flocking estimate for general coupling strength was further investigated in Ref. 124. In the literature, the flocking estimate for the kinetic CS equation is done using the following Lyapunov functional:

$$\Lambda(\mu(t)) := \int_{\mathbb{R}^{2d}} \|\mathbf{v} - \mathbf{v}_c\|^2 d\mu(t),$$

where

$$\mathbf{v}_c := \int_{\mathbb{R}^{2d}} \mathbf{v}\, d\mu = \int_{\mathbb{R}^{2d}} \mathbf{v}\, d\mu_0,$$

which is usually assumed to be 0 thanks to the Galilean invariance. We provide the result in Ref. 124, which is the most general result about flocking estimate in the kinetic level.

Theorem 4.5. (Ref. 124) *Suppose that initial measure $\mu_0 \in \mathcal{M}(\mathbb{R}^{2d})$ has a compact support in $\mathbb{R}^{2d}$, and the coupling strength and communication weight satisfy*

$$\kappa > \frac{D_{\mathbf{v}}(0)}{\int_{D_{\mathbf{x}}(0)}^{\infty} \psi(s) ds},$$

and let $\mu = \mu(t)$ be a measure-valued solution to (4.3) with initial data μ_0. Then, there exists $\mathbf{x}_M$ and positive constant C such that

$$\Lambda(\mu(t)) \leq Ce^{-\kappa\psi(\mathbf{x}_M)t}\Lambda(\mu_0).$$

The kinetic CS equation coupled with various types of fluids is also studied in the serial works.[13,14] The coupling between particle ensemble with fluid, such as a flow of red blood cells in blood vessel or movement of birds through the air flow, can be easily found. In this context, it is highly desirable to study the coupling between particle ensemble with fluid. In the aforementioned works, a global existence of solutions of the kinetic CS model coupled with compressible/incompressible Navier–Stokes equation, as well as the large time behavior of the coupled system are also developed (see a recent survey article).[72] The numerical simulation for the kinetic CS equation coupled with fluid is also done.[15]

4.1.3. *The hydrodynamic CS model*

In this part, we briefly discuss the macroscopic description of the CS flocking. In order to compare with results from the kinetic level, the hydrodynamic model

for the CS model can be formally derived from the kinetic CS model using the moment system and the mono-kinetic ansatz[134] and this formal procedure was rigorously justified under the help of local alignment term in Ref. 109. The resulting hydrodynamic equation is a pressure-less Euler system with non-local source term:

$$\begin{aligned}
&\rho_t + \nabla_{\mathbf{x}} \cdot (\rho u) = 0, \quad t > 0, \ \ \mathbf{x} \in \mathbb{R}^d, \\
&(\rho u)_t + \nabla_{\mathbf{x}} \cdot (\rho u \otimes u) = \kappa \int_{\mathbb{R}^d} \psi(\|\mathbf{x} - \mathbf{x}^*\|)(u(\mathbf{x}^*) - u(\mathbf{x}))\rho(\mathbf{x})\rho(\mathbf{x}^*)d\mathbf{x}^*, \qquad (4.13) \\
&(\rho, u)|_{t=0} = (\rho_0, u_0).
\end{aligned}$$

We also refer to Sec. 6 for another hydrodynamic model arising from a suitable scaling limit. The well-posedness theory for system (4.13) was studied in Refs. 125 and 126. To describe flocking behavior in the macroscopic level, we again introduce the following Lyapunov functionals:

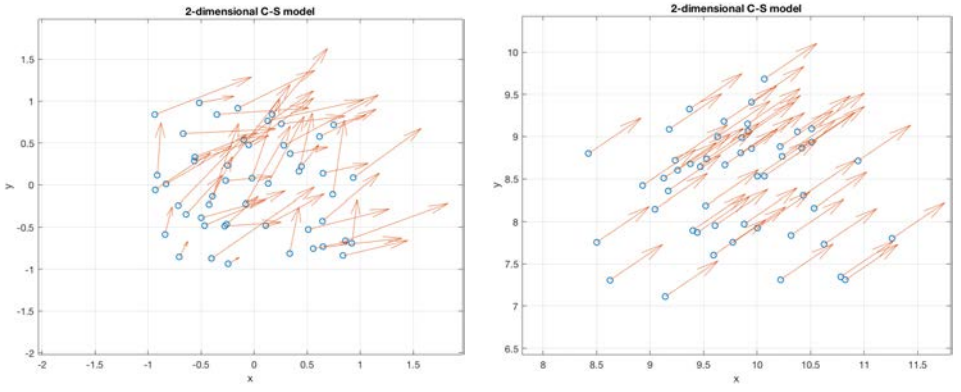
$$D_{\mathbf{x}}(t) := \sup_{\mathbf{x},\mathbf{x}^* \in \mathrm{supp}\rho(t)} \|\mathbf{x} - \mathbf{x}^*\|, \quad D_{\mathbf{u}}(t) := \sup_{\mathbf{x},\mathbf{x}^* \in \mathrm{supp}\rho(t)} \|u(\mathbf{x}, t) - u(\mathbf{x}^*, t)\|.$$

Theorem 4.6. (Ref. 72) *Suppose that the initial data, coupling strength and communication weight satisfy*

$$D_{\mathbf{u}}(0) < \kappa \int_{D_{\mathbf{x}}(0)}^{\infty} \psi(s)ds,$$

and let (ρ, u) *be any smooth solution to* (4.13) *with compactly supported initial data* (ρ_0, u_0). *Then, there exists a positive constant* C *such that*

$$\sup_{t \geq 0} D_{\mathbf{x}}(t) \leq C, \quad \lim_{t \to \infty} D_{\mathbf{u}}(t) = 0.$$



(a) Initial CS particle (b) CS particle at $t = 20$

Fig. 4. Particle CS model.

4.1.4. *Numerical simulations*

In this section, we provide several numerical simulations on the particle and kinetic CS model presenting flocking behaviors. Figure 4 shows the numerical simulation for the particle CS model. We did not take a mean zero velocity assumption to show flocking behavior much more clearly. Initially, velocity of each particle seems not aligned with each other (Fig. 4(a)). However, after sufficiently large time, all particles become aligned with the same velocity, which clearly shows the flocking behavior.

The kinetic CS equation is also simulated here (see Fig. 2). For a clear visualization, we tested the 1D CS equation with a spatially periodic domain. We take an initial data with an oval shape on the phase space. Then, as we can found in Fig. 5, velocity support shrinks and finally, CS ensemble is concentrated on the single velocity value.

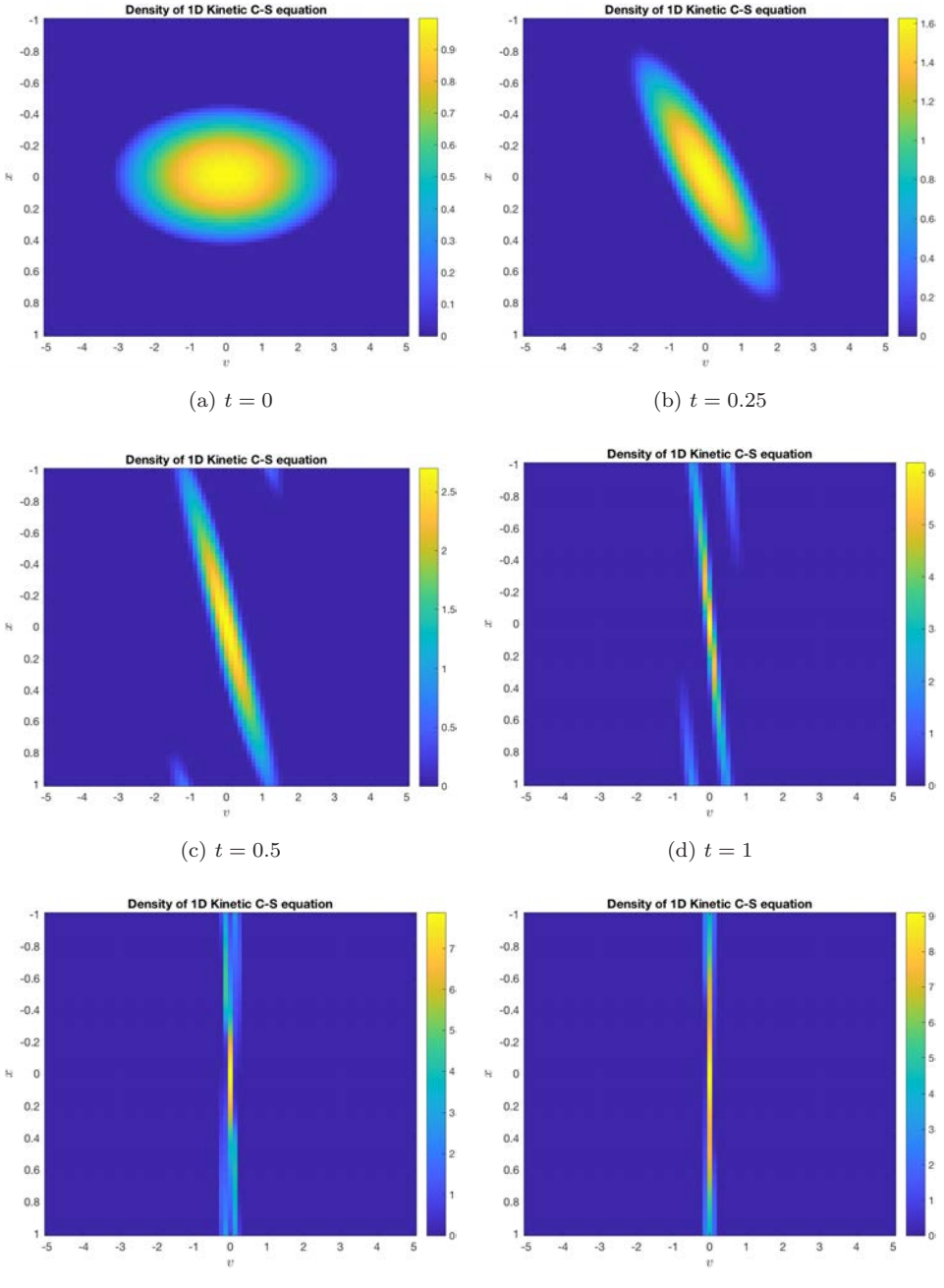
4.2. *Emergent dynamics of CS ensemble with temperature*

In this subsection, we will briefly survey the emergent dynamics of the thermomechanical CS flocking models, particle model (4.4), kinetic model (4.5) and hydrodynamic model (4.19). As we mentioned, the motivation for the TCS model is to consider a flocking model with *internal variable* such as a temperature. A new flocking model was introduced in Ref. 133 motivated by the analogous theory of gas mixtures in the context of rational thermodynamics. Below, we briefly review the heuristic derivation of the TCS model. To fix the idea, let $\rho_\alpha, \mathbf{v}_\alpha, \varepsilon_\alpha, q_\alpha$ and $\mathbf{t}_\alpha$ be the density, velocity, specific internal energy, heat flux and stress tensor of fluid α, respectively. Then, the dynamics of the mixture can be described by the following balance laws:

$$\begin{aligned}
&\partial_t \rho_\alpha + \nabla \cdot (\rho_\alpha \mathbf{v}_\alpha) = \tau_\alpha, \quad (\mathbf{x}, t) \in \mathbb{R}^d \times \mathbb{R}_+, \quad \alpha = 1, 2, \ldots, N, \\
&\partial_t (\rho_\alpha \mathbf{v}_\alpha) + \nabla \cdot (\rho_\alpha \mathbf{v}_\alpha \otimes \mathbf{v}_\alpha - \mathbf{t}_\alpha) = \mathbf{m}_\alpha, \\
&\partial_t \left(\frac{1}{2} \rho_\alpha \|\mathbf{v}_\alpha\|^2 + \rho_\alpha \varepsilon_\alpha \right) + \nabla \cdot \left\{ \left(\frac{1}{2} \rho_\alpha \|\mathbf{v}_\alpha\|^2 + \rho_\alpha \varepsilon_\alpha \right) \mathbf{v}_\alpha - \mathbf{t}_\alpha \mathbf{v}_\alpha + \mathbf{q}_\alpha \right\} = e_\alpha,
\end{aligned} \tag{4.14}$$

where τ_α, $\mathbf{m}_\alpha$ and e_α denote production terms from the interaction between constituents. We assume that the total mass, momentum and energy are conserved so that the sums of production terms are vanish.

$$\sum_{\alpha=1}^{N} \tau_\alpha = 0, \quad \sum_{\alpha=1}^{N} \mathbf{m}_\alpha = 0, \quad \sum_{\alpha=1}^{N} e_\alpha = 0.$$



(a) $t = 0$ (b) $t = 0.25$

(c) $t = 0.5$ (d) $t = 1$

(e) $t = 2$ (f) $t = 3$

Fig. 5. Kinetic CS model.

Under this condition, the global mixture quantities $\rho, \mathbf{v}, \varepsilon, \mathbf{t}$ and $\mathbf{q}$ defined as below

$$\rho := \sum_{\alpha=1}^{N} \rho_\alpha, \quad \mathbf{v} := \frac{1}{\rho}\sum_{\alpha=1}^{N} \rho_\alpha \mathbf{v}_\alpha,$$

$$\varepsilon := \frac{1}{\rho}\sum_{\alpha=1}^{N} \rho_\alpha \varepsilon_\alpha + \frac{1}{2\rho}\sum_{\alpha=1}^{N} \rho_\alpha \|\mathbf{u}_\alpha\|^2, \quad \mathbf{u}_\alpha := \mathbf{v}_\alpha - \mathbf{v},$$

$$\mathbf{t} := \sum_{\alpha=1}^{N} (\mathbf{t}_\alpha - \rho_\alpha \mathbf{u}_\alpha \otimes \mathbf{u}_\alpha), \quad \mathbf{q} := \sum_{\alpha=1}^{N} \left\{ \mathbf{q}_\alpha + \rho_\alpha \left(\varepsilon_\alpha + \frac{1}{2}\|\mathbf{u}_\alpha\|^2 \right) \mathbf{u}_\alpha - \mathbf{t}_\alpha \mathbf{u}_\alpha \right\}$$

satisfy conservation laws

$$\begin{aligned}
&\partial_t \rho + \nabla \cdot (\rho \mathbf{v}) = 0, \quad (\mathbf{x}, t) \in \mathbb{R}^d \times \mathbb{R}_+, \\
&\partial_t (\rho \mathbf{v}) + \nabla \cdot (\rho \mathbf{v} \otimes \mathbf{v} - \mathbf{t}) = 0, \\
&\partial_t \left(\frac{1}{2}\rho \|\mathbf{v}\|^2 + \rho\varepsilon \right) + \nabla \cdot \left\{ \left(\frac{1}{2}\rho \|\mathbf{v}\|^2 + \rho\varepsilon \right) \mathbf{v} - \mathbf{t}\mathbf{v} + \mathbf{q} \right\} = 0.
\end{aligned} \tag{4.15}$$

We combine two descriptions of dynamics of gas mixture (4.14) and (4.15) as

$$\begin{aligned}
&\partial_t \rho + \nabla \cdot (\rho \mathbf{v}) = 0, \quad (\mathbf{x}, t) \in \mathbb{R}^d \times \mathbb{R}_+, \\
&\partial_t (\rho \mathbf{v}) + \nabla \cdot (\rho \mathbf{v} \otimes \mathbf{v} - \mathbf{t}) = 0, \\
&\partial_t \left(\frac{1}{2}\rho \|\mathbf{v}\|^2 + \rho\varepsilon \right) + \nabla \cdot \left\{ \left(\frac{1}{2}\rho \|\mathbf{v}\|^2 + \rho\varepsilon \right) \mathbf{v} - \mathbf{t}\mathbf{v} + \mathbf{q} \right\} = 0, \\
&\partial_t \rho_i + \nabla \cdot (\rho_i \mathbf{v}_i) = \tau_i, \quad i = 1, 2, \ldots, N-1, \\
&\partial_t (\rho_i \mathbf{v}_i) + \nabla \cdot (\rho_i \mathbf{v}_i \otimes \mathbf{v}_i - \mathbf{t}_i) = \mathbf{m}_i, \\
&\partial_t \left(\frac{1}{2}\rho_i \|\mathbf{v}_i\|^2 + \rho_i \varepsilon_i \right) + \nabla \cdot \left\{ \left(\frac{1}{2}\rho_i \|\mathbf{v}_i\|^2 + \rho_i \varepsilon_i \right) \mathbf{v}_i - \mathbf{t}_i \mathbf{v}_i + \mathbf{q}_i \right\} = e_i.
\end{aligned} \tag{4.16}$$

We close the above system according to universal principles of thermodynamics, such as Galilean invariance and entropy principle.[133] The Galilean invariance enforces the production term as

$$\tau_b = \hat{\tau}_b, \quad \mathbf{m}_b = \hat{\tau}_b \mathbf{v} + \hat{\mathbf{m}}_b, \quad e_b = \hat{\tau}_b \frac{\|\mathbf{v}\|^2}{2} + \hat{\mathbf{m}}_b \cdot \mathbf{v} + \hat{e}_b,$$

where the variables with hat denote variables that are independent of the velocity. Here, we consider the simplest case where there is no chemical reaction, i.e. $\tau_\alpha = 0$. In this case, the entropy principle requires

$$\hat{\mathbf{m}}_i = \frac{1}{N}\sum_{j=1}^{N-1} \phi_{ij} \left(\frac{\mathbf{u}_N}{\theta_N} - \frac{\mathbf{u}_j}{\theta_j} \right), \quad \hat{e}_i = \frac{1}{N}\sum_{j=1}^{N-1} \zeta_{ij} \left(\frac{1}{\theta_j} - \frac{1}{\theta_N} \right).$$

Now, for simplicity, we consider the case when mixtures are spatially homogeneous and assume $\rho_\alpha = 1$, $\varepsilon_\alpha = \theta_\alpha$, and $\mathbf{v} = 0$. Then, system (4.16) becomes

$$\frac{d\mathbf{v}_i}{dt} = \frac{1}{N}\sum_{j=1}^{N-1}\phi_{ij}\left(\frac{\mathbf{v}_j}{\theta_j} - \frac{\mathbf{v}_i}{\theta_i}\right), \quad t > 0, \ \ i = 1, 2, \ldots, N-1,$$

$$\frac{d\left(\theta_i + \frac{1}{2}\|\mathbf{v}_i\|^2\right)}{dt} = \frac{1}{N}\sum_{j=1}^{N-1}\zeta_{ij}\left(\frac{1}{\theta_i} - \frac{1}{\theta_N}\right),$$

$$\mathbf{v}_N = -\sum_{i=1}^{N-1}\mathbf{v}_i, \quad \theta_N = \sum_{\alpha=1}^{N}\left(\theta_\alpha^0 + \frac{1}{2}\|\mathbf{v}_\alpha^0\|^2\right) - \sum_{i=1}^{N-1}\theta_i - \frac{1}{2}\sum_{\alpha=1}^{N}\|\mathbf{v}_\alpha\|^2.$$

This dynamics of N-constituents can be recovered to the original form:

$$\begin{aligned}\frac{d\mathbf{v}_i}{dt} &= \frac{\kappa_1}{N}\sum_{j=1}^{N}\phi_{ij}\left(\frac{\mathbf{v}_j}{\theta_j} - \frac{\mathbf{v}_i}{\theta_i}\right),\\ \frac{d\left(\theta_i + \frac{1}{2}\|\mathbf{v}_i\|^2\right)}{dt} &= \frac{\kappa_2}{N}\sum_{j=1}^{N}\zeta_{ij}\left(\frac{1}{\theta_i} - \frac{1}{\theta_j}\right),\end{aligned} \tag{4.17}$$

after simple transformation. This is the very original form of the TCS model derived in Ref. [133], and in the same reference, the simplified model is also introduced in which the quadratic term in the dynamics of second equation in (4.17) is neglected. Although the coefficients ϕ_{ij} and ζ_{ij} are originally independent of spatial variable $\mathbf{x}_i$, considering the similarity of formulation with the CS model, it is natural to consider that these interacting coefficients possibly depend on the spatial difference. Then, together with the definition of velocity $\frac{d\mathbf{x}_i}{dt} = \mathbf{v}_i$, we obtain the TCS model (4.4) under the small diffusion velocity assumption. In the sequel, we present emergent dynamics of particle, kinetic and fluid TCS models.

4.2.1. *The particle TCS model*

In this part, we briefly recall emergent dynamics of the particle TCS model. We first present the concept of flocking for the TCS model which is almost similar to that of the CS model, except that the temperature alignment is added in Definition 4.1.

Definition 4.2. Let $\bar{\mathcal{P}} := \{(\mathbf{x}_i, \mathbf{v}_i, \theta_i)\}_{i=1}^N$ be a solution to the TCS model (4.4). Then, the ensemble $\bar{\mathcal{P}}$ exhibits a global(mono-cluster) flocking, if and only if the following three relations hold.

(1) The relative positions are uniformly bounded:

$$\sup_{0\le t<\infty}\max_{1\le i\ne j\le N}\|\mathbf{x}_i(t) - \mathbf{x}_j(t)\| < \infty.$$

(2) The relative velocities and temperatures decay to zero asymptotically:

$$\lim_{t\to\infty}\max_{1\le i\ne j\le N}\|\mathbf{v}_i(t) - \mathbf{v}_j(t)\| = 0, \quad \lim_{t\to\infty}\max_{1\le i\ne j\le N}|\theta_i(t) - \theta_j(t)| = 0.$$

Different from the CS model, the TCS model has a singularity on the right-hand side when the temperature variable θ becomes zero. Therefore, it is highly needed that the temperature variable does not touch zero during its evolution in order to get a well-posedness of the model. Fortunately, this is easily guaranteed for the TCS model (4.4). In fact, for $0 \le s \le t$,

$$\max_{1\le i\le N} \theta_i(t) \le \max_{1\le i\le N} \theta_i(s), \quad \min_{1\le i\le N} \theta_i(t) \ge \min_{1\le i\le N} \theta_i(s).$$

In other words, the support of temperature variable shrinks over time. In particular, this guarantees that if the initial temperatures are positive, then all temperature variables cannot touch zero on whole time interval.

$$\theta_{m0} := \min_{1\le j\le N} \theta_{j0} \le \theta_i(t) \le \max_{1\le j\le N} \theta_{j0} =: \theta_{M0}, \quad t \ge 0.$$

This monotonicity property of the particle TCS model can also be extended to the kinetic TCS equation as well. Before we move on to the kinetic TCS equation, we provide several recent results on the particle TCS model related to the mean-field limit of the TCS model.

Theorem 4.7. (Ref. 128) *Suppose that there exist positive constants C and x^∞ such that the initial data and coupling strength satisfy the following relations*:

$$D_\theta(0) \le \frac{\theta_{m0}^2 \phi(D_{\mathbf{x}}(0))}{2C\theta_{M0}}, \quad D_{\mathbf{v}}(0) \le \frac{\kappa_1}{2\theta_{M0}} \int_{D_{\mathbf{x}}(0)}^{x^\infty} \phi(s)ds,$$

$$\frac{8}{3C}\phi(D_{\mathbf{x}}(0)) \le \phi(x^\infty) \le \phi(D_{\mathbf{x}}(0)),$$

and let $\{(\mathbf{x}_i, \mathbf{v}_i, \theta_i)\}_{i=1}^N$ be a solution to the TCS model with initial data $\{(\mathbf{x}_{i0}, \mathbf{v}_{i0}, \theta_{i0})\}_{i=1}^N$. Then, the TCS model exhibits the flocking behavior in the sense of Definition 4.2:

$$\sup_{0\le t<\infty} D_{\mathbf{x}}(t) \le x^\infty, \quad D_\theta(t) \le D_\theta(0) e^{-\frac{\kappa_2}{\theta_{M0}^2}\zeta(x^\infty)t},$$

$$D_{\mathbf{v}}(t) \le D_{\mathbf{v}}(0) \exp\left[-\frac{\kappa_1\phi(x^\infty)}{\theta_{M0}}t + \frac{2\kappa_1\theta_{M0}^2 D_\theta(0)}{\kappa_2\theta_{m0}^2\zeta(x^\infty)}\right].$$

4.2.2. *The kinetic and hydrodynamic TCS models*

The derivation of kinetic equation (4.5) can be done by performing exactly identical procedure with that of the CS model. The only difference is that the TCS model has an additional temperature variable, which does not cause extra difficulty in deriving the kinetic equation (4.5). The first interest on this model is again rigorous justification of the mean-field limit. Since the kinetic TCS model (4.5) is a transport equation as the kinetic CS equation, finite-in-time mean-field limit can be made as in Theorem 4.3. Moreover, with extra flocking estimate in Theorem 4.7, depending only on the diameter of initial data and independent of the number of particles,

we can also derive the uniform stability estimate[128] at particle level similar to Theorem 4.2. Again, we use this to derive the uniform-in-time mean-field limit for the TCS model for some class of initial data. This is summarized in the following theorem.

Theorem 4.8. (Ref. 128) *Suppose that there exist positive constants C and x^∞ such that the initial measure $\mu_0 \in \mathcal{M}(\mathbb{R}^{2d}\times\mathbb{R}_+)$ is compactly supported in $\mathbb{R}^{2d}$, and the coupling strength and communication weight satisfy*

$$D_\theta(0) \le \frac{\theta_{m0}^2\phi(D_{\mathbf{x}}(0))}{2C\theta_{M0}}, \quad D_{\mathbf{v}}(0) \le \frac{\kappa_1}{2\theta_{M0}}\int_{D_{\mathbf{x}}(0)}^{x^\infty}\phi(s)ds,$$

$$\frac{8}{3C}\phi(D_{\mathbf{x}}(0)) \le \phi(\mathbf{x}^\infty) \le \phi(D_{\mathbf{x}}(0)),$$

where $D_{\mathbf{x}}$, $D_{\mathbf{v}}$ and D_θ are defined parallel to those of the particle system:

$$D_{\mathbf{x}}(t) := \sup_{\mathbf{x},\mathbf{x}^*\in\mathrm{supp}_{\mathbf{x}}\mu(t)}\|\mathbf{x}-\mathbf{x}^*\|, \quad D_{\mathbf{v}}(t) := \sup_{\mathbf{v},\mathbf{v}^*\in\mathrm{supp}_{\mathbf{v}}\mu(t)}\|\mathbf{v}-\mathbf{v}^*\|,$$

$$D_\theta(t) := \sup_{\theta,\theta^*\in\mathrm{supp}_\theta\mu(t)}|\theta-\theta^*|.$$

Then there exists the unique measure-valued solution $\mu = \mu(t)$ to (4.5) with initial data μ_0. Moreover, $\mu(t)$ is approximated by the empirical measure $\mu^N(t)$ uniformly in time:

$$\lim_{N\to\infty}\sup_{0\le t<\infty} d_{\mathrm{BL}}(\mu^N(t),\mu(t)) = 0.$$

The coupling of the kinetic TCS model with compressible/incompressible fluids is also studied.[70,71] For example, the coupling between the kinetic TCS equation and incompressible Navier–Stokes (NS) equation can be done via the drag force:

$$\begin{aligned}
&\partial_t f + \mathbf{v}\cdot\nabla_{\mathbf{x}} f + \nabla_{\mathbf{v}}\cdot(F[f,\mathbf{u}]f) + \partial_\theta(G[f]f) = 0,\\
&\partial_t\mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} + \nabla p - \Delta\mathbf{u} = -\int_{\mathbb{R}^d\times\mathbb{R}_+}(\mathbf{u}-\mathbf{v})f\,d\mathbf{v}\,d\theta,
\end{aligned} \tag{4.18}$$

where adjusted nonlocal force term $F[f,\mathbf{u}]$ is defined as

$$F[f,\mathbf{u}](t,\mathbf{x},\mathbf{v},\theta) := F[f](t,\mathbf{x},\mathbf{v},\theta) + \mathbf{u}(t,\mathbf{x}) - \mathbf{v}.$$

Global well-posedness theories of weak and strong solutions to the TCS-NS system (4.18) for the 3-dimensional spatially periodic domain, as well as the asymptotic flocking behavior are done in Ref. 70. More precisely, we define the following Lyapunov functional $\mathcal{L}$ to measure the degree of flocking

$$\begin{aligned}
\mathcal{L}(t) := &\int_{\mathbb{T}^3\times\mathbb{R}^3\times\mathbb{R}_+}\|\mathbf{v}-\mathbf{v}_c\|^2 f\,d\mathbf{z} + \int_{\mathbb{T}^3\times\mathbb{R}^3\times\mathbb{R}_+}|\theta-\theta_c|^2 f\,d\mathbf{z}\\
&+\int_{\mathbb{T}^3}\|\mathbf{u}-\mathbf{u}_c\|^2\,d\mathbf{x} + \frac{1}{2}\|\mathbf{u}_c-\mathbf{v}_c\|^2,
\end{aligned}$$

where

$$\mathbf{v}_c := \int_{\mathbb{T}^3\times\mathbb{R}^3\times\mathbb{R}_+} \mathbf{v} f \, d\mathbf{z}, \quad \mathbf{u}_c := \int_{\mathbb{T}^3} \mathbf{u} \, d\mathbf{x}, \quad \theta_c := \int_{\mathbb{T}^3\times\mathbb{R}^3\times\mathbb{R}_+} \theta f \, d\mathbf{z}.$$

Here, $d\mathbf{z}$ is an abbreviated notation for $d\mathbf{x}\, d\mathbf{v}\, d\theta$. Then, the decay of Lyapunov functional $\mathcal{L}$ was proved,[70] which implies that flocking behavior emerges in the coupled TCS-NS system for some well-prepared initial data.

We close this section by briefly reviewing the hydrodynamic description of the TCS model. Similar to the CS model, mono-kinetic ansatz yields a hydrodynamic description of the TCS model[127]:

$$\begin{aligned}
&\rho_t + \nabla_{\mathbf{x}} \cdot (\rho u) = 0, \quad t > 0, \quad \mathbf{x} \in \mathbb{R}^d, \\
&(\rho u)_t + \nabla_{\mathbf{x}} \cdot (\rho u \otimes u) = \kappa_1 \int_{\mathbb{R}^d} \phi(\|\mathbf{x} - \mathbf{x}^*\|) \left(\frac{u(\mathbf{x}^*)}{e(\mathbf{x}^*)} - \frac{u(\mathbf{x})}{e(\mathbf{x})} \right) \rho(\mathbf{x}) \rho(\mathbf{x}^*) d\mathbf{x}^*, \\
&(\rho e)_t + \nabla_{\mathbf{x}} \cdot (\rho u e) = \kappa_2 \int_{\mathbb{R}^d} \zeta(\|\mathbf{x} - \mathbf{x}^*\|) \left(\frac{1}{e(\mathbf{x})} - \frac{1}{e(\mathbf{x}^*)} \right) \rho(\mathbf{x}) \rho(\mathbf{x}^*) d\mathbf{x}^*, \\
&(\rho, u, e)|_{t=0} = (\rho_0, u_0, e_0).
\end{aligned} \tag{4.19}$$

A well-posedness theory of classical solution to (4.19) under the spatially periodic domain and smallness assumption of initial data was done in Ref. 127. The asymptotic flocking estimate is also done in the same literature. Before we state the result, we define mean temperature and energy functional as follows.

$$e_c(t) := \frac{\int_{\mathbb{T}^d} \rho(t,\mathbf{x}) e(t,\mathbf{x}) d\mathbf{x}}{\int_{\mathbb{T}^d} \rho(t,\mathbf{x}) d\mathbf{x}}, \quad \mathcal{E}(t) := \frac{1}{2} \int_{\mathbb{T}^d} \rho(t,\mathbf{x}) \|\mathbf{u}(t,\mathbf{x})\|^2 \, d\mathbf{x}, \quad t \geq 0.$$

Theorem 4.9. (Ref. 127) *Let (ρ, u, e) be a classical solution to (4.19) with initial data (ρ_0, u_0, e_0) satisfying*

$$0 < e_m < \min_{x\in\mathbb{T}^d} e_0(x) < \max_{x\in\mathbb{T}^d} e_0(x) =: e_M.$$

Then, there exist positive constants C_1, C_2 and C_3 such that

$$|e(\mathbf{x},t) - e_c| \leq \max\{(e_M - e_c), (e_c - e_m)\} e^{-C_1 t}, \quad \mathcal{E}(t) \leq \mathcal{E}(0) e^{-C_2 t + C_3}.$$

4.3. *Research perspective*

In this subsection, we discuss some possible research perspectives. The flocking is one of interesting collective phenomena which appears in biological and social complex systems. Despite of extensive research on the CS-type models in last 10 years, there are still lots of open problems in the particle/kinetic CS models, not to mention the TCS model. In the sequel, we briefly mention two open problems. First, the complete clustering predictability problem of the CS model in multi-dimensions is one of the interesting open problems in relation with the clustering problem of big data. Recently, this clustering predictability problem is completely

resolved in one-dimension[122] using the first-order reduction of the CS model. However, when the physical dimension is larger than or equal to two, it is completely open to predict how many clusters will emerge from given initial data and system parameters. Second, it is still not known whether the kinetic CS model with a singular communication weight is globally well-posed or not. Although the local-in-time well-posedness for the kinetic CS model with a singular communication is studied in Ref. 59, extension of this local well-posedness result to a global one is a nontrivial matter. The singular communication weight can exclude collisions in microscopic level, when the singularity is sufficiently strong. Thus, investigation of singular interactions will be an interesting research topic.

Of course, aforementioned two problems can also be posed for the TCS type models. So far, all flocking results for the TCS model were focused on the flocking of both velocity and temperature at the same time. Thus, it is still not clear whether the additional temperature variable in the TCS model generates different behaviors compared to the CS model. This remains as a future work.

Beside the analytic point of view, the flocking algorithm inherent in the CS model can be applied to more practical research. For example, the analytical results for the CS model can be employed to the robust design of meta-heuristic optimization algorithms such as particle swarm optimization.[62] The flocking algorithm can also be applied to the control theory where the desired patterns emerge using decentralized formation control.[41,77] The engineering applications of flocking dynamics has just begun in the control theory of multi-agent systems, and it is expected to be applied to broader fields in future.

5. Macroscopic Equations of Swarms from Microscopic Interactions

As explained above, macroscopic fluid-type models for large crowds or swarms can always be proposed in two different ways. On the one hand, we can always close the continuity equation of the continuous density through a phenomenological speed-density relation, giving rise to first-order models. On the other hand, we can also propose second-order models by closing the balance equation of momentum via the addition of appropriate terms to the Euler equations. Such terms are aimed to explain, in an indirect way, the dynamics of the agents toward their goal along with the effect of the environment and obstacles. All these two approaches are phenomenological and do not care at all about the real inter-particle interactions of agents within the given population.

In general, tackling the latter approach is a convoluted task. Notice that, depending on the nature of the population, such inter-particle interactions may not be clear at first glance and might require an interdisciplinary point of view in order to elucidate which are the fundamental variables and laws governing the social relations between agents. However, in some particular situations we may give light to the problem with part of the dynamics. Specifically, a real crowd is actually

a complex system and, as such, it involves plenty of physical, social, biological and cognitive variables. However, we may be able to simplify the complete dynamics of the system by disregarding the secondary variables and just looking at the effect of the main variables related to a given feature of the population. Although this will not be a valid description for every scenario, in this way, we can obtain a simpler and more manageable model that still recasting some of the most important features of the population: for instance, emergence of some type of *global collective behavior*.

The interest on collective dynamics models has notably raised during the recent years. From the applied side, this is specially interesting since simple rules governing pairwise interactions between agents, leads to global emergent behavior of the total population as a whole. As it can be seen for instance in Refs. 33, 58, 63, 69, 110, 120, 130, 173, 191, 200, 201, 215, 221 and 222 and references therein, collective dynamics models have proved relevant in several areas of *soft active*. From the theoretical point of view, collective dynamics can be regarded as a rich source of problems in mathematics. Indeed, many recent strong techniques in mathematics have arisen as a consequence of such models, what highlights that the feedback and exchange of ideas between the applied and theoretical communities are undoubtedly positive.

In this part, we will focus on a few collective dynamics models that obey a similar structure and have been analyzed in the literature during the last years. At the agent-based level they consist of a system of N first-order coupled ODEs, one for each agent. Although the kinetic description is the objective of this work, together with its macroscopic limit, we highlight in this part its relation with the microscopic description of which the mesoscopic description inherits a large part of its properties. Specifically, assume that each subject is located at a given position $\mathbf{x}_i = \mathbf{x}_i(t)$. Then, all these first-order systems take the following form:

$$\begin{aligned} \frac{d\mathbf{x}_i}{dt} &= \boldsymbol{\nu}_i + \mathbf{F}_{\mathbf{e}}(\mathbf{x}_i) + \frac{\kappa}{N}\sum_{j=1}^{N} m_j \mathbf{F}(\mathbf{x}_i, \mathbf{x}_j), \\ \mathbf{x}_i(0) &= \mathbf{x}_{i,0}, \end{aligned} \tag{5.1}$$

for $i = 1, \ldots, N$. Here, $\mathbf{x}_i$ are regarded as positions in $\mathbb{R}^d$ or $\mathbb{T}^d$ (for periodic domains). Nevertheless, $\mathbf{x}_i$ do not necessarily restrict to positions, but they can rather represent any other physical, social, state or internal variable of the agents that we are interested in. The parameter m_i is the mass of the ith agent, $\boldsymbol{\nu}_i \in \mathbb{R}$ are often called *natural velocities* and introduce some heterogeneities in the population, $\mathbf{F}_{\mathbf{e}} = \mathbf{F}_{\mathbf{e}}(\mathbf{x})$ represents an external force acting on the system and $\mathbf{F} = \mathbf{F}(\mathbf{x}, \mathbf{x}^*)$ is some interaction kernel governing the force that any subject at $\mathbf{x}^*$ exerts on a subject at $\mathbf{x}$.

To start, let us now list some of the main collective behavior models of first-order type that embed into the above formulation (5.1). In all of them, we will disregard the effect of a possible external force, i.e. $\mathbf{F}_{\mathbf{e}} \equiv 0$ with the sake of focusing on the communication part between agents:

(1) If the internal variables are chosen to be the phases of oscillators in the unit circle, $\mathbf{x}_i \equiv \theta_i \in \mathbb{R}$, $\mathbf{F}(\mathbf{x}, \mathbf{x}^*) = \sin(\mathbf{x}^* - \mathbf{x})$, $\nu_i \equiv \Omega_i$ are the natural frequencies and $m_i = 1$, then (5.1) is nothing but the Kuramoto model for coupled oscillators

$$\begin{aligned} \frac{d\theta_i}{dt} &= \Omega_i + \frac{\kappa}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_i), \\ \theta_i(0) &= \theta_{i,0}. \end{aligned} \tag{5.2}$$

This is a classical model that was proposed by Kuramoto as a prototype system exhibiting emergence of *synchronization*.[97,119,123,130,171] It is intimately related to opinion dynamics and the heterogeneities ν_i which play the role of a biased tendency of agents to move at their own frequency while being influenced by their neighbors. The values θ_i sometimes represent the phases of the neuronal signals, and $\dot{\theta}_i$ represents the firing frequencies of neurons in the brain. In this setting, the Kuramoto model consists in a first approach towards a mathematical description of *neuronal synchronization*,[209,226,231,232] that is known to rule many cognitive processes of the brain that are activated when a specific group of neurons fire together forming a cluster. Of course, this model can be made more realistic by adding coupling weights governing the plasticity of connections via learning mechanisms,[85,132,189,196,213] inertia terms and delays in time,[73,76] noise or many other features like singular couplings.[191,200] We will review some of this associated models later on.

(2) If the heterogeneities are neglected $\boldsymbol{\nu}_i \equiv 0$, $m_i = 1$ and $\mathbf{F}(\mathbf{x}, \mathbf{x}^*) = -\nabla_{\mathbf{x}} W(\mathbf{x} - \mathbf{x}^*)$, then (5.1) agrees with the *aggregation equation* with a given potential function W

$$\begin{aligned} \frac{d\mathbf{x}_i}{dt} &= -\frac{\kappa}{N}\sum_{j=1}^{N}\nabla W_{\mathbf{x}}(\mathbf{x}_i - \mathbf{x}_j), \\ \mathbf{x}_i(0) &= \mathbf{x}_{i,0}. \end{aligned} \tag{5.3}$$

This is probably one of the best known models in this family. It represents *swarming* of a population of bacteria or other entities, that try to aggregate and form a unique group or cluster. Depending on the nature of potential W, one might include both attractive and repulsive interactions, then enriching the dynamics. This is one of the reason why this model has specially pulled the attention of the scientific community during the last years, indeed it has the ability to generate dynamics converging to equilibria that exhibit relevant patterns in biological contexts, see Refs. 38, 49, 47, 63, 66, 100, 101, 180, 179, 221 and 222.

(3) If $d = 2$, we neglect heterogeneities, $m_i \equiv \omega_i \in \mathbb{R}$ and $\mathbf{F}(\mathbf{x}, \mathbf{x}^*) = \frac{(\mathbf{x}-\mathbf{x}^*)^\perp}{2\pi\|\mathbf{x}-\mathbf{x}^*\|^2}$ (where $(z_1, z_2)^\perp = (-z_2, z_1)$ for any $\mathbf{z} = (z_1, z_2) \in \mathbb{R}^2$), then (5.1) is the N

vortex system

$$\begin{aligned} \frac{d\mathbf{x}_i}{dt} &= \frac{1}{N}\sum_{\substack{1\le j\le N\\ j\neq i}} \omega_j \frac{(\mathbf{x}_i-\mathbf{x}_j)^\perp}{2\pi\|\mathbf{x}_i-\mathbf{x}_j\|^2}, \\ \mathbf{x}_i(0) &= \mathbf{x}_{i,0}. \end{aligned} \tag{5.4}$$

Here, the values ω_i denote the strength of the vortices. This model has been widely studied in the fluid mechanics community. Specially, its associated PDE macroscopic model agrees with the $2D$ *Euler equation in vorticity form* for an perfect incompressible fluid. When, white noise is also added, we recover the well-known NSs system for viscous fluids.

Before we talk about the macroscopic counterparts of (5.1), let us link the above system (5.1) with the classical well-known second-order models arising from Newton's second law. It is a very well-known fact in classical mechanics with applications to statistical mechanics called *overdamped or Smoluchowski limit*, specially when some white nose is added to the system. Here, we will work with second-order deterministic systems as follows:

$$\begin{aligned} \frac{d\mathbf{x}_i}{dt} &= \mathbf{v}_i, \\ \frac{d\mathbf{v}_i}{dt} &= -\frac{1}{\tau}\mathbf{v}_i + \mathbf{F}_{\mathbf{e}}(\mathbf{x}_i) + \frac{\kappa}{N}\sum_{j=1}^{N} m_j \mathbf{F}(\mathbf{x}_i,\mathbf{x}_j), \\ \mathbf{x}_i(0) &= \mathbf{x}_{i,0}, \quad \mathbf{v}_i(0) = \mathbf{v}_{i,0}, \end{aligned} \tag{5.5}$$

where we have included a friction term with the environment with a relaxation time τ.

The overdamped dynamics is the regime where damping dominates inertia. It can be recast in the following scaled equation:

$$\varepsilon\frac{d\mathbf{v}_i}{dt} = -\mathbf{v}_i + \mathbf{F}_{\mathbf{e}}(\mathbf{x}_i) + \frac{\kappa}{N}\sum_{j=1}^{N} m_j \mathbf{F}(\mathbf{x}_i,\mathbf{x}_j). \tag{5.6}$$

Taking limits $\varepsilon \searrow 0$ in (5.6), the inertia term vanishes and make all the second-order dependence of the system disappear. Then, the first-order system (5.1) arises naturally. Such arguments can be made rigorous for Lipschitz-continuous forces via Tikhonov's theorem.[156] See Refs. [100], [101], [107], [108], [117], [198] and [201] for some recent advances in this line both for smooth and singular kernels at the microscopic and macroscopic levels.

Of course, such models (5.1) give rise to a large family of variants when, instead of smooth, kernels are singular at the origin (as in the classical Vlasov–Poisson or Euler systems) or, instead of isotropic, they include some more realistic anisotropy. Also, the addition of some noise that distorts the deterministic dynamics can be relevant. In this part, we will be mostly interested in non-smooth forces. Indeed, the lack of smoothness in force fields not only is expected in real system, but it

also has proved strongly relevant in order to describe real scenarios with *finite-time clustering*. Specifically, such "sticky" behavior of particles in general, and crowds' opinion in particular, allows pushing subjects into the formation of distinguished groups in finite time with an eventual global collapse into a final unique cluster (or in several ones), see for example Refs. 191 and 200. Here, we will be mostly interested in system (5.2), but some vague ideas and references will be provided for the readers' convenience, specially regarding (5.3) and some anisotropic versions.

5.1. *Mean field limit and propagation of chaos*

There are two classical approaches in the literature in order to derive rigorously the large crowd limit, or mean-field limit, as $N \to \infty$ in (5.1) or (5.5): the *empirical measure* approach and the *BBGKY hierarchy* approach. The second method is stronger and harder to follow, but it has proved a strong method as it is intimately related to propagation of chaos in many particle systems. As it will be seen, the later has to do with a control of the fall-off of inter-particle correlations as the amount of agents N becomes large. For more accurate descriptions, see Refs. 136, 137, 150, 151, 152, 153, 155, 177, 178, 187 and 219. Let us assume that agents are all identical and masses are normalized to 1. Since the natural velocities in the discrete model (5.1) are constant parameters, we can equivalently restate the system as follows:

$$\begin{aligned} \frac{d\mathbf{x}_i}{dt} &= \boldsymbol{\nu}_i + \mathbf{F}_{\mathbf{e}}(\mathbf{x}_i) + \frac{\kappa}{N}\sum_{j=1}^{N} \mathbf{F}(\mathbf{x}_i, \mathbf{x}_j), \\ \dot{\boldsymbol{\nu}}_i &= 0, \\ (\mathbf{x}_i(0), \boldsymbol{\nu}_i(0)) &= (\mathbf{x}_{i,0}, \boldsymbol{\nu}_{i,0} \equiv \boldsymbol{\nu}_i). \end{aligned}$$

Then, both $\mathbf{x}_i$ and $\boldsymbol{\nu}_i$ are regarded as variables, although the dynamics of $\boldsymbol{\nu}_i$ is trivial.

On the one hand, for the first *empirical measures approach* let us recover the sequence of empirical measure of such N agents

$$\mu^N(t, \mathbf{x}, \mathbf{v}) := \frac{1}{N}\sum_{i=1}^{N} \delta_{(\mathbf{x}_i(t), \mathbf{v}_i(t))}, \tag{5.7}$$

for every $t \geq 0$. Then, it is clear for Lipschitz forces that μ^N solves the following Vlasov equation in the sense of distributions

$$\partial_t \mu^N + \nabla_{\mathbf{x}} \cdot \left[\left(\boldsymbol{\nu} + \mathbf{F}_{\mathbf{e}}(\mathbf{x}) + \kappa \int_{\mathbb{R}^d \times \mathbb{R}} \mathbf{F}(\mathbf{x}, \mathbf{x}^*)\, \mu^N(t, d\mathbf{x}^*, d\boldsymbol{\nu}^*) \right) \mu^N \right] = 0. \tag{5.8}$$

We now take the initial data such that μ_0^N weakly-* converges toward the initial probability distribution $f_0 \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R})$ by virtue of some law of large numbers, e.g. Ref. 224. Then, our goal is to show that the compactness is propagated uniformly

in compact intervals of time and we get

$$\mu^N \to f \in C([0,T], \mathcal{P}(\mathbb{T}\times\mathbb{R}) - \text{narrow}).$$

In such case, we may pass to the limit in (5.8) to recover the Vlasov equation

$$\partial_t f + \nabla_{\mathbf{x}} \cdot \left[\left(\boldsymbol{\nu} + \mathbf{F_e}(\mathbf{x}) + \kappa \int_{\mathbb{R}^d\times\mathbb{R}} \mathbf{F}(\mathbf{x},\mathbf{x}^*) f(t, d\mathbf{x}^*, d\boldsymbol{\nu}^*) \right) f \right] = 0. \tag{5.9}$$

Such ideas are widely known for $W^{1,\infty}$ kernels and can be found in detail in Ref. 187. Indeed, the following *Dobrushin-type inequality* with respect to the bounded-Lipschitz distance d_{BL} holds true for any two measure-valued solutions $f, g \in C([0,+\infty), \mathcal{P}_1(\mathbb{R}^d\times\mathbb{R}) - \text{narrow})$

$$d_{\mathrm{BL}}(f(t), g(t)) \le e^{t(\|\mathbf{F_e}\|_{W^{1,\infty}} + \kappa\|\mathbf{F}\|_{W^{1,\infty}})} d_{\mathrm{BL}}(f_0, g_0), \quad t \ge 0. \tag{5.10}$$

That provides an explicit control justifying how the discrete system (5.1) can be approximated by (5.9) for large N

$$d_{\mathrm{BL}}(\mu^N(t), f(t)) \lesssim e^{Ct} N^{-\frac{1}{d+1}}. \tag{5.11}$$

Indeed, under such Lipschitz condition the characteristic flow is well defined and bi-Lipschitz, thus guaranteeing that absolutely continuous initial data f_0 propagates the same L^1 integrability for all times. Then, no way that Dirac masses emerge from smooth initial data. Of course, the lack of Lipschitz-continuity breaks most of the above arguments down.

On the other hand, the *BBGKY approach* departs from the hierarchy of Liouville equations for the joint laws $f^N = f^N(t, \mathbf{x}_1, \dots, \mathbf{x}_N, \boldsymbol{\nu}_1, \dots, \boldsymbol{\nu}_N) \in \mathcal{P}_{\mathrm{sym}}(\mathbb{R}^{dN}\times\mathbb{R}^{dN})$ reads

$$\partial_t f^N + \sum_{i=1}^N \nabla_{\mathbf{x}_i} \cdot \left[\left(\boldsymbol{\nu}_i + \mathbf{F_e}(x_i) + \frac{\kappa}{N} \sum_{j=1}^N \mathbf{F}(\mathbf{x}_i, \mathbf{x}_j) \right) f^N \right] = 0. \tag{5.12}$$

Notice that since $f_0^N \in \mathcal{P}_{\mathrm{sym}}(\mathbb{R}^{dN}\times\mathbb{R}^{dN})$, in the sense that interchanging ith and jth positions and natural velocities let the measure invariant, then the same continues happening for all times by virtue of the properties of the system. Define the projection onto the first $k \in \{1, \dots, N\}$ variables,

$$\begin{aligned} \boldsymbol{\pi}^{k,N} : \mathbb{R}^{dN}\times\mathbb{R}^{dN} &\to \mathbb{R}^{dk}\times\mathbb{R}^{dk}, \\ (\mathbf{X}^N, \boldsymbol{\nu}^N) &\mapsto (\mathbf{X}^{k,N}, \boldsymbol{\nu}^{k,N}), \end{aligned}$$

where for any $(\mathbf{X}^N = (\mathbf{x}_1, \dots, \mathbf{x}_N),\ \boldsymbol{\nu}^N = (\boldsymbol{\nu}_1, \dots, \boldsymbol{\nu}_N)) \in \mathbb{R}^{dN}\times\mathbb{R}^{dN}$ we are denoting

$$\mathbf{X}^{k,N} := (\mathbf{x}_1, \dots, \mathbf{x}_k) \quad \text{and} \quad \boldsymbol{\nu}^{k,N} := (\boldsymbol{\nu}_1, \dots, \boldsymbol{\nu}_k).$$

Consider the marginal measures $f^{k,N}(t) := \boldsymbol{\pi}^{k,N}_{\#}(f^N(t)) \in \mathcal{P}_{\mathrm{sym}}(\mathbb{R}^{dk} \times \mathbb{R}^{dk})$. Thanks to the assumed symmetry in the system, integration in (5.12) yields

$$\partial_t f^{k,N} + \sum_{i=1}^{k} \nabla_{\mathbf{x}_i} \cdot \left[\left(\boldsymbol{\nu}_i + \mathbf{F}_{\mathbf{e}}(\mathbf{x}_i) + \frac{\kappa}{N} \sum_{j=1}^{k} \mathbf{F}(\mathbf{x}_i, \mathbf{x}_j) \right) f^{k,N} + \kappa \frac{N-k}{N} \right.$$

$$\left. \int_{\mathbb{R}^d \times \mathbb{R}} \mathbf{F}(\mathbf{x}_i, \mathbf{x}_{k+1}) f^{k+1,N}(t, \mathbf{x}_1, \ldots, \mathbf{x}_k, d\mathbf{x}_{k+1}, \boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_k, d\boldsymbol{\nu}_{k+1}) \right] = 0. \tag{5.13}$$

Via a diagonal argument we can obtain weak limits of an appropriate subsequence

$$f^{k,\infty} := \text{weak } * - \lim_{N \to \infty} f^{k,N}.$$

Although (5.13) is not closed for fixed N, some way to try to close it in the limit is via *propagation of chaos* as follows. For simplicity, denote $f := f^{1,\infty}$ and assume that all the initial values are tensorized, that is $f_0^{k,\infty} = f_0^{\otimes^k}$. Then, propagation of chaos means that such tensorization remains true for all times, i.e.

$$f^{k,\infty}(t) = f(t)^{\otimes^k}, \quad \text{for all } t \geq 0.$$

Such property can be shown to happen for this Lipschitz kernels. Then, we can pass to the limit as $N \to \infty$ in (5.13) in the equation of $f^{1,N}$ and recover (5.9). The key problem is to derive estimates for the distance between the marginals $f^{k,N}$ and the solution $f^{\otimes^k}$. Notice that the case $k = 1$ would yield a similar bound to the Dobrushin inequality (5.11). Such bound has been studied by using Wasserstein distances as well in Refs. 137, 177 and 178 and others references. However, in Refs. 151–153 a different strategy has been followed to quantify chaoticity for a large class of models with non-smooth forces. Nevertheless, some presence of noise guaranteeing existence of entropy solutions of (5.12) is required. In other case, for the deterministic situation some condition close to Lipschitz is required. Assume for the moment that the heterogeneities ν_i are neglected and we focus on forces $\mathbf{F}(\mathbf{x}, \mathbf{x}^*) = \mathbf{F}(\mathbf{x} - \mathbf{x}^*)$. Then, for $f^N = f^N(t, \mathbf{x}_1, \ldots, \mathbf{x}_N)$ entropy solution to (5.12) and $f = f(t, \mathbf{x}_1, \ldots, \mathbf{x}_N)$ solution to (5.9) respectively, one can measure their closeness in the sense of entropy by defining the scaled entropies

$$\mathcal{H}_N(f^N(t) | f(t)^{\otimes^N}) := \frac{1}{N} \int_{\mathbb{R}^{dN}} f^N(t) \log \left(\frac{f^N(t)}{f(t)^{\otimes N}} \right) d\mathbf{x}_1, \ldots, d\mathbf{x}_N.$$

Then,

Theorem 5.1. (Ref. 153) *Assume that* $\nabla \cdot \mathbf{F}_{\mathbf{e}} \in L^\infty(\mathbb{T}^d)$, $\nabla \cdot \mathbf{F} \in L^\infty(\mathbb{T}^d)$ *and that either* $\mathbf{F} \in L^\infty(\mathbb{T}^d)$ *or for* $d \geq 2$, $\mathbf{F}$ *is an odd kernel and* $\|\mathbf{x}\|\mathbf{F} \in L^\infty(\mathbb{T}^d)$. *Then, there exists a constant* $M > 0$ *depending on* κ, f_0 *and* $\|\nabla \cdot \mathbf{F}_{\mathbf{e}}\|_{L^\infty(\mathbb{T}^d)}$ *such that*

$$\mathcal{H}_N(f^N(t) | f(t)^{\otimes^N}) \leq e^{M\|\mathbf{F}\|_\infty t} \left(\mathcal{H}_N(f_0^N | f_0^{\otimes^N}) + \frac{1}{N} \right),$$

for every $N \in \mathbb{N}$ *and* $t \geq 0$. *Here,*

$$\|\mathbf{F}\|_\infty := \begin{cases} \|\mathbf{F}\|_{L^\infty(\mathbb{T}^d)} + \|\nabla \cdot \mathbf{F}\|_{L^\infty(\mathbb{T}^d)}, \\ \qquad \textit{if } \mathbf{F} \in L^\infty(\mathbb{T}^d), \\ \||\mathbf{x}|\mathbf{F}\|_{L^\infty(\mathbb{T}^d)} + \|\nabla \cdot \mathbf{F}\|_{L^\infty(\mathbb{T}^d)}, \\ \qquad \textit{if } d \geq 2,\ \mathbf{F} \textit{ is odd and } |\mathbf{x}|\mathbf{F} \in L^\infty(\mathbb{T}^d). \end{cases}$$

Notice that for every $k \leq N$ one has the relation

$$\mathcal{H}_k(f^{k,N}(t)|f(t)^{\otimes^k}) \leq \mathcal{H}_N(f^N(t)|f(t)^{\otimes^N}),$$

then Theorem 5.1 amounts to quantitative estimates of propagation of chaos of system (5.1). Indeed, from it we can recover the standard propagation of chaos in L^1 and Wasserstein distances by virtue of *Csiszár–Kullback–Pinsker* and *Talagrand inequalities*, respectively

$$\|f^{k,N}(t) - f(t)^{\otimes^k}\|_{L^1(\mathbb{T}^{dk})} \leq \sqrt{2k\mathcal{H}_k(f^{k,N}(t)|f(t)^{\otimes^k})},$$

$$W_p(f^{k,N}(t), f(t)^{\otimes^k}) \leq C(f(t), p)\big(k\mathcal{H}_k(f^{k,N}(t)|f(t)^{\otimes^k})\big)^{1/2p},$$

for any $p \geq 1$, $N \geq k$ and $t \geq 0$.

In the following, we will come back to Kuramoto model (5.2) where all the above theory works. In addition, we will review a singularly weighted version, that has been proposed by some of the authors in Refs. 191 and 200. In such model, $\nabla \cdot \mathbf{F}$ is not bounded anymore but it can even become a singular measure at $\mathbf{x} = \mathbf{x}^*$. We will also review an adapted Dobrushin-type inequality like (5.10) that has been proposed in Ref. 200 and will skip propagation of chaos, that will not be studied here.

5.2. *The classical Kuramoto model of coupled oscillators*

The Kuramoto model with mean-field coupling (5.2) (see Refs. 166 and 167) has been extensively studied during the last years as a first approach for synchronization of agents. Their eventual applications are well-known and many of them are addressed in the review.[2] Although Kuramoto initially proposed it for synchronization in chemical reactions, these captivating cooperative phenomena are also observed in biological, physical, and social systems and they have attracted the interest of scientists for centuries. Such mechanism governs the synchronization of flashing of fireflies,[46] chorusing of crickets, beating of cardiac cells, metabolic synchrony in yeast cell suspension, etc. Here, we are mainly interested in the above-mentioned application on synchronization of the frequencies of synaptic firing of neurons in the brain. In particular, it allows explaining phase transitions from disordered to ordered states at a critical coupling strength, that is one of the main

features of this model and will be slightly addressed later for the readers' convenience. For some applications to the human connective network and how the realistic connectome maps that are available in the literature affect the emergence of synchronization, see Ref. [226] and references therein. Those ideas exploit that the human connectome turns out to be organized in modula (characterized by a much larger intra than inter connectivity) structured in a hierarchical nested fashion across many scales, affecting to the neural dynamics.[209,231,232]

From a mathematical point of view, there have been important contribution in the analysis of phase and frequency synchronization in the system. The reader may want to look in Refs. [97], [119], [123], [130] and [171] and references therein. Here, we will sketch some of the most relevant results in the study of synchronization. On the one hand, when the heterogeneities disappear and all the agents are identical, there is complete phase synchronization.

Theorem 5.2. (Ref. [119]) *Let* $\Theta(t) = (\theta_1(t), \dots, \theta_N(t))$ *be a smooth solution to* (5.2) *with* $\Omega_i = 0$ *and assume that* $D_\Theta(0) < \pi$. *Then, we have an asymptotic complete phase synchronization and, moreover,*

$$e^{-\kappa t} D_\Theta(0) \leq D_\Theta(t) \leq e^{-\kappa C t} D_\Theta(0), \quad t \geq 0,$$

where $C = \frac{\sin(D_\Theta(0))}{D_\Theta(0)}$ *and, again, we denote the phase diameter*

$$D_\Theta(t) := \max_{1 \leq i \neq j \leq N} |\theta_i(t) - \theta_j(t)|.$$

Although some other approaches had been explored, most of them produced N-dependent rates that do not fit the mean-field scaling and, to our best knowledge, this is the first result that avoids that issue. Of course, complete frequency synchronization of identical oscillators is the clear from the above result with a new exponential rate on the diameter. Regarding non-identical oscillators, one cannot expect global phase synchronization. However, one still expects frequency synchronization when the coupling strength κ is large enough. Indeed, when the decay rate is fast enough it implies emergence of phase-locked states. Those equilibria are characterized by the fact that the inter-particle distances remain constant while rotating in the unit circle. There are several approaches to that we will shortly sketch. On the one hand, the first approach is based on an uniform bound of the phase diameter under appropriate conditions, that can be used to achieve an explicit exponential decay of the frequency diameter of the system. Specifically, we have the following.

Theorem 5.3. (Ref. [119]) *Assume that* $\frac{1}{N}\sum_{i=1}^N \Omega_i = 0$, $D_\Omega > 0$, $\kappa > D_\Omega$ *and let* $\Theta(t) = (\theta_1(t), \dots, \theta_N(t))$ *be a smooth solution to* (5.2) *such that* $D_\Theta(0) < D^\infty$ *and* $\theta_{i,0} \neq \theta_{j,0}$ *for every* $i \neq j$, *where* $D^\infty \in (0, \frac{\pi}{2})$ *is the unique root of*

$$\frac{D_\Omega}{\kappa} = \sin x, \quad x \in \left(0, \frac{\pi}{2}\right).$$

Then, the phase diameter keeps bounded by D^∞ *for all times and, in addition, we have asymptotic complete frequency synchronization*

$$D_{\dot\Theta}(t) \leq D_{\dot\Theta}(0) e^{-\kappa \cos(D^\infty) t}, \quad t \geq 0.$$

Again, we denote the frequency diameter by

$$D_{\dot\Theta}(t) := \max_{1\le i\neq j\le N} |\dot\theta_i(t) - \dot\theta_j(t)|.$$

Some other improvements are given in Ref. [97] and later in Ref. [123], where emergence of phase locked states was proved for any initial configuration. The main restriction in such extensions is the lack of estimate for the frequency decay, that in the general case is still an open problem.

The second approach exploits the gradient system structure of (5.2). Specifically, notice that for

$$V(\Theta) = -\sum_{i=1}^{N} \theta_i\Omega_i + \frac{\kappa}{2N}\sum_{i,j=1}^{N}(1-\cos(\theta_i-\theta_j)), \quad \Theta = (\theta_1,\dots,\theta_N)\in\mathbb{R}^N, \tag{5.14}$$

system (5.2) is nothing but the gradient flow of V, i.e.

$$\dot\Theta = -\nabla_\Theta V(\Theta). \tag{5.15}$$

Indeed, V is an analytic potential, what in particular implies Łojasiewicz gradient inequality. Namely, for every $\Theta^*\in\mathbb{R}^N$ there exist $\gamma\in\left[\frac{1}{2},1\right)$, $L>0$ and a ball $B_R(\Theta)$ centered at Θ^* such that

$$|V(\Theta)-V(\Theta^*)|^{1-\gamma} \le C\|\nabla_\Theta V(\Theta)\|^2,\ \Theta\in B_R(\Theta^*).$$

That can be used to prove that whenever one has a bounded trajectory $\Theta=\Theta(t)$, frequency has to converge to zero as $t\to\infty$ and a phase-locked state emerges, see Ref. [130]. However, the explicit rate is not given since it is known to depend on the explicit Łojasiewicz exponent γ of the phase-locked state Θ^*. See Ref. [171] where some decay rates have been given in particular cases.

Regarding the macroscopic model, notice that according to the above part, Neuzert's techniques[187] yield the rigorous mean-field limit of (5.2), thanks to the regularity of the kernel $F(x,y)=\sin(y-x)$. Indeed, the Vlasov equation agrees with the well-known Kuramoto–Sakaguchi equation

$$\partial_t f + \partial_\theta\left(\left(\Omega+\kappa\int_{\mathbb{T}\times\mathbb{R}}\sin(\theta^*-\theta)f(t,d\theta^*,d\Omega^*)\right)f\right)=0. \tag{5.16}$$

Such idea was first proposed in Ref. [168], where L^1 solutions were obtained. However, a more recent approach in Ref. [58] also addresses measure-valued solutions and a contractivity estimate was given in a sort of Dobrushin inequality (5.10) with negative exponential decay. The authors used such information to transfer the above dynamical properties of agent-based system to the macroscopic system.

On the one hand, the mean-field limit allows transferring complete phase synchronization in the identical case.

Theorem 5.4. (Ref. 58) *Suppose* $f_0 \in \mathcal{P}(\mathbb{T}\times\mathbb{R})$ *such that all the oscillators are identical, i.e.* $g = \delta_0(\Omega)$ *where* $g = (\pi_\Omega)_\# f_0$ *is the* Ω*-marginal, or distribution of natural frequencies. Assume that*

$$\int_{[0,2\pi)\times\mathbb{R}} \theta f_0(\theta,\Omega)d\theta\, d\Omega = \pi, \quad D_\theta(f_0) < \pi \quad \text{and} \quad \kappa > 0.$$

Then, the measure-valued solution f *to* (5.16) *issued at* f_0 *satisfies*

$$D_\theta(f(t)) \le D_\theta(f_0)e^{-\kappa C t},$$

where $C = \frac{\sin(D_\theta(f_0))}{D_\theta(f_0)}$. *In particular,*

$$\lim_{t\to\infty} d_{\mathrm{BL}}(f(t), f_\infty) = 0,$$

where the equilibrium reads $f_\infty = \delta_\pi(\theta)\otimes\delta_0(\Omega)$.

Here, $D_\theta(f(t)) = \mathrm{diam}(\mathrm{supp}_\theta f(t))$ is the diameter of the θ-support of $f(t)$, that is, $\mathrm{supp}_\theta(f(t)) = \pi_\theta(\mathrm{supp}\, f(t))$, where π_θ stands for the projection onto the variable θ, i.e.

$$\begin{aligned}\pi_\theta : \mathbb{T}\times\mathbb{R} &\to \mathbb{T}\\ (\theta,\Omega) &\mapsto \theta.\end{aligned}$$

The Ω-support of $f(t)$ and $D_\Omega(f(t))$ can be similarly defined. Regarding complete frequency synchronization of non-identical oscillators, the necessary result is the aforementioned contractivity estimate.

Theorem 5.5. (Ref. 58) *Suppose that two initial measures* $f_0, \widetilde{f}_0 \in \mathcal{P}(\mathbb{T}\times\mathbb{R})$ *and* $\kappa > 0$ *satisfy*

(1) $0 < D_\theta(\widetilde{f}_0) \le D_\theta(f_0) < \pi$.
(2) $\displaystyle\int_{[0,2\pi)\times\mathbb{R}} \theta f_0(\theta,\Omega)d\theta\, d\Omega = \int_{[0,2\pi)\times\mathbb{R}} \theta \widetilde{f}_0(\theta,\Omega)d\theta\, d\Omega = \pi$.
(3) $\kappa > D_\Omega(f_0)\cdot\max\left\{\frac{1}{\sin(D_\theta(f_0))}, \frac{1}{\sin(D_\theta(\widetilde{f}_0))}\right\}$.

Let $f, \widetilde{f}$ *be two measure-valued solutions to* (5.16) *corresponding to the initial data* f_0 *and* $\widetilde{f}_0$, *respectively. Then, there exists* $t_0 > 0$ *and some* $D^\infty \in (0, \frac{\pi}{2})$ *such that*

$$\widetilde{W}_p(f(t),\widetilde{f}(t)) \le \exp\left(-\frac{2\kappa\cos D^\infty}{\pi}(t-t_0)\right)\widetilde{W}_p(f_{t_0},\widetilde{f}_{t_0}),$$

for every $t \ge t_0$ *and* $1 \le p \le \infty$.

Using such result, frequency synchronization of non-identical oscillators takes place.

Corollary 5.6. (Ref. 58) *There exists a unique stationary state f_∞ in the set of probability measures fulfilling the properties in Theorem 5.5 such that for any other $f_0 \in \mathcal{P}(\mathbb{T}\times\mathbb{R})$ in such set, the solution f of (5.16) issued at f_0 verifies*

$$\widetilde{W}_p(f(t), f_\infty) \leq \exp\left(-\frac{2\kappa\cos D^\infty}{\pi}(t-t_0)\right)\widetilde{W}_p(f(t_0), f_\infty).$$

In the above result, $\widetilde{W}_p$ is not the usual Wasserstein distance in $\mathbb{T}\times\mathbb{R}$ as one might expect. Instead, it is the Wasserstein distance in $[0, 2\pi)\times\mathbb{R}$ after one unwraps $\mathbb{T}$ into the segment $[0, 2\pi)$. Of course, depending on the point in the torus where we unwrap $\mathbb{T}$, a different Wasserstein distance in an interval arises. The real Wasserstein distance in $\mathbb{T}$ is exactly the infimum of all them. In the next part, we will use the real Wasserstein distance in $\mathbb{T}$ and will introduce some results in the singular coupling case without needing such "artificial" Wasserstein distances $\widetilde{W}_p$ on intervals.

In most of the above result, the cornerstone is the phase diameter. Then, such results are nothing but a transference toward the continuous level of the results at the discrete level. However, as explored in Ref. 123 and other papers, there is a classical quantity that simplifies the understanding of the dynamics, namely, the order parameter. For the continuous case (the reader can easily adapt to the agent-based system), the order parameters $R = R(t)$ and $\phi = \phi(t)$ are given by the relation

$$R(t)e^{i\phi(t)} = \int_{\mathbb{T}\times\mathbb{R}} e^{i\theta}\, f(t, d\theta, d\Omega), \quad t \geq 0. \tag{5.17}$$

The parameter R is a measure of order in the system, ranging from disordered states with $R = 0$ to global phase synchronized states with $R = 1$. In fact, such parameters allow restating (5.16) as follows

$$\partial_t f + \partial_\theta(\Omega - \kappa R\sin(\theta - \phi)f) = 0.$$

Therefore, we may want to analyze the dynamics of (5.16) in terms of the above macroscopic order parameters (5.17). This approach has been addressed in particular in Refs. 33 and 120. The aim in such ideas is to get rid of the diameter assumption in preceding results and, it has been successfully achieved in certain cases. Let us comment on the above improvements of such idea.

On the one hand, the identical case $g = \delta_0$ is much simpler since (5.16) can be restated in terms of the macroscopic density $\rho_t = (\pi_\theta)_\# f_t$ (that is, the θ-marginal of ρ or zeroth-order moment of f_t)

$$\partial_t\rho + \partial_\theta(-\kappa R\sin(\theta - \phi)\rho) = 0. \tag{5.18}$$

In addition, straightforward computations show that R is non-decreasing and, as shown in Ref. 33, $\phi(t)$ converges to an asymptotic value. On the other hand, the stationary states read

$$\rho^*(\theta) = \sigma\delta_{\phi^*}(\theta) + (1-\sigma)\delta_{\phi^*+\pi}(\theta), \tag{5.19}$$

for a parameter $\sigma \in [\frac{1}{2}, 1)$. With all these ingredients, the following result holds true.

Theorem 5.7. (Ref. 33) *Consider* $\rho_0 \in \mathcal{P}(\mathbb{T})$ *and the measure-valued solution* ρ *to* (5.18)*. Then, there exist* $\phi^* \in \mathbb{R}$ *and* $\sigma \in [\frac{1}{2}, 1)$ *such that*

$$\rho(t) \to \rho^* \quad \textit{narrow in } \mathcal{P}(\mathbb{T}),$$

as $t \to \infty$*, where* ρ^* *is defined in* (5.19)*. In addition, if* ρ_0 *is non-atomic then* $\sigma = 1$*, i.e. there is complete phase synchronization.*

Such result also has also been proved in Ref. 120 , showing concentration of mass around $\phi(t)$ with exponential fall-off of the L^2 norm around $\phi(t) + \pi$. Regarding non-identical oscillators, the distribution of natural frequencies $g = (\pi_\Omega)_\# f$ plays a role. Specifically, for compactly supported g and large enough κ compared to its support, there are stationary solutions that play an analogue role to the above two-delta functions

$$f^*(\theta, \Omega) = g^+(\Omega)\delta_{\vartheta^+(\Omega)}(\theta) + g^-(\Omega)\delta_{\vartheta^-(\Omega)}(\theta), \tag{5.20}$$

where $g = g^+ + g^-$, $\vartheta^+(\Omega) = \phi^+ + \arcsin\left(\frac{\Omega}{\kappa R}\right)$ and $\vartheta^-(\Omega) = \phi^+ + \pi + \arcsin\left(\frac{\Omega}{\kappa R}\right)$. In addition, the order parameter R has to verify the consistency relation

$$\kappa R^2 = \int_{\mathbb{R}} \sqrt{(\kappa R)^2 - \Omega^2}\,(g^+ - g^-)(d\Omega). \tag{5.21}$$

However, in such case, the results in Ref. 33 are just conditional and do success on proving any convergence of the order parameters but only characterize the possible equilibria.

Proposition 5.8. (Ref. 33) *Consider* $f_0 \in \mathcal{P}(\mathbb{T} \times \mathbb{R})$ *non-atomic and let* f_t *be the solution to* (5.16) *issued at* f_0*. If* $R(t) \to R^*$ *with* $\operatorname{supp} g \subseteq [-\kappa R^*, \kappa R^*]$ *and* $\phi(t) \to \phi^*$ *as* $t \to \infty$*, then*

$$f(t) \overset{*}{\rightharpoonup} f^* \textit{ weakly } * \quad \textit{in } \mathcal{P}(\mathbb{T} \times \mathbb{R}),$$

where f^* *is given by* (5.20) *with* $g^- \equiv 0$*.*

To the authors' best knowledge, the first unconditional result was analyzed in Ref. 120. In that result, emergence of phase concentration for non-identical oscillators was detected independently on the size of the diameter $D_\theta(f_0)$ of the initial configuration as long as $R(0) > 0$ and κ larger that a large enough critical value depending on $R(0)$ and the size of $\operatorname{supp} g$.

Theorem 5.9. (Ref. 120) *Let* f *be a classical solution to* (5.16) *with* $R(0) > 0$ *and assume that the distribution of natural frequencies has compact support* $\operatorname{supp} g \subseteq [-M, M]$*. Then, for large enough* κ *compared to* $\frac{1}{R_0}$ *and* M

$$\liminf_{t\to\infty} R(t) \geq R_\infty := 1 + \frac{M}{\kappa} - \sqrt{\frac{M^2}{\kappa^2} + 4\frac{M}{\kappa}},$$

and

$$\lim_{t\to\infty} \|f(t)\chi_{(\mathbb{T}\setminus L_\infty(t))\times\mathbb{R}}\| = 0.$$

Here $L_\infty(t)$ is the interval centered at $\phi(t)$ with constant width larger, but arbitrarily close to

$$\arccos\left(\sqrt{1-\left[\frac{M}{\kappa}\frac{(1+R_\infty)}{R_\infty^2}+\frac{1-R_\infty}{R_\infty}\right]^2}\right).$$

Notice that as $\kappa\to\infty$, R_∞ tends to one.

The lower bound of the order parameter was essential and is the first result in this line. Also, it is reminiscent of *practical synchronization* at the agent-based level. That is, κ has to be large enough, in order for the order parameter R to oscillate arbitrarily close to 1.

As mentioned before, one of the most interesting features of the Kuramoto model in the mean-field limit is the presence of a phase transition at a given critical coupling strength from disordered to ordered states. This was initially conjectured by Kuramoto, and was later rigorously obtained by several authors by analyzing the bifurcation diagram, see Ref. 69.

Theorem 5.10. (Ref. 69) *Assume that $g=g(\Omega)$ is the Gaussian distribution or a rational function which is even, unimodal and bounded. Consider the Kuramoto transition point $\kappa_c:=\frac{2}{\pi g(0)}$ and let $f_{\rm inc}(\theta,\Omega)=\frac{g(\Omega)}{2\pi}$ be the incoherent state. Then, the following results hold true*:

(1) (*Instability of the incoherent state*) *If $\kappa>\kappa_c$, then $f_{\rm inc}$ is linearly unstable.*
(2) (*Local stability of the coherent state*) *If $0<\kappa<\kappa_c$, there exists $\delta>0$ such that if f_0 has distribution of natural frequencies equals g, i.e. $(\pi_\Omega)_\# f_0=g$ and*

$$\left|\int_{\mathbb{T}\times\mathbb{R}} e^{in\theta}\, d_{(\theta,\Omega)} f_0\right| < \delta, \quad \text{for all } n\in\mathbb{N}, \tag{5.22}$$

then, $R(t)$ decays to zero exponentially fast.
(3) (*Bifurcation*) *There exist $\varepsilon,\delta>0$ such that if $\kappa_c<\kappa<\kappa_c+\varepsilon$ and f_0 fulfils (5.22) then*

$$R(t)=\sqrt{\frac{-16}{\pi\kappa_c^4 g''(0)}}\sqrt{\kappa-\kappa_c}+\mathcal{O}(\kappa-\kappa_c) \quad \text{as } t\to\infty.$$

Similar results were also obtained in Refs. 34 and 92. Notice that the second item can be regarded as *Landau damping* in the vicinity of the incoherent state, a phenomenon that was first observed in Vlasov equation. Third result actually states that a pitchfork bifurcation arises after $\kappa=\kappa_c$, thus generating stable inhomogeneous *partially locked states*. In relation with it, Landau damping toward those partially locked states was introduced in Refs. 93 and 94.

5.3. *The Kuramoto model with singular couplings*

The above Kuramoto–Sakaguchi equation is still subject of deep study due to the complicatedness of the dynamics and its applications in many areas of Science. In

particular, recall that in Refs. 209, 226, 231 and 232 such model has been applied to model neuronal synchronization. Each node represents neurons in a specific area of the brain and the firing frequencies evolve through the coupled laws (5.2) (or (5.16) for macroscopic description consisting of many nodes). However, uniform coupling weights between neurons are unrealistic in general. Specifically, connections should change with time and adapt to the dynamics itself:

$$\begin{aligned} \frac{d\theta_i}{dt} &= \Omega_i + \frac{\kappa}{N}\sum_{j=1}^{N} a_{ij}\sin(\theta_j - \theta_i), \\ \theta_i(0) &= \theta_{i,0}, \end{aligned} \tag{5.23}$$

that is, $a_{ij} = a_{ij}(t)$ are time-evolving and coupled with the dynamics of phases. This is called *plasticity* and can be modeled via a *learning rule*,[85,132,189,196,213] for instance

$$\frac{da_{ij}}{dt} = \eta(\Gamma(\theta_i - \theta_j) - a_{ij}). \tag{5.24}$$

The function $\Gamma = \Gamma(\theta)$ is called plasticity function and $\eta > 0$ determines the learning parameter. According to the neuroscientist Hebb,[139] *any two cells or systems of cells that are repeatedly active at the same time will tend to become associated, so that activity in one facilitates activity in the other*. In our setting, it means that Γ must achieve a maximum at the origin so that neurons in phase become associated and increase the coupling weights. The choice $\Gamma(\theta) = \cos(\theta)$ was proposed in the above references as a prototype of *Hebbian learning*. However, negative (unrealistic) coupling weights might eventually arise. As a modification of such choice, the following plasticity function was proposed in Ref. 191:

$$\Gamma(\theta) := \frac{\sigma^{2\alpha}}{\left(\sigma^2 + |\theta|_o^2\right)^{\alpha}}, \tag{5.25}$$

where $\sigma \in (0, \pi)$, and $|\theta|_o$ is the geodesic distance of $e^{i\theta}$ to 1 along the unit circle, that is

$$|\theta|_o := |\bar{\theta}| \quad \text{for } \bar{\theta} \equiv \theta \mod 2\pi, \quad \bar{\theta} \in (-\pi, \pi].$$

In order to achieve a new dynamics that is relevant in this paradigm of the Kuramoto model, the following regime with fast learning and singular weights was proposed. By scaling

$$\eta \to \varepsilon^{-1}, \quad \sigma \to \varepsilon, \quad \kappa \to \kappa\varepsilon^{-2\alpha},$$

the formal limit $\varepsilon \searrow 0$ in (5.23)–(5.25) produces the following model with singular weights

$$\begin{aligned} \frac{d\theta_i}{dt} &= \Omega_i + \frac{\kappa}{N}\sum_{j=1}^{N} h(\theta_j - \theta_i), \\ \theta_i(0) &= \theta_{i,0}, \end{aligned} \tag{5.26}$$

where the interaction kernel reads

$$h(\theta) := \frac{\sin\theta}{|\theta|_o^{2\alpha}}, \quad \theta \in \mathbb{R}. \tag{5.27}$$

This is a sort of Kuramoto–Daido model with three regimes of singularity: subcritical for $\alpha \in (0, \frac{1}{2})$, critical for $\alpha = \frac{1}{2}$ and supercritical for $\alpha \in (\frac{1}{2}, 1)$. The well-posedness of global-in-time absolutely continuous solutions of the agent-based system was analyzed in Ref. 191; specially in the critical and supercritical regime, where the kernel is discontinuous and solutions in the sense of Filippov were proposed. Filippov solutions are nothing but solutions to the differential inclusion into the Filippov set-valued map of the system. For the more singular cases, the Filippov set-valued map at $\Theta \in \mathbb{R}^N$ consist of the values $(\omega_1, \dots, \omega_N) \in \mathbb{R}^N$ parameterized by

$$\omega_i = \Omega_i + \frac{\kappa}{N} \sum_{\substack{1 \le j \le N \\ \bar{\theta}_j \neq \bar{\theta}_i}} h(\theta_j - \theta_i) + \frac{\kappa}{N} \sum_{\substack{1 \le j \le N \\ \bar{\theta}_j = \bar{\theta}_i}} y_{ij}, \tag{5.28}$$

for some skew-symmetric matrix $Y = (y_{ij})_{1 \le i,j \le N}$ with general items in $\mathbb{R}$ if $\alpha \in (\frac{1}{2}, 1)$ or items in $[-1, 1]$ if $\alpha = \frac{1}{2}$. Although one-sided uniqueness follows for $\alpha \in (0, \frac{1}{2}]$ (because the Filippov set-valued map (5.28) is one-sided Lipschitz), it is not clear yet for the supercritical regime $\alpha \in (\frac{1}{2}, 1)$. Indeed, two different methods to obtain solutions were proposed in such paper: rigorous limit $\varepsilon \searrow 0$ and continuation criterion of classical solutions after collision times. Checking whether they agree is an open problem yet that has only be solved in Ref. 191 for two oscillators. In this part, we are interested in the rigorous derivation of macroscopic equations from the underlying description (5.26)–(5.27) at the microscopic scale. Before reviewing such results in Ref. 200, let us briefly emphasize the main novelties that this new model introduces: *finite-time sticking* and *clustering into groups*.

Theorem 5.11. (Ref. 191) *Consider $\Theta = (\theta_1, \dots, \theta_N)$, the global-in-time classical solution to (5.26)–(5.27) for $\alpha \in (0, \frac{1}{2})$. Assume that two oscillators collide at t^*, i.e. $\bar{\theta}_i(t^*) = \bar{\theta}_j(t^*)$ for some $i \neq j$. Then, the following two statements are equivalent*:

(1) *θ_i and θ_j stick together for all $t \ge t^*$.*
(2) *Their natural frequencies agree, i.e. $\Omega_i = \Omega_j$.*

However, some richer phenomena arise in the critical regime.

Theorem 5.12. (Ref. 191) *Consider $\Theta = (\theta_1, \dots, \theta_N)$ the global-in-time Filippov solution to (5.26)–(5.27) for $\alpha = \frac{1}{2}$. Assume that t^* is some collision time and fix any formed cluster with indices in the set $E \subseteq \{1, \dots, N\}$ and size $\#E = n$. Then, the following two statements are equivalent*:

(1) *The n oscillators in the cluster E stick all together after $t = t^*$.*

(2) *The next condition takes place*

$$\left|\frac{1}{n}\sum_{i\in E}\Omega_i - \frac{1}{m}\sum_{i\in I}\Omega_i\right| \leq \frac{\kappa}{N}(n-m), \tag{5.29}$$

for every $1 \leq m \leq n$ *and every* $I \subseteq E$ *such that* $\#I = m$.

Then, oscillators with different natural frequencies are still allowed to stick in finite time after a collision takes place as long as condition (5.29) holds true. This is a conditional result, since we have not shown yet that finite-time collision can take place. However, explicit sufficient conditions for global phase synchronization in finite time were also obtained in Ref. [191] for identical oscillators ($\Omega_i = 0$) initially confined to the half circle both in the subcritical and critical regimes. It is an analogue to the above asymptotic complete phase synchronization of identical oscillators in Theorem 5.2.

Theorem 5.13. *Let* $\Theta = (\theta_1, \ldots, \theta_N)$ *be the classical solution to* (5.26)–(5.27) *with* $\alpha \in \left(0, \frac{1}{2}\right)$ *for identical oscillators* ($\Omega_i = 0$). *Assume that the initial configuration* Θ_0 *is confined in a half circle, i.e.* $0 < D_\Theta(0) < \pi$. *Then, there is complete phase synchronization at a finite time not larger than* T_c, *where*

$$T_c = \frac{D_\Theta(0)^{1-2\alpha}}{2\alpha\kappa h(D_\Theta(0))}.$$

Theorem 5.14. *Let* $\Theta = (\theta_1, \ldots, \theta_N)$ *be the Filippov solution to* (5.26)–(5.27) *with* $\alpha = \frac{1}{2}$ *for identical oscillators* ($\Omega_i = 0$). *Assume that the initial configuration* Θ_0 *is confined in a half circle, i.e.* $0 < D_\Theta(0) < \pi$. *Then, there is complete phase synchronization in a finite time not larger than* T_c, *where*

$$T_c = \frac{D_\Theta(0)}{\kappa h(D_\Theta(0))}.$$

Indeed, such result would remain true in the supercritical regime if the Filippov solutions obtained through the above two methods agreed [191, Remark 5.3]. However, there are still some open problems regarding the non-identical case in the critical and supercritical regimes. Such problem was easily solved in the subcritical regime, yielding purely asymptotic complete frequency synchronization of non-identical oscillators and emergence of phase-locked states [191, Theorem 5.3].

Since in this paper we are interested in the kinetic description, we will not give more unnecessary details about the microscopic scale other than the above-mentioned tools that are required in order to properly understand the kinetic scale. Notice that the kernel h in (5.27) is no longer Lipschitz-continuous and the above techniques by Neunzert[187] that were used in Refs. 58 and 168 do not work for these more singular regimes. Also, the approach in Theorem 5.1 does not yield any result since the divergence (derivative in 1D) of the coupling force is not bounded anymore in any of the regimes $\alpha \in (0, 1)$. A new approach was introduced in Ref. 200 to deal with this sort of kernels for $\alpha \in (0, \frac{1}{2}]$. It produced weak measure-valued solutions

(in the sense of the Filippov flow for $\alpha \in \frac{1}{2}$) to the kinetic singular Kuramoto model. Here, we will first recall the subcritical and critical regimes whilst the supercritical case will be sketched later. On the one hand, it is clear that the kinetic singular Kuramoto model stands for

$$\partial_t f + \partial_\theta \left(\left(\Omega + \kappa \int_{\mathbb{T}\times\mathbb{R}} h(\theta^* - \theta) f(t, d\theta^*, d\Omega^*) \right) f \right) = 0,$$

for all $t \geq 0,\ (\theta, \Omega) \in (-\pi, \pi] \times \mathbb{R}$, with periodic boundary conditions $f(t, -\pi, \Omega) = f(t, \pi, \Omega)$. In order to avoid it, a standard trick was used in Ref. 200 to identify a transport equation along $\mathbb{T} \times \mathbb{R}$

$$\begin{cases} \dfrac{\partial f}{\partial t} + \nabla \cdot (\boldsymbol{\mathcal{V}}[f] f) = 0, \quad (z, \Omega) \in \mathbb{T} \times \mathbb{R}, \\ f(0) = f_0, \end{cases} \tag{5.30}$$

where the divergence is considered along $\mathbb{T} \times \mathbb{R}$ and the nonlinear transport field reads

$$\boldsymbol{\mathcal{V}}[f](z, \Omega) := (\mathcal{P}[f](z, \Omega) i z, 0), \tag{5.31}$$

$$\mathcal{P}[f](z, \Omega) := \Omega + \kappa \int_{\mathbb{T}\setminus\{z\}} \int_{\mathbb{R}} h(\theta^* - \theta) f(d\theta^*, d\Omega^*), \tag{5.32}$$

for any $(z = e^{i\theta}, \Omega) \in \mathbb{T} \times \mathbb{R}$. Obviously, the transport field only makes sense for $\alpha \in (0, \frac{1}{2}]$ because f_t is merely measure-valued. Indeed, the integral is intentionally considered off $\{z\}$ to avoid concentration issues for $\alpha = \frac{1}{2}$. Notice that it is totally consistent with the microscopic dynamics as $y_{ii} = 0$ in (5.28). It does not make any sense for $\alpha \in (\frac{1}{2}, 1)$ unless f_t enjoys some extra integrability, that we do not expect to propagate due to concentration phenomena at the microscopic scale.

Existence and sided-uniqueness of classical flow for $\alpha \in (0, \frac{1}{2})$ (respectively Filippov flow for $\alpha = \frac{1}{2}$) of the transport field $\boldsymbol{\mathcal{V}}[f]$ was guaranteed in Ref. 200 due to the fact that the transport field is continuous with linear growth at infinity for $\alpha \in (0, \frac{1}{2})$ (respectively, it is locally bounded with linear growth at infinity for $\alpha = \frac{1}{2}$) and it is one-sided Lipschitz continuous. In addition, the mean-field limit approach was shown to work despite the fact that Theorem 5.1 does not apply. Specifically, empirical measures supported on classical (respectively Filippov) solutions to (5.26) were proved to be solutions to (5.26) that converge to the unique weak measure-valued solution (respectively, solution in the sense of the Filippov flow) to (5.30) as $N \to \infty$. Indeed, a similar Dobrushin-type estimate to (5.10) was obtained, thus quantifying the mean-field limit.

Theorem 5.15. (Ref. 200) *Consider* $\alpha \in (0, \frac{1}{2}]$, $\kappa > 0$ *and two time-dependent probability measures* $f, \widetilde{f} \in AC_{\mathrm{loc}}([0, \infty), C_c^\infty(\mathbb{T} \times \mathbb{R})^* - weak*)$, *solving* (5.30) *weakly* (*respectively in the sense of Filippov flow*) *with initial data* $f_0, \widetilde{f}_0 \in \mathcal{P}_2(\mathbb{T}\times\mathbb{R})$.

Then,

$$W_2(f(t),\widetilde{f}(t)) \leq e^{\left(\frac{1}{2}+2\kappa L_0\right)t} W_2(f_0,\widetilde{f}_0),$$

for every $t \geq 0$, *where* L_0 *is the one-sided Lipschitz constant of* $-h$.

Here, $\mathcal{P}_2(\mathbb{T}\times\mathbb{R})$ represents the metric space of probability measures on $\mathbb{T}\times\mathbb{R}$ with finite second-order moment endowed with the 2-Wasserstein distance W_2. The extra tightness in $\mathcal{P}_2$ is required in order for W_2 to make sense. It allows obtaining a quantitative mean-field limit when $\widetilde{f} = \mu^N$ and the initial empirical measures approximate the initial datum f_0, i.e.

$$\lim_{N\to\infty} W_2(\mu_0^N, f_0) = 0.$$

Nevertheless, the mean-field limit is still valid in $\mathcal{P}_1(\mathbb{T}\times\mathbb{R})$ without a control on rates. In addition, notice that uniqueness follows by simply choosing $f_0 = \widetilde{f}_0 \in \mathcal{P}_2(\mathbb{T}\times\mathbb{R})$. One can still obtain uniqueness for general probability measures $f_0 \in \mathcal{P}(\mathbb{T}\times\mathbb{R})$ by virtue of a similar Dobrushin inequality with respect to a modified Wasserstein distance. Specifically, if $f, \widetilde{f} \in \mathcal{P}(\mathbb{T}\times\mathbb{R})$ are probability measures with $g := (\pi_\theta)_\# f = (\pi_\theta)_\#\widetilde{f}$ and $\{f(\cdot|\Omega)\}_{\Omega\in\mathbb{T}}$, $\{\widetilde{f}(\cdot|\Omega)\}_{\Omega\in\mathbb{R}}$ are the families of conditional probabilities or disintegrations, then we can define the fibered 2-Wasserstein distance

$$W_{2,g}(f,\widetilde{f}) := \left(\int_{\mathbb{R}} W_2\big(f(\cdot|\Omega),\widetilde{f}(\cdot|\Omega)\big)^2\, g(d\Omega)\right)^{1/2}.$$

The following result directly implies uniqueness in the full sense.

Theorem 5.16. (Ref. 200) *Consider* $\alpha \in (0,\frac{1}{2}]$, $\kappa > 0$ *and two time-dependent probability measures* $f, \widetilde{f} \in AC_{\mathrm{loc}}([0,\infty), C_c^\infty(\mathbb{T}\times\mathbb{R})^* - weak*)$, *solving* (5.30) *weakly* (*respectively in the sense of Filippov flow*) *with initial data* $f_0, \widetilde{f}_0 \in \mathcal{P}(\mathbb{T}\times\mathbb{R})$. *Consider the distribution of natural frequencies* $g = (\pi_\Omega)_\# f_0$ *and* $\widetilde{g} = (\pi_\Omega)_\#\widetilde{f}_0$. *If both distributions of natural frequencies agree* $g = \widetilde{g}$, *then*

$$W_{2,g}(f(t),\widetilde{f}(t)) \leq W_{2,g}(f_0,\widetilde{f}_0)e^{2\kappa L_0 t},$$

for every $t \geq 0$, *where* L_0 *is the one-sided Lipschitz constant of* $-h$.

In addition, the mean-field limit allows transferring Theorems 5.13 and 5.14 toward the macroscopic equation (5.30).

Theorem 5.17. (Ref. 200) *Set* $\alpha \in (0,\frac{1}{2})$ *and consider any initial datum* $f_0 \in \mathcal{P}(\mathbb{T}\times\mathbb{R})$ *with identical distribution of natural frequencies, namely,* $g = (\pi_\Omega)_\# f_0 = \delta_0$. *Let* f *be the unique global-in-time weak measure-valued solution to* (5.30) *issued at* f_0 *and assume that* $0 < D_\theta(f_0) < \pi$. *Then,*

$$f(t) = f_\infty \quad \text{for all } t \geq T_c,$$

where $T_c = \frac{D_\theta(f_0)^{1-2\alpha}}{2\alpha\kappa h(D_\theta(f_0))}$, *the equilibrium* f_∞ *is given by the monopole* $f_\infty := \delta_{z_{\mathrm{av}}(0)}(z)\otimes\delta_0(\Omega)$ *and* z_{av} *is the average phase of the oscillators.*

Theorem 5.18. (Ref. 200) *Set* $\alpha = \frac{1}{2}$ *and consider any initial datum* $f_0 \in \mathcal{P}(\mathbb{T}\times\mathbb{R})$ *with identical distribution of natural frequencies, namely,* $g = (\pi_\Omega)_\# f_0 = \delta_0$. *Let* f *be the unique global-in-time measure-valued solution to* (5.30) *in the sense of the Filippov flow issued at* f_0 *and assume that* $0 < D_\theta(f_0) < \pi$. *Then,*

$$f(t) = f_\infty \quad \textit{for all } t \geq T_c,$$

where $T_c = \frac{D_\theta(f_0)}{\kappa h(D_\theta(f_0))}$, *the equilibrium* f_∞ *is given by the monopole* $f_\infty := \delta_{z_{\rm av}(0)}(z) \otimes \delta_0(\Omega)$ *and* $z_{\rm av}$ *is the average phase of the oscillators.*

5.4. *From Kuramoto to CS*

Recall that the CS model was introduced in Sec. 4 as an adequate second-order model for swarms.[84] At the microscopic level, the CS model describes a system of N interacting agents under the influence of self-alignment effects through the dynamical system (4.1). The idea is that each agent should modify its velocity to mimic some average relative velocity of the remaining $N-1$ agents. The most natural setting is the homophilous case (4.2)

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

where weights depend decreasingly on the distance between agents (the closer the agents, the stronger the influence). Here, ψ plays the role of the *communication kernel* or *influence function* and κ stands for the strength of the influence. The non-negative exponent β controls the asymptotic fall-off of long range interactions.

Here, we will briefly mention a link between Kuramoto and CS models. It does not only work to relate the classical Kuramoto to CS model with smooth influence function, but also between the above singular Kuramoto model and the CS model with weakly singular influence functions. In general, consider a Kuramoto–Daido model as follows

$$\begin{aligned} \dot\theta_i &= \Omega_i + \frac{\kappa}{N}\sum_{j=1}^N \phi(\theta_j - \theta_i), \\ \theta_i(0) &= \theta_{i,0}, \end{aligned} \tag{5.33}$$

for some periodic force ϕ. If it is smooth, differentiation implies

$$\begin{aligned} \dot\theta_i &= \omega_i, \\ \dot\omega_i &= \frac{\kappa}{N}\sum_{j=1}^N \psi(\theta_j - \theta_i)(\omega_j - \omega_i), \\ (\theta_i(0), \omega_i(0)) &= (\theta_{i,0}, \omega_{i,0}), \end{aligned} \tag{5.34}$$

where $\psi := \phi'$. Notice that the initial and natural frequencies are related through the rule

$$\omega_{i,0} = \Omega_i + \frac{\kappa}{N}\sum_{j=1}^N \phi(\theta_{j,0} - \theta_{i,0}).$$

This amounts to say that Kuramoto–Daido model (5.33) agrees with CS model (5.34) for well-prepared initial data. In other words, the Kuramoto–Daido model (5.33) implicitly describes the evolution of swarms. In particular, when $\phi(\theta) = \sin\theta$, one obtains that the classical Kuramoto model agrees with the CS model with an influence function $\psi(\theta) = \cos\theta$. This can be used to understand frequency synchronization as a flocking phenomenon and allows transferring techniques between both models in 1D. See Ref. 121, where it was introduced and used to derive an alternative kinetic description for synchronization of oscillators

$$\partial_t F + \omega\partial_\theta F + \partial_\omega\left[\kappa\left(\int_{\mathbb{T}\times\mathbb{R}} \cos(\theta^* - \theta)(\omega^* - \omega)F(t, d\theta^*, d\omega^*, d\Omega^*)\right)F\right] = 0.$$

Regarding the singular cases, in the subcritical regime frequencies enjoy a minimum regularity required to describe the augmented second-order equation. Namely, we have the following.

Theorem 5.19. (Ref. 191) *Consider a classical solution* $\Theta = (\theta_1, \dots, \theta_N)$ *to* (5.26) *with* $\alpha \in (0, \frac{1}{2})$. *Then, the frequencies verify* $\dot\theta_i \in W^{1,p}([T_{k-1}, \tau])$, *for* $1 \le p < \frac{1}{2\alpha}$, *every* $k \in \mathbb{N}$ *and every* $\tau \in (T_{k-1}, T_k)$. *In addition, they verify the following equation in weak sense*

$$\ddot\theta_i = \frac{\kappa}{N}\sum_{j\notin S(i)(T_{k-1})} h'(\theta_j - \theta_i)(\dot\theta_j - \dot\theta_i), \tag{5.35}$$

for all $t \in [T_{k-1}, \tau]$. *Here,* $\{T_k\}_{k\in N}$ *are the new collision times after some oscillators have stick together and* $S_i(T_{k-1})$ *means the set of indices* j *of oscillators that stick with the* i*th one at* $t = T_{k-1}$.

In other words, the singular Kuramoto model in the weakly singular regime agrees with the CS model with influence function

$$\psi(\theta) = h'(\theta) = \frac{1}{|\theta|_o^{2\alpha}}\left[\cos\theta - 2\alpha\frac{\sin|\theta|_o}{|\theta|_o}\right] \sim \frac{1}{|\theta|_o^{2\alpha}}.$$

Similar piecewise solutions were obtained in Refs. 194 and 195 for the genuine CS model in general dimensions endowed with the standard weakly singular kernel

$$\psi(r) = \frac{1}{r^\beta}, \tag{5.36}$$

for the range $\beta \in (0, 1)$. Notice that here, the parameter β for the weakly singular CS model (4.1)–(5.36) agrees with the parameter 2α in the singular Kuramoto model (5.26)–(5.27). Since such piecewise regularity becomes global in the smaller range $\beta \in (0, \frac{1}{2})$, the authors used it to produce measure-valued solutions to the associated kinetic CS model (4.3) with weakly singular weights (5.36),

namely,

$$\begin{aligned}
&\partial_t f + \mathbf{v}\cdot\nabla_{\mathbf{x}} f + \nabla_{\mathbf{v}}\cdot[\mathbf{F}[f]f] = 0, \quad (\mathbf{x},\mathbf{v})\in\mathbb{R}^{2d}, \ t\geq 0,\\
&\mathbf{F}[f](t,\mathbf{x},\mathbf{v}) := \kappa\int_{\mathbb{R}^{2d}} \psi(\|\mathbf{x}-\mathbf{x}^*)(\mathbf{v}^*-\mathbf{v})f(t,\mathbf{x}^*,\mathbf{v}^*)d\mathbf{x}^*\,d\mathbf{v}^*, \qquad (5.37)\\
&\psi(r) = r^{-\beta}, \quad r>0.
\end{aligned}$$

in such regime $\beta\in(0,\frac{1}{2})$ via the mean-field limit approach, see Ref. 184. This is the content of the next result.

Theorem 5.20. (Ref. 184) *Let us consider* $0<\beta<\frac{1}{2}$. *For any compactly supported initial data* $f_0\in\mathcal{P}(\mathbb{R}^{2d})$ *and any* $T>0$, (5.37) *admits at least one weak measure-valued solution* $f\in L^\infty(0,T;\mathcal{M}(\mathbb{R}^{2d}))$ *with* $\partial_t f\in L^p(0,T;C_c^1(\mathbb{R}^{2d})^*)$ *for some* $p>1$. *Moreover, if* f_0 *is of the form*

$$f_0(\mathbf{x},\mathbf{v}) := \sum_{i=1}^N m_i\delta_{(\mathbf{x}_i(0),\mathbf{v}_i(0))},$$

with $\sum_{i=1}^N m_i = 1$, *then* f *remains atomic of the form*

$$f(t,\mathbf{x},\mathbf{v}) := \sum_{i=1}^N m_i\delta_{(\mathbf{x}_i(t),\mathbf{v}_i(t))}.$$

In particular, it is a unique measure-valued solution to (5.37) (*weak atomic uniqueness*).

Notice that the main difference between the (singular) Kuramoto and (singular) CS models is the periodicity assumption along with the fact that for the Kuramoto model, the influence function is not always positive. Indeed, it also attains negative values near $\theta=\pi$. This means that Kuramoto oscillators with far apart phases are pushed away from the flock. Nevertheless, as shown in Refs. 33 and 120, the periodicity conditions recover the unique flock when the natural frequencies agree.

5.5. *Hydrodynamic limits in the Cucker–Smale model*

In this subsection, we will introduce several hyperbolic hydrodynamic limits that have been analyzed in the literature for the kinetic CS model (4.3). In the first part, we will focus on Lipschitz influence functions ψ, e.g. (4.2), which corresponds to the classical CS model. It produces a method to derive the well-known Euler-alignment model when the influence function is smooth. Later, since the full hydrodynamic limit in the singular regimes is still a hard open problem, we will introduce a singular hyperbolic hydrodynamic limit of vanishing inertia type for the weakly singular case. Such method yields a reduced first-order fluid model where inertia in the balance equation of momentum has been neglected in the flavor of the overdamped or Smoluchowski limit $\varepsilon\searrow 0$ in (5.6). Finally, due to its relation to this last case,

we will show that a similar approach can be done to derive weak measure-valued solutions of the kinetic singular Kuramoto model (5.30) in the supercritical regime $\alpha \in (\frac{1}{2}, 1)$. We will sketch a similar singular hyperbolic hydrodynamic limit of vanishing inertia type on an augmented Kuramoto model with inertia and regularized weights.

5.5.1. *Hydrodynamic limits for Lipschitz influence function*

When the weight function is Lipschitz (e.g. given by (4.2)) and if the masses m_i are normalized to one, the BBGKY approach get succeed[131,134] and we can rigorously deduce the kinetic CS model (4.3) with regular weights (4.2). Regarding the weakly singular cases (5.37) the only results that are available in the literature were introduced in Ref. 184 and, more specifically, they take the form of the preceding theorem.

To derive macroscopic hydrodynamic models from the classical (Lipschitz interactions) kinetic model of CS, new terms were introduced in Ref. 159. Specifically, the following (hyperbolic) scaled model was considered

$$\begin{aligned} &\partial_t f_\varepsilon + \mathbf{v}\cdot\nabla_{\mathbf{x}} f_\varepsilon + \nabla_{\mathbf{v}}\cdot[\mathbf{F}(f_\varepsilon)f_\varepsilon] = \frac{1}{\varepsilon}\nabla_{\mathbf{v}}\cdot[\nabla_{\mathbf{v}} f_\varepsilon + (\mathbf{v}-\mathbf{u}_\varepsilon)f_\varepsilon], \\ &\mathbf{F}(f_\varepsilon)(t,\mathbf{x},\mathbf{v}) := \kappa\int_{\mathbb{R}^{2d}} \psi(\|\mathbf{x}-\mathbf{x}^*\|)(\mathbf{v}^*-\mathbf{v})f(t,\mathbf{x}^*,\mathbf{v}^*)d\mathbf{x}^*\,d\mathbf{v}^*. \end{aligned} \tag{5.38}$$

Notice that such model includes velocity noise (through a Fokker–Planck term) and a *local alignment effect* of the velocity toward the mean velocity field

$$\mathbf{u}_\varepsilon(t,\mathbf{x}) = \frac{\int_{\mathbb{R}^d}\mathbf{v} f_\varepsilon(t,\mathbf{x},\mathbf{v})\,d\mathbf{v}}{\int_{\mathbb{R}^d} f_\varepsilon(t,\mathbf{x},\mathbf{v})\,d\mathbf{v}}.$$

The hyperbolic scaling sets a regime with large noise and strong local alignment but weak nonlocal alignment of CS type. Notice that such local alignment term can be regarded as linear damping toward the macroscopic velocity field and provides no effect on the balance equation of momentum by virtue of its cancelations

$$\int_{\mathbb{R}^N}\mathbf{v}\nabla_{\mathbf{v}}\cdot[(\mathbf{v}-\mathbf{u}_\varepsilon)f_\varepsilon]d\mathbf{v} = -\int_{\mathbb{R}^N}(\mathbf{v}-\mathbf{u}_\varepsilon)f_\varepsilon\,d\mathbf{v} = 0.$$

This local alignment term $(\mathbf{v}-\mathbf{u})f$ was introduced in Ref. 181 as the singular limit $\psi \to \delta_0$ in the Mostch–Tadmor nonlinear alignment term

$$\mathbf{F}_{MT}(f)(t,\mathbf{x},\mathbf{v}) = \frac{\kappa\int_{\mathbb{R}^{2d}}\psi(\|\mathbf{x}-\mathbf{x}^*\|)(\mathbf{v}^*-\mathbf{v})f(t,\mathbf{x}^*,\mathbf{v}^*)d\mathbf{x}^*\,d\mathbf{v}^*}{\int_{\mathbb{R}^{2d}}\psi(\|\mathbf{x}-\mathbf{x}^*\|)f(t,\mathbf{x},\mathbf{v}^*)d\mathbf{x}^*\,d\mathbf{v}^*}. \tag{5.39}$$

The main idea in (5.39) is to normalized the pairwise interactions $\psi(\|\mathbf{x}_i-\mathbf{x}_j\|)$ between agents in terms of a relative influence. Of course, it breaks the symmetry of the initial CS model, what in particular causes severe problems to recover such kinetic model as mean-field limit of the corresponding agent-based problem.

When $\varepsilon \to 0$, relative entropy methods were used in Ref. 159 to obtain the hydrodynamic limit of (5.38), that takes the form

$$\begin{aligned}
&\partial_t \rho + \nabla_{\mathbf{x}} \cdot (\rho\, \mathbf{u}) = 0,\\
&\partial_t (\rho\, \mathbf{u}) + \nabla_{\mathbf{x}} \cdot (\rho\, \mathbf{u} \otimes \mathbf{u})\\
&\qquad = -\nabla_{\mathbf{x}} \rho + \int_{\mathbb{R}^d} \psi(\|\mathbf{x} - \mathbf{x}^*\|)(\mathbf{u}(t, \mathbf{x}^*) - \mathbf{u}(t, \mathbf{x}))\rho(t, \mathbf{x})\rho(t, \mathbf{x}^*) d\mathbf{x}^*.
\end{aligned}$$

Note that such models maintain nonlocal alignment effects but does not include any local damping, but it disappeared in the limit. Indeed, the strong local alignment in (5.38) was only introduced as an extra term that helps the system to reach the hydrodynamic regime. Also, notice that a pressure term $-\nabla_{\mathbf{x}} \rho$ has appeared as a consequence of the velocity noise in the Fokker–Planck term of the right-hand side of (5.38).

In relation with such scaling, the method was very recently improved in Ref. 109 to remove the velocity term noise. Specifically, the following system was considered

$$\begin{aligned}
&\partial_t f_\varepsilon + \mathbf{v} \cdot \nabla_{\mathbf{x}} f_\varepsilon + \nabla_{\mathbf{v}} \cdot [\mathbf{F}(f_\varepsilon) f_\varepsilon] = \frac{1}{\varepsilon} \nabla_{\mathbf{v}} \cdot [(\mathbf{v} - \mathbf{u}_\varepsilon) f_\varepsilon],\\
&\mathbf{F}(f_\varepsilon)(t, \mathbf{x}, \mathbf{v}) := \kappa \int_{\mathbb{R}^{2d}} \psi(\|\mathbf{x} - \mathbf{x}^*\|)(\mathbf{v}^* - \mathbf{v}) f(t, \mathbf{x}^*, \mathbf{v}^*) d\mathbf{x}^*\, d\mathbf{v}^*.
\end{aligned} \tag{5.40}$$

Again, similar relative methods allow recovering the well-known pressureless Euler-alignment model

$$\begin{aligned}
&\partial_t \rho + \nabla_{\mathbf{x}} \cdot (\rho\, \mathbf{u}) = 0,\\
&\partial_t (\rho\, \mathbf{u}) + \nabla_{\mathbf{x}} \cdot (\rho\, \mathbf{u} \otimes \mathbf{u})\\
&\qquad = \int_{\mathbb{R}^d} \psi(\|\mathbf{x} - \mathbf{x}^*\|)(\mathbf{u}(t, \mathbf{x}^*) - \mathbf{u}(t, \mathbf{x}))\rho(t, \mathbf{x})\rho(t, \mathbf{x}^*) d\mathbf{x}^*,
\end{aligned} \tag{5.41}$$

where the above pressure term in the right-hand side of the momentum equation has disappeared. Such model arises as the monokinetic ansatz of the CS model but the rigorous hydrodynamic limit toward such monokinetic distribution has not been proved yet without the help of any extra damping or strong local alignment terms. The pressureless Euler-alignment system has been analyzed by many authors during the recent years.[61,96,138,162,163,215–217,220] Specially, one of the main topics of interest is the study of the phenomenon of critical threshold on the initial data, that discriminates initial data enjoying blow-up from data enjoying global regularity depending on the properties of the "magic quantity"

$$G = \partial_x u + \psi * \rho.$$

Another close approach was given in Ref. 158. The velocity noise and nonlocal CS alignment term in (5.38) were neglected but the strong local alignment was kept and linear damping was also added to the system

$$\partial_t f_\varepsilon + \mathbf{v} \cdot \nabla_{\mathbf{x}} f_\varepsilon - \lambda \nabla_{\mathbf{v}} \cdot (\mathbf{v} f_\varepsilon) = \frac{1}{\varepsilon} \nabla_{\mathbf{v}} \cdot ((\mathbf{v} - \mathbf{u}_\varepsilon) f_\varepsilon).$$

In this case, a similar analysis provides the limiting system

$$\partial_t \rho + \nabla_{\mathbf{x}} \cdot (\rho\, \mathbf{u}) = 0,$$
$$\partial_t (\rho\, \mathbf{u}) + \nabla_{\mathbf{x}} \cdot (\rho\, \mathbf{u} \otimes \mathbf{u}) = -\lambda \mathbf{u}.$$

Again, the strong local alignment is lost in the macroscopic system, but a linear damping has been recovered in the limit. In the same line, when agents are driven by a fluid, the following coupled system with fluids has been considered in Ref. 60

$$\partial_t f_\varepsilon + \mathbf{v} \cdot \nabla_{\mathbf{x}} f_\varepsilon + \nabla_{\mathbf{v}} \cdot ((\mathbf{U}_\varepsilon - \mathbf{v}) f_\varepsilon) = \frac{1}{\varepsilon} \nabla_{\mathbf{v}} \cdot (\nabla_{\mathbf{v}} f_\varepsilon + (\mathbf{v} - \mathbf{u}_\varepsilon) f_\varepsilon),$$

where $\mathbf{U}_\varepsilon$ is the velocity field of the fluid, which evolves according to the incompressible NS system, i.e.

$$\partial_t \mathbf{U}_\varepsilon + (\mathbf{U}_\varepsilon \cdot \nabla_{\mathbf{x}}) \mathbf{U}_\varepsilon = -\nabla_{\mathbf{x}} p_\varepsilon + \nu \Delta \mathbf{U}_\varepsilon + \int_{\mathbb{R}^N} (\mathbf{v} - \mathbf{U}_\varepsilon) f_\varepsilon \, d\mathbf{v},$$
$$\nabla_{\mathbf{x}} \cdot \mathbf{U}_\varepsilon = 0,$$

where $\nu \geq 0$ is the viscosity and p_ε stands for the pressure of the fluid. In such paper, an entropy method in the spirit of Ref. 159 was derived to pass to the limit $\varepsilon \to 0$ and the following limiting macroscopic system was obtained

$$\partial_t \rho + \nabla_{\mathbf{x}} \cdot (\rho\, \mathbf{u}) = 0,$$
$$\partial_t (\rho\, \mathbf{u}) + \nabla_{\mathbf{x}} \cdot (\rho\, \mathbf{u} \otimes \mathbf{u}) + \nabla_{\mathbf{x}} \rho = \rho(\mathbf{U} - \mathbf{u}),$$

coupled with the limiting NS system

$$\partial_t \mathbf{U} + (\mathbf{U} \cdot \nabla_{\mathbf{x}}) \mathbf{U} = -\nabla_{\mathbf{x}} p + \nu \Delta \mathbf{U} + \rho(\mathbf{u} - \mathbf{U}),$$
$$\nabla_{\mathbf{x}} \cdot \mathbf{U} = 0.$$

For the readers' convenience, let us mention another alternative to the above hydrodynamic limits in which the scaling lead to a vanishing inertia effect on the macroscopic limit, thus reducing the second-order dynamics to the Smoluchoski first-order dynamics. In particular, this line has been developed in Refs. 31, 117, 188 and 198 for the Vlasov–Poisson–Fokker–Planck system, that give rise to the aggregation equation with Newtonian interactions, i.e.

$$\partial_t \rho + \nabla_{\mathbf{x}} \cdot (\rho\, \mathbf{u}) = 0,$$
$$\mathbf{u} = -\nabla_{\mathbf{x}} \varphi,$$

where the potential $\varphi = \varphi(t, x)$ can be recovered from the density ρ through the Poisson equation

$$\Delta \varphi = \theta \rho$$

and $\theta = 1$ or $\theta = -1$ depending on the attractive of repulsive character of the Newtonian interactions. In the parabolic case, the limiting system changes the velocity field from $\mathbf{u} = -\nabla_{\mathbf{x}} \varphi$ to $\mathbf{u} = -\nabla_{\mathbf{x}} \varphi + \frac{\nabla_{\mathbf{x}} \rho}{\rho}$, that includes viscosity on the continuity equation for ρ.

5.5.2. *Hydrodynamic singular limits of vanishing inertia type*

Mimicking the preceding ideas we can consider the kinetic singular CS model, with linear damping, velocity noise (Fokker–Planck term) and the effect of an external force $-\nabla_{\mathbf{x}} V$. A dimensionless analysis was proposed in Ref. [201], leading to the following scaled system:

$$\begin{aligned} &\partial_t f_\varepsilon + \mathbf{v}\cdot\nabla_{\mathbf{x}} f_\varepsilon + \frac{1}{\varepsilon}\nabla_{\mathbf{v}}\cdot(\mathbf{F}_\varepsilon(f_\varepsilon) f_\varepsilon) - \frac{1}{\varepsilon}\nabla_{\mathbf{x}} V_\varepsilon\cdot\nabla_{\mathbf{v}} f_\varepsilon \\ &\quad = \frac{1}{\varepsilon}\nabla_{\mathbf{v}}\cdot(\mathbf{v} f_\varepsilon + \nabla_{\mathbf{v}} f_\varepsilon), \\ &\mathbf{F}_\varepsilon(f_\varepsilon)(t,\mathbf{x},\mathbf{v}) := \int_{\mathbb{R}^{2d}} \psi_\varepsilon(\|\mathbf{x}-\mathbf{x}^*\|)(\mathbf{v}^*-\mathbf{v}) f_\varepsilon(t,\mathbf{x}^*,\mathbf{v}^*) d\mathbf{x}^*\, d\mathbf{v}^*. \end{aligned} \tag{5.42}$$

Here, the singular influence function (5.36) in (5.37) has been regularized as follows

$$\psi_\varepsilon(r) = \frac{1}{(\varepsilon^2 + c_\beta r^2)^{\beta/2}}, \quad r > 0, \tag{5.43}$$

for some $\beta \in (0,1]$ and a β-dependent coefficient c_β. Notice that the scaled kernel (5.43) converges toward the singular one (5.36) as $\varepsilon \to 0$. Given a sufficiently regular initial data $f_\varepsilon(0) = f_\varepsilon(0,\mathbf{x},\mathbf{v})$ and the corresponding smooth solution $f_\varepsilon(t) = f_\varepsilon(t,\mathbf{x},\mathbf{v})$ to the regularized system (5.42), one can associate the macroscopic quantities:

$$\begin{aligned} \text{Density:} \quad & \rho_\varepsilon(t,\mathbf{x}) := \int_{\mathbb{R}^d} f_\varepsilon(t,\mathbf{x},\mathbf{v}) d\mathbf{v}, \\ \text{Current:} \quad & \mathbf{j}_\varepsilon(t,\mathbf{x}) := \int_{\mathbb{R}^d} \mathbf{v}\, f_\varepsilon(t,\mathbf{x},\mathbf{v}) d\mathbf{v}, \\ \text{Velocity field:} \quad & \mathbf{u}_\varepsilon(t,\mathbf{x}) := \frac{\mathbf{j}_\varepsilon(t,\mathbf{x})}{\rho_\varepsilon(t,\mathbf{x})}, \\ \text{Stress tensor:} \quad & \boldsymbol{\mathcal{S}}_\varepsilon(t,x) := \int_{\mathbb{R}^d} \mathbf{v}\otimes\mathbf{v}\, f_\varepsilon(t,\mathbf{x},\mathbf{v}) d\mathbf{v}, \\ \text{Stress flux tensor:} \quad & \boldsymbol{\mathcal{T}}_\varepsilon(t,\mathbf{x}) := \int_{\mathbb{R}^d} (\mathbf{v}\otimes\mathbf{v})\otimes\mathbf{v} f_\varepsilon(t,\mathbf{x},\mathbf{v}) d\mathbf{v}, \end{aligned}$$

which verifies the following conservation laws

$$\partial_t \rho_\varepsilon + \nabla_{\mathbf{x}}\cdot\mathbf{j}_\varepsilon = 0, \tag{5.44}$$

$$\varepsilon\,\partial_t \mathbf{j}_\varepsilon + \varepsilon\,\nabla_{\mathbf{x}}\cdot\boldsymbol{\mathcal{S}}_\varepsilon + \rho_\varepsilon\nabla_{\mathbf{x}} V_\varepsilon + \mathbf{j}_\varepsilon + (\psi_\varepsilon * \rho_\varepsilon)\,\mathbf{j}_\varepsilon - (\psi_\varepsilon * \mathbf{j}_\varepsilon)\,\rho_\varepsilon = 0, \tag{5.45}$$

$$\begin{aligned} &\varepsilon\,\partial_t \boldsymbol{\mathcal{S}}_{\boldsymbol{\varepsilon}} + \varepsilon\,\nabla_{\mathbf{x}}\cdot\boldsymbol{\mathcal{T}}_\varepsilon + 2\,\mathrm{Sym}(\mathbf{j}_\varepsilon\otimes\nabla_{\mathbf{x}} V_\varepsilon) \\ &\qquad + 2((1+\psi_\varepsilon*\rho_\varepsilon)\boldsymbol{\mathcal{S}}_\varepsilon - \rho_\varepsilon\mathbf{I}) - 2\,\mathrm{Sym}((\psi_\varepsilon*\mathbf{j}_\varepsilon)\otimes\mathbf{j}_\varepsilon) = 0 \end{aligned} \tag{5.46}$$

being $\mathrm{Sym}(\mathbf{M})$ the symmetric part of a square matrix $\mathbf{M}$, i.e. $\mathrm{Sym}(\mathbf{M}) := \frac{1}{2}(\mathbf{M} + \mathbf{M}^{\mathsf{T}})$.

The analysis of these equations allows obtaining the following estimates

$$\begin{aligned}
\|\rho_\varepsilon\|_{L^\infty(0,T;L^1(\mathbb{R}^d))} &= \|\rho_\varepsilon(0)\|_{L^1(\mathbb{R}^d)},\\
\|\|\mathbf{v}\|^2 f_\varepsilon\|_{L^1(0,T;L^1(\mathbb{R}^{2d}))} &\le \varepsilon\|\|\mathbf{v}\|^2 f_\varepsilon(0)\|_{L^1(\mathbb{R}^{2d})}\\
&\quad + \big(2dT + \|\nabla_{\mathbf{x}} V_\varepsilon\|^2_{L^2(0,T;L^\infty(\mathbb{R}^d))}\big)\|f_\varepsilon(0)\|_{L^1(\mathbb{R}^{2d})},\\
\|\mathbf{j}_\varepsilon\|_{L^2(0,T;L^1(\mathbb{R}^{2N}))} &\le (\|f_\varepsilon(0)\|_{L^1(\mathbb{R}^{2d})})^{1/2}\big(\|\|\mathbf{v}\|^2 f_\varepsilon\|_{L^1(0,T;L^1(\mathbb{R}^{2d}))}\big)^{1/2},
\end{aligned}$$

together with the dissipation of kinetic energy due to interactions

$$\begin{aligned}
&\int_0^T\int_{\mathbb{R}^{4d}} \psi_\varepsilon(\|\mathbf{x}-\mathbf{x}^*\|)\|\mathbf{v}-\mathbf{v}^*|^2 f_\varepsilon(t,\mathbf{x},\mathbf{v}) f_\varepsilon(t,\mathbf{x}^*,\mathbf{v}^*) d\mathbf{x}\, d\mathbf{x}^*\, d\mathbf{v}\, d\mathbf{v}^*\, dt\\
&\quad \le \varepsilon\|\|\mathbf{v}\|^2 f_\varepsilon(0)\|_{L^1(\mathbb{R}^{2d})} + \big(2dT + \|\nabla_{\mathbf{x}} V_\varepsilon\|^2_{L^2(0,T;L^\infty(\mathbb{R}^d))}\big)\|f_\varepsilon(0)\|_{L^1(\mathbb{R}^{2d})}.
\end{aligned} \tag{5.47}$$

This gives us the compactness for current and density sequences when $\varepsilon \searrow 0$

$$\begin{aligned}
\rho_\varepsilon &\overset{*}{\rightharpoonup} \rho, \quad \text{in } L^\infty(0,T;\mathcal{M}(\mathbb{R}^d)),\\
\mathbf{j}_\varepsilon &\overset{*}{\rightharpoonup} \mathbf{j}, \quad \text{in } L^2(0,T;\mathcal{M}(\mathbb{R}^d))^d.
\end{aligned}$$

Passing to the limit in all the linear terms (5.44)–(5.45) is clear. However, the main problem is the nonlinear term in (5.45). To achieve this goal, it is necessary to take into account the kindness of the commutator that defines the nonlinear term, the symmetry of the influence function, the range of values $\beta \in (0,1)$, as well as additional properties of convergence in time for ρ, see Ref. 201 for the details. The above properties allow us to identify the limit

$$\rho_\varepsilon \otimes \mathbf{j}_\varepsilon \overset{*}{\rightharpoonup} \rho \otimes \mathbf{j} \quad \text{in } L^2(0,T;\mathcal{M}(\mathbb{R}^{2d}))^d,$$

and prove that the limiting current vector $\mathbf{j} = \mathbf{j}(t,\mathbf{x})$ satisfies

$$\mathbf{j} = (\psi * \mathbf{j})\rho - (\psi * \rho)\mathbf{j} - \rho\nabla_{\mathbf{x}} V,$$

in the sense of distributions for $0 < \beta \le 1$.

Theorem 5.21. (Ref. 201) *Assume that the initial data f_ε^0 verify the set of hypothesis*

$$\begin{cases}
f_\varepsilon^0 = f_\varepsilon^0(\mathbf{x},\mathbf{v}) \ge 0 \quad \textit{and} \quad f_\varepsilon^0 \in C_c^\infty(\mathbb{R}^d \times \mathbb{R}^d),\\
\|f_\varepsilon^0\|_{L^1(\mathbb{R}^{2d})} \le M_0 \quad \textit{and} \quad \rho_\varepsilon^0 \overset{*}{\rightharpoonup} \rho^0 \ \textit{in } \mathcal{M}(\mathbb{R}^d),\\
\|\|\mathbf{v}\|^2 f_\varepsilon^0\|_{L^1(\mathbb{R}^{2d})} \le E_0,
\end{cases}$$

for every $\varepsilon > 0$ and some ε-independent constants $M_0, E_0 > 0$. Also, assume that the external forces $-\nabla_{\mathbf{x}} V_\varepsilon$ satisfy appropriate mild assumptions. Let $f_\varepsilon = f_\varepsilon(t,\mathbf{x},\mathbf{v})$ be the smooth solutions to (5.42) with $\beta \in (0,1]$. Then, ρ_ε and $\mathbf{j}_\varepsilon$ converge in a

weak sense to some finite Radon measure ρ *and* $\mathbf{j}$ *that solve the Cauchy problem associated with the following Euler-type system in the distributional sense*

$$\begin{aligned} &\partial_t \rho + \nabla_{\mathbf{x}} \cdot \mathbf{j} = 0, && \mathbf{x} \in \mathbb{R}^d, \ \ t \in [0,T), \\ &\rho \nabla_{\mathbf{x}} V + \mathbf{j} = (\phi * \mathbf{j})\rho - (\phi * \rho)\mathbf{j}, && \mathbf{x} \in \mathbb{R}^d, \ \ t \in [0,T) \\ &\rho(0,\cdot) = \rho^0, && \mathbf{x} \in \mathbb{R}^d. \end{aligned} \tag{5.48}$$

The problem of taking limits in the commutator $\rho_\varepsilon \otimes \mathbf{j}_\varepsilon - \mathbf{j}_\varepsilon \otimes \rho_\varepsilon$ was solved for $\beta \in (0,1)$ by cancelling the full singularity through the use of appropriate (Lipschitz-continuous) test functions and the properties of the commutator. The endpoint case $\beta = 1$ can be compared with the 2D Euler equations in vorticity formulation

$$\begin{aligned} &\partial_t \omega + \mathbf{u} \cdot \nabla_{\mathbf{x}} \omega = 0, && \mathbf{x} \in \mathbb{R}^2, \ \ t > 0, \\ &\mathbf{u} = \mathbf{K} * \omega, && \mathbf{x} \in \mathbb{R}^2, \ \ t > 0, \end{aligned} \tag{5.49}$$

where $\mathbf{K}$ is the so-called Biot–Savart kernel, that reads

$$\mathbf{K}(\mathbf{x}) = \frac{\mathbf{x}^\perp}{2\pi \|\mathbf{x}\|^2}, \quad \mathbf{x} \in \mathbb{R}^2.$$

This is the mean-field equation associated with the N vortex problem (5.4). In this context, a well-known bound of the vorticity in some logarithmic Morrey space is all we need to guarantee the absence of concentrations on the diagonal and to pass to the limit. Notice that in 2D Euler (5.49) the Biot–Savart kernel $\mathbf{K}$ is odd. However, the Riesz-type ψ in (5.36) for the weakly singular CS model is even and does not admit similar cancellations. Fortunately, the extra estimate for the dissipation of kinetic energy due to alignment interactions (5.47) gives rise to the required non-concentration estimate that allows the kinetic nonlinear term to be bounded for $\beta = 1$. Clearly, it allows obtaining a measure-valued solution to the asymptotic system also in the limiting case $\beta = 1$.

5.5.3. *Hydrodynamic limits in the singular Kuramoto model*

As mentioned before, similar hydrodynamic limits of vanishing inertia type have been considered in recent literature for other related systems like the Vlasov–Poisson–Fokker–Planck, the aggregation equation, the alignment-aggregation system and some other anisotropic versions of the aggregation equation, see Refs. 100, 101, 107, 108, 117 and 198 and last Subsec. 5.6. Before ending this part, we will sketch the idea for a different suitable system where one can apply such method. It was proposed in Ref. 200 as a completely different approach that yield solutions to the kinetic singular Kuramoto model (5.30) in the remaining regime $\alpha \in (\frac{1}{2}, 1)$. The cornerstone is again the cancellation property of the nonlinear term. We will show that indeed such cancellation works in the case of identical oscillators, i.e. $g = \delta_0$.

Specifically, one can consider the next scaled kinetic equation for the distribution function $F_\varepsilon = F_\varepsilon(t,\theta,\omega)$ at time t with phase $\theta \in \mathbb{T}$ and frequency $\omega \in \mathbb{R}$:

$$\begin{aligned}&\partial_t F_\varepsilon + \omega\partial_\theta F_\varepsilon + \frac{1}{\varepsilon}\partial_\omega\left[\kappa\left(\int_{\mathbb{T}\times\mathbb{R}\times\mathbb{R}} h_\varepsilon(\theta^*-\theta)F_\varepsilon(t,d\theta^*,d\omega^*)\right)F_\varepsilon\right]\\ &= \frac{1}{\varepsilon}\partial_\omega\left(\omega F_\varepsilon + \frac{\partial F_\varepsilon}{\partial\omega}\right).\end{aligned} \tag{5.50}$$

This is nothing but the Vlasov–McKean equation associated with the stochastic agent-based model

$$\begin{aligned}d\theta_i &= \omega_i\,dt,\\ \varepsilon d\omega_i &= \frac{\kappa}{N}\sum_{j=1}^{N} h_\varepsilon(\theta_j-\theta_i)dt - \omega_i\,dt + \sqrt{2\varepsilon}\,dW_t^i,\\ \theta_i(0) &= \theta_{i,0}, \quad \omega_i(0) = \omega_{i,0}.\end{aligned} \tag{5.51}$$

Such second-order system is a Kuramoto–Daido model with identical oscillators, regularized kernel h_ε, endowed with inertia, white noise W_t^i and frequency damping. The inertia term and noise have been scaled and disappeared as $\varepsilon \searrow 0$ while the scaled regularized kernel reads

$$h_\varepsilon(\theta) := \frac{\sin\theta}{(\varepsilon^2 + |\theta|_o^2)^\alpha},$$

and converges toward the singular kernel h in (5.27). Notice that the formal limit $\varepsilon \searrow 0$ in (5.51) recovers the singular first-order system (5.26). Then, we expect that the hydrodynamic limit in (5.50) can be closed and yields rigorous weak solutions to (5.30). Starting with smooth initial data F_ε^0, the above system (5.51) produces smooth solutions due to the regularizing effect of the diffusion and the regularized kernels h_ε. Again, in Ref. 200 the following ω moments were considered as the analogues of those in the above subsection:

$$\begin{aligned}\text{Phase density:} \quad & \rho_\varepsilon(t,\theta) := \int_{\mathbb{R}} F_\varepsilon(t,\theta,\omega)d\omega,\\ \text{Phase current:} \quad & j_\varepsilon(t,\theta) := \int_{\mathbb{R}} \omega F_\varepsilon(t,\theta,\omega)d\omega,\\ \text{Frequency field:} \quad & \omega_\varepsilon(t,\theta) := \frac{j_\varepsilon(t,\theta)}{\rho_\varepsilon(t,\theta)},\\ \text{Phase stress:} \quad & \mathcal{S}_\varepsilon(t,\theta) := \int_{\mathbb{R}} \omega^2 F_\varepsilon(t,\theta,\omega)d\omega,\\ \text{Stress flux tensor:} \quad & \mathcal{T}_\varepsilon(t,\theta) := \int_{\mathbb{R}} \omega^3 F_\varepsilon(t,\theta,\omega)d\omega.\end{aligned}$$

The corresponding conservation laws read

$$\partial_t \rho_\varepsilon + \partial_\theta j_\varepsilon = 0, \tag{5.52}$$

$$\varepsilon\, \partial_t j_\varepsilon + \varepsilon\, \partial_\theta \mathcal{S}_\varepsilon + j_\varepsilon + \kappa(h_\varepsilon * \rho_\varepsilon)\rho_\varepsilon = 0, \tag{5.53}$$

$$\varepsilon\, \partial_t \mathcal{S}_\varepsilon + \varepsilon\, \partial_\theta \mathcal{T}_\varepsilon + 2\mathcal{S}_\varepsilon - 2\rho_\varepsilon + 2\kappa(h_\varepsilon * \rho_\varepsilon) j_\varepsilon = 0. \tag{5.54}$$

Now, the key fact is the following equation

$$\frac{d}{dt}\left(\varepsilon \int_{\mathbb{T}} \mathcal{S}_\varepsilon\, d\theta + \kappa \int_{\mathbb{T}} (W_\varepsilon * \rho_\varepsilon)\rho_\varepsilon\, d\theta\right) + 2\int_{\mathbb{T}} \mathcal{S}_\varepsilon\, d\theta = 2,$$

coming from integration in (5.54), where

$$W_\varepsilon(\theta) = \int_0^\theta h_\varepsilon(\theta^*) d\theta^* = \int_0^{\bar\theta} h_\varepsilon(\theta^*) d\theta^*.$$

Then, integration in time show that under mild assumptions on the initial data one achieves *a priori* estimates for $\rho_\varepsilon, j_\varepsilon$ and $\mathcal{S}_\varepsilon$. In particular, up to a subsequence we achieve

$$\rho_\varepsilon \overset{*}{\rightharpoonup} \rho \quad \text{in } L^\infty(0,T;\mathcal{M}(\mathbb{T})),$$
$$j_\varepsilon \overset{*}{\rightharpoonup} j \quad \text{in } L^2(0,T;\mathcal{M}(\mathbb{T})).$$

This allows passing to the limit in all the linear terms in (5.52)–(5.53). Then, the main issue is again passing to the limit in the nonlinear term in (5.53). On the one hand, stronger time properties of ρ can be derived (check Ref. 200 for the details), thus showing

$$\rho_\varepsilon \otimes \rho_\varepsilon \overset{*}{\rightharpoonup} \rho \otimes \rho \quad \text{in } L^\infty(0,T;\mathcal{M}(\mathbb{T}^2 \times \mathbb{R}^2)).$$

On the other hand, the nonlinear terms read, in weak form, as follows

$$\frac{\kappa}{2}\int_{\mathbb{T}^2} (\varphi(\theta) - \varphi(\theta^*)) h_\varepsilon(\theta - \theta^*)\rho_\varepsilon(t,\theta)\rho_\varepsilon(t,\theta^*),$$

for any test function $\varphi \in C_c^\infty(\mathbb{T})$. Here, we have used the standard symmetrization trick that allows cancelling the singularity of the kernel in the whole range $\alpha \in (0,1)$ and allows concluding the following result.

Theorem 5.22. (Ref. 200) *For any $\alpha \in (0,1)$, consider the strong solution $F_\varepsilon = F_\varepsilon(t,\theta,\omega)$ to (5.50) with smooth initial data F_ε^0 verifying $\rho_\varepsilon^0 \to \rho^0$ in $\mathcal{P}(\mathbb{T})$-narrow and*

$$\sup_{\varepsilon>0} \|\omega^2 F_\varepsilon^0\|_{L^1(\mathbb{T})} < \infty.$$

Then, for every $T > 0$ there exists $\rho \in AC([0,T], C^\infty(\mathbb{T})^* - weak*)$ uniformly-in-time tight and a subsequence, that we denote in the same way, such that $\rho_\varepsilon \to \rho$ in $C([0,T],\mathcal{P}(\mathbb{T}) - narrow)$ and ρ verifies the kinetic singular Kuramoto model for identical oscillators in the sense of distributions

$$\partial_t \rho - \kappa \partial_\theta((h * \rho)\rho) = 0,$$
$$\rho(0) = \rho^0.$$

The above result provides global existence in the supercritical regime, but as for the microscopic scale, uniqueness is only guaranteed for $\alpha \in (0, \frac{1}{2}]$, the case $\alpha \in (\frac{1}{2}, 1)$ being an open problem. Similarly, emergence of phase synchronization is not guaranteed for these sort of (very) weak measure-valued solutions because the existence technique is not supported by the mean-field limit approach this time.

5.6. *Other models*

Apart from the preceding models that have been exhibited as prototype of first- and second-order agent-based models, where one can study its kinetic and macroscopic counterparts, there are a few more that have been proposed and analyzed in the literature. Although we will not enter into details, we will mention some of them and its main features.

Related to the last technique in Subsec. 5.3, the classical Kuramoto model with inertia has been analyzed at the microscopic and kinetic scales in Refs. [73–76]:

$$\partial_t f + \omega \partial_\theta f + \partial_\omega \left[\kappa \left(-\omega f + \int_{\mathbb{T}\times\mathbb{R}\times\mathbb{R}} \sin(\theta^* - \theta) f(t, d\theta^*, d\omega^*) \right) f \right] = 0, \tag{5.55}$$

where $f = f(t, \theta, \omega)$ is the distribution of identical oscillators at time t, phase $\theta \in \mathbb{T}$ and frequency $\omega \in \mathbb{R}$. The dynamics introduces a transient regime due to the inertial, that can be used to model certain physical situations. Nevertheless, the final dynamics essentially agrees with the starting model without inertia, as depicted in the above references. Although a hydrodynamic has not been proposed yet, the same vanishing inertia limit in Theorem 5.22 can be achieved mutatis mutandis, thus recovering the Kuramoto–Sakaguchi equation (5.16) with identical oscillators.

Another interesting swarming model in $\mathbb{R}^3$ arises when one consider constant speed. Then, only positions and orientations play a role. This is known as the *Couzin–Vicsec model* that in its kinetic version reads

$$\partial_t f + c\,\boldsymbol{\omega} \cdot \nabla_{\mathbf{x}} f + \nabla_{\boldsymbol{\omega}} \cdot [\nu(\mathbf{Id} - \boldsymbol{\omega} \otimes \boldsymbol{\omega})\boldsymbol{\Omega}(t, \mathbf{x}, \boldsymbol{\omega}) - D\nabla_{\boldsymbol{\omega}} f] = 0, \tag{5.56}$$

where $f = f(t, \mathbf{x}, \boldsymbol{\omega})$ is the distribution of particles at time t, position $\mathbf{x} \in \mathbb{R}^3$ and orientation $\boldsymbol{\omega} \in \mathbb{S}^2$. Here, $c > 0$ is the constant speed of particles, $D > 0$ is the strength of the orientation noise and $\boldsymbol{\Omega} = \boldsymbol{\Omega}(t, \mathbf{x}, \boldsymbol{\omega})$ is a normalized momentum vector $\boldsymbol{\Omega} = \frac{\boldsymbol{J}}{\|\boldsymbol{J}\|}$ and

$$\boldsymbol{J}(t, \mathbf{x}, \boldsymbol{\omega}) := \int_{\mathbb{R}^3 \times \mathbb{S}^2} K_R(|\mathbf{x} - \mathbf{x}^*|)\boldsymbol{\omega}^* f(t,\, d\mathbf{x}^*, d\boldsymbol{\omega}^*).$$

K_R is the "observation kernel" and typically stands for the characteristic function of the ball centered at the origin with radius R but one can consider general kernels modeling the fact that the influence of the particles falls off with distance. The global case $R \to \infty$ is classically considered, that is, J is the global momentum

when all particles are taken into account. The model was rigorously derived via the mean-field limit approach in the discrete Couzin–Vicsec model[90] and some numerical simulations were obtained in Ref. 115. The strong nonlinearity, that gives rise to degenerate terms when the momentum vanishes, has proved a strong obstruction and makes well posedness a hard issue. The *a priori* assumption of positivity of momentum[116] provides well posedness of solutions in the full space-inhomogeneous case. The only unconditional results have been obtained for the space-homogeneous case, see Refs. 110 and 157. Hydrodynamic limits toward macroscopic limits have been derived in Refs. 90 and 113. Other corrections of (5.56) that smooth the momentum term have been studied in Refs. 39 and 89.

Regarding the aggregation equation (5.3), there is a huge literature, see Refs. 38, 47, 57, 180, 179, 221 and 222 and related references. Indeed, similar estimates to those in Theorem 5.15 have been proved for such family of *gradient-flow systems*. When W is λ-convex for some $\lambda \geq 0$, the same estimate was derived in the Euclidean space $\mathbb{R}^d$ for the associated kinetic equation[63,66]

$$\partial_t \rho + \nabla_{\mathbf{x}} \cdot [-(\nabla_{\mathbf{x}} W * \rho)\rho] = 0, \tag{5.57}$$

where $\rho = \rho(t, \mathbf{x})$ is the probability density of particles. The main differences between (5.30) and (5.57) are: the absence of heterogeneities ν_i in (5.57), the gradient-flow structure of (5.57) with potential

$$\mathcal{W}[\rho(t)] := \int_{\mathbb{R}^d \times \mathbb{R}^d} W(\mathbf{x} - \mathbf{x}^*)\rho(t, d\mathbf{x})\rho(t, d\mathbf{x}^*),$$

and the underlying manifold $\mathbb{R}^d$ compared to $\mathbb{T} \times \mathbb{R}$ in (5.30).

The same ideas as in Subsec. 5.5.2 were analyzed in Ref. 57 in order to derive a hydrodynamic limit of vanishing-inertia-type (or large friction) of the second-order kinetic aggregation equation toward the first-order aggregation equation (5.57). Also, explicit convergence rates were measured in Wasserstein distances, adding the above-mentioned strong local alignment, but not noise (to avoid pressure terms). To such end, the same scaling as in Ref. 201 was considered to make inertia small.

Related to the aggregation equations, several more realistic variants have been proposed in order to include anisotropies in the interactions. For instance, in Refs. 100 and 101, the force $\mathbf{F}(\mathbf{x}, \mathbf{x}^*) = \nabla_{\mathbf{x}} W(\mathbf{x} - \mathbf{x}^*)$ in (5.3) is replaced by a velocity-dependent force

$$\mathbf{F}(\mathbf{x}, \mathbf{v}, \mathbf{x}^*, \mathbf{v}^*) := \nabla_{\mathbf{x}} W(\mathbf{x} - \mathbf{x}^*)\eta\left(\frac{\mathbf{x} - \mathbf{x}^*}{|\mathbf{x} - \mathbf{x}^*|} \cdot \frac{\mathbf{v}}{\|\mathbf{v}\|}\right).$$

The function η is considered a cut-off function at the origin so that it fades the effect of the aggregation force when the velocity of the particle and the director vector of such particle with respect to any other test particle are not sufficiently aligned. That anisotropic term arises as a sort of "cone of vision" that has been included in many other settings. In such paper, the anisotropies are velocity induced. However, there

are some other related models where no velocity dependence appears. Specifically, in Ref. 49 the authors proposed a 2D model where the force reads

$$\begin{aligned}\mathbf{F}(\mathbf{x},\mathbf{x}^*) &:= \mathbf{F_A}(\mathbf{x}-\mathbf{x}^*,\mathbf{T}(\mathbf{x})) + \mathbf{F_R}(\mathbf{x}-\mathbf{x}^*)\\ &= f_A(\|\mathbf{x}-\mathbf{x}^*\|)\mathbf{T}(\mathbf{x})(\mathbf{x}-\mathbf{x}^*) + f_R(\|\mathbf{x}-\mathbf{x}^*\|)(\mathbf{x}-\mathbf{x}^*).\end{aligned}$$

Then, there is an isotropic repulsive part and an anisotropic part dependent on the tensor

$$\mathbf{T}(\mathbf{x}) = \chi \mathbf{s}(\mathbf{x}) \otimes \mathbf{s}(\mathbf{x}) + (1-\chi)\mathbf{l}(\mathbf{x}) \otimes \mathbf{l}(\mathbf{x}),$$

for $\chi \in [0,1]$, being $\{\mathbf{s}(\mathbf{x}), \mathbf{l}(\mathbf{x})\}$ an orthonormal frame of $\mathbb{R}^2$. The constant χ is regarded as the anisotropy parameter, being $\chi = \frac{1}{2}$ the isotropic case. Such model was proved useful to describe the formation of fingerprints and becomes a generalization of the Kücken–Champod model.[165]

6. Computational Methods and Control Problems

In this section, we review some numerical methods for the solution of kinetic equations for interacting particle systems. In particular, we will discuss stochastic algorithms based on the techniques developed in Refs. 9 and 192. Subsequently, we will consider these algorithms in the context of optimal control problems with applications to swarming, herding behavior and crowd dynamics.

6.1. *Stochastic numerical techniques*

Realistic numerical simulations of swarms account the numerical solution of a large number of ODEs, which can constitute a serious difficulty leading to a prohibitive computational cost proportional to the square of the number of interacting individuals. As discussed in Sec. 4, a first step toward a possible numerical treatment is to consider a non-negative distribution function $f(t,x,v)$ describing the number density of individuals at time $t \geq 0$ in position $x \in \mathbb{R}^d$ with velocity $v \in \mathbb{R}^d$.

To illustrate the numerical approaches, we consider as a prototype kinetic equations describing the evolution of $f(t,x,v)$ the following nonlinear integro-differential PDEs

$$\partial_t f + v \cdot \nabla_x f = -\lambda \nabla_v \cdot (F[f]f) + \sigma \Delta_v (D[f]f), \tag{6.1}$$

where the force term $-\lambda \nabla_v \cdot (F[f]f)$ describes the forces acting on the agents dynamics, whereas the random effects are included in the diffusion, which may depend nonlinearly on the density f itself through the function $D[f] \equiv D[f](t,x,v)$. In what follows, we will concentrate in particular on the construction of stochastic numerical algorithms which permits to approximate the microscopic dynamic at a cost directly proportional to the number of sample particles involved in the computation, thus avoiding the quadratic computational cost of the force term (see Fig. 6 for a sketch of the approximation technique).

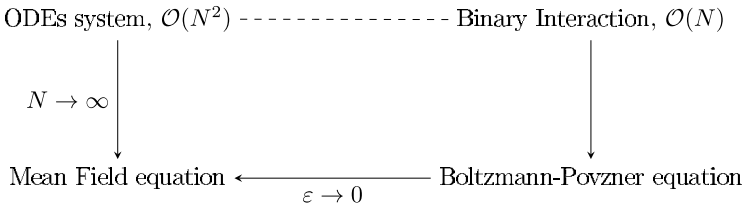


Fig. 6. Diagram of the approximation method: For a fixed time $t \geq 0$ a direct computation of the particle system requires $\mathcal{O}(N^2)$ evaluations to compute the agents' interactions. The approximation through binary interactions reduces the computational cost to $\mathcal{O}(N)$. Consistency of the Boltzmann–Povzner model for binary interactions with the mean-field kinetic equations is retrieved using the grazing collision scaling.

6.1.1. *Binary interaction methods*

In order to develop a stochastic algorithm capable to avoid the quadratic cost, we approximate the interaction dynamics by considering a binary interaction among two individuals, and resorting on their statistical description by means of Boltzmann-type equations. To exemplify these method we concentrate on CS type dynamics as introduced in (4.1), and we consider a binary interaction between two agents with positions and velocities $(x, v) \in \mathbb{R}^{2d}$ and $(y, w) \in \mathbb{R}^{2d}$ according to

$$\begin{cases} v^* = v + \eta\psi(x, y)(w - v) + \sqrt{2\varsigma D(x, v)}\xi, \\ w^* = w + \eta\psi(y, x)(v - w) + \sqrt{2\varsigma D(y, w)}\zeta, \end{cases} \tag{6.2}$$

where v^*, w^* are the *post-interaction* velocities and η a parameter that measures the strength of the interaction. Apart from the deterministic dynamics of (4.1), we also introduce the contribution of d-dimensional random variables ξ, ζ independent and identically distributed (i.i.d.) according to $\mathcal{Z}$ with zero mean and unitary variance. These noise terms are weighted by a nonlinear function $D(\cdot)$ which depends on the position and velocity state of each agent. Analogous binary interaction models can be introduced for other particle system in swarming, traffic or crowd dynamics.[5,65,106,164,192,227]

The evolution of the system is described by the following integro-differential equation of Boltzmann type

$$\begin{aligned} (\partial_t f + v \cdot \nabla_x f)(t, x, v) &= \frac{1}{\varepsilon} Q(f, f)(t, x, v), \\ Q(f, f) &= \mathbb{E}\left[\int_{\mathbb{R}^{2d}} \left(\frac{1}{J} f(x, v_*, t) f(y, w_*, t) \right.\right. \\ &\qquad \left.\left. - f(x, v, t) f(y, w, t) \right) dw dy \right], \end{aligned} \tag{6.3}$$

where (v_*, w_*) are the *pre-interacting* velocity that generate the couple (v, w) according to (6.2), J is the Jacobian of the transformation of (v, w) to (v_*, w_*). For

the standard CS model the Jacobian reads $J = (1 - 2\eta\psi(\|x - y\|))^d$. Moreover, to average the interactions we consider the expectation value $\mathbb{E}$ of ξ defined as follows

$$\mathbb{E}(\xi) = \int_{\mathbb{R}^d} x\mathcal{Z}(x)dx.$$

Note that, at variance with classical Boltzmann equation the interaction is non-local as in the Boltzmann–Povzner kinetic model.[199]

Remark 6.1. The averaged dynamics (6.2) preserves the momentum for symmetric communication function, i.e. $\psi(x, y) = \psi(y, x)$,

$$v^* + w^* = v + w - \eta(\psi(x, y) - \psi(y, x))(w - v) = v + w, \tag{6.4}$$

in general this is not the case for non-symmetric $\psi(\cdot)$, e.g. in presence of visual limitation.[65,185] Moreover under the assumptions $|\psi(r)| \leq 1$ and $\eta \leq 1/2$, it is easy to see that the support of the post interaction velocity is limited by the interacting velocities

$$v^* = (1 - \eta\psi(x, y))v + \eta\psi(x, y)w \leq \max\{|v|, |w|\}. \tag{6.5}$$

Consistency with mean-field kinetic model. Let us introduce the following scaling

$$t \to t/\varepsilon, \quad \eta = \lambda\varepsilon, \quad \varsigma = \varepsilon\sigma \tag{6.6}$$

where λ is a constant and ε a small parameter. The scaling corresponds to assume that the parameter η characterizing the strength of the microscopic interactions is small, thus the frequency of interactions has to increase otherwise the collisional integral will vanish. This corresponds to large scale interaction frequencies and small interaction strengths, in agreement with a classical mean-field limit and similarly to the so-called *grazing collision limit* of the Boltzmann equation for granular gases.[172]

Definition 6.1. (Test functions) We denote with $\mathcal{T}_\delta$ the set of compactly supported functions ϕ from $\mathbb{R}^{2d}$ to $\mathbb{R}$ such that for any multi-index $\beta \in \mathbb{N}^d$ we have:

(1) if $|\beta| < 2$, then $\partial_v^\beta \phi(x, \cdot)$ is continuous for every $x \in \mathbb{R}^d$;
(2) if $|\beta| = 2$, then there exists $M > 0$ such that:
   (a) $\partial_v^\beta \phi(x, \cdot)$ is uniformly Hölder continuous of order δ for every $x \in \mathbb{R}^d$ with Hölder bound M, that is for every $x \in \mathbb{R}^d$ and for every $v, w \in \mathbb{R}^d$ it holds

$$\left\|\partial_v^\beta \phi(x, v) - \partial_v^\beta \phi(x, w)\right\| \leq M\|v - w\|^\delta;$$

   (b) $\|\partial_v^\beta \phi(x, v)\| \leq M$ for every $(x, v) \in \mathbb{R}^{2d}$.

Notice that $\mathcal{C}_c^\infty(\mathbb{R}^d \times \mathbb{R}^d; \mathbb{R}) \subseteq \mathcal{T}_\delta$ for every $0 < \delta \leq 1$.

Definition 6.2. For a fixed $T > 0$, $\delta > 0$, then by a *δ-weak solution* of the initial value problem for Eq. (6.3) we consider $f \in L^2([0, T], \mathcal{M}_0(\mathbb{R}^{2d}))$ such that there

exists $R_T > 0$ with $\mathrm{supp}(f(t)) \subset B_{R_T}(0)$ for every $t \in [0,T]$ and that satisfies the weak form of (6.3)

$$\frac{d}{dt}\int_{\mathbb{R}^{2d}} \phi(x,v) f(x,v,t) dvdx + \int_{\mathbb{R}^{2d}} (v \cdot \nabla_x \phi(x,v)) f(x,v,t) dvdx$$
$$= \frac{1}{\varepsilon}\mathbb{E}\left[\int_{\mathbb{R}^{4d}} (\phi(x,v^*) - \phi(x,v)) f(x,v,t) f(y,w,t) dvdxdwdy\right], \tag{6.7}$$

for $t > 0$ and for all $t \in (0,T]$ and all $\phi \in \mathcal{T}_\delta$, with

$$\lim_{t\to 0}\int_{\mathbb{R}^{2d}} \phi(x,v) f(x,v,t) dvdx = \int_{\mathbb{R}^{2d}} \phi(x,v) f_0(x,v) dvdx, \tag{6.8}$$

where $f_0(x,v) = f(0,x,v)$ is the initial datum.

Based on these definitions, we state the following theorem.

Theorem 6.1. *For a fixed $T > 0$, $\delta > 0$, for every $\varepsilon > 0$, let f^ε be a δ-weak solution of (6.3) corresponding to the initial condition $f^0 = f(0,x,v)$, and where the quantities η and ς are rescaled with respect to ε according to (6.6). Assuming that*

(1) *the $2+\delta$ moment of ξ is finite, i.e. $\mathbb{E}(\|\xi\|^{2+\delta}) < \infty$,*
(2) *the functions $\psi(\cdot), D(\cdot)$ are in L^p_{loc} for $p = 2, 2+\delta$,*

then, as $\varepsilon \to 0$, the solutions f^ε converge pointwise, up to a subsequence, to f, where f satisfies the nonlinear Fokker–Planck-type equation

$$\partial_t f + v \cdot \nabla_x f = -\lambda \nabla_v \cdot (F[f]f) + \sigma \Delta_v (D(x,v) f), \tag{6.9}$$

with initial datum $f^0 = f(0,x,v)$ and

$$F[f](x,v) = \int_{\mathbb{R}^{2d}} \psi(x,y)(w-v) f(y,w) dydw.$$

We sketch briefly the passages of the proof, we refer to Ref. 192 for full details. The proof is based on two fundamental steps. The first step is based on a Taylor expansion of $\phi(x,v^*)$, since for small values of ε we have $v^* \approx v$. Hence expanding around v up to the second-order we obtain the following weak formulation for the collisional integral in (6.7),

$$\frac{1}{\varepsilon}\mathbb{E}\left[\int_{\mathbb{R}^{4d}} (\phi(x,v^*) - \phi(x,v)) f(x,v,t) f(y,w,t) dvdxdwdy\right]$$
$$= \underbrace{\frac{1}{\varepsilon}\,\mathbb{E}\left[\int_{\mathbb{R}^{4d}} (\nabla_v \phi(x,v) \cdot (v^* - v)) f(x,v,t) f(y,w,t) dvdxdwdy\right]}_{:= I_1^\varepsilon(f,f)}$$

$$+\frac{1}{2\varepsilon}\underbrace{\mathbb{E}\left[\int_{\mathbb{R}^{4d}}\left[\sum_{i,j=1}^{d}\partial_v^{(i,j)}\phi(x,v)(v^*-v)_i(v^*-v)_j\right]\right.}_{:=I_2^\varepsilon(f,f)}$$
$$\left.\times\ f(x,v,t)f(y,w,t)dvdxdwdy\right]$$

$$+\frac{1}{2\varepsilon}\underbrace{\mathbb{E}\left[\int_{\mathbb{R}^{4d}}\left[\sum_{i,j=1}^{d}(\partial_v^{(i,j)}\phi(x,v)-\partial_v^{(i,j)}\phi(x,\tilde{v}))(v^*-v)_i(v^*-v)_j\right]\right.}_{:=R_\phi^\varepsilon(f,f)}$$
$$\left.\times\ f(x,v,t)f(y,w,t)dvdxdwdy\right]$$

for some $\tilde{v}=\tau v+(1-\tau)v^*$, $0\le\tau\le 1$ and where the last term $R_\phi(f,f)$ represents the reminder of the Taylor expansion. In the limit $\varepsilon\to 0$ we estimate the contribution of the three terms by using the binary relation (6.2) under the grazing collision scaling (6.6) as follows

$$v^*-v=\lambda\varepsilon\psi(x,y)(w-v)+\sqrt{2\varepsilon\sigma D(x,v)}\xi. \tag{6.10}$$

Thus, we have for the first-order term

$$I_1^\varepsilon(f,f)=\lambda\int_{\mathbb{R}^{4d}}\phi(x,v)\cdot(w-v)\psi(x,y)f(y,w,t)f(x,v,t)dvdxdwdy,$$

and for the second-order term

$$I_2^\varepsilon(f,f)=\sigma\int_{\mathbb{R}^{4d}}\Delta\phi(x,v)D(x,v)f(x,v)f(y,w)dvdxdwdy$$
$$+\frac{\varepsilon}{2}\int_{\mathbb{R}^{4d}}\left[\sum_{i,j=1}^{d}\partial_v^{(i,j)}\phi(x,v)(w-v)_i(w-v)_j\right]$$
$$\times\psi(x,y)^2f(x,v)f(y,w)dvdxdwdy,$$

where we used the fact that there is no cross-correlation among the components of ξ and it has zero mean. Note that the second integral term is bounded thanks to Remark 6.1, and the assumptions of Theorem 6.1. Hence, considering the contribution of $I_1^\varepsilon(f,f)$, $I_2^\varepsilon(f,f)$ and neglecting the contribution of the reminder $R_\phi^\varepsilon(f,f)$, in the limit $\varepsilon\to 0$ the scaled weak equation (6.7) reads

$$\frac{d}{dt}\int_{\mathbb{R}^{2d}}\phi(x,v)f(x,v,t)dvdx+\int_{\mathbb{R}^{2d}}(v\cdot\nabla_x\phi(x,v))f(x,v,t)dvdx$$
$$=\int_{\mathbb{R}^{4d}}(\lambda\phi(x,v)\cdot(w-v)\psi(x,y)+\sigma\Delta\phi(x,v)D(x,v))$$
$$\times f(x,v,t)f(y,w,t)dvdxdwdy. \tag{6.11}$$

Thus, reverting to the strong form we recover model (6.9). To conclude the proof we need to estimate the reminder term $R_\phi^\varepsilon(f,f)$ and in particular to show that $R_\phi^\varepsilon(f,f)/\varepsilon\to 0$. We refer to Refs. 192 and 9 for a detailed proof.

6.1.2. *Asymptotic binary interaction algorithms*

Following Ref. 9, we introduce different stochastic algorithms for the above kinetic equations based on Monte Carlo methods. The main idea, similarly to Ref. 37 is to approximate the dynamic by solving the Boltzmann-like models for small value of ε. For the sake of simplicity we describe the algorithms in the case of the collision operator (6.3). As we will see, thanks to the structure of the equations, the resulting algorithms are fully meshless.

As in most Monte Carlo methods for kinetic equations, see Ref. 190, the starting point is a splitting method based on evaluating in two different steps the transport and collisional part of the scaled Boltzmann–Povzner equation

$$\text{(T)}: \quad \frac{\partial f}{\partial t} = -v \cdot \nabla_x f, \quad \text{(C)}: \quad \frac{\partial f}{\partial t} = \frac{1}{\varepsilon} Q_\varepsilon(f, f)$$

where we used the notation $Q_\varepsilon(f, f)$ to denote the scaled Boltzmann operator (6.3).

By decomposing the collisional operator in its gain and loss parts we can rewrite the collision step (C) as

$$\frac{\partial f}{\partial t} = \frac{1}{\varepsilon}[Q_\varepsilon^+(f, f) - \rho f], \tag{6.12}$$

where $\rho = \int_{\mathbb{R}^{2d}} f(x, v, t) dx dv > 0$ represent the total mass and Q_ε^+ the *gain* part of the collisional operator. Without loss of generality in the sequel we assume $\rho = 1$.

In order to solve the transport step we use the exact free flow of the sample particles $(x_i(t), v_i(t))$ in a time interval Δt

$$x_i(t + \Delta t) = x_i(t) + v_i(t)\Delta t, \tag{6.13}$$

and describe the different schemes used for the solution of (6.12).

Remark 6.2. We emphasize that the solution to the collision step for small values of ε has very little in common with the classical fluid-limit of the Boltzmann equation. Here in fact the whole collision process depends on space and on the small scaling parameter ε. In particular, in the small ε limit the solution is expected to converge toward the solution of the mean-field model (6.9).

Nanbu-like asymptotic methods. Let us now consider a time interval $[0, T]$ discretized in n_{tot} intervals of size Δt. We denote by f^n the approximation of $f(x, v, n\Delta t)$.

The forward Euler scheme writes

$$f^{n+1} = \left(1 - \frac{\Delta t}{\varepsilon}\right) f^n + \frac{\Delta t}{\varepsilon} Q_\varepsilon^+(f^n, f^n), \tag{6.14}$$

where since f^n is a probability density, thanks to mass conservation, also $Q_\varepsilon^+(f^n, f^n)$ is a probability density. Under the restriction $\Delta t \leq \varepsilon$ then also f^{n+1} is a probability density, since it is a convex combination of probability densities.

Since we aim at small values of e the maximum admissible time step is a $\Delta t = e$. We refer to Ref. 37 for further discussion on this choice. The major difference compared to standard Nanbu algorithm here is the way particles are sampled from

$Q_\varepsilon^+(f^n, f^n)$ which does not require the introduction of a space grid. A simple algorithm for the solution of (6.14) in a time interval $[0, T]$, $T = n_{\text{tot}}\Delta t$, $\Delta t = \varepsilon$ is sketched in the sequel.

Algorithm 6.2 (Asymptotic Nanbu).

(1) Given N_s samples (x_k^0, v_k^0), with $k = 1, \dots, N$ from the initial distribution $f_0(x, v)$;
(2) `for` $n = 0$ `to` $n_{\text{tot}} - 1$

`for` $i = 1$ `to` N_s;

(a) select an index j uniformly among all possible individuals (x_k^n, v_k^n) except i;
(b) evaluate $\psi(x_i^m, x_j^n)$ and $\psi(x_j^m, x_i^n)$;
(c) compute the velocity change v_i^* using the first relation in (6.2) with $\eta = \varepsilon$;
(d) set $(x_i^{n+1}, v_i^{n+1}) = (x_i^n, v_i^*)$.

`end for`

`end for`

A symmetric version of the previous algorithms which preserves at a microscopic level other interaction invariants, like momentum in standard CS model, is obtained by selecting particles by pairs. We refer to Ref. 9 and to the next paragraph for more details.

Bird-like asymptotic method. The most popular Monte Carlo approach to solve the collision step in Boltzmann-like equations is due to Bird.[36] The major differences are that the method simulate the time continuous equation and that individuals are allowed to interact more than once in a single time step. As a result the method achieves a higher time accuracy.[190]

In its simplest form the method is based on a constant time counter Δt_c corresponding to the average time between interactions.

Using the symmetric formulation and the time counter $\Delta t_c = 2\varepsilon/Ns$, we obtain the following method in a time interval $[0, T]$, $T = n_{\text{tot}}\Delta t_c$.

Algorithm 6.3 (Asymptotic Bird I).

(1) Given N samples (x_k, v_k), with $k = 1, \dots, Ns$ from the initial distribution $f_0(x, v)$
(2) `for` $n = 0$ `to` $n_{\text{tot}} - 1$

(a) select a random pair (i, j) uniformly among all possible pairs;
(b) evaluate $\psi(x_i^m, x_j^n)$ and $\psi(x_j^m, x_i^n)$;
(c) compute the velocity changes v_i^*, v_j^* using relations (6.2) with $\eta = \varepsilon$;
(d) set $v_i = v_i^*$ and $v_j = v_j^*$;

`end for`

As a result in the limit of large numbers of individuals the method converges toward the time continuous Boltzmann equation (6.3) and not to its time discrete counterpart (6.14), as it happens for Nanbu formulation.

Mean-field interaction algorithms. Let us finally tackle directly the limiting mean-field equation. The interaction step now corresponds to solve

$$\partial_t f = -\nabla_v \cdot \left(f \int_{\mathbb{R}^{2d}} \psi(x,y)(v-w) f(y,w,t) dw dy \right).$$

As already observed, in a particle setting this corresponds to compute the original $O(N^2)$ dynamic. We can reduce the computational cost using a Monte Carlo evaluation of the summation term as described in the following simple algorithm.

Algorithm 6.4 (Mean Field Monte Carlo).

(1) Given N samples v_k^0, with $k = 1, \ldots, N$ computed from the initial distribution $f_0(x,v)$ and $M \leq N$;
(2) `for` $n = 0$ `to` $n_{\text{tot}} - 1$
  (a) `for` $i = 1$ to N
  (b) sample M particles $j_1, \ldots, j_M$ uniformly without repetition among all particles;
  (c) compute the velocity change

$$v_i^{n+1} = \frac{1}{M} \sum_{k=1}^{M} \left[\left(1 - \Delta t \psi(x_i^n, x_{j_k}^n)\right) v_i^n + \Delta t \psi(x_i^n, x_{j_k}^n) v_{j_k}^n \right].$$

  `end for`
`end for`

The overall cost of the above simple algorithm is $O(MN)$, clearly for $M = N$ we obtain the explicit Euler scheme for the original N particle system. As shown in Refs. 67 and 154 the convergence rate with respect to the original particle system is in fact $\mathcal{O}(\sqrt{N^{-1} - M^{-1}})$.

In this formulation the method is closely related to asymptotic Nanbu's Algorithm 6.2. It is easy to verify that taking $M = 1$ leads exactly to the same numerical method. On the other hand for $M > 1$ the above algorithm can be interpreted as an averaged asymptotic Nanbu method over M runs. The only difference is that averaging the result of Algorithm 6.2 does not guarantee the absence of repetitions in the choice of the indexes $j_1, \ldots, j_M$. Thus the choice $\Delta t = \varepsilon$ in Algorithm 6.2 originates a numerical method consistent with the limiting mean-field kinetic equation. Following this description we can construct other Monte Carlo methods for the mean-field limit taking suitable averaged versions of the corresponding algorithms for the Boltzmann models. Here we omit for brevity the details.

Finally, in Fig. 7, we report the L_2-norm of the error for $ANMC$, $ABMC$ and $MFMC_M$ for various M as a function of $\Delta t = \varepsilon$.

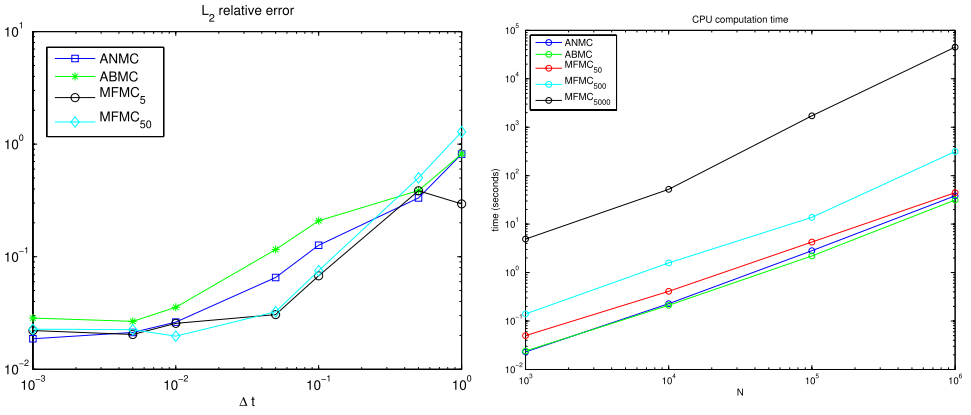


Fig. 7. Left: Relative errors in the L_2 norm at $T = 1$ for the different Monte Carlo type methods as a function of $\Delta t = \varepsilon$. Simulations are performed with $N_s = 50000$ particles. Right: Computational time with respect to the number of particles N_s used to perform the simulation.

Note that the convergence rate of the schemes is rather close and for $\varepsilon = \Delta t < t^*$, $t^* \approx 0.01$, the statistical error dominates the time error so that we observe a saturation effect.

6.2. *Control problems*

As shown in the first part of this survey, simple rules within groups of interacting agents can lead to the spontaneous formation of global behavior and this is a typical characteristic of real-life social systems. However, it is also common experience to observe that such formations are either unstable or imperfect. For instance, when we think to a herd of sheep, we know that these animals tend to move altogether, but it is also known that one needs a shepherd dog to keep the group well-ordered and cohesive. In the context of swarming, traffic and human behaviors this can be mathematically expressed, by the action of an additional external control which drives the ensemble of agents toward a desired goal[8,43,55,223] or by introducing the interaction with additionally independent agents, in terms of a coupling of the PDEs of the followers with the ODEs of the independent agents.[5,79,83,111]

6.2.1. *Mean-field optimal control problems*

We concentrate on CS type models where we want to enforce a flocking condition. To this end, we aim to design a control strategy $u = (u_1, \dots, u_N) \in \mathbb{R}^{d\times N}$ in the space of the admissible controls $\mathcal{U}$, as solution of the minimization problem,

$$\min_{u(\cdot)\in\mathcal{U}} J_T^N(\mathbf{u}(\cdot)) = \frac{1}{2}\int_0^T \frac{1}{N}\sum_{j=1}^{N}\left(\|v_j - \bar{v}\|^2 + \|u_j\|^2\right)dt, \tag{6.15}$$

where $\bar{v}$ is a desired velocity, and subject to the dynamics

$$\begin{aligned} \dot{x}_i &= v_i, \\ \dot{v}_i &= \frac{1}{N}\sum_{j=1}^{N} \psi(x_i, x_j)(v_j - v_i) + u_i S(t, x_i, v_i), \quad i = 1, \dots, N. \end{aligned} \tag{6.16}$$

From the modeling view point, the control problem (6.15)–(6.16) is equivalent to assume the presence of a *policy maker* able to exert an action on every agent. This assumption is rather unrealistic, in particular when the size of the system is large $N \gg 1$. For this reason, we introduce a function $S(\cdot)$ to account the limited amount of resources, or localized control action.[5,10,55,56] From the numerical view point, the direct solution of the nonlinear control problem (6.15)–(6.16) may be prohibitive even for moderate dimensions and a relatively small number of agents N, see Ref. 40. As soon as these parameters increase we run into the *curse of dimensionality*.[16]

A first step toward a dimensional reduction is to derive an equivalent mean-field optimal control problem which is consistent with (6.15)–(6.16) when $N \gg 1$. The resulting control problem reads

$$\min_{\mathbf{u}(\cdot)\in\mathcal{U}} J_T(\mathbf{u}(\cdot)) = \frac{1}{2}\int_0^T \int_{\mathbb{R}^{2d}} \left(\|v - \bar{v}\|^2 + \|u\|^2\right) f(t, x, v) dx\, dv\, dt, \tag{6.17}$$

where $u = u(t, x, v)$, subject to the mean-field kinetic dynamics

$$\partial_t f + v \cdot \nabla_x f = -\lambda \nabla_v \cdot ((F[f](t, x, v) + uS(t, x, v))f). \tag{6.18}$$

In Ref. 112, the authors show that the optimal control problems are consistent with the original microscopic problem. In particular convergence of minimizers among (6.15)–(6.16) and (6.17)–(6.18) when $N \to \infty$, is based on a Γ-convergence argument.

These results have been extended in several directions, for example when the control is applied in a parsimonious way in a localized set, or for PDEs-ODEs system where the control is concentrated only on few agents.[6,10,11,42,111,197]

Model Predictive Control. In what follows, we concentrate on the numerical solution of these optimal control problems, introducing a Model Predictive Control (MPC) technique, and some modeling examples showing how we can control agent systems with a moderate computational cost.

We introduce a numerical technique based on model predictive control, also called receding horizon strategy in the engineering and mathematical literature.[54,118,147,176] Our goal is to reduce the computational cost of optimal control problems of the type

$$\min_{\mathbf{u}(\cdot)\in\mathcal{U}} J_T^N(\mathbf{u}(\cdot)) = \frac{1}{2}\int_0^T \ell^N(x(t), v(t), u(t))dt. \tag{6.19}$$

To this purpose, we consider a time sequence $0 = t_0 < t_1 < \cdots < t_{N_T} = T$, as a discretization of the time interval $[0, T]$ with $\Delta t = t_n - t_{n-1}$, for all $n = 1, \dots, N_T$

and $t_{N_T} = N_T \Delta t$. We then consider the reduced minimization problem over $N_{\mathtt{mpc}}$ time steps, starting from a fixed time step $t_{\bar{n}}$,

$$\min_{u(\cdot)\in\mathcal{U}} \sum_{n=\bar{n}}^{\bar{n}+N_{\mathtt{mpc}}-1} \ell^N(x(n\Delta t), v(n\Delta t), u(n\Delta t)), \tag{6.20}$$

generating an optimal sequence of controls $\{u(\bar{n}\Delta t), \dots, u((\bar{n}+N_{\mathtt{mpc}}-1)\Delta t)\}$. Only the first $N_p < N_{\mathtt{mpc}}$ terms are taken to evolve the dynamics for a time $N_p \Delta t$, to recast the minimization problem over an updated time frame $\bar{n} \leftarrow \bar{n} + N_p$.

Note that for $N_p = 2$, the MPC approach recovers an instantaneous controller, whereas for $N_{\mathtt{mpc}} = T/\Delta t$ it solves the full time frame problem. Such flexibility is complemented with a robust behavior, as the optimization is re-initialized every time step, allowing to address perturbations along the optimal trajectory.

Instantaneous control. We provide explicit derivation of instantaneous feedback control strategies for the optimal control of CS-type models (6.15)–(6.16). We then consider the following one-step first-order approximation

$$\begin{aligned} &\min_{u^n} J_{\Delta t} = \frac{1}{2N}\sum_{j=1}^{N}\Big(\|v_j^{n+1} - \bar{v}\|^2 + \frac{\nu}{2}\|u_j^n\|^2\Big), \\ &\quad \text{s.t.} \\ &x_i^{n+1} = x_i^n + \Delta t v_i^n, \\ &v_i^{n+1} = v_i^n + \frac{\Delta t}{N}\sum_{j=1}^{N}\psi\big(x_i^n, x_j^n\big)\big(v_j^n - v_i^n\big) + \Delta t u_i^n S\big(x_i^n, v_i^n, t^n\big), \end{aligned} \tag{6.21}$$

for all $i = 1, \dots, N$, and $u^n \in \mathbb{R}^d$. The MPC aims at determining the value of the control u^n by solving for the known state (x_i^n, v_i^n) a (reduced) optimization problem on $[t_n, t_{n+1}]$ in order to obtain the new state (x_i^{n+1}, v_i^{n+1}). This procedure is reiterated until $n\Delta t = T$ is reached. In this way, it is possible to reduce the complexity of the initial problem (6.16)–(6.15), to an optimization problem in a single variable u^n. Therefore, we introduce the compact notation $\psi_{ij}^n = \psi(x_i^n, x_j^n)$, and $S_i^n = S(t^n, x_i^n, v_i^n)$, where for every i, p_i is the associated Lagrangian multiplier of v_i, and we define the discrete Lagrangian $\mathcal{L}_{\Delta t} = \mathcal{L}_{\Delta t}(v^{n+1}, u^n, p^{n+1})$, such that

$$\begin{aligned} \mathcal{L}_{\Delta t} &= J_{\Delta t}(v^{n+1}, u^n) \\ &\quad + \frac{1}{N}\sum_{j=1}^{N} p_j^{n+1} \cdot \left(v_j^{n+1} - v_j^n - \frac{\Delta t}{N}\sum_{\ell=1}^{N}\psi_{j\ell}^n\big(v_\ell^n - v_j^n\big) - \Delta t u_j^n S_j^n\right). \end{aligned} \tag{6.22}$$

Computing the gradient of (6.22) with respect to each component of v_i^{n+1} and u^n for every $i = 1, \dots, N$, we obtain the following first-order optimality conditions

$$v_i^{n+1} - \bar{v} + p_i^{n+1} = 0, \quad \nu u_i^n - \Delta t p_i^{n+1} S_i^n = 0. \tag{6.23}$$

This approach allows to express explicitly the control as feedback term of the state variable. We have that for every $n = 0, \dots, N_T - 1$

$$u_i^n = \frac{\Delta t}{\nu}(\bar{v} - v_i^{n+1})S_i^n. \tag{6.24}$$

Substituting in the discretized system (6.21) expression (6.24), the feedback controlled system results

$$v_i^{n+1} = v_i^n + \frac{\Delta t}{N}\sum_{j=1}^{N}\psi_{ij}^n(v_j^n - v_i^n) + \frac{\Delta t^2}{\nu}(\bar{v} - v_i^{n+1})(S_i^n)^2, \quad i = 1, \dots, N, \tag{6.25}$$

where the action of the control is substituted by an implicit term representing the relaxation toward the desired velocity $\bar{v}$.

Note that in this implicit formulation the action of the control is lost for $\Delta t \to 0$, since it is expressed in terms of $O(\Delta t^2)$. Thus, in order to rewrite the system as a consistent time discretization of the original control problem it is necessary to assume the following scaling on the regularization parameter, $\nu = \Delta t\kappa$. In addition, we can write the system in explicit form as

$$v_i^{n+1} = v_i^n + \frac{\Delta t}{N}\sum_{j=1}^{N}\psi_{ij}^n(v_j^n - v_i^n) + \frac{\Delta t}{\kappa + \Delta t(S_i^n)^2}\big(\bar{v} - v_i^n\big)(S_i^n)^2 + O(\Delta t^2). \tag{6.26}$$

System (6.26) represents a consistent discretization of the following dynamical system

$$\begin{aligned} \dot{x}_i &= v_i, \\ \dot{v}_i &= \frac{1}{N}\sum_{j=1}^{N}\psi(x_i, x_j)(v_j - v_i) + \frac{1}{\kappa}(\bar{v} - v_i)S(x_i, v_i, t)^2, \end{aligned} \tag{6.27}$$

where the control term is expressed by a steering factor toward the average weighted by the selective function $S(\cdot, \cdot)$.

Finally for $N \to \infty$ we can derive the mean-field model for the controlled swarming dynamic (6.27) which reads

$$\partial_t f + v \cdot \nabla_x f = -\nabla_v \cdot (F[f]f) - \frac{1}{\kappa}\nabla_v \cdot ((\bar{v} - v)S(t, x, v)^2 f). \tag{6.28}$$

We refer to Ref. 10 for the analytical details and to Ref. 145 for the general MPC approach for kinetic models with longer horizon.

Remark 6.3. Equivalent computation can be carried out for isotropic control, which assume the same value for each $u_i = u$ for every $i = 1, \dots, N$. In this case, the resulting instantaneous control reads

$$u_i^n = -\frac{\Delta t}{\nu N}\sum_{j=1}^{N}(v_j^{n+1} - \bar{v})S_j^n, \quad i = 1, \dots, N. \tag{6.29}$$

Hence, the microscopic model for the N agent system can be derived as follows

$$\begin{aligned} \dot{x}_i &= v_i, \\ \dot{v}_i &= \frac{1}{N}\sum_{j=1}^{N}\psi(x_i,x_j)(v_j-v_i)+\frac{1}{\kappa}(\bar{v}-v_j)S(x_j,v_j,t)S(x_i,v_i,t), \end{aligned} \tag{6.30}$$

which lead to the mean-field controlled equation

$$\partial_t f + v\cdot\nabla_x f = -\nabla_v\cdot(F[f]f) - \nabla_v\cdot((K[f]f), \tag{6.31}$$

where the instantaneous control term writes

$$K[f](t,x,v) = \frac{1}{\kappa}\int_{\mathbb{R}^{2d}}(\bar{v}-w)S(t,y,w)S(t,x,v)f(t,y,w)dy\ dw. \tag{6.32}$$

Example 1. (Radial spreading) We consider the computational domain $(x,v)\in\mathbb{R}^2\times\mathbb{R}^2$, defining an initial data $f(0,x,v)=f_0(x,v)$ normally distributed in space, with center in zero and unitary variance, and in velocity, uniformly distributed on a circumference of radius 5. Our goal is to enforce alignment with respect to the desired velocity $\bar{v}=(1,1)^T$. System (6.31) is solved numerically by means of the asymptotic binary algorithm, with final time $T=4$, time step $\Delta t=0.01$, considering $N_s=5\times10^5$ sampled particles and scaling parameter $\varepsilon=\Delta t$.

We consider the mean-field model (6.31), with the standard communication function,

$$\psi(x,y)=(1+|x-y|^2)^{-\gamma},$$

with $\gamma=10$ and consequently the unconditional flocking is not guaranteed *a priori*. We report in Fig. 8 the initial data and the final state reached at time $T=4$, depicting the spatial density $\rho(x,t)=\int_{\mathbb{R}^d}f(x,v,t)\,dv$ and showing at each point $x\in\mathbb{R}^2$ the value of the flux $\rho u(x,t)=\int_{\mathbb{R}^d}f(x,v,t)\,dv$. Note that, in the right-hand

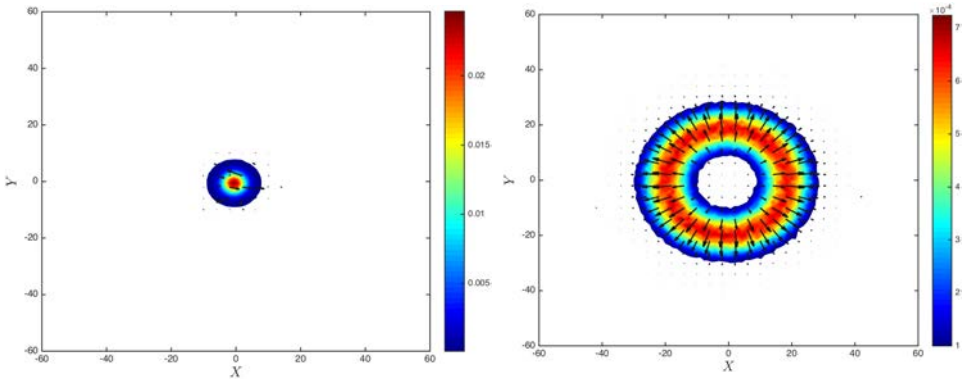


Fig. 8. On the left-hand side initial data, on the right-hand side configuration of the solution of (6.31) at time $T=4$. In absence of control the density spreads radially in the domain without reaching a flocking state.

side figure the flocking state is not reached, and the density is spreading around the domain following the initial radial symmetric distribution of the velocity field.

Next, we study the evolution of the system in presence of a selective control, where the selective function is $S(x,v) = \chi_{B_R}(x)$, and the instantaneous control (6.32) defined by

$$K[f](t,x,v) = \frac{\chi_{B_R}(x)}{\kappa} \int_{B_R} (\bar{v} - w) f(t,y,w) dy dw. \tag{6.33}$$

Moreover, in order to compare the behavior of the action of the selective control we define total cost C_T as

$$J_T := \int_0^T \int_{\mathbb{R}^4} \|v - \bar{v}\|^2 f(t,x,v) dx dv + \kappa \|u(t)\|^2. \tag{6.34}$$

In Fig. 9, as expected we observe that a flocking state is reached more easily for a stronger action of the control (i.e. decreasing values of the parameter κ). Moreover, comparing the first row ($R = 5$), with the second row ($R = 10$), we observe that the control is able to steer the velocity field more coherently toward the desired direction, when it acts for larger times, namely for larger radius.

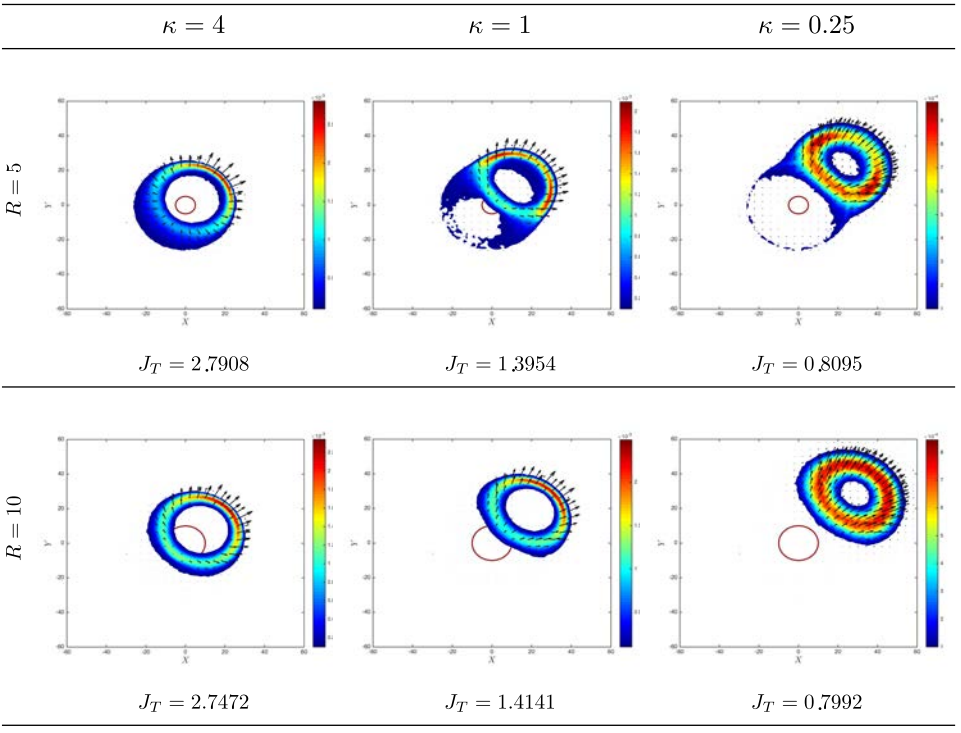


Fig. 9. Final solution at time $T = 4$ with control acting through a selective function $S(x,v) = B_R(x)$. The top and bottom pictures represent the action of the control, respectively for $R = 5$, and $R = 10$, and for different values of the penalization parameter κ. Value of the cost functional (6.34) is reported below each simulation.

Example 2. (Following a desired trajectory) We can extend the previous methodology to the control problem expressed by the constraint of the following a desired trajectory $\bar{\gamma}(t)$ considering a desired speed as a function of time $\bar{v}(t) = \bar{\gamma}'(t)$. In this way, the control action at time t forces the system to converge to the corresponding desired velocity at time t. Note that, in this case the choice of κ and Δt are of paramount importance to reconstruct exactly the trajectory.

We simulate the evolution of an aligned density solution forced to follow a desired trajectory $\bar{\Gamma}(t) = (R\cos(t), R\sin(2t))$ with $R = 1$, which corresponds to a *lemniscate*. In Fig. 10, we show the numerical solution for $\Delta t = 0.01$ with control action, $\kappa = 0.1$. In this case, the selective function $S(t, x, v) \equiv 1$.

6.2.2. *Control by independent agents*

A different approach to the control of multi-agent systems is to influence the system through a few independent agents, which are coordinated according to a centralized action, in order to promote a certain behavior for the larger group of agents.[5,79,83,111]

A general model is described by the following systems of ODEs:

$$\dot{x}_i = v_i, \quad i = 1, \dots, N$$
$$\dot{v}_i = \frac{1}{N}\sum_{j=1}^{N} \Psi_1(x_i, x_j, v_i, v_j) + \frac{1}{M}\sum_{\ell=1}^{M} \Phi_1(x_i, y_\ell, v_i, w_\ell),$$

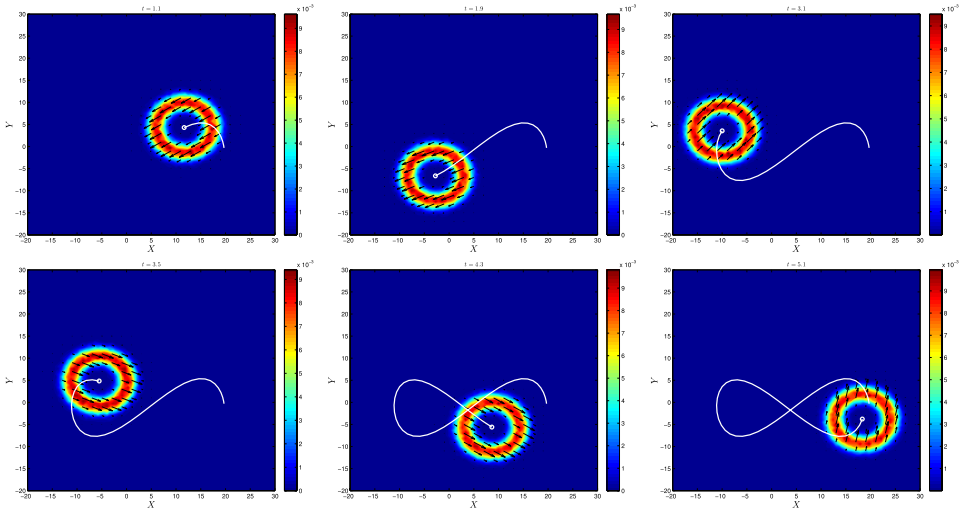


Fig. 10. The flock density is forced to follow a desired trajectory $\bar{\Gamma}(t) = (\cos(t), \sin(2t))$, described by a *lemniscate*. The regularization parameter is $\kappa = 0.1$ and the scaling parameter $\varepsilon = 0.01$.

$$\begin{aligned}\dot{y}_k &= w_k, \quad k = 1, \dots, M \\ \dot{w}_k &= \frac{1}{N}\sum_{j=1}^{N} \Psi_2(y_k, x_j, w_k, v_j) + \frac{1}{M}\sum_{\ell=1}^{M} \Phi_2(y_k, y_\ell, w_k, w_\ell) + u_k.\end{aligned} \tag{6.35}$$

where $N \gg M$, Ψ_1, Ψ_2, Φ_1 and Φ_2 are the communication functions, the control $u = (u_1, \dots, u_M) \in \mathbb{R}^{d\times M}$ is solution to the minimization problem

$$\min_{\mathbf{u}(\cdot)\in\mathcal{U}} J_T^N(\mathbf{u}(\cdot)) = \frac{1}{2}\int_0^T \ell^N(x(t), v(t), y(t), w(t), u(t))dt, \tag{6.36}$$

and the running cost $\ell^N(\cdot)$ has to be designed according to the specific application. The mean-field approximation $N \to \infty$ of the microscopic model (6.35) reads as a coupled system of PDEs and ODEs as follows

$$\begin{aligned}\partial_t f + v\cdot\nabla_x f &= -\nabla_v\cdot(F_1[f]f) - \nabla_v\cdot\left(\frac{1}{M}\sum_{\ell=1}^{M}\Phi_1(x, y_\ell, v, w_\ell)f\right), \\ \dot{y}_k &= w_k, \\ \dot{w}_k &= F_2[f](y_k, w_k) + \frac{1}{M}\sum_{\ell=1}^{M}\Phi_2(y_k, y_\ell, w_k, w_\ell) + u_k,\end{aligned} \tag{6.37}$$

where the operators $F_1[f], F_2[f]$ are defined by

$$\begin{aligned}F_1[f](t, x, v) &= \int_{\mathbb{R}^{2d}} \Psi_1(x, z, v, s) f(t, z, s) dzds, \\ F_2[f](t, y_k, w_k) &= \int_{\mathbb{R}^{2d}} \Psi_2(y_k, z, w_k, s) f(t, z, s) dzds.\end{aligned} \tag{6.38}$$

Example 1. (Swarm attacked by a predator) First, we consider an external agent approaching a swarm as a predator. In this case, the independent agent adopts a control strategy which drives its motion toward the center of mass of the preys, in other word its action can be modeled by the following running cost

$$\ell^N(x(t), v(t), y(t), w(t), u(t)) = \|\bar{x} - y_1\|^2 + \nu\|u_1\|^2, \tag{6.39}$$

where $\bar{x}(t) = \frac{1}{N}\sum_{j=1}^{N} x_j(t)$ is the center of mass of the swarm. The dynamics of the swarm is characterized by

$$\Psi_1(x, v, y, w) = \psi_1(\|x - y\|)(w - v) + \nabla_x W_1(\|x - y\|), \tag{6.40}$$

where $\psi(r) = (1 + r^2)^{-\gamma}$ is the CS alignment kernel, whereas $W(\cdot)$ is an attraction-repulsion kernel with power-law structure, $W(r) = r^a/a - r^b/b$. Interaction with the predator is driven by pure repulsion as follows:

$$\Psi_2(x, v, y, w) = \nabla_x W_2(\|x - y\|), \tag{6.41}$$

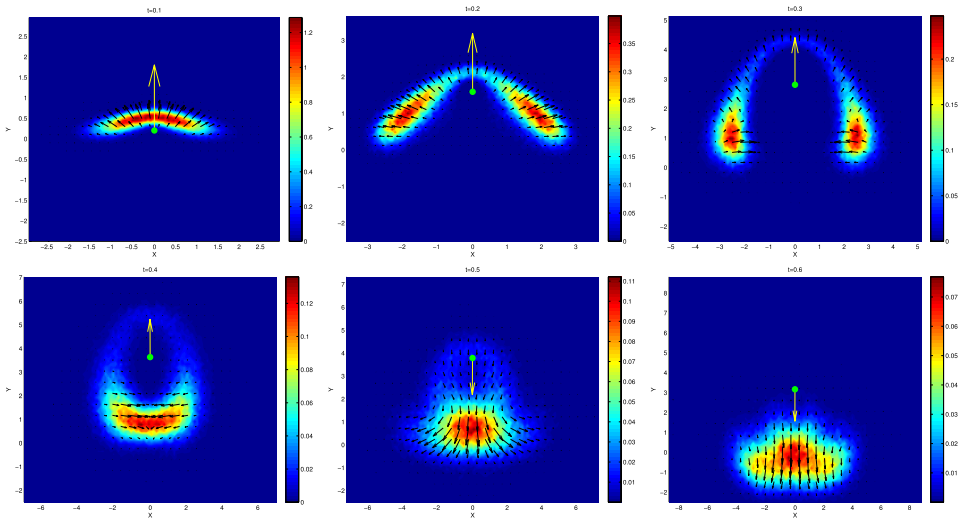


Fig. 11. Swarm attacked by a predator with parameters $a = 4$, $b = 2$, $\gamma = 0.45$.

with $W_2(r) = -r^c/c$ with $c > 0$. Finally, we assume that the predator's dynamics is ruled only by the control u_1 computed as an instantaneous control with respect to (6.39) and $\Psi_2 \equiv \Phi_2 \equiv 0$.

In Fig. 11, we report the evolution of the swarm which undergoes the action of the predator $(y_1, w_1) \in \mathbb{R}^{2d}$, $d = 2$. It is evident how the predator attack splits the flock in two groups which subsequently merge again together.

Example 2. (Confinement via shepherd dogs strategies) Next, we consider two external agents acting with the aim to confine the spatial spread of a swarm with the same interaction rules defined in (6.40) and (6.41) (see Refs. 9 and 79). Unlike the previous case we use a first-order differential model for the external agents, namely $w_k = u_k$. Thus we adopt a designed strategies which enforce a rotation around the local center of mass of the swarm as follows

$$u_k = V_k \frac{s_k^{\perp}}{\sqrt{1 + |s_k|^2}}; \quad V_k = 300, \quad r_k = 5, \quad k = 1, 2.$$

$$\eta(x) = \frac{3}{\pi r_k^6}(\max\{0, r_k^2 - |x|^2\})^2, \quad s_k := \int_{\mathbb{R}^{2d}} \nabla\eta(y_k - z) f(z, s) dz ds.$$

As we can see, in Fig. 12, the action of the leaders is able to force the confinement of the swarm. More refined strategies can be designed to solve the confinement problem, see for example.[79]

Application. Crowd control in an unknown environment. Finally we consider the realistic application of improving the evacuation of a crowd.[3,5,229] We identify a large ensemble of agents (*followers*) in an unknown environment, $\Omega \subset \mathbb{R}^2$,

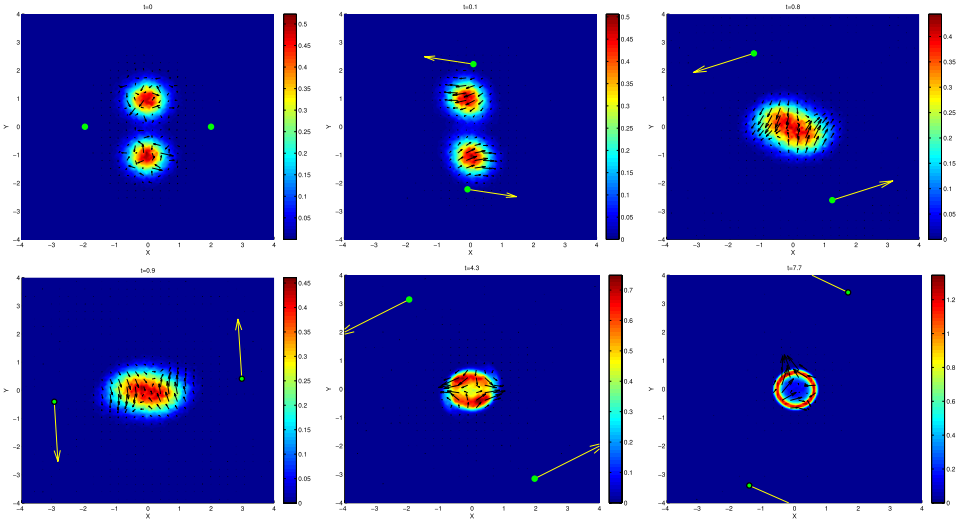


Fig. 12. Confinement via shepherd dogs. With parameters $a = 2.5$, $b = 0.1$, $\gamma = 0.45$.

influenced by the intervention of few "informed" agents (*leaders*). We consider a soft application of the control by assuming that the external agents are not recognized, namely we have in (6.35), that $\Psi_1 \equiv \Phi_1$. The dynamics of the pedestrian changes according to the relative position with respect to the exit x^τ: agents inside the visibility region of the exit, Σ, are driven toward x^τ; agents outside the visibility area, in $\Omega \backslash \Sigma$ align their velocity with respect to the velocity of the closest N^* neighbors (topological alignment), and a random direction ξ sampled from a normal distribution $\mathcal{N}(0, \sigma^2)$. Additionally every agent relaxes its speed toward the characteristic value of 1 m/s, and has a repulsive force against too close agents regulated by the potential $-\nabla R$. These social forces can be expressed in terms of the communication function Ψ_1, introducing the characteristic function of the unknown environment $\Theta(x) = \chi_{\Omega\backslash\Sigma}(x)$ and setting

$$\Psi_1(x, y, v, w) = S(x, y, v, w) + \Theta(x) H(x, v, y, w) + (1 - \Theta(x)) \left(\frac{x^\tau - x}{\|x^\tau - x\|} - v \right), \tag{6.42}$$

where

$$\begin{aligned} S(x, y, v, w) &= C_a(1 - \|v\|)v - C_r \nabla R(\|x - y\|), \\ H(x, v, y, w) &= C_\xi(\xi - v) - C_h \frac{N}{N^*}(w - v)\chi_{\mathcal{B}^*(x,t)}(y), \end{aligned}$$

with C_a, C_r, C_ξ, C_h non-negative constants, and $B^*(x, t)$ the topological set containing the N^* closest neighbors to the agent with position x. The dynamics of the

leaders agents is driven by a first-order differential equation as follows

$$\dot{y}_k = w_k = -\frac{1}{N}\sum_{j=1}^{N}\nabla R(\|y_k - x_j\|) - \frac{1}{M}\sum_{\ell=1}^{M}\nabla R(\|y_k - y_\ell\|) + u_k, \qquad (6.43)$$

where the control action u_k is computed by minimizing a cost functional of the form

$$\ell^N(x(t), y(t), u(t)) = \lambda\sum_{i=1}^{N}\|x_i - x^\tau\|^2 + \mu\sum_{i=1}^{N}\sum_{k=1}^{M}\|x_i - y_k\|^2 + \nu\sum_{k=1}^{M}\|u_k\|^2, \qquad (6.44)$$

for some positive constants μ, λ, and ν. The first term promotes the fact that followers have to reach the exit while the second forces leaders to keep contact with the crowd. The last term penalizes excessive velocities. This minimization is performed via MPC strategy at every instant (instantaneous control), or along a fixed time frame $[t_n, t_{n+N_{\text{mpc}}}]$. For the full details of the modeling choice we refer to Ref. 5.

At the mean-field level the following system of equations can be derived

$$\begin{aligned} \partial_t f + v\cdot\nabla_x f &= -\nabla_v\cdot\big(F[f, g^M]f\big) + \frac{1}{2}(C_\xi\sigma)^2\Theta(x)\Delta_v f, \\ \dot{y}_k &= G[f, g^M](y_k, w_k) + u_k, \end{aligned} \qquad (6.45)$$

where g^M is the empirical measure of leaders' state, and the operators $F[f, g^M], G[f, g^M]$ include the deterministic interaction forces of the microscopic dynamics (6.42) and (6.43). Note that this system is slightly different from model (6.45) due to the presence of the diffusion term.

We consider the exit as a point located at $x^\tau = (30, 10)$ which can be reached from any direction. We set $\Sigma = \{x \in \mathbb{R}^2 : \|x - x^\tau\| < 4\}$. Followers are initially randomly distributed in the domain $[17, 29]\times[6.5, 13.5]$ with velocity $(0, 0)$. Leaders, if present, are located to the left of the crowd. We report this setting in Fig. 13 for the microscopic and mesoscopic setting.

Figure 14 shows the evolution of the agents computed, without leaders. Followers having a direct view of the exit immediately point toward it, and some group mates

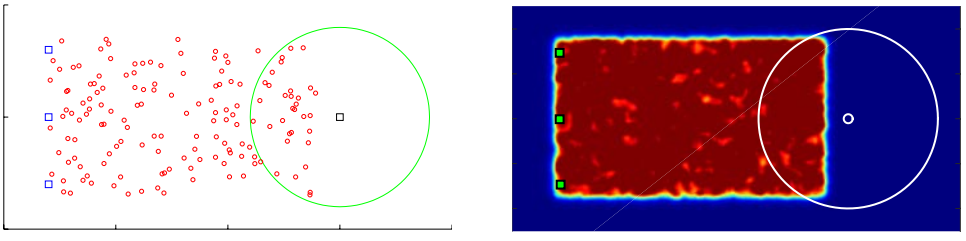


Fig. 13. Left: initial positions of followers (circles) and leaders (squares). Right: uniform density of followers and the microscopic leaders (squares).

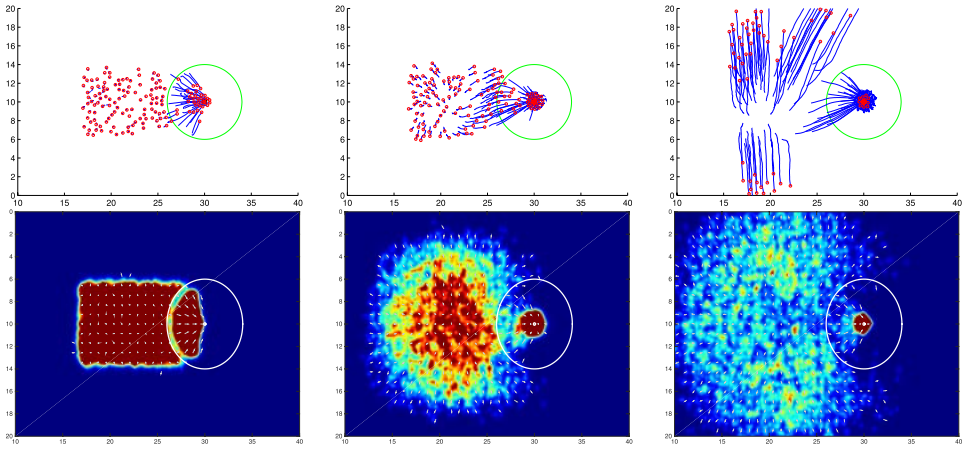


Fig. 14. Microscopic dynamics versus mean-field dynamics. Evolution with no controlling leaders.

close to them follow thanks to the alignment force. On the contrary, farthest people split in several but cohesive groups with random direction and never reach the exit. Second row shows the mean-field equivalent dynamics where the spread of the solution is generated by the diffusion operator acting on $\Omega \backslash \Sigma$.

Figure 15 shows the evolution of the agents with three leaders. The leaders' strategy is defined manually as an instantaneous strategy. More precisely, at any time the control is equal to the unit vector pointing toward the exit from the current position. Note that the final leaders' trajectories are not straight lines because of the additional repulsion force. As it can be seen, the crowd behavior changes completely since, this time, the whole crowd reaches the exit. However, followers form a heavy congestion around the exit. Note that the congestion notably delays the evacuation. Second row shows the equivalent mean-field dynamics. We finally

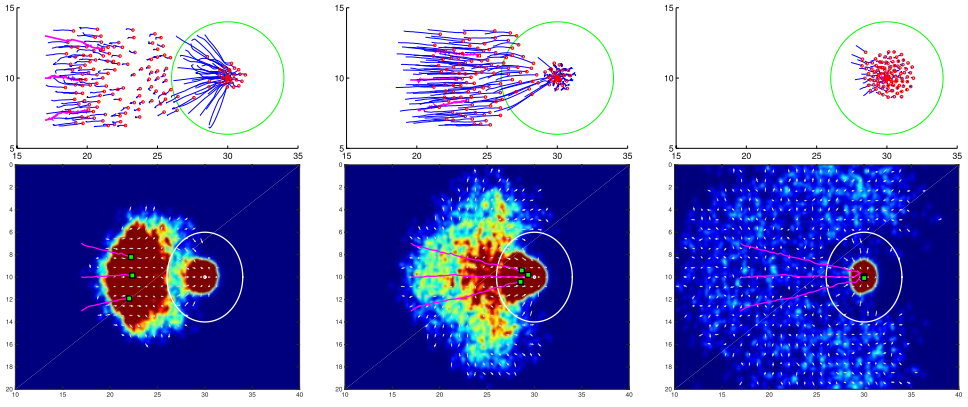


Fig. 15. Microscopic dynamics versus Mean-field dynamics. Evolution with controlling leaders with strategy computed via MPC with instantaneous control.

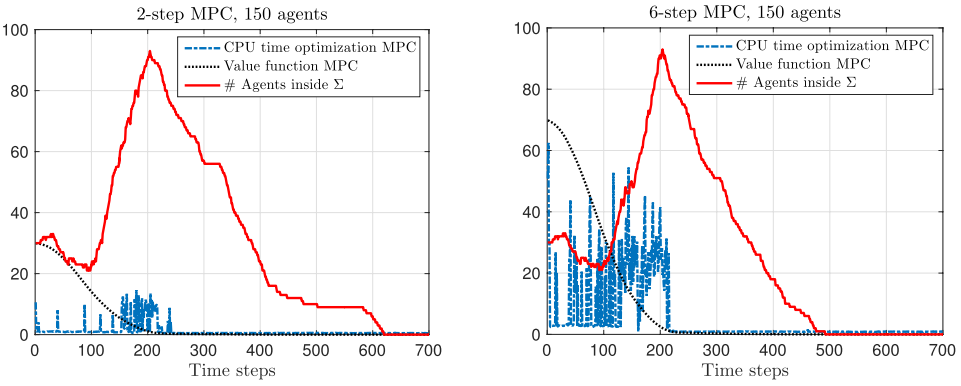


Fig. 16. Optimization of the microscopic dynamics by MPC. Occupancy of the exit's visibility zone Σ as a function of time, CPU time of the optimization call embedded in the MPC solver, and the evolution of the corresponding value (2-step and 6-step MPC).

compute MPC optimization, including a box constraint $u_k(t) \in [-1, 1]$. We choose $\lambda = 1$, $\mu = 10^{-5}$, and $\nu = 10^{-5}$. MPC results are consistent in the sense that for $N_{\texttt{mpc}} = 2$, the algorithm recovers a controlled behavior similar to the application of the instantaneous controller. Increasing the time frame up to $N_{\texttt{mpc}} = 6$ improves both congestion and evacuation times. In Fig. 16, we compare the occupancy of the exit's visibility zone as a function of time for "go-to-target" strategy and optimal strategies computed with 2-step, and 6-step MPC. We also show the decrease of the value function as a function of time.

It is worth to mention that better results can be obtained introducing a different functional, with respect to (6.44), and using an heuristic optimization method where non-trivial strategies are recovered. We refer to Ref. 5 for further details.

7. Looking Forward to Research Perspectives

A review of the kinetic theory approach to the modeling of vehicular traffic, human crowds, and swarms has been presented in our paper with main focus on modeling, applications and computing by the kinetic theory approach, namely by differential systems which describe the dynamics in time and space, of the probability distribution function over the microscopic state of the interacting entities.

The presentation has included, for each one of the aforementioned systems, the following topics: Modeling kinetic equations from the underlying description delivered by individual-based models, derivation of macroscopic equations from the underlying description by kinetic type models, analytic problems related to the qualitative analysis of the solutions of problems generated by the application of models, and computational methods for kinetic type equations. A specific feature of our paper, which enriches the overall contents, is a detailed presentation of computational tools to obtain simulations for equations which need stochastic, rather than deterministic, methods.

Hence, our paper aims at covering the whole path from modeling to computing passing through multiscale and analytical problems. If we look ahead to research perspectives, we remark that some specific hints have been already given in the preceding sections. However, these hints do not claim to cover the overall variety of possible research perspectives. Therefore we have selected, according to the authors' bias, four key research perspectives which are somewhat related to Secs. 3–6. These are proposed as the final hint of this paper.

Let us focus on the modeling crowd dynamics with propagation of social behaviors.[27,35,228] It is an important topic as the presence of stress, as well as of high emotional states, can induce significant modification in the overall self-organization, and hence on the collective dynamics. The modeling approach proposed in Ref. 27 has shown that the derivation of computational models and a systems approach to social dynamics[4] should march together as a deep understanding of possible social dynamics which can contribute to account for the interactions of different groups and for specific exchanges of social behaviors, for instance including the presence of leaders and even of antagonist groups. This modeling can contribute to the derivation of new models of swarms along the approach proposed in Refs. 28 and 30. Models of criminal actions,[99] which can play a role on the onset of sudden incidents, might be taken into account in the dynamics of crowds.[135,169,170]

As discussed in Sec. 4.1.3, the rigorous justification of the hydrodynamic CS model from the kinetic CS equation has been obtained by combining the relative entropy method and 2-Wasserstein metric in a strong local alignment regime in Ref. 109. Similarly, one can attempt to derive hydrodynamic TCS model from the kinetic TCS equation. However, due to the lack of suitable free energy for the kinetic TCS equation and strong nonlinearity structure of model, it is not easy to implement the same strategy in aforementioned work as it is, unless the initial temperature support is concentrated to the constant initial average temperature as scaling parameter tends to 0. This will be an interesting problem need to be done in future work.

As highlighted in Sec. 5, the rigorous derivation of macroscopic limits through a kinetic approach based on microscopic descriptions is a hot topic in the study of crowds and swarms, both for regular and singular interaction rules. Firstly, including time-dependent coupling weights via plasticity rules opens new perspectives in the modeling from micro to macro scales. Secondly, the dissipation processes of kinetic energy due to interactions are sometimes responsible for the convergence towards macroscopic models, and its characterization is essential to complete the fundamentals of this theory. Finally, the inclusion of time-dependent coupling weights via plasticity rules opens new perspectives in the modeling from micro to macro scales. Lastly, studying clustering dynamics is the basis for elucidating the internal structure of collective evolution of species, as well as combining several effects (e.g. aggregation and synchronization) can give rise to new relevant phenomena.

Let us now focus on numerical methods for kinetic equations presented in Sec. 6, focused on a class of binary collision algorithms first developed in Ref. 9, and inspired by classical direct simulation schemes for rarefied gas dynamic. The resulting algorithms are fully meshless, and permit to reduce the intrinsic quadratic cost of the model to a linear evaluation of statistical samples. Further research directions include applications of these methods to crowd and traffic dynamics Refs. 106 and 227. In the second part of Sec. 6, we have discussed the approximation of control problems for large particle systems.[40] Numerical techniques typical of control theory, such as MPC methods, can be efficiently applied at the microscopic level and allow to derive feedback controlled kinetic models. These problems are strongly connected to mean-field optimal control problems.[42,111,112] Further improvements are possible by considering other feedback controllers for the binary interaction system. Some works have been devoted to this direction starting from Refs. 8 and 11 and more recently by computing feedback controllers via Dynamic Programming and Hamilton–Jacobi–Bellman equations for the binary system.[61,7]

Acknowledgments

The work of G. Albi has been partially supported by the project Computer Science for Industry 4.0, 'MIUR Departments of Excellence 2018-2022'. G. Albi and L. Pareschi acknowledge the partial support of GNCS-INdAM 2019 project "Numerical approximation of hyperbolic problems and applications". The work of Luisa Fermo has been partially supported by the research project "Algorithms for Approximation with Applications [Acube], Fondazione di Sardegna — year 2017". The work of S.-Y. Ha and J. Kim has been partially supported by National Research Foundation of Korea (NRF-2017R1A2B2001864). The work of D. Poyato and J. Soler has been partially supported by the MINECO-Feder (Spain) research grant number MTM2014-53406-R and RTI2018-098850-B-I00 (N.B., J.S, D.P.), by the Junta de Andalucía (Spain) Project FQM 954 (J.S, D.P.), and the MECD (Spain) research grant FPU2014/06304 (D.P.).

References

1. M. Akbarzadeh and E. Estrada, Communicability geometry captures traffic flows in cities, *Nature Human Behav.* **2** (2018) 645–652.
2. J. A. Acebrón, L. L. Bonilla, C. J. P. Vicente, F. Ritort and R. Spigler, The Kuramoto model: A simple paradigm for synchronization phenomena, *Rev. Mod. Phys.* **77** (2005) 137–185.
3. J. P. Agnelli, F. Colasuonno and D. Knopoff, A kinetic theory approach to the dynamics of crowd evacuation from bounded domains, *Math. Models Methods Appl. Sci.* **25** (2015) 109–129.
4. G. Ajmone Marsan, N. Bellomo and L. Gibelli, Stochastic evolutionary differential games toward a systems theory of behavioral social dynamics, *Math. Models Methods Appl. Sci.* **26** (2016) 1051–1093.
5. G. Albi, M. Bongini, E. Cristiani and D. Kalise, Invisible control of self-organizing agents leaving unknown environments, *SIAM J. Appl. Math.* **7** (2016) 1219–1763.

6. G. Albi, Y. P. Choi, M. Fornasier and D. Kalise, Mean field control hierarchy, *App. Math. Optim.* **76** (2017) 93–135.
7. G. Albi, M. Fornasier and D. Kalise, A Boltzmann approach to mean-field sparse feedback control, *IFAC* **50** (2017) 2898–2903.
8. G. Albi, M. Herty and L. Pareschi, Kinetic description of optimal control problems and applications to opinion consensus, *Commun. Math. Sci.* **13** (2015) 1407–1429.
9. G. Albi and L. Pareschi, Binary interaction algorithms for the simulation of flocking and swarming dynamics, *Multiscale Model. Simul.* **11**(1) (2013) 1–29.
10. G. Albi and L. Pareschi, Selective model-predictive control for flocking systems, *Commun. Appl. Ind. Math.* **13** (2018) 1–18.
11. G. Albi, L. Pareschi and M. Zanella, Boltzmann-type control of opinion consensus through leaders, *Philos. Trans. R. Soc. A* **372** (2014) 1–18.
12. B. Aylaj, N. Bellomo, L. Gibelli and A. Reali, On a unified multiscale vision of behavioral crowds, *Math. Models Methods Appl. Sci.* **29** (2019), to appear.
13. H.-O. Bae, Y.-P. Choi, S.-Y. Ha and M.-J. Kang, Time-asymptotic interaction of flocking particles and an incompressible viscous fluid, *Nonlinearity* **25** (2012) 1155–1177.
14. H.-O. Bae, Y.-P. Choi, S.-Y. Ha and M.-J. Kang, Asymptotic flocking dynamics of Cucker–Smale particles immersed in compressible fluids, *Discr. Cont. Dyn. Syst. Ser. A* **34** (2014) 4419–4458.
15. H.-O. Bae, S.-Y. Ha and Y. Kim, Simulation of interaction of flocking particles and an incompressible fluid, *Comp. Math. Appl.* **71** (2016) 2020–2033.
16. R. Bellman and R.-E. Kalaba, *Dynamic Programming and Modern Control Theory*, Vol. 81 (Academic Press, New York, 1965).
17. N. Bellomo and A. Bellouquid, On multiscale models of pedestrian crowds from mesoscopic to macroscopic, *Comm. Math. Sci.* **13** (2015) 1649–1664.
18. N. Bellomo, A. Bellouquid, L. Gibelli and N. Outada, *A Quest Towards a Mathematical Theory of Living Systems* (Birkhäuser, 2017).
19. N. Bellomo, A. Bellouquid and D. Knopoff, From the micro-scale to collective crowd dynamics, *Multiscale Model. Simul.* **11** (2013) 943–963.
20. N. Bellomo, A. Bellouquid, J. Nieto and J. Soler, On the asymptotic theory from microscopic to macroscopic growing tissue models: An overview with perspectives, *Math. Models Methods Appl. Sci.* **22** (2012) 1130001 (37 pages), doi:10.1142/S0218202512005885.
21. N. Bellomo, A. Bellouquid, J. Nieto and J. Soler, On the multiscale modeling of vehicular traffic: From kinetic to hydrodynamics, *Discr. Cont. Dyn. Syst. Ser. B* **19** (2014) 1869–1888.
22. N. Bellomo, D. Clarke, L. Gibelli, P. Townsend and B. J. Vreugdenhil, Human behaviours in evacuation crowd dynamics: From modeling to "big data" toward crisis management, *Phys. Life Rev.* **18** (2016) 1–21.
23. N. Bellomo, P. Degond and E. Tadmor eds., *Active Particles*, Vol. 1 (Birkhäuser, 2017).
24. N. Bellomo and C. Dogbè, On the modeling of traffic and crowds: A survey of models, speculations, and perspectives, *SIAM Rev.* **53** (2011) 409–463.
25. N. Bellomo and L. Gibelli, Toward a mathematical theory of behavioral-social dynamics for pedestrian crowds, *Math. Models Methods Appl. Sci.* **25** (2015) 2417–2437.
26. N. Bellomo and L. Gibelli, Behavioral crowds: Modeling and Monte Carlo simulations toward validation, *Comput. Fluids* **141** (2016) 13–21.

27. N. Bellomo, L. Gibelli and N. Outada, On the interplay between behavioral dynamics and social interactions in human crowds, *Kinet. Relat. Models* **12** (2019) 397–409.
28. N. Bellomo and Seung-Y. Ha, A quest toward a mathematical theory of the dynamics of swarms, *Math. Models Methods Appl. Sci.* **27** (2017) 745–770.
29. N. Bellomo, D. Knopoff and J. Soler, On the difficult interplay between life, "complexity", and mathematical sciences, *Math. Models Methods Appl. Sci.* **23** (2013) 1861–1913.
30. N. Bellomo and J. Soler, On the mathematical theory of the dynamics of swarms viewed as a complex system, *Math. Models Methods Appl. Sci.* **22** (2012) 1140006 (29 pages) doi: 10.1142/S0218202511400069.
31. A. Bellouquid, J. Calvo, J. Nieto and J. Soler, Hyperbolic versus parabolic asymptotics in kinetic theory towards fluid dynamic models, *SIAM J. Appl. Math.* **73** (2013) 1327–1346.
32. A. Bellouquid, E. De Angelis and L. Fermo, Towards the modeling of vehicular traffic as a complex system: A kinetic approach, *Math. Models Methods Appl. Sci.* **22** (2012) 1140006 (29 pages).
33. D. Benedetto, E. Caglioti and U. Montemagno, On the complete phase concentration for the Kuramoto model in the mean field limit, *Commun. Math. Sci.* **13** (2015) 1775–1786.
34. D. Benedetto, E. Caglioti and U. Montemagno, Exponential dephasing of oscillators in the kinetic Kuramoto model, *J. Stat. Phys.* **162** (2016) 813–823.
35. A. L. Bertozzi, J. Rosado, M. B. Short and L. Wang, Contagion shocks in one dimension, *J. Stat. Phys.* **158** (2015) 647–664.
36. G. A. Bird, Direct simulation and the Boltzmann equation, *Physics of Fluids* **13**(11) (1970) 2676–2681, https://doi.org/10.1063/1.169284.
37. A. Bobylev and K. Nanbu, Theory of collision algorithms for gases and plasmas based on the Boltzmann equation and the Landau–Fokker–Planck equation, *Phys. Rev. E* **61** (2000) 4576.
38. S. Boi, V. Capasso and D. Morale, Modeling the aggregative behavior of ants of the species Polyergus rufescens, in *Spatial Heterogeneity in Ecological Models, Nonlinear Anal–Real.* **1** (2000) 163–176.
39. F. Bolley, J. A. Cañizo and J. A. Carrillo, Mean-field limit for the stochastic Vicsek model, *Appl. Math. Lett.* **25** (2012) 339–343.
40. M. Bongini, R. Bailo, J-A. Carrillo and D. Kalise, Optimal consensus control of the Cucker–Smale model, *IFAC* **51** (2018) 1–6.
41. M. Bongini, M. Fornasier and D. Kalise, (Un)conditional consensus emergence under perturbed and decentralized feedback controls, *Discr. Contin. Dyn. Syst. Ser. A* **35** (2015) 4071–4094.
42. M. Bongini, M. Fornasier, F. Rossi and F. Solombrino, Mean-field Pontryagin maximum principle, *J. Optim. Theory Appl.* **75** (2017) 1–38.
43. A. Borzì and S. Wongkaew, Modeling and control through leadership of a refined flocking system, *Math. Models Methods Appl. Sci.* **25** (2015) 255–282.
44. R. Borsche, A. Klar, S. Köhn and A. Meurer, Coupling traffic flow networks to pedestrian motion, *Math. Models Methods Appl. Sci.* **24** (2014) 359–380.
45. A. Bressan, S. Canic, M. Garavello, M. Herty and B. Piccoli, Flows on networks: Recent results and perspectives, *EMS Surv. Math. Sci.* **1** (2014) 47–111.
46. J. Buck and E. Buck, Biology of synchronous flashing of fireflies, *Nature* **211** (1966) 562–564.

47. M. Burger, V. Capasso and D. Morale, On an aggregation model with long and short range interactions, *Nonlinear Anal. Real World Appl.* **8** (2007) 939–958.
48. M. Burger, M. Di Francesco, P. A. Markowich and M. T. Wolfram, Mean field games with nonlinear mobilities in pedestrian dynamics, *Discr. Cont. Dyn. Syst. Ser. B* **19** (2014) 1311–1333.
49. M. Burger, B. Düring, L. M. Kreusser, P. A. Markowich and C.-B. Schönlieb, Pattern formation of a nonlocal, anisotropic interaction model, *Math. Models Methods Appl. Sci.* **28** (2018) 409–451.
50. D. Burini and N. Chouhad, Hilbert method toward a multiscale analysis from kinetic to macroscopic models for active particles, *Math. Models Methods Appl. Sci.* **27** (2017) 1327–1353.
51. D. Burini and N. Chouhad, A multiscale view of nonlinear diffusion in biology: From, cells to tissues, *Math. Models Methods Appl. Sci.* **29** (2019) 791–823.
52. D. Burini, S. De Lillo and G. Fioriti, Influence of drivers ability in a discrete vehicular traffic model, *Int. J. Mod. Phys. C.* (2017) **28**(3) (2017) 1750030.
53. D. Burini, S. De Lillo and L. Gibelli, Stochastic differential "nonlinear" games modeling collective learning dynamics, *Phys. Life Rev.* **16** (2016) 123–139.
54. E. Camacho and C. Bordons, *Model Predictive Control* (Springer, 2004).
55. M. Caponigro, M. Fornasier, B. Piccoli and E. Trélat, Sparse stabilization and optimal control of the Cucker–Smale model, *Math. Control Relat. Fields* **3** (2013) 447–466.
56. M. Caponigro, B. Piccoli, F. Rossi and E. Trélat, Mean-field sparse Jurdjevic–Quinn control, *Math. Models Methods Appl. Sci.* **6** (2017) 1–31.
57. J. A. Carrillo and Y.-P. Choi, Quantitative error estimates for the large friction limit of Vlasov equation with nonlocal forces, preprint (2019), arXiv:1901.07204v1.
58. J. A. Carrillo, Y.-P. Choi, S.-Y. Ha, M.-J. Kang and Y. Kim, Contractivity of transport distances for the kinetic Kuramoto equation, *J. Stat. Phys.* **156** (2014) 395–415.
59. J.-A. Carrillo, Y.-P. Choi and M. Hauray, Mathematical Modelling of Complex Systems, in *ESAIM Proc. Surveys*, Vol. 47, EDP Sci., Les Ulis (2014), pp. 17–35.
60. J. A. Carrillo, Y.-P. Choi and T. K. Karper, On the analysis of a coupled kinetic-fluid model with local alignment forces, *Ann. I. H. Poincareé* **33** (2016) 273–307.
61. J. A. Carrillo, Y.-P. Choi, E. Tadmor and C. Tan, Critical thresholds in 1D Euler equations with nonlocal forces, *Math. Models Methods Appl. Sci.* **26** (2016) 185–206.
62. J.-A. Carrillo, Y.-P. Choi, C. Totzeck and O. Tse, An analytical framework for consensus-based global optimization method, *Math. Models Methods Appl. Sci.* **28** (2018) 1037–1066.
63. J. A. Carrillo, M. DiFrancesco, A. Figalli, T. Laurent and D. Slepčev, Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations, *Duke Math. J.* **156** (2011) 229–271.
64. J. A. Carrillo, M. Fornasier, J. Rosado and G. Toscani, Asymptotic flocking dynamics for the kinetic Cucker–Smale model, *SIAM J. Math. Anal.* **42** (2010) 218–236.
65. J. A. Carrillo, M. Fornasier, G. Toscani and F. Vecil, Particle, kinetic, and hydrodynamic models of swarming, in *Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences* (Springer, 2010), pp. 297–336.
66. J. A. Carrillo, F. James, F Lagoutière and N. Vauchelet, The Filippov characteristic flow for the aggregation equation with mildly singular potentials, *J. Differential Equations* **260** (2016) 304–338.
67. J. A. Carrillo, L. Pareschi and M. Zanella, Particle based gPC methods for mean-field models of swarming with uncertainty, *Commun. Comput. Phys.* **25** (2018) 508–531.

68. C. Cercignani, R. Illner and M. Pulvirenti, *The Kinetic Theory of a Diluted Gas* (Springer, 1993).
69. H. Chiba, A proof of the Kuramoto conjecture for a bifurcation structure of the infinite-dimensional Kuramoto model, *Ergod. Theor. Dyn. Syst.* **35** (2015) 762–834.
70. Y.-P. Choi, S.-Y. Ha, J. Jung and J. Kim, A global existence of strong solutions and large time behavior of kinetic thermomechanical Cucker–Smale model with incompressible Navier–Stokes flow, *Nonlinearity* **32** (2019) 1597–1640.
71. Y.-P. Choi, S.-Y. Ha, J. Jung and J. Kim, The global Cauchy problem for a coupled system of kinetic thermomechanical Cucker–Smale model and compressible Navier–Stokes flow, Submitted.
72. Y.-P. Choi, S.-Y. Ha and Z. Li, Emergent dynamics of the Cucker–Smale flocking model and its variants, *Active particles*, Vol. 1. Advances in theory, models, and applications, Model. Simul. Sci. Eng. Technol. (Birkhäuser/Springer, Cham, 2017), pp. 299–331.
73. Y.-P. Choi, S.-Y. Ha, Z. Li, X. Xue and S.-B. Yun, Complete entrainment of Kuramoto oscillators with inertia on networks via gradient-like flow, *J. Differential Equations* **257** (2014) 2591–2621.
74. Y.-P. Choi, S.-Y. Ha and J. Morales, Emergent dynamics of the Kuramoto ensemble under the effect of inertial, *Discr. Cont. Dyn. Syst. Ser. A* **38** (2018) 4875–4913.
75. Y.-P. Choi, S.-Y. Ha and S.-B. Yun, Complete synchronization of Kuramoto oscillators with finite inertia, *Physica D* **240** (2011) 32–44.
76. Y.-P. Choi, S.-Y. Ha and S.-B. Yun, Global existence and asymptotic behavior of measure valued solutions to the kinetic Kuramoto–Daido model with inertia, *Netw. Heterog. Media* **8** (2013) 943–968.
77. Y.-P. Choi, D. Kalise, J. Peszek and A. A. Peters, A collisionless singular Cucker–Smale model with decentralized formation control, preprint (2018), arXiv:1807.05177.
78. G. Coclite, M. Garavello and B. Piccoli, Traffic flow on road networks, *SIAM J. Math. Anal.* **36** (2004) 1882–1886.
79. R. Colombo and M. Lécureux-Mercier, An analytical framework to describe the interactions between individuals and a continuum, *J. Nonlinear Sci.* **22** (2012) 39–61.
80. A. Corbetta, L. Bruno, A. Mountean and F. Yoschi, High statistics measurements of pedestrian dynamics, models via probabilistic method, *Transp. Res. Proc.* **2** (2014) 96–104.
81. A. Corbetta, A. Mountean and K. Vafayi, Parameter estimation of social forces in pedestrian dynamics models via probabilistic method, *Math. Biosci. Eng.* **12** (2015) 337–356.
82. V. Coscia, M. Delitala and P. Frasca, On the mathematical theory of vehicular traffic flow models II. Discrete velocity kinetic models, *Int. J. Non-linear Mech.* **42** (2007) 411–421.
83. E. Cristiani, B. Piccoli and A. Tosin, Multiscale modeling of granular flows with application to crowd dynamics, *Multiscale Model. Simul.* **9** (2011) 155–182.
84. F. Cucker and S. Smale, Emergent behavior in flocks, *IEEE Trans. Automat. Control* **52** (2007) 852–862.
85. D. Cumin and C. P. Unsworth, Generalising the Kuramoto model for the study of neuronal synchronisation in the brain, *Physica D* **226** (2007) 181–196.
86. C. F. Daganzo, Requiem for second-order fluid approximations of traffic flow, *Transp. Res. B* **29** (1995) 277–286.

87. P. Degond, C. Appert-Rolland, M. Moussaid, J. Pettré and G. Theraulaz, A hierarchy of heuristic-based models of crowd dynamics, *J. Stat. Phys.* **152** (2013) 1033–1068.
88. P. Degond, C. Appert-Rolland, J. Pettré and G. Theraulaz, Vision based macroscopic pedestrian models, *Kinet. Relat. Models* **6** (2013) 809–839.
89. P. Degond, A. Frouvelle and J.-G. Liu, Phase transitions, hysteresis, and hyperbolicity for self-organized alignment dynamics, *Arch. Ration. Mech. Anal.* **216** (2015) 63–115.
90. P. Degond and S. Motsch, Continuum limit of self-driven particles with orientation interaction, *Math. Models Methods Appl. Sci.* **18** (2008) 1193–1217.
91. M. Delitala and A. Tosin, Mathematical modelling of vehicular traffic: A discrete kinetic theory approach, *Math. Models Methods Appl. Sci.* **17** (2007) 901–932.
92. H. Dietert, Stability and bifurcation for the Kuramoto model, *J. Math. Pures Appl.* **105** (2016) 451–489.
93. H. Dietert, Stability of partially locked states in the Kuramoto model through Landau damping with Sobolev regularity, preprint (2017), arXiv:1707.03475.
94. H. Dietert, B. Fernandez and D. Gérard-Varet, Landau damping to partially locked states in the Kuramoto model, *Commun. Pure Appl. Math.* **71** (2018) 953–993.
95. G. Dimarco and L. Pareschi, Numerical methods for kinetic equations, *Acta Numer.* **23** (2014) 369–520.
96. T. Do, A. Kiselev, L. Ryzhik and C. Tan, Global regularity for the fractional Euler alignment system, *Arch. Ration. Mech. Anal.* **228** (2018) 1–37.
97. J.-G. Dong and X. Xue, Synchronization analysis of Kuramoto oscillators, *Commun. Math. Sci.* **11** (2013) 465–480.
98. A. Elmoussaoui, P. Argoul, M. ElRhabi and A. Hakim, Discrete kinetic theory for 2-Dmodeling of a moving crowd: Application to the evacuation of a non-connected bounded domain, *Comp. Math. Appl.* **75** (2018) 1159–1180.
99. J.-M. Epstein J. M., Modeling civil violence: An agent based computational approach, *Proc. Natl. Acad. Sci.* **99** (2002) 7243–7250.
100. J. Evers, R. Fetecau and L. Ryzhik, Anisotropic interactions in a first-order aggregation model, *Nonlinearity* **28** (2015) 2847–2871.
101. J. Evers, R. Fetecau and W. Sun, Small inertia regularization of an anisotropic aggregation model, *Math. Models Methods Appl. Sci.* **27** (2017) 1795–1842.
102. S. Faure and B. Maury, Crowd motion from the granular standpoint, *Math. Models Methods Appl. Sci.* **25** (2015) 463–493.
103. L. Fermo and A. Tosin, A fully-discrete-state kinetic theory approach to modeling vehicular traffic, *SIAM J. Appl. Math.* **73** (2013) 1533–1556.
104. L. Fermo and A. Tosin, Fundamental diagrams for kinetic equations of traffic flow, *Dyn. Syst. Ser. S* **7** (2014) 449–462.
105. L. Fermo and A. Tosin, A fully-discrete-state kinetic theory approach to traffic flow on road networks, *Math. Models Methods Appl. Sci.* **25** (2015) 423–461.
106. A. Festa, A. Tosin and M. T. Wolfram, Kinetic description of collision avoidance in pedestrian crowds by sidestepping, *Kinet. Relat. Models* **11** (2018) 491–520.
107. R. Fetecau and W. Sun, First-order aggregation models and zero inertial limits, *J. Differential Equations* **259** (2015) 6774–6802.
108. R. C. Fetecau, W. Sun and C. Tan, First-order aggregation models with alignment, *Physica D* **325** (2016) 146–163.
109. A. Figalli and M.-J. Kang, A rigorous derivation from the kinetic Cuker–Smale model to the pressureless Euler system with nonlocal alignment, to appear in *Anal. Partial Differential Equation.*

110. A. Figalli, M.-J. Kang and J. Morales, Global well-posedness of spatially homogeneous Kolmogorov–Vicsek model as a gradient flow, *Arch. Ration. Mech. Anal.* **227** (2018) 869–896.
111. M. Fornasier, B. Piccoli and F. Rossi, Mean-field sparse optimal control, *Philos. Trans. R. Soc. A* **372** (2014) 1–21.
112. M. Fornasier and F. Solombrino, Mean-field optimal control, *ESAIM Control Optim. Calc. Var.* **20** (2014) 1123–1152.
113. 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.
114. Z. Fu, L. Luo, Y. Yang, Y. Zhuang, P. Zhang, L. Yang, H. Yang, J. Ma, K. Zhu and Y. Li, Effect of speed matching on fundamental diagram of pedestrian flow, *Physica A* **458** (2016) 31–42.
115. I. M. Gamba, J. R. Haack and S. Motsch, Spectral method for a kinetic swarming model, *J. Comput. Phys.* **297** (2015) 32–46.
116. I. M. Gamba and M.-J. Kang, Global weak solutions for Kolmogorov–Vicsek type equations with orientational interactions, *Arch. Ration. Mech. Anal.* **222** (2016) 317–342.
117. T. Goudon, J. Nieto, F. Poupaud and J. Soler, Multidimensional high-field limit of the electrostatic Vlasov–Poisson–Fokker–Planck system, *J. Differential Equations* **213** (2005) 418–442.
118. L. Grüne and J. Pannek, *Nonlinear Model Predictive Control* (Springer, 2011).
119. S.-Y. Ha, T. Ha and J.-H. Kim, On the complete synchronization of the Kuramoto phase model, *Physica D* **239** (2010) 1692–1700.
120. S.-Y. Ha, Y.-H. Kim, J. Morales and J. Park, Emergence of phase concentration for the Kuramoto–Sakaguchi equation, preprint (2016), arXiv:1610.01703.
121. S.-Y. Ha, J. Kim, J. Park and X. Zhang, Uniform stability and mean-field limit for the augmented Kuramoto model, *Netw. Heterog. Media* **13** (2018) 297–322.
122. S.-Y. Ha, J. Kim, J. Park and X. Zhang, Complete cluster predictability of the Cucker–Smale flocking model on the real line, *Arch. Ration. Mech. Anal.* **231** (2019) 319–365.
123. S.-Y. Ha, H. K. Kim and S. W. Ryoo, Emergence of phase-locked states for the Kuramoto model in a large coupling regime, *Commun. Math. Sci.* **14** (2016) 1073–1091.
124. S.-Y. Ha, J. Kim and X. Zhang, Uniform stability of the Cucker–Smale model and its application to the mean-field limit, *Kinet. Relat. Models* **11** (2018) 1157–1181.
125. S.-Y. Ha, B. Kwon and M.-J. Kang, A hydrodynamic model for the interaction of Cucker–Smale particles and incompressible fluid, *Math. Models Methods Appl. Sci.* **24** (2014) 2311–2359.
126. S.-Y. Ha, B. Kwon and M.-J. Kang, Emergent dynamics for the hydrodynamic Cucker–Smale system in a moving domain, *SIAM J. Math. Anal.* **47** (2015) 3813–3831.
127. S.-Y. Ha, J. Kim, C. Min, T. Ruggeri and X. Zhang, A global existence of classical solutions to the hydrodynamic Cucker–Smale model in presence of a temperature field, *Anal. Appl.* **16** (2018) 757–805.
128. S.-Y. Ha, J. Kim, C. Min, T. Ruggeri and X. Zhang, Uniform stability and mean-field limit of thermodynamic Cucker–Smale model, *Quart. Appl. Math.* **77** (2019) 131–176.
129. S.-Y. Ha, D. Ko, J. Park and X. Zhang, Collective synchronization of classical and quantum oscillators, *EMS Surv. Math. Sci.* **3** (2016) 209–267.

130. S.-Y. Ha, Z. Li and X. Xue, Formation of phase-locked states in a population of locally interacting Kuramoto oscillators, *J. Differential Equations* **255** (2013) 3053–3070.
131. S.-Y. Ha and J.-G. Liu, A simple proof of Cucker–Smale flocking dynamics and mean field limit, *Commun. Math. Sci.* **7** (2009) 297–325.
132. S.-Y. Ha, S. E. Noh and J. Park, Synchronization of Kuramoto oscillators with adaptive couplings, *SIAM J. Appl. Dyn. Syst.* **15** (2016) 162–194.
133. S.-Y. Ha and T. Ruggeri, Emergent dynamics of a thermodynamically consistent particle model, *Arch. Ration. Mech. Anal.* **223** (2017) 1397–1425.
134. S.-Y. Ha and E. Tadmor, From particle to kinetic and hydrodynamic description of flocking, *Kinet. Relat. Models* **1** (2008) 415–435.
135. M. Haghani and M. Sarvi, Social dynamics in emergency evacuations: Disentangling crowds attraction and repulsion effects, *Physica A* **475** (2017) 24–34.
136. M. Hauray and P. E. Jabin, Particles approximations of Vlasov equations with singular forces: Propagation of chaos, *Ann. Sci. Ec. Norm. Super.* **48** (2015) 891–940.
137. M. Hauray and S. Mischler, On Kac's chaos and related problems, *J. Funct. Anal.* **266** (2014) 6055–6157.
138. S. He and E. Tadmor, Global regularity of two-dimensional flocking hydrodynamics, *C. Rend. Math.* **355** (2017) 795–805.
139. D. O. Hebb, *The Organization of Behavior* (Wiley, 1949).
140. D. Helbing and A. Johansson, Pedestrian crowd and evacuation dynamics, in *Encyclopedia of Complexity and System Science* (2009) 6476–6495.
141. D. Helbing, A. Johansson and H.-Z. Al-Abideen, Dynamics of crowd disasters: An empirical study, *Phys. Rev. E* **75** (2007).
142. M. Herty and L. Pareschi, A Fokker–Planck asymptotics for traffic flow models, *Kinet. Relat. Models* **13** (2010) 165–179.
143. M. Herty, L. Pareschi and M. Seaid, Enskog-like discrete velocity models for vehicular traffic flow, *Netw. Heterog. Media*, **2** (2007) 481–496.
144. M. Herty, A. Tosin, G. Visconti and M. Zanella, Hybrid stochastic kinetic description of two-dimensional traffic dynamics, *SIAM J. Appl. Math.* (2018) **78**(5), 27372762 (26 pages).
145. M. Herty and M. Zanella. Performance bounds for the mean-field limit of constrained dynamics, *Discr. Cont. Dyn. Syst. Series A*, **37** (2017) 2023–2043.
146. D. Hilbert, Mathematical problems, *Bull. Amer. Math. Soc.* **8** (1902) 437–479.
147. M. Hinze and S. Volkwein, Instantaneous control for the instationary Burgers equation — convergence analysis and numerical implementation, *Nonlinear Anal. T.M.A.* **50** (2002) 1–26.
148. S. P. Hoogendoorn, F. van Wageningen-Kessels, W. Daamen and D.C. Duives, Continuum modelling of pedestrian flows: From microscopic principles to self-organised macroscopic phenomena, *Physica A* **416** (2014) 684–694.
149. M. L. L. Iannini and R. Dickman, Kinetic theory of vehicular traffic, *Amer. J. Phys.* **84** (2016) 135–145.
150. P. E. Jabin, A review of the mean field limits for Vlasov equations, *Kinet. Relat. Models* **7** (2014) 661–711.
151. P. E. Jabin and Z. Wang, Mean field limit and propagation of chaos for Vlasov systems with bounded forces, *J. Funct. Anal.* **271** (2016) 3588–3627.
152. P. E. Jabin and Z. Wang, Mean field limit for stochastic particle systems. In Active Particles, Volume 1. *Modeling and Simulation in Science, Engineering and Technology Active Particle*, Vol. 1 N. Bellomo, P. Degond and E. Tadmor (eds.) (Birkhäuser,

2017), pp. 379–402.
153. P. E. Jabin and Z. Wang, Quantitative estimates of propagation of chaos for stochastic systems with $W^{1,\infty}$ kernels, *Invent. Math.* **214** (2018) 523–591.
154. S. Jin, L. Li and J-G. Liu, Random batch methods (RBM) for interacting particle systems, preprint (2019), arXiv:1812.10575.
155. M. Kac, Foundations of kinetic theory, in *Proc. 3rd Berkeley Symp. on Mathematical Statistics and Probability*, 1954–1955, Vol. III, pp. 171–197 (University of California Press, 1956), pp. 171–197.
156. B. S. Kalitin and T. Sari, B-stability and its applications to the tikhonov and Malkin–Gorshin Theorems, *Differential Equations*, **37** (2001) 11–16. translated from *Differentsial'nye Uravneniya*, **37** (2001) 12–17.
157. M.-J. Kang and J. Morales, Dynamics of a spatially homogeneous Vicsec model for oriented particles in the plane, preprint (2016), arXiv:1608.00185.
158. M.-J. Kang and A. Vasseur, Asymptotic analysis of Vlasov-type equations under strong local alignment regime, *Math. Models Methods Appl. Sci.* **25** (2015) 2153–2173.
159. T. K. Karper, A. Mellet and K. Trivisa, Hydrodynamic limit of the kinetic Cucker–Smale flocking model, *Math. Models Methods Appl. Sci.* **25** (2015) 131–163.
160. D. Kim and A. Quaini, A kinetic theory approach to model pedestrian dynamics in bounded domains with obstacles, preprint (2019), arXiv:1901.07620v2 [mathNA].
161. M. Kinateder, T. D. Wirth and W. H. Warren, Crowd dynamics in virtual reality, in *Crowd Dynamics Voume 1 — Theory Models and Safety Problems*, eds. L. Gibelli and N. Bellomo (Birkhäuser, 2018), pp. 11–62.
162. A. Kiselev and C. Tan, Finite time blow up in the hyperbolic Boussinesq system, *Adv. Math.* **325** (2018) 34–55.
163. A. Kiselev and C. Tan, Global regularity for 1D Eulerian dynamics with singular interaction forces, *SIAM J. Numer. Anal.* **50** (2018) 6208–6229.
164. A. Klar and R. Wegener, Enskog-like kinetic models for vehicular traffic, *J. Stat. Phys.* **87** (1997) 91–114.
165. M. Kücken and C. Champod, Merkel cells and the individuality of friction ridge skin, *J. Theor. Biol.* **317** (2013) 229–237.
166. Y. Kuramoto, *Chemical Oscillations, Waves and Turbulence* (Springer, 1984).
167. Y. Kuramoto, International symposium on mathematical problems in mathematical physics, *Lect. Notes Theor. Phys.* **30** (1975) 420.
168. C. Lancellotti, On the Vlasov limit for systems of nonlinearly coupled oscillators without noise, *Transp. Theory Stat. Phys.* **34** (2005) 523–535.
169. X. Li, F. Guo, H. Kuang, Z. Geng and Y. Fan, An extended cost potential field cellular automaton model for pedestrian evacuation considering the restriction of visual field, *Physica A*, **515** (2019) 47–56.
170. X. Li, F. Guo, H. Kuang and H. Zhou, Effect of psychological tension on pedestrian counter flow via an extended cost potential field cellular automaton model, *Physica A*, **487** (2017) 47–56.
171. Z. Li, X. Xue and D. Yu, On the Łojasiewicz exponent of Kuramoto model, *J. Math. Phys.* **56** (2015) 022704.
172. S. McNamara and W. Young, Kinetics of a one-dimensional granular medium in the quasi-elastic limit, *Phys. Fluids A* **5** (1995) 34–45.
173. M. C. Marchetti, J. F. Joanny, S. Ramaswamy, T. B. Liverpool, J. Prost, M. Rao and R. Aditi Simha, *Rev. Mod. Phys.* **85** (2013) 1143–1189.
174. W. Marques Jr. and A. R. Méndez, On the kinetic theory of vehicular traffic flow: Chapman–Enskog expansion versus Grad's moment method, *Physica A: Stat. Mech.*

Appl. **392** (2013) 3430–3440.
175. B. Maury and J. Venel, A discrete contact model for crowd motion, *ESAIM: M2AN* **45** (2011) 145–168.
176. D. Q. Mayne and H. Michalska, Receding horizon control of nonlinear systems, *IEEE Trans. Automat. Control* **35** (1990) 814–824.
177. S. Mischler and C. Mouhot, Kac's program in kinetic theory, *Invent. Math.* **193** (2013) 1–147.
178. S. Mischler, C. Mouhot and B. Wennberg, A new approach to quantitative chaos propagation for drift, diffusion and jump process, *Probab. Theory Related Fields* **161** (2015) 1–59.
179. A. Mogilner and L. Edelstein-Keshet, A non-local model for a swarm, *J. Math. Bio.* **38** (1999) 534–570.
180. D. Morale, V. Capasso and K. Oelschläger, An interacting particle system modelling aggregation behavior: From individuals to populations, *J. Math. Biol.* **50** (2005) 49–66.
181. S. Motsch and E. Tadmor, A New Model for Self-Organized Dynamics and Its Flocking Behavior, *J. Stat. Phys.* **144** (2011) 923–947.
182. M. Moussaïd, D. Helbing, S. Garnier, A. Johansson, M. Combe and G. Theraulaz, Experimental study of the behavioural mechanisms underlying self-organization in human crowds, *Proc. Roy. Soc. B*, **276** (2009) 2755–2762.
183. M. Moussaïd and G. Theraulaz, Comment les piétons marchent dans la foule, *La Recherche* **450** (2011) 56–59.
184. P. B. Mucha and J. Peszek, The Cucker–Smale equation: Singular communication weight, measure solutions and weak-atomic uniqueness, *Arch. Ration. Mech. Anal.* **227** (2018) 273–308.
185. G. Naldi, L. Pareschi and G. Toscani, *Mathematical Modeling of Collective Behavior in Socio-economic and Life Sciences* (Springer Science & Business Media, 2010).
186. P. Nelson, A kinetic model of vehicular traffic and its associated bimodal equilibrium solution, *Transp. Theory Statist. Phys.* **24** (1995) 383–409.
187. H. Neunzert, An introduction to the nonlinear Boltzmann–Vlasov equation, in *Kinetic Theories and the Boltzmann Equation*, Lecture Notes in Mathematics, Vol. 1048 (Springer-Verlag, 1984).
188. J. Nieto, F. Poupaud and J. Soler, High-field limit for the Vlasov–Poisson–Fokker–Planck system, *Arch. Ration. Mech. Anal.* **158** (2001) 29–59.
189. R. K. Niyogi and L. Q. English, Learning-rate-dependent clustering and self-development in a network of coupled phase oscillators, *Phys. Rev. E.* **80** (2009) 066213.
190. L. Pareschi and G. Russo, An introduction to Monte Carlo methods for the Boltzmann equation, in *CEMRACS 1999*, ESAIM Proceedings, Vol. 10 (2018), (Society for Industrial and Applied Mathematics, 1999), pp. 35–76.
191. J. Park, D. Poyato and J. Soler, Filippov trajectories and clustering in the Kuramoto model with singular couplings, preprint (2018), arXiv:1809.04307.
192. L. Pareschi and G. Toscani, *Interacting Multiagent Systems: Kinetic Equations and Monte Carlo Methods* (Oxford Univ. Press, 2013).
193. S. Paveri Fontana, On Boltzmann like treatments for traffic flow, *Transp. Res.* **9** (1975) 225–235.
194. J. Peszek, Existence of piecewise weak solutions of a discrete Cucker–Smale's flocking model with a singular communication weight, *J. Differential Equations* **257** (2014) 2900–2925.

195. J. Peszek, Discrete Cucker–Smale flocking model with a weakly singular weight, *SIAM J. Math. Anal.* **47** (2015) 3671–3686.
196. C. B. Picallo and H. Riecke, Adaptive oscillator networks with conserved overall coupling: Sequential firing and near-synchronized states, *Phys. Rev. E.* **83** (2011) 036206.
197. B. Piccoli, F. Rossi and E. Trélat, Control to Flocking of the Kinetic Cucker–Smale Model, *SIAM J. Math. Anal.* **47** (2015) 4685–4719.
198. F. Poupaud and J. Soler, Parabolic limit and stability of the Vlasov–Fokker–Planck system, *Math. Models Methods Appl. Sci.* **10** (2000) 1027–1045.
199. A. J. Povzner, On the Boltzmann equation in the kinetic theory of gases, *Mat. Sb. (N.S.)* **58** (1962) 65–86.
200. D. Poyato, Filippov flows and mean-field limits in the kinetic singular Kuramoto model, preprint (2019), arXiv:1903.01305.
201. D. Poyato and J. Soler, Euler-type equations and commutators in singular and hyperbolic limits of kinetic Cucker–Smale models, *Math. Models Methods Appl. Sci.* **27** (2017) 1089–1152.
202. I. Prigogine and R. Herman, *Kinetic Theory of Vehicular Traffic* (Elsevier, 1971).
203. G. Puppo, M. Semplice, A. Tosin and G. Visconti, Fundamental diagrams in traffic flow: The case of heterogeneous kinetic models, *Comm. Math. Sci.* **14** (2016) 643–669.
204. G. Puppo, M. Semplice, A. Tosin and G. Visconti, Analysis of a multi-population kinetic model for traffic flow, *Comm. Math. Sci.* **15** (2017) 379–412.
205. G. Puppo, M. Semplice, A. Tosin and G. Visconti, Kinetic models for traffic flow resulting in a reduced space of microscopic velocities, *Kinet. Relat. Models* **10** (2017) 823–854.
206. E. Ronchi, E. D. Kuligowski, D. Nilsson, R. D. Peacock and P. A. Reneke, Assessing the verification and validation of building fire evacuation models, *Fire Technol.* **52** (2016) 197–219.
207. F. Ronchi, F. Nieto Uriz, X. Criel and P. Reilly, Modelling large-scale evacuation of music festival, *Fire Safety* **5** (2016) 11–19.
208. E. Ronchi and D. Nilsson, Pedestrian movement in smoke, Data and Modeling approaches, in *Crowd Dynamics Voume 1 — Theory Models and Safety Problems*, eds. L. Gibelli and N. Bellomo (Birkhäuser, 2018), pp. 37–62.
209. M. Rubinov, O. Sporns, J. P. Thivierge and M. Breakspear, Neurobiologically realistic determinants of self-organized criticality in networks of spiking neurons, *PLoS Comput. Biol.* **7** (2011) e1002038.
210. A. Schadschneider, M. Chraibi, A. Seyfried, A. Tordeux and J. Zhang, Pedestrian dynamics: From empirical results to modeling, in *Crowd Dynamics Volume 1 — Theory Models and Safety Problems* eds. L. Gibelli and N. Bellomo (Birkhäuser, 2018), pp. 63–102.
211. A. Schadschneider, W. Klingsch, H. Kläpfel, T. Kretz, C. Rogsch and A. Seyfried, Evacuation dynamics: Empirical results, modeling and applications, in *Encyclopedia of Complexity and System Science* (Springer, 2009), pp. 3142–3176.
212. A. Schadschneider and A. Seyfried, Empirical results for pedestrian dynamics and their implications for modeling, *Netw. Heterog. Media* **6** (2011) 545–560.
213. P. Seliger, S. C. Young and L. S. Tsimring, Plasticity and learning in a network of coupled phase oscillators, *Phys. Rev. E.* **65** (2002) 041906.
214. A. Seyfried, B. Steffen, W. Klingsch and M. Boltes, The fundamental diagram of pedestrian movement revisited, *J. Stat. Mech.: Theory Exper.* **360** (2006) 232–238.

215. R. Shvydkoy and E. Tadmor, Eulerian dynamics with a commutator forcing, *Trans. Math. Appl.* **1** (2017) 1–26.
216. R. Shvydkoy and E. Tadmor, Eulerian dynamics with a commutator forcing II: Flocking, *Discr. Contin. Dyn. Syst. Ser. A* **37** (2017) 5503–5520.
217. R. Shvydkoy and E. Tadmor, Eulerian dynamics with a commutator forcing III: Fractional diffusion of order $0 < \alpha < 1$, *Physica D* **376–377** (2018) 131–137.
218. R. Shvydkoy and E. Tadmor, Topological models for emergent dynamics with short-range interactions, preprint (2018), arXiv:1806.01371.
219. A.-S. Sznitman, Topics in propagation of chaos, in *Ecole d'été de Probabilités de Saint-Flour XIX-1989*, Lecture Notes in Maths., Vol. 1464 (Springer, Berlin, 1991), pp. 165–251.
220. E. Tadmor and C. Tan, Critical thresholds in flocking hydrodynamics with nonlocal alignment, *Proc. R. Soc. A* **372** (2014) 20130401.
221. C. M. Topaz and A. L. Bertozzi, Swarming patterns in a two-dimensional kinematic model for biological groups, *SIAM J. Appl. Math.* **65** (2004) 152–174.
222. C. M. Topaz, A. L. Bertozzi and M. A. Lewis, A nonlocal continuum model for biological aggregation, *Bull. Math. Biol.* **68** (2006) 1601–1623.
223. A. Tosin and M. Zanella, Control strategies for road risk mitigation in kinetic traffic modelling, IFAC **51**(9) (2018) 67–72.
224. V. S. Varadarajan, On the convergence of sample probability distributions, *Sankhyā: Indian J. Statist.* **19** (1958) 23–26.
225. T. Vicsek, A. Czirók, 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) 1226–1229.
226. P. Villegas, P. Moretti and M. A. Muñoz, Frustrated hierarchical synchronization and emergent complexity in the human connectome network, *Sci. Rep.* **4** (2014) 5990.
227. G. Visconti, M. Herty, G. Puppo and A. Tosin, Multivalued Fundamental Diagrams of Traffic Flow in the Kinetic Fokker–Planck Limit, *Multiscale Model. Simul.* **15** (2017) 1267–1293.
228. L. Wang, M. B. Short and A. L. Bertozzi, Efficient numerical methods for multiscale crowd dynamics with emotional contagion, *Math. Models Methods Appl. Sci.* **27** (2017) 205–230.
229. N. Wijermans, C. Conrado, M. van Steen, C. Martella and J.-L. Li, A landscape of crowd management support: An integrative approach, *Safety Sci.* **86** (2016) 142–164.
230. B. Zhan, D.-N. Monekosso, P. Remagnino, S. A. Velastin and L.-Q. Xu, Crowd analysis: A survey, *Mach. Vis. Appl.* **19** (2008) 345–357.
231. C. Zhou, L. Zemanová, G. Zamora-López, C. C. Hilgetag and J. Kurths, Hierarchical organization unveiled by functional connectivity in complex brain networks, *Phys. Rev. Lett.* **97** (2006) 238103.
232. C. Zhou, L. Zemanová, G. Zamora-López, C. C. Hilgetag and J. Kurths, Structure-function relationship in complex brain networks expressed by hierarchical synchronization, *New J. Phys.* **9** (2007) 178.