Mean Field Limit for Stochastic Particle Systems

Pierre-Emmanuel Jabin and Zhenfu Wang

Abstract We review some classical and more recent results for the derivation of mean field equations from systems of many particles, focusing on the stochastic case where a large system of SDE's leads to a McKean-Vlasov PDE as the number N of particles goes to infinity. Classical mean field limit results require that the interaction kernel be essentially Lipschitz. To handle more singular interaction kernels is a longstanding and challenging question but which has had some recent successes.

1 Introduction

Large systems of interacting particles are now fairly ubiquitous. The corresponding microscopic models are usually conceptually simple, based for instance on Newton's 2nd law. However they are analytically and computationally complicated since the number N of particles is very large (with N in the range of $10^{20}-10^{25}$ for typical physical settings).

Understanding how this complexity can be reduced is a challenging but critical question with potentially deep impact in various fields and a wide range of applications: in Physics where particles can represent ions and electrons in plasmas, or molecules in a fluid and even galaxies in some cosmological models; in the Biosciences where they typically model micro-organisms (cells or bacteria); in Economics or Social Sciences where particles are individual "agents".

P-E Jabin is partially supported by NSF Grant 1312142 and by NSF Grant RNMS (Ki-Net) 1107444.
CSCAMM and Dept. of Mathematics, University of Maryland, College Park, MD, 20742, USA.
e-mail: pjabin@cscamm.umd.edu

Z. Wang is supported by NSF Grant 1312142.
CSCAMM and Dept. of Mathematics, University of Maryland, College Park, MD, 20742, USA.
e-mail: zwang423@math.umd.edu

The classical strategy to reduce this complexity is to derive a mesoscopic or macroscopic system, *i.e.* a continuous description of the dynamics where the information is embedded in densities typically solving non-linear PDE's.

The idea of such a kinetic description of large systems of particles goes back to the original derivation of statistical mechanics and the works of Maxwell and Boltzmann on what is now called the Boltzmann equation which describes the evolution of dilute gases.

We consider here a different (and in several respects easier) setting: The mean field scaling is in a collisionless regime meaning that collisions seldom occur and particles interact with each other at long range.

The first such mean field equation was introduced in galactic dynamics by Jeans in 1915 [49]. The Vlasov equation was introduced in plasma physics by Vlasov in [75, 76] as the mean field equation for large particle systems of ions or electrons interacting through the Coulomb force, ignoring the effect of collisions.

The rigorous derivation of the Vlasov equation with the Coulomb or Newtonian potential from Newton dynamics is still a major open question in the topic. For some recent progress in that direction, we refer to [42, 54, 55]. However we focus here on the stochastic case and refer to [35, 46] for a review of the mean field limit for deterministic systems.

In the rest of this introduction, we present some of the classical models that one typically considers.

1.1 Classical 2nd Order dynamics

The most classical model is the Newton dynamics for N indistinguishable point particles driven by 2-body interaction forces and Brownian motions. Denote by $X_i \in \Gamma$ and $V_i \in \mathbb{R}^d$ the position and velocity of particle number i. The evolution of the system is given by the following SDE's,

$$\mathrm{d}X_i = V_i\,\mathrm{d}t, \quad \mathrm{d}V_i = \frac{1}{N}\sum_{j\neq i} K(X_i - X_j)\,\mathrm{d}t + \sqrt{2\sigma}\,\mathrm{d}W_t^i, \tag{1}$$

where $i = 1, 2, \cdots, N$. The W^i are N independent Brownian motions or Wiener processes, which may model various types of random phenomena: For instance random collisions against a given background. If $\sigma = 0$, the system (1) reduces to the classical deterministic Newton dynamics. We always assume here that $\sigma > 0$ but we may consider cases where σ scales with N. We then denote the coefficient σ_N and assume that $\sigma_N \to \sigma \geq 0$.

Observe that the Wiener processes are only present in the velocity equations which will have several important consequences.

The space domain Γ may be the whole space $\mathbb{R}^d$, the flat torus $\mathbb{T}^d$ or some bounded domain. The analysis of a bounded, smooth domain Γ is strongly dependent on the type of boundary conditions but can sometimes be handled in a manner similar to the other cases with some adjustments. Thus for simplicity we typically limit ourselves to $\Gamma = \mathbb{R}^d$, $\mathbb{T}^d$. Even if Γ is bounded, there is no hard cap on velocities so that the actual domain in position and velocity, $\Gamma \times \mathbb{R}^d$ is always unbounded.

The critical scaling in (1) (and later in (3)) is the factor $\frac{1}{N}$ in front of the interaction terms. This is *the mean field scaling* and it keeps, at least formally, the total strength of the interaction of order 1.

At least formally, one expects that as the number N of particles goes to infinity, (1) will be replaced by a continuous PDE. In the present case, the candidate is the so-called McKean-Vlasov equation (or sometimes Vlasov-Fokker-Planck) which reads

$$\partial_t f + v \cdot \nabla_x f + (K \star \rho) \cdot \nabla_v f = \sigma \Delta_v f, \tag{2}$$

where the unknown $f = f(t,x,v)$ is the phase space density or 1-particle distribution and $\rho = \rho(t,x)$ is the spatial (macroscopic) density obtained through

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

This type convergence is what we call *mean field limit* and it is connected to the important property of *propagation of chaos*.

We point out that Eq. (2) is of degenerate parabolic type as the diffusion $\Delta_v f$ only acts on the velocity variable.

1.2 First Order Systems

As the companion of (1), we also consider the 1st order stochastic system

$$\mathrm{d}X_i = \frac{1}{N}\sum_{j\neq i} K(X_i - X_j)\,\mathrm{d}t + \sqrt{2\sigma}\,\mathrm{d}W_t^i, \quad i = 1, \cdots, N, \tag{3}$$

with the same assumptions as for the system (1). As before, one expects that as the number N of particles goes to infinity the system (3) will converge to the following PDE

$$\partial_t f + \mathrm{div}_x(f\ (K \star f)) = \sigma \Delta_x f, \tag{4}$$

where the unknown $f = f(t,x)$ is now the spatial density.

The model (3) can be regarded as the small mass limit (Smoluchowski-Kramers approximation) of Langevin equations in statistical physics. However, the model (3) has its own important applications.

The best known classical application is in Fluid dynamics with the Biot-Savart kernel

$$K(x) = \frac{1}{2\pi}\left(\frac{-x_2}{|x|^2}, \frac{x_1}{|x|^2}\right).$$

This leads to the well-know vortex model which is widely used to approximate the 2D Navier-Stokes equation written in vorticity form. See for instance [14, 15, 29, 59, 67].

Systems (1) or (3) can be written in the more general form of

$$\mathrm{d}Z_i = \frac{1}{N}\sum_{j=1}^{N} H(Z_i, Z_j)\,\mathrm{d}t + \sqrt{2\sigma}\,\mathrm{d}\tilde{W}_t^i, \quad i = 1, 2, \cdots, N, \tag{5}$$

where we denote $\tilde{W}_t^i = W_t^i$ in the case of 1st order models and $\tilde{W}_t^i = (0,\, W_t^i)$ for 2nd order models where there is only diffusion in velocity.

It is easy to check that taking

$$H(Z_i, Z_j) = K(Z_i - Z_j)$$

with the convention that $K(0) = 0$, the system of SDE's (5) becomes (3). Furthermore, if we replace Z_i by the pair (X_i, V_i) as in (1), and set H as

$$H((X_i, V_i), (X_j, V_j)) = (V_i,\; K(X_i - X_j))$$

and again with convention that $K(0) = 0$, then the SDE's (5) can also represent the Newton like systems (1). We write particles systems in the forms as (1) and (3) simply because that is enough for most interesting applications.

1.3 Examples of Applications

As mentioned above, the best know example of interaction kernel is the Poisson kernel, that is

$$K(x) = \pm C_d \frac{x}{|x|^d}, \quad d = 2, 3, \cdots,$$

where $C_d > 0$ is a constant depending on the dimension and the physical parameters of the particles (mass, charges...). This corresponds to particles under gravitational interactions for the case with a minus sign and electrostatic interactions (ions in a plasma for instance) for the case with a positive sign.

The description by McKean-Vlasov PDE's goes far beyond plasma physics and astrophysics. Large systems of interacting particles are now widely used in the Biosciences and social sciences. Since, many individual based particle models are formulated to model collective behaviors of self-organizing particles or agents. Given the considerable literature, we only give a few limited examples and the references that are cited have no pretension to be exhaustive.

- Biological systems modeling the collective motion of micro-organisms (bacteria, cells...). The canonical example is again the case of the Poisson kernel for K where System (3) coincides with the particle models to approximate the Keller-

Segel equation of chemotaxis. We refer mainly to [30] for the mean field limit, together with [34, 56] (see also [68, 69, 72] for general modeling discussions).

- Aggregation models correspond to System (3), typically with $K = -\nabla W$ and an extra potential term $-\nabla V(X_i)$,

$$\mathrm{d}X_i = -\frac{1}{N}\sum_{j\neq i}\nabla W(X_i - X_j)\,\mathrm{d}t - \nabla V(X_i)\,\mathrm{d}t + \sqrt{2\sigma}\,\mathrm{d}W_t^i, \quad i = 1, \cdots, N. \tag{6}$$

  They are used in many settings (in biology, ecology, for space homogeneous granular media as in [3]...). See for instance [9, 10, 13, 57, 58] for the mathematical analysis of the particle system (6), [19, 20, 25] for the analysis of the limiting PDE.
- Since the pioneering works in [74] and [24], 2nd order systems like (1) have been used to model flocks of birds, schools of fishes, swarms of insects, ... One can see [17, 18, 40] and the references therein for a more detailed discussion of flocking or swarming models in the literature. We emphasize that presence of the noise in the models is important since we cannot expect animals to interact with each other or the environment in a completely deterministic way. We in particular refer to [39] for stochastic Cucker-Smale model with additive white noise as in (1) and to [1] for multiplicative white noise in velocity variables respectively. The rigorous proof of the mean field limit was given in [7] for systems similar to (1) with locally Lipschitz vector fields; the mean-field limit for stochastic Vicsek model where the speed is fixed is given in [8].
- First order models are quite popular to model opinion dynamics among a population (such as the emergence of a common belief in a pricing system). We refer for instance to [43, 52, 65, 77]. Individual-based models are even used for coordination or consensus algorithm in control engineering for robots and unmanned vehicles, see [21].

There are many other interesting questions that are related to mean field limit for stochastic systems but that are out of the scope of the present article. For instance

- *The derivation of collisional models and Kac's Program in kinetic theory.* After the seminal in [53] and later [23], the rigorous derivation of the Boltzmann equation was finally achieved in [31] but only for short times (of the order of the average time between collisions). The derivation for longer times is still widely open in spite of some critical progress when close to equilibrium in [4, 5]. Many tools and concepts that are used for mean field limits were initially introduced for collisional models, such as the ideas in the now famous Kac's program. Kac first introduced a probabilistic approach to simulate the spatially homogeneous Boltzmann equation in [50] and formulated several related conjecture. After some major contributions in the 90's, see in particular [26, 37, 38], significant progress was again achieved recently in solving these conjectures, see [41, 63, 64], and the earlier [16].
- Stochastic vortex dynamics with multiplicative (instead of additive) noise leading to Stochastic 2D Euler equation. In [28], the authors showed that the point vortex

dynamics becomes fully well-posed for every initial configuration when a generic stochastic perturbation (in the form of multiplicative noises) compatible with the Euler description is introduced. The SDE systems in [28] will converge to the stochastic Euler equation, rather than Navier-Stokes equation as the number N of point vortices goes to infinity. However, the rigorous proof of the convergence is difficult and still open.

- Scaling limit (hydrodynamic limit) of random walks on discrete spaces, for instance on lattice $\mathbb{Z}^d$ for which we refer to [51]. In this setting, one also tries to obtain a continuum model, usually a deterministic PDE, from a discrete particle model on a lattice, as $N \to \infty$ and of course the mesh size h converges to 0. An interesting observation is that we can use a stochastic PDE as a correction to the limit deterministic PDE, see [27].

This article is organized as follows: In section 2 we introduce the basic concepts and tools in this subject. We define the classical notion of mean field limit and propagation of chaos as well as more recent notions of chaos. Then in section 3 we review some classical results under the assumptions that K is globally Lipschitz and more recent results for singular kernels K. Both qualitative and quantitative results will be presented. Finally, in section 4 we briefly review the authors' recent results [47] and [48] for very rough interaction kernels K with the relative entropy method.

2 The basic concepts and main tools

In order to compare the particle systems (1) and (3) with the expected mean field equations (2) and (4) respectively, we need to introduce several concepts and tools to capture the information on both levels of descriptions.

Those tools have often been introduced in different contexts (and in particular for the derivation of collisional models such as the Boltzmann equation). The main classical references here are [6, 36, 50, 71] and [45] for the stochastic aspects.

2.1 The empirical measure

In the following, to make the presentation simple, we sometimes use the unified, one variable formulation (5). Therefore, for the 2nd order model,

$$Z_i = (X_i, V_i) \in E := \Gamma \times \mathbb{R}^d$$

and for the 1st order model,

$$Z_i = X_i \in E := \Gamma.$$

One defines the empirical measure as

$$\mu_N(t,z) = \frac{1}{N}\sum_{i=1}^{N} \delta(z - Z_i(t)), \tag{7}$$

where $z = (x,v) \in E = \Gamma \times \mathbb{R}^d$ or $z = x \in \Gamma$.

The empirical measure is a random probability measure, *i.e.* $\mu_N(t) \in \mathscr{P}(E)$, whose law lies in the space $\mathscr{P}(\mathscr{P}(E))$. Recall that $\mathscr{P}(E)$ represents the set of all Borel probability measures on E. Since all particles Z_i are assumed to be indistinguishable, $\mu_N(t,z)$ gives the full information of the solution

$$Z^N(t) = (Z_1(t), \cdots, Z_N(t)),$$

to the particle system (1) or (3).

One uses a slight variant in the case of 2D stochastic vortex model for approximating Navier-Stokes equation where for convenience one usually defines

$$\mu_N(t,x) = \frac{1}{N}\sum_{i=1}^{N} \alpha_i \delta(x - X_i(t)), \tag{8}$$

see for instance [29]. In that case, α_i models the strength of the circulation which can be positive or negative and hence the empirical measure is a signed measure and not a probability measure.

In the deterministic setting, that is provided $\sigma = 0$ in (1) or (3) (and therefore in (2) and (4)), the empirical measure μ_N is also deterministic. Furthermore in this special case it actually solves exactly the limiting PDE (2) or (4) in the sense of distribution. However, this cannot be true anymore for the stochastic setting: the stochastic behavior can only vanish when the number N of particles goes to infinity.

Systems (1) and (3) need to be supplemented with initial conditions. A first possibility is to choose a deterministic sequence of initial data $Z^N(t=0)$.

It is considered more realistic though to use random initial conditions in which case $Z^N(t=0)$ is taken according to a certain law

$$\mathrm{Law}(Z_1^0, \cdots, Z_N^0) = F^N(0) \in \mathscr{P}(E^N).$$

One often assumes some sort of independence (or almost independence) in the law $F^N(0)$. A typical example is *chaotic law* for which $F^N(0, z_1, \ldots, z_N) = \Pi_{i=1}^N f^0(z_i)$ for a given function f^0.

No matter which type of initial condition is chosen, $\mu_N(t,z)$ is always random for any $t > 0$.

The concept of mean field limit is defined for a particular choice of sequence of initial data

Definition 1 (Mean field limit). Consider a sequence of deterministic initial data $(Z_1^0, \cdots, Z_N^0)$ as $N \to \infty$ or equivalently a sequence of deterministic initial empirical measures μ_N^0, such that

$$\mu_N^0 \to f_0$$

in the tight topology of measures. Then the mean field limit holds for this particular sequence iff for *a.e.* t

$$\mu_N(t) \to f_t$$

with convergence in law, where f_t denotes the solution to (2) or (4) with initial data f_0.

Here the convergence in law of μ_N means that for any ϕ bounded continuous function from the set of probability measures on E, $\mathscr{P}(E)$ to $\mathbb{R}$, one has that

$$\phi(\mu_N(t,.)) \longrightarrow \phi(f_t(.)), \tag{9}$$

as $N \to \infty$, where $t \geq 0$. This in particular implies the convergence of the first moment (the case where ϕ is linear in μ_N) s.t. for any $\varphi \in C_b(E)$

$$\mathbb{E}\int_E \varphi(z) \mathrm{d}\mu_N(t,z) := \mathbb{E}\frac{1}{N}\sum_{i=1}^N \varphi(Z_i(t)) \to \int_E \varphi(z) f_t(z)\,\mathrm{d}z, \tag{10}$$

A few important points follow from this definition:

- The mean field limit may hold for one particular choice of sequence and not hold for another. In fact for many singular kernels, this is very likely as it is easy to build counterexamples where the particles are initially already concentrated. This means that the right questions are *for which sequence of initial data the mean field limit holds* and whether *this set of initial data is somehow generic*.
- If one chooses random initial data for instance according to some chaotic law $F^N(0) = (f^0)^{\otimes N}$ then the question of the mean field limit can be asked for each instance of the initial data. Hence the mean field limit could *a priori* holds with a certain probability that one would hope to prove equal to 1. This will lead to the important notion of Propagation of chaos.
- Because the limit f is deterministic (as a solution to a PDE), the mean field limit obviously only holds if μ_N becomes deterministic at the limit. This hints at possible connection with some sort of law of large numbers.

2.2 The Liouville equation

While the empirical measure follows the trajectories of the system, it can be useful to have statistical informations as given by the joint law

$$F^N(t, z_1, \cdots, z_N) = \mathrm{Law}(Z_1(t), \cdots, Z_N(t))$$

of the particle systems (1) or (3). F^N is not experimentally measurable for practical purposes and instead the observable statistical information of the system is contained in its marginals. One hence defines $k-$marginal distribution as

$$F_k^N(t,z_1,\cdots,z_k) = \int_{E^{N-k}} F^N(t,z_1,\cdots,z_N)\mathrm{d}z_{k+1}\cdots\mathrm{d}z_N. \tag{11}$$

The 1-marginal is also known as the 1 particle distribution while the 2-marginal contains the information about pairwise correlations between particles.

It is possible to write a closed equation, usually called the Liouville equation, governing the evolution of the law F^N. For 2nd order systems, it reads

$$\partial_t F^N + \sum_{i=1}^N v_i\cdot\nabla_{x_i}F^N + \frac{1}{N}\sum_{i=1}^N\sum_{j\neq i}K(x_i-x_j)\cdot\nabla_{v_i}F^N = \sigma\sum_{i=1}^N \Delta_{v_i}F^N. \tag{12}$$

Similarly, for the 1st order systems (3), one has

$$\partial_t F^N + \frac{1}{N}\sum_{i=1}^N\sum_{j\neq i}K(x_i-x_j)\cdot\nabla_{x_i}F^N = \sigma\sum_{i=1}^N \Delta_{x_i}F^N. \tag{13}$$

In the deterministic case, those equations had been derived by Gibbs, see [32, 33]. In the present stochastic setting, they follow from Itô's formula, see [45], applied to $\phi(Z^N(t))$ for any test function.

The fact that particles are indistinguishable implies that F^N is a symmetric probability measure on the space E^N, that is for any permutation of indices $\tau\in S_N$,

$$F^N(t,z_1,\cdots,z_N) = F^N(t,z_{\tau(1)},\cdots,z_{\tau(N)}).$$

We write it as $F^N\in\mathscr{P}_{\mathrm{Sym}}(E^N)$. Similarly, it is easy to check that the $k-$marginal distribution is also symmetric $F_k^N\in\mathscr{P}_{\mathrm{Sym}}(E^k)$ for $2\le k\le N$.

2.3 The BBGKY hierarchy

For simplicity we only focus on the 1st order system (1) as the 2nd order dynamics (1) case can be dealt with similarly by adding the corresponding free transport terms.

From the Liouville equation (13), it is possible to deduce equations on each marginal F_k^N. Noticing the fact that particles are indistinguishable and using the appropriate permutation, one obtains

$$\begin{aligned}\partial_t F_k^N &+ \frac{1}{N}\sum_{i=1}^k\sum_{j=1,j\neq i}^k K(X_i-X_j)\cdot\nabla_{x_i}F_k^N \\ &+ \frac{N-k}{N}\sum_{i=1}^k\int_\Gamma \mathrm{div}_{x_i}\left(K(x_i-y)F_{k+1}^N(t,x_1,\cdots,x_k,y)\right)\mathrm{d}y = \sigma\sum_{i=1}^k\Delta_{x_i}F_k^N.\end{aligned} \tag{14}$$

The equation (14) is not closed as it involves the next marginal F_{k+1}^N; thus the denomination of hierarchy.

On the other hand, each marginal F_k^N is defined on a fixed space E^k contrary to F^N which is defined on a space depending on N. Therefore one may easily consider the limit of F_k^N as $N \to \infty$ for a fixed k. Formally one obtains the limiting hierarchy

$$\partial_t F_k^\infty + \sum_{i=1}^k \int_\Gamma \mathrm{div}_{x_i} \left(K(x_i - y) F_{k+1}^\infty(t, x_1, \cdots, x_k, y) \right) \mathrm{d}y = \sigma \sum_{i=1}^k \Delta_{x_i} F_k^\infty.$$

Each equation is still not closed and *a priori* the hierarchy would have to be considered for all k up to infinity.

However if $F_k^\infty(t)$ is tensorized, $F_k^\infty(t) = f_t^{\otimes k}$, then each equation is closed and they all reduce to the limiting mean field equation (4). This leads us to the important notion of propagation of chaos.

2.4 Various notions of chaos

The original notion of propagation of chaos goes as far back as Maxwell and Boltzmann. The classical notion of propagation of chaos was formalized by Kac in [50], see also the famous [71]. The other, stronger notions of chaos presented here were investigated more recently in particular in [41], [63] and [64] (see also [16]).

Let us begin with the the simplest definition that we already saw

Definition 2. A law F^N is tensorized/factorized/chaotic if

$$F^N(z_1, \cdots, z_N) = \Pi_{i=1}^N F_1^N(z_i).$$

Unfortunately for N fixed the law $F^N(t)$ solving the Liouville Eq. (12) or (13) cannot be chaotic. Indeed for a fixed N some measure of dependence necessarily exists between particles and strict independence is only possible asymptotically. This leads to Kac's chaos.

Definition 3. Let E be a measurable metric space (here $E = \Gamma \times \mathbb{R}^d$ or Γ). A sequence $(F^N)_{N\in\mathbb{N}}$ of symmetric probability measures on E^N is said to be $f-$chaotic for a probability measure f on E, if one of the following equivalent properties holds:
i) For any $k \in \mathbb{N}$, the $k-$marginal F_k^N of F^N converges weakly towards $f^{\otimes k}$ as N goes to infinity, *i.e.* $F_k^N \rightharpoonup f^{\otimes k}$;
ii) The second marginal F_2^N converges weakly towards $f^{\otimes 2}$ as N goes to infinity: $F_2^N \rightharpoonup f^{\otimes 2}$;
iii) The empirical measure associate to F^N, that is $\mu_N(z) \in \mathscr{P}(E)$ as in (7) with $F^N = \mathrm{Law}(Z_1, \cdots, Z_N)$, converges in law to the deterministic measure f as N goes to infinity.

Here the weak convergence $F_k^N \rightharpoonup f^{\otimes k}$ simply means that for any test functions $\phi_1, \cdots, \phi_k \in C_b(E)$,

$$\lim_{N\to 0} \int_{E^k} \phi_1(z_1) \cdots \phi_k(z_k) F_k^N(z_1, \cdots, z_k) \mathrm{d}z_1 \cdots \mathrm{d}z_k = \Pi_{i=1}^k \int_E f(z_i) \phi_1(z_i) \, \mathrm{d}z_i,$$

and μ_N converges in law to f is as before in the sense of (9).

We refer to [71] for the classical proof of equivalence between the three properties. A version of the equivalence has recently been obtained in [41], quantified by the 1 Monge-Kantorovich-Wasserstein (MKW) distance between the laws.

We now can define the corresponding notion of propagation of chaos.

Definition 4 (Propagation of chaos). Assume that the sequence of the initial joint distribution $(F^N(0))_{N\geq 2}$ is f_0-chaotic. Then propagation of chaos holds for systems (1) or (3) up to time $T>0$ iff for any $t\in[0,T]$, the sequence of the joint distribution at time t $(F^N(t))_{N\geq 2}$ is also f_t-chaotic, where f_t is the solutions to the mean field PDE (2) or (4) respectively with initial data f_0.

If initially $(Z_1^0,\cdots,Z_N^0)$ was chosen according to the law $F^N(0)$ with the sequence $F^N(0)$ to be f_0-chaotic, then by property iii) of Definition 3 propagation of chaos implies that the mean field limit holds with probability one.

Kac's chaos and propagation of chaos is rather weak and thus does not allow a very precise control on the initial data. For this reason, it can be useful to consider stronger notions of chaos.

There are two natural physical quantities that can help quantify such stronger notions of chaos: the (Boltzmann) entropy and the Fisher information. The scaled entropy of the law F^N is defined as

$$H_N(F^N)=\frac{1}{N}\int_{E^N}F^N\log F^N\mathrm{d}z_1\cdots\mathrm{d}z_N,$$

where we recall that $E=\Gamma\times\mathbb{R}^d$ for the 2nd order system and $E=\Gamma$ for the 1st order system. The Fisher information is

$$I_N(F^N)=\frac{1}{N}\int_{E^N}\frac{|\nabla F^N|^2}{F^N}\mathrm{d}z_1\cdots\mathrm{d}z_N.$$

We normalized both quantities by factor $\frac{1}{N}$ such that for any $f\in\mathscr{P}(E)$,

$$H_N(f^{\otimes N})=H_1(f),\quad I_N(f^{\otimes N})=I_1(f).$$

The use of those quantities leads to alternative and stronger definitions of a $f-$chaotic sequence, *entropy chaos* and *Fisher information chaos*.

Definition 5 (Definition 1.3 in [41]). Consider $f\in\mathscr{P}(E)$ and a sequence $(F^N)_{N\geq 2}$ of $\mathscr{P}_{\mathrm{Sym}}(E^N)$ such that for some $k>0$ the $k-$th moment $M_k(F_1^N)=\int|z|^k\,dF_1^N$ of F_1^N is uniformly bounded in N. We say that
i) the sequence (F^N) is $f-$Fisher information chaotic if

$$F_1^N\rightharpoonup f,\quad I_N(F^N)\to I_1(f),\quad I_1(f)<\infty$$

ii) the sequence (F^N) is $f-$entropy chaotic if

$$F_1^N\rightharpoonup f,\quad H_N(F^N)\to H_1(f),\quad H_1(f)<\infty.$$

There are even intermediary notions that we omit for simplicity. There exists a strict hierarchy between these two definitions as per

Theorem 1 (Theorem 1.4 in [41]). *Consider $f \in \mathscr{P}(E)$ and $(F^N)_{N\geq 2}$ a sequence of $\mathscr{P}_{Sym}(E^N)$ such that the $k-$th moment $M_k(F_1^N)$ is bounded, for some $k > 2$. In the list of assertions below, each one implies the assertion which follows:*
i) (F^N) is $f-$Fisher information chaotic;
ii) (F^N) is $f-$Kac's chaotic (as in Definition 3) and $I_N(F^N)$ is bounded;
iii) (F^N) is $f-$entropy chaotic;
iv) (F^N) is $f-$Kac's chaotic.

For some recent results on propagation of chaos in strong sense, we refer to [29] where the convergence is in the sense of entropy as in ii) in Definition 5 and to [34] for a similar argument for the sub-critical Keller-Segel model.

We will finish this section by considering the relation between the entropy of the full joint law and the entropy of the marginals

Proposition 1. *For any $F^N \in \mathscr{P}_{Sym}(E^N)$ and $f \in \mathscr{P}(E)$, one has*

$$H_k(F_k^N) \leq H_N(F^N), \quad H_k(F_k^N | f^{\otimes k}) \leq H_N(F^N | f^{\otimes N}). \tag{15}$$

The scaled relative entropy is defined by

$$H_N(F^N | f^{\otimes N}) = \frac{1}{N} \int_{E^N} F^N \log \frac{F^N}{f^{\otimes N}}.$$

By Proposition 1 and the Liouville equation (12) or (13), we can control the entropy of any marginal at any time $t > 0$ given the uniform bound $\sup_{N\geq 2} H_N(F^N(0))$ initially. This is really surprising since for instance the free bound

$$\sup_{N\geq 2} H_k(F_k^N(t,\cdot)) < \infty$$

will *a priori* ensure that the weak limit f_1 of F_1^N belongs to $L^1(E)$ at any time $t > 0$ by Theorem 3.4 in [41], without having to know or prove anything about the mean field limit.

3 Some of the main results on mean field limits

In this section, we review some main results on mean field limits for stochastic particle systems. The classical results, such as the famous [60] or [71] (see also the very nice presentation in [61]), will normally require that the interaction kernel K be Lipschitz; we also refer to [70]. For singular (not locally Lipschitz) kernels, only a few results are available, mostly in the context of 2D stochastic vortex model, see for instance [29, 59, 62, 67]. We also refer to [14, 15, 22] and [30, 34, 56] for the singular kernel cases.

3.1 The classical approach: Control on the trajectories

The results we present here are taken mainly from [71] and [61]. Note though that the diffusion processes considered in [61] are much more general than what we use here

$$\mathrm{d}X_i = \frac{1}{N}\sum_{i=1}^{N} b(X_i, X_j)\,\mathrm{d}t + \frac{1}{N}\sum_{i=1}^{N} \sigma(X_i, X_j)\,\mathrm{d}W_t^i$$

where $i = 1, 2, \cdots, N$, and the vector field b and matrix field σ are Lipschitz continuous with respect to both variables. As before the W_t^i are mutually independent $d-$dimensional Brownian motions.

However, to make the presentation simple, we only focus on the 1st order system (3) assuming that the kernel K is *Lipschitz*. 2nd order systems with the same Lipschitz assumption can be treated in a similar manner.

If the interaction kernel K is globally Lipschitz, a standard method to show the mean field limit was popularized by Sznitman [71] (see also the more recent [58]). We also refer to [7, 9, 12] and the reference therein for recent developments and to [64] for the quantitative Grunbaum's duality method.

The basic idea of the method is as follows: For system (3) endowed with initial data

$$X_i(0) = X_i^0, \quad i.i.d. \text{ with } \operatorname{Law} X_i^0 = f_0,$$

we construct a symmetric particle system coupled to (3), that is

$$\mathrm{d}\bar{X}_i = K \star f_t(\bar{X}_i) + \sqrt{2\sigma}\,\mathrm{d}W_t^i, \quad \bar{X}_i(0) = X_i^0, \quad i = 1, 2, \cdots, N, \tag{16}$$

where the W_t^i are the same Brownian motions as in (3) and $f_t = \operatorname{Law}(\bar{X}_i(t))$. The coupling between (3) and (16) is only due to the fact that they have the same initial data and share the same Brownian motions.

Observe that Eq. (16) is not anymore an SDE system: The dynamics between particles is now coupled through the law f_t. That law is the same for each $\bar{X}_i$ and in that sense one only considers N independent copies of the same system given by

$$\mathrm{d}X_t = K \star f_t(X_t) + \sqrt{2\sigma}\,\mathrm{d}W_t, \quad \operatorname{Law}(X_t) = f_t. \tag{17}$$

It is straightforward to check that f_t is just the (weak) solution to the mean field PDE (4) by Itô's formula. Therefore the law of large numbers in the right function space will give us that the system (16) is close to the mean field PDE (4).

Under some stability estimates that follow the traditional Gronwall type of bounds for SDE's, it can be shown that the symmetric system (16) is close to the original one (3).

On the other hand, (17) and (4) are well-posed with existence and uniqueness, which finally implies the mean field limit. We summarize the results of the above discussion with the following two theorems

Theorem 2 (Theorem 1.1 in [71] and Theorem 2.2 in [61]). *Assume that K is globally Lipschitz and f_0 is a Borel probability measure on $\mathbb{R}^d$ with finite second* *moment. Then there is existence and uniqueness of the solutions to* (17) *as well as to* (4).

We refer to [71] for a detailed proof in the case where K is also bounded and to [61] for a proof in the general case. We remark that the existence and uniqueness hold both trajectory-wise and in law for (17).

With Theorem 2 in mind, as long as we can show that the systems (3) and (16) are close, we will obtain the mean field limit. This is provided by

Theorem 3 (Theorem 2.3 in [61] and Theorem 1.4 in [71]). *Assume that K is globally Lipschitz and f_0 is a Borel probability measure on $\mathbb{R}^d$ with finite second moment. Then for any $i = 1, \cdots, N$, one has*

$$\mathbb{E}\left(\sup_{0\le t\le T} |X_i(t) - \bar{X}_i(t)|^2\right) \le \frac{C}{N}, \tag{18}$$

where C is independent of N, but depends on the time interval T and the Lipschitz constant $\|K\|_{Lip}$.

We omit the proof and instead remark that Theorem 2 obviously implies the mean field limit or propagation of chaos. Recall that the $p-$MKW distance between two probability measures μ and ν with finite $p-$th moments is defined by

$$W_p(\mu,\nu) = \inf_{(X,Y)} \left(\mathbb{E}|X-Y|^p\right)^{\frac{1}{p}},$$

where the infimum runs over all all couples of random variables (X,Y) with Law$X = \mu$ and Law$Y = \nu$ (see for instance [73]).

From Theorem 3, one obtains an estimate on the distance between the $1-$marginal F_1^N and f_t as $N \to \infty$,

$$W_2^2(F_1^N, f_t) \le \mathbb{E}|X_i(t) - \bar{X}_i(t)|^2 \le \frac{C}{N}.$$

More generally, one has a quantitative version of propagation of chaos. For any fixed k, the $k-$marginal distribution F_k^N converges to $f^{\otimes k}$ as N goes to infinity

$$W_2^2(F_k^N(t), (f_t)^{\otimes k}) \le \mathbb{E}\left|(X_1(t),\cdots,X_k(t)) - (\bar{X}_1(t),\cdots,\bar{X}_k(t))\right|^2 \le \frac{kC}{N}.$$

Similarly, we can obtain the convergence in law of the empirical measure towards the limit f_t. Indeed, for a test function $\phi \in C_b^1(\mathbb{R}^d)$, one has

$$\begin{aligned}
&\mathbb{E}\left|\frac{1}{N}\sum_{i=1}^N \phi(X_i(t)) - \int_{\mathbb{R}^d} \phi(x) f_t(x)\,\mathrm{d}x\right|^2 \\
&\le 2\mathbb{E}|\phi(X_1(t)) - \phi(\bar{X}_1(t))|^2 + 2\mathbb{E}\left|\frac{1}{N}\sum_{i=1}^N \phi(\bar{X}_i(t)) - \int_{\mathbb{R}^d} \phi(x) f_t(x)\,\mathrm{d}x\right|^2 \le \frac{C}{N}\|\phi\|_{C^1}.
\end{aligned} \tag{19}$$

3.2 Large Deviation and time uniform estimates

The results referred to here were mostly obtained in [9, 10, 13, 57, 58]. The average estimates for instance (18) and (19) guarantee that the particle system (1) or (3) is an approximation to the mean field PDE (2) or (4) respectively for a fixed time interval.

However, uniform in time estimates are sometimes necessary: To make sure that the equilibrium for the discrete system is close to the equilibrium of the continuous model for instance.

Furthermore, it can be critical to be able to estimate precisely how likely any given instance of the discrete system is to be far from the limit. When large stochastic particle systems are used to do numerical simulations, one may want to make sure that the numerical method has a very small probability to give wrong results. Those are usually called concentration estimates.

Unfortunately the bound (18) can only give very weak concentration estimates by Chebyshev's inequality, for instance

$$\mathbb{P}\left\{|X_i(t)-\bar{X}_i(t)| \geq \frac{\sqrt{C}L}{\sqrt{N}}\right\} \leq \frac{1}{L^2}. \tag{20}$$

Uniform in time estimates in particular cannot hold for any system. For this reason we consider here a particular variant of 1st order particle system (3), namely System (6) with $\sigma = 1$ and with V and W convex. In that case the mean field equation corresponding to system (6) is

$$\partial_t f = \Delta_x f + \mathrm{div}_x(f\,\nabla_x(V + W\star f)). \tag{21}$$

One similarly constructs a symmetric coupling system

$$\mathrm{d}\bar{X}_i = -\nabla W \star f_t(\bar{X}_i) - \nabla V(\bar{X}_i) + \sqrt{2}\,\mathrm{d}W_t^i, \quad \bar{X}_i(0) = X_i^0, \tag{22}$$

where $f_t = \mathrm{Law}(\bar{X}_i(t))$ and is hence the solution to (21) with initial data f_0. Then one can obtain the following theorem

Theorem 4 (Theorem 1.2, Theorem 1.3 and Proposition 3.22 in [57]). *Assume that the interaction potential W is convex, even and with polynomial growth and V is uniformly convex* i.e. $D^2V(x) \geq \beta I$ *for some* $\beta > 0$. *Assume in addition that initially,*

$$X_i(0) = X_i^0, \quad i.i.d., \quad Law X_i^0 = f_0,$$

where f_0 is smooth. Then there exists a constant C such that for any $N \geq 2$,

$$\sup_{t>0} \mathbb{E}\left(|X_i(t) - \bar{X}_i(t)|^2\right) \leq \frac{C}{N}, \tag{23}$$

and for any $\varepsilon > 0$,

$$\sup_{\|\phi\|_{Lip}\le 1} \mathbb{P}\left[\left|\frac{1}{N}\sum_{i=1}^{N}\phi(X_i(t)) - \int \phi(x) f_t(x)\,\mathrm{d}x\right| > \frac{C}{\sqrt{N}} + \varepsilon\right] \le 2\exp(-\frac{\beta}{2}N\varepsilon^2). \quad (24)$$

Compare the exponential control on the tail in (24) to the polynomial estimate in (20). This large deviation type estimate (24) for the empirical measure is obtained by the use of logarithmic Sobolev inequalities.

Under the above convexity assumptions on the potentials V and W, the solution f_t to the granular media equation (21) converges to a unique equilibrium exponentially fast. It is in this context that a uniform in time estimate can be expected.

In [10], a stronger version of (24) was achieved but at the same time more restrictions were imposed on V and W and the initial law f_0. At least if W and V do not grow too fast, the following theorem holds

Theorem 5 (Theorem 2.9 in [10]). *Assume that V and W are both uniformly convex and have appropriate growth at infinity. Assume that initially,*

$$X_i(0) = X_i^0, \quad i.i.d., \quad Law X_i^0 = f_0,$$

for a smooth f_0 with a finite square exponential moment, i.e. there exists $\alpha_0 > 0$, such that

$$\int \exp(\alpha_0 |x|^2) f_0(x)\,\mathrm{d}x < \infty.$$

One has for any $T > 0$, there exists a constant $C = C(T)$ such that for any $d' > d$, there exist some constants N_0 and C' such that for all $\varepsilon > 0$, if $N \ge N_0 \max(\varepsilon^{-(d'+2)}, 1)$, then

$$\mathbb{P}\left[\sup_{0\le t\le T} W_1(\mu_N(t), f_t) > \varepsilon\right] \le C'(1 + T\varepsilon^{-2})\exp(-CN\varepsilon^2). \quad (25)$$

In the above theorem W_1 denotes the 1 MKW distance. While the constants are now time dependent, the result is more precise. It is even possible to estimate the deviation on the empirical measure on pairs of particles, so that

$$\mu_N^2(t) = \frac{1}{N(N-1)}\sum_{i\neq j}\delta_{(X_i(t), X_j(t))}$$

is close to $f_t^{\otimes 2}$ in the sense of (25). See Theorem 2.10 in [10] and also Theorem 2.12 there for uniform in time estimates in the spirit of (25).

3.3 Singular kernels: Stochastic vortex model leading to 2D Navier-Stokes equation

In this subsection, we review some results on the mean field limit for stochastic systems with singular kernels K: K is smooth on $\mathbb{R}^d \setminus \{0\}$, but in general $|K(x)| \to \infty$

when $x \to 0$. Therefore when two particles are close the interaction between them becomes extremely large.

As mentioned in the introduction, the 2nd order systems are generally harder, as diffusion is now degenerate. For the Poisson kernel and particles interacting under gravitational force or Coulomb force, both the deterministic case (leading to Vlasov equation) and the stochastic case (leading to McKean-Vlasov equation) are still open.

However the 1st order systems are usually easier. There are several mean field limit results for 1st order systems with

$$K(x) \sim \frac{1}{|x|^{\alpha}}, \quad \text{when } x \sim 0$$

for some $\alpha > 0$. We refer for instance [30, 34] for Keller-Segel equation and [56] for random particle blob method also for Keller-Segel equation. For other singular kernels, one can also see [22] about the Coulomb gas model in 1D, where the singularity is repulsive and strong, behaving like $\frac{1}{|x|}$.

We focus in the rest of this subsection on the Biot-Savart kernel K which leads to the vortex model in Fluid Mechanics in dimension 2. This case is now better understood, thanks for instance to [59, 62, 67] and more recently [29]. The results presented here are mainly based on [67] and [29].

In general, the stochastic or viscous vortex model is written as

$$\mathrm{d}X_i = \frac{1}{N}\sum_{j\neq i} \alpha_j K(X_i - X_j)\,\mathrm{d}t + \sqrt{2\sigma}\,\mathrm{d}W_t^i, \tag{26}$$

and the empirical measure is defined as in (8) to more easily allow for a negative vorticity. However, for simplicity we here assume that all $\alpha_i \equiv 1$ and hence (26) reduces to our classical 1st order system (3) with

$$K(x) = \frac{1}{2\pi}\left(\frac{-x_2}{|x|^2}, \frac{x_1}{|x|^2}\right).$$

The expected mean field PDE given by (4) is now the 2D Navier-Stokes equation in vorticity formulation with positive viscosity σ.

It is not initially obvious that there even exists solutions for a fixed N because of the singularity in K. In [66] (see also Theorem 2.10 in [29]), it was showed that almost surely for all $t \geq 0$, and all $i \neq j$, $X_i(t) \neq X_j(t)$. Hence the system (26) is well-posed since the singularity of K is almost surely never visited.

The main result from [67] is the propagation of chaos

Theorem 6. *Assume that $(F^N(0))_{N\geq 2}$ is f_0-chaotic and*

$$\limsup_{N\to\infty} \left\| \int_{(\mathbb{R}^2)^{N-i}} F^N(0)\,\mathrm{d}x_{i+1}\cdots\mathrm{d}x_N \right\|_{L^\infty((\mathbb{R})^i)} < \infty, \quad \textit{for } i = 1, 2. \tag{27}$$

Then there exists $\sigma_0 > 0$ such that $(F^N(t))_{N\geq 2}$ is f_t-chaotic for any $\sigma > \sigma_0$. In an equivalent way, we have $\mu_N(t) \to f(t)$ in law as $N \to \infty$ for $\sigma > \sigma_0$.

More recently, the above result was improved in [29]: No assumption is required on the viscosity and the initial vorticity f_0 belongs to $L^1(\mathbb{R}^2)$ while in [67] it is essentially required that $f_0 \in L^\infty$. We state a simplified version of the main theorem from [29]

Theorem 7 (Theorem 2.12 and Theorem 2.13 in [29]). *Consider any $f_0 \geq 0$, an initial data for* (4) *with*

$$\int_{\mathbb{R}^2} f_0 \left(1 + |x|^k + |\log f_0|\right) \mathrm{d}x < \infty, \quad \textit{for some } k \in (0,2).$$

Assume that for $N \geq 2$, the law $F^N(0)$ of the initial distribution of particles is f_0-chaotic and

$$\sup_{N\geq 2} \frac{1}{N} \int_{(\mathbb{R}^2)^N} \left(1 + |X|^2\right)^{\frac{k}{2}} F^N(0,X)\,\mathrm{d}X < \infty,$$
$$\sup_{N\geq 2} H_N(F^N(0)) < \infty.$$

Then both the particle system (26) *and the 2D Navier-Stokes Eq.* (4) *are globally well-posed and $(F^N(t))_{N\geq 2}$ is f_t-chaotic.*

If we assume furthermore that initially $(F^N(0))$ is f_0- entropy chaotic as in Definition 5, then $(F^N(t))_{N\geq 2}$ is also f_t- entropy chaotic and $F_k^N \to f_t^{\otimes k}$ strongly in $L^1((\mathbb{R}^2)^k)$.

The proof of Theorem 7 follows the classical tightness/consistency/uniqueness arguments. The dissipation of the entropy

$$H_N(F^N) + \sigma \int_0^t I_N(F^N(s))\,\mathrm{d}s = H_N(F^N(0)),$$

and a control on a moment of $F^N(t)$ will give us a bound on

$$\int_0^T I_N(F^N(t))\,\mathrm{d}t.$$

This will in turn essentially bound several quantities for instance

$$\mathbb{E}\left[\sup_{0<s<t<T} \frac{|X_i(t) - X_i(s)|}{|t-s|^\alpha}\right],$$

which finally helps to complete the tightness/consistency argument. The uniqueness argument is based on [2].

3.4 Other extensions

We only very briefly mention some other recent extensions to the classical theory.

The first such case concerns kernels which are only locally Lipschitz. Such models have been studied in the context of flocking models, in [7] for example, and models of neuron dynamics, see [11] for instance where the model is even more general with the diffusion coefficients possibly depending on the law.

Those models include interaction terms that are all locally Lipschitz but with a Lipschitz constant which grows to infinity when the region considered grows to the whole $\mathbb{R}^d$.

The classical method has to be adapted with typically a faster decay assumed on the limit law f_t. One key ingredient in the proof is then to show that trajectories cannot escape to infinity, typically because the model includes confining forces. In the absence of such assumptions, the problem can become ill-posed as shown in [70].

Finally we mention that in the coming article [44], a new coupling strategy and a Glivenko-Cantelli theorem are used to show the mean field limits for systems (1) or (3) with global Hölder continuous interaction kernels $K \in C^{0,\alpha}$. For 1st order system, $\alpha > 0$ is enough. But it requires $\alpha > \frac{1}{3}$ for 2nd systems in order to ensure the existence of a differentiable stochastic flow.

4 A new statistical approach

In the authors' recent articles [47] and [48], we proposed a new relative entropy method to deal with mean field limit for very rough interaction kernels K.

The idea is to directly compare the distance between the joint distribution $F^N(t)$ solving the Liouville equation (12) and the tensor product of the limit law $f_t^{\otimes N}$ through the relative entropy

$$H_N(F^N|f^{\otimes N}) = \frac{1}{N}\int_{E^N} F^N \log\frac{F^N}{f^{\otimes N}}\,\mathrm{d}z_1\cdots\mathrm{d}z_N.$$

The main theorem for the 2nd order systems with $\sigma > 0$ in [47] can be stated simply as follows

Theorem 8 (Main Theorem in [47]). *Assume that $K \in L^\infty$ and that there exists $f_t \in L^\infty([0,T],\ L^1(E)\cap W^{1,p}(E))$ for every $1 \le p \le \infty$ which solves the limiting equation* (2) *with in addition*

$$\theta_f = \sup_{t\in[0,\,T]}\int_{\Gamma\times\mathbb{R}^d} e^{\lambda_f\,|\nabla_v \log f_t|}\, f_t\,\mathrm{d}x\,\mathrm{d}v < \infty,$$

for some $\lambda_f > 0$. Furthermore assume initially that

$$\sup_{N\geq 2} H_N(F^N(0)) < \infty, \quad H_N(F^N(0)|f_0^{\otimes N}) \to 0, \ \textit{as}\ N \to \infty.$$

and

$$\sup_{N\geq 2} \frac{1}{N} \int_{E^N} \sum_{i=1}^N \left(1+|z_i|^2\right) F^N(0, z_1, \cdots, z_N)\, \mathrm{d}z_1 \cdots \mathrm{d}z_N < \infty.$$

Then there exists a universal constant $C > 0$ s.t. for any $t \in [0,T]$,

$$H_N(F^N(t)|f_t^{\otimes N}) \leq e^{C\,\|K\|_{L^\infty}\,\theta_f\, t/\lambda_f} \left(H_N(F^N(0)|f_0^{\otimes N}) + \frac{C}{N}\right).$$

This results implies a strong form of propagation of chaos as the $k-$marginal converges to $f^{\otimes k}$ in L^1. Indeed if for instance $H_N(F^N(0)|f_0^{\otimes N}) \lesssim N^{-1}$ then the classical Csiszár-Kullback-Pinsker inequality (see chapter 22 in [73] for instance) implies that

$$\|F_k^N(t) - f_t^{\otimes k}\|_{L^1} \leq \sqrt{2kH_k(F_k^N(t)|f_t^{\otimes k})} \lesssim \frac{1}{\sqrt{N}}.$$

The argument is in essence a weak-strong estimate comparing a very smooth solution to the limiting equation with a weak solution F_t^N to the Liouville equation (12). The heart of the proof consists of precise combinatorics estimates which lead to a new type of law of large numbers.

In a coming article [48], we extend the result to the 1st order system (3) with $K \in W^{-1,\infty}$, *i.e.* K is the derivative of a bounded function but with the restriction that $\mathrm{div}_x K \in L^\infty$. By a careful truncation of the Biot-Savart kernel K and repeating our procedure, we can also provide an explicit convergence rate for stochastic vortex model (26) approximating 2D Navier-Stokes equation.

References

1. Ahn, S. M., Ha, S.-Y.: Stochastic flocking dynamics of the Cucker-Smale model with multiplicative white noise. J. Math. Physics, **51**, 103301 (2010)
2. Ben-Artzi, M.: Global solutions of two-dimensional Navier-Stokes and Euler equations. Arch. Rational Mech. Anal. **128**, 329-358 (1994)
3. Benedetto, D., Caglioti, E., Carrillo, J.A., Pulvirenti, M.: A non Maxwellian steady distribution for one-dimensional granular media. J. Stat. Phys. **91**, 979-990 (1998)
4. Bodineau, T., Gallagher, I., Saint-Raymond, L.: The Brownian motion as the limit of a deterministic system of hard-spheres. Invent. Math. **203**, 493-553 (2016)
5. T. Bodineau, T., Gallagher, I., Saint-Raymond, L.: From hard spheres dynamics to the Stokes-Fourier equations: an L2 analysis of the Boltzmann-Grad limit. C. R. Math. Acad. Sci. Paris **353**, 623-627 (2015)
6. Bogoliubov, N. N.: Kinetic equations. Journal of Physics USSR **10**, 265-274 (1946)
7. Bolley, F., Cañizo, J. A., Carrillo, J. A.: Stochastic mean-field limit: non-Lipschitz forces and swarming. Math. Mod. Meth. App. S. **21**, 2179-2210 (2011)
8. Bolley, F., Cañizo, J. A., Carrillo, J. A.: Mean-field limit for the stochastic Vicsek model. Appl. Math. Lett. **25**, 339-343 (2012)

9. Bolley, F., Guillin, A., Malrieu, F.: Trend to equilibrium and particle approximation for a weakly self-consistent Vlasov-Fokker-Planck equation. Math. Model. Numer. Anal. **44**, 867-884 (2010)
10. Bolley, F., Guillin, A., Villani, C.: Quantitative concentration inequalities for empirical measures on non-compact space. Probab. Theory Relat. Fields **137**, 541-593 (2007)
11. Bossy, M., Faugeras, O., Talay, D.: Clarification and complement to "Mean-field description and propagation of chaos in networks of Hodgkin-Huxley and FitzHugh-Nagumo neurons". J. Math. Neurosci. **5** Art. 19, 23 pp. (2015)
12. Bossy, M., Jabir, J. F., Talay, D.: On conditional McKean Lagrangian stochastic models. Probab. Theory Relat. Fields **151**, 319-351 (2011)
13. Cattiaux, P., Guillin, A., Malrieu, F.: Probabilistic approach for granular media equations in the non-uniformly convex case. Probab. Theory Relat. Fields **140**, 19-40 (2008)
14. Caglioti, E., Lions, P.L., Marchioro, C., Pulvirenti, M.: A special class of two-dimensional Euler Equations: A statistical mechanics description. Commun. Math. Phys. **143**, 501-525 (1992)
15. Caglioti, E., Lions, P.L., Marchioro, C., Pulvirenti, M.: A special class of two-dimensional Euler Equations: A statistical mechanics description. Part II. Commun. Math. Phys. **174**, 229-260 (1995)
16. Carlen, E.A., Carvalho, M.C., Le Roux, J., Loss, M., Villani, C.: Entropy and chaos in the Kac model. Kinet. Relat. Models **3**, 85-122 (2010)
17. Carrillo, J. A., Choi, Y.-P., Hauray, M.: The derivation of swarming models: Mean Field limit and Wasserstein distances. In: Collective Dynamics from Bacteria to Crowds, volume 553 of CISM International Centre for Mechanical Sciences, pages 1-46. Springer Vienna, (2014)
18. Carrillo, J. A., Fornasier, M., Toscani, G., Vecil, F.: Particle, kinetic, and hydrodynamic models of swarming. In: Mathematical modeling of collective behavior in socio-economic and life sciences. pp. 297-336, Model. Simul. Sci. Eng. Technol., Birkhauser Boston, Inc., Boston, MA, (2010)
19. Carrillo, J. A., DiFrancesco, M., Figalli, A., Laurent, T., Slepcev, D.: Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations. Duke Math. J. **156**, 229-271 (2011)
20. Carrillo, J. A., Lisini, S., Mainini, E.: Gradient flows for non-smooth interaction potentials. Nonlinear Anal. **100** 122-147 (2014)
21. Chuang, Y.L., Huang, Y.R., D'Orsogna, M.R., Bertozzi, A.L.: Multi-vehicle flocking: scalability of cooperative control algorithms using pairwise potentials. IEEE Int. Conf. Robotics. Automation, 2292-2299 (2007)
22. Cépa, E., Lépingle, D.: Diffusing particles with electrostatic repulsion. Probab. Theory Relat. Fields **107**, 429-449 (1997)
23. Cercignani,C., Illner, R., Pulvirenti, M.: The Mathematical Theory of Dilute Gases. Springer-Verlag, New York, (1994)
24. Cucker, F., Smale, S.: Emergent behavior in flocks. IEEE Trans. Automat. Control **52**, 852-862 (2007)
25. Degond, P., Frouvelle, A., Liu, J.-G.: Macroscopic limits and phase transition in a system of self-propelled particles. J. Nonlinear Sci. **23** 427-456 (2013)
26. Desvillettes, L., Graham, C., Méléard, S.: Probabilistic interpretation and numerical approximation of a Kac equation without cutoff. Stochastic Process. Appl. **84**, 115-135 (1999)
27. Dirr, N., Stamatakis, M., Zimmer, J.: Entropic and gradient flow formulations for nonlinear diffusion. ArXiv: 1508.00549 (2016)
28. Flandolia, F., Gubinellib, M., Priolac, E.: Full well-posedness of point vortex dynamics corresponding to stochastic 2D Euler equations. Stoch. Process. Appl. **121**, 1445-1463 (2011)
29. Fournier, N., Hauray, M., Mischler, S.: Propagation of chaos for the 2d viscous vortex model. J. Eur. Math. Soc. **16**, 1425-1466 (2014)
30. Fournier, N., Jourdain, B.: Stochastic particle approximation of the Keller-Segel Equation and two-dimensional generalization of Bessel process. ArXiv: 1507.01087 (2015)
31. Gallagher, I., Saint-Raymond, L., Texier, B.: From newton to Boltzmann : hard spheres and short-range potentials. In: Zurich Advanced Lectures in Mathematics Series, (2014)

32. Gibbs, J. W.: On the Fundamental Formulae of Dynamics. Amer. J. Math. **2** 49-64 (1879)
33. Gibbs, J. W.: Elementary principles in statistical mechanics: developed with especial reference to the rational foundation of thermodynamics. Dover publications, Inc., New York, (1960)
34. Godinho, D., Quininao, C.: Propagation of chaos for a sub-critical Keller-Segel Model. Ann. Inst. H. Poincaré Probab. Statist. **51**, 965-992 (2015)
35. Golse, F.: On the dynamics of large particle systems in the mean field limit. In: Macroscopic and Large Scale Phenomena: Coarse Graining, Mean Field Limits and Ergodicity. Volume 3 of the series Lecture Notes in Applied Mathematics and Mechanics, pp. 1-144. Springer, (2016)
36. Grad, H.: On the kinetic theory of rarefied gases. Comm. on Pure and Appl. Math. **2**, 331-407 (1949)
37. Graham, C., Méléard, S.: Stochastic particle approximations for generalized Boltzmann models and convergence estimates. Ann. Probab. **25**, 115-132 (1997)
38. Graham, C., Méléard, S.: Existence and regularity of a solution of a Kac equation without cutoff using the stochastic calculus of variations. Comm. Math. Phys. **205**, 551-569 (1999)
39. Ha, S.-Y., Lee, K., Levy, D.: Emergence of time-asymptotic flocking in a stochastic Cucker-Smale system. Commun. Math. Sci. **7**, 453-469 (2009)
40. Ha, S.-Y., Tadmor, E.: From particle to kinetic and hydrodynamic description of flocking. Kinet. Relat. Models **1**, 415-435 (2008)
41. Hauray, M., Mischler, S.: On Kac's chaos and related problems. J. Funct. Anal. **266**, 6055-6157 (2014)
42. Hauray, M., Jabin, P. E.: Particle Approximation of Vlasov Equations with Singular Forces. Ann. Scient. Ecole Norm. Sup. **48** 891-940 (2015)
43. Hegselmann, R., Krause, U.: Opinion dynamics and bounded confidence models, analysis, and simulation. Journal of Artifical Societies and Social Simulation (JASSS). **5**, no. 3, (2002)
44. Holding, T.: Propagation of chaos for Hölder continuous interaction kernels via Glivenko-Cantelli. Manuscript, Personal communication.
45. K. Itô: On stochastic differential equations. Memoirs of the American Mathematical Society. **4**, 1-51 (1951)
46. Jabin, P.E.: A review for the mean field limit for Vlasov equations. Kinet. Relat. Models **7**, 661-711 (2014)
47. Jabin, P.E., Wang, Z.: Mean field limit and propagation of chaos for Vlasov systems with bounded forces. ArXiv: 1511.03769 (2015)
48. Jabin, P.E., Wang, Z.: Mean filed limit for stochastic 1st order systems with almost bounded stream functions. In preparation.
49. Jeans, J. H.: On the theory of star-streaming and the structure of the universe. Monthly Notices of the Royal Astronomical Society **76**, 70-84 (1915)
50. Kac, M.: Foundations of kinetic theory. In: Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954-1955, Vol. III, pp. 171-197. University of California Press, Berkeley (1956)
51. Kipnis, C., Landim, C.: Scaling limit of interacting particle systems. In: Grundlehren der mathematischen Wissenschaften 320. Springer, (1999)
52. Krause, U.: A discrete nonlinear and non-autonomous model of consensus formation. Communications in difference equations, pp. 227-236 (2000)
53. Lanford, O. E. III: Time evolution of large classical systems. In Dynamical systems, theory and applications (Recontres, Battelle Res. Inst., Seattle, Wash., 1974), pp 1-111. Lecture Notes in Phys., Vol. 38. Springer, Berlin, (1975)
54. Lazarovici, D: The Vlasov-Poisson dynamics as the mean-field limit of rigid charges. ArXiv:1502.07047 (2015)
55. Lazarovici, D., Pickl, P.: A Mean-field limit for the Vlasov-Poisson system. ArXiv 1502.04608 (2015)
56. Liu, J.-G., Yang, R.: A random particle blob method for the Keller-Segel equation and convergence analysis. Math. Comp., to appear.

57. Malrieu, F.: Logarithmic Sobolev inequalities for some nonlinear PDE's. Stoch. Process. Appl. **95**, 109-132 (2001)
58. Malrieu, F.: Convergence to equilibrium for granular media equations and their Euler schemes. Ann. Appl. Probab. **13**, 540-560 (2003)
59. Marchioro, C., Pulvirenti, M.: Hydrodynamics in two dimensions and vortex theory. Commun. Math. Phys. **84**, 483-503 (1982)
60. McKean, H.P. Jr.: Propagation of chaos for a class of non-linear parabolic equations. In: Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967), pp. 41-57. Air Force Office Sci. Res., Arlington, VA, (1967)
61. Méléard, S.: Asymptotic behavior of some interacting particle systems; McKean-Vlasov and Boltzmann models. In: Probabilistic Models for Nonlinear Partial Differential Equations (Lecture Notes in Mathematics), Vol. 1627, Springer, (1996)
62. Méléard, S.: Monte-Carlo approximation for 2d Navier-Stokes equations with measure initial data. Probab. Theory Relat. Fields **121**, 367-388 (2001)
63. Mischler, S., Mouhot, C.: Kac's Program in Kinetic Theory. Invent. Math. **193**, 1-147 (2013)
64. Mischler, S., Mouhot, C., Wennberg, B.: A new approach to quantitative chaos propagation for drift, diffusion and jump process. Probab. Theory Relat. Fields **161**, 1-59 (2015)
65. Motsch, S., Tadmor, E.: A new model for self-organized dynamics and its flocking behavior. *J. Stat. Phys.*, **144**, 923-947 (2011)
66. Osada, H.: A stochastic differential equation arising from the vortex problem. Proc. Japan Acad. Ser. A Math. Sci. **62**, 333-336 (1986)
67. Osada, H.: Propagation of chaos for the two-dimensional Navier-Stokes equation. In: Probabilistic methods in mathematical physics (Katata/Kyoto, 1985), pp. 303-334. Academic Press, Boston, MA, (1987)
68. Othmer, H. G., Stevens, A.: Aggregation, blowup, and collapse: the ABCs of taxis in reinforced random walks. SIAM J. Appl. Math. **57**, 1044-1081 (1997)
69. Perthame, B.: Transport equations in biology. Frontiers in Mathematics. Birkhäuser Verlag, Basel, (2007)
70. Scheutzow, M: Uniqueness and non-uniqueness of solutions of Vlasov-McKean equations. J. Austral. Math. Soc. Series A, **43**, 246-256 (1987)
71. Sznitman, A.-S.: Topics in propagation of chaos. In: Ecole d'été de probabilités de Saint-Flour XIX-1989, pp. 165-251. Springer, (1991)
72. Topaz, C. M., Bertozzi, A. L., Lewis, M. A.: A nonlocal continuum model for biological aggregation. Bull. Math. Biol. **68**, 1601-1623 (2006)
73. Villani, C.: Optimal Transport, Old and New. In: Grundlehren der mathematischen Wissenschaften 338. Springer Science & Business Media, (2008)
74. Vicsek, T., Czirok, E., Ben-Jacob, E., Cohen, I., Shochet., O.: Novel type of phase transition in a system of self-driven particles. Phys. Rev. Lett. **75**, 1226-1229 (1995)
75. Vlasov, A. A.: On vibration properties of electron gas. *J. Exp. Theor. Phys.* (in Russian), **8** (3):291, (1938)
76. Vlasov, A. A.: The vibrational properties of an electron gas. Sov. Phys. Usp. **10**, 721-733 (1968)
77. Xia, H., Wang, H., Xuan, Z.: Opinion dynamics: A multidisciplinary review and perspective on future research. International Journal of Knowledge and Systems Science (IJKSS) **2**, 72-91 (2011)