

COMPLEXITY REDUCTION IN MANY PARTICLES SYSTEMS WITH RANDOM INITIAL DATA*

LEONID BERLYAND[†], PIERRE-EMMANUEL JABIN[‡], AND MYKHAILO POTOMKIN[§]

Abstract. We consider the motion of interacting particles governed by a coupled system of ODEs with random initial conditions. Direct computations for such systems are prohibitively expensive due to a very large number of particles and randomness requiring many realizations in their locations in the presence of strong interactions. While there are several approaches that address the above difficulties, none addresses all three simultaneously. Our goal is to develop such a computational approach in order to capture the experimentally observed emergence of correlations in the collective state (patterns due to strong interactions). Our approach is based on the truncation of the BBGKY hierarchy that allows one to go beyond the classical Mean Field limit and capture correlations while drastically reducing the computational complexity. Our theoretical approach is illustrated by numerical examples. The numerical simulations in these examples are highly non-trivial due to the combination of non-locality and non-linearity in the PDE resulting from the truncation.

Key words. Mean Field, correlation, systems of a large number of particles.

1. Motivation and Settings. Systems of interacting particles described by a coupled system of a large number of ODEs with random initial conditions appear in many problems of physics, cosmology, chemistry, biology, social science and economics:

$$(1.1) \quad \dot{X}_i = F(X_i) + \frac{\alpha}{N} \sum_{j=1}^N K(X_j - X_i), \quad i = 1, \dots, N.$$

Here $X_i(t)$ denotes the position of i th particle and X_i belongs to D , where D throughout this paper can stand for \mathbb{R}^d , a d -dimensional torus Π^d , or a compact domain in \mathbb{R}^d in which case boundary conditions must be added. The scalar function K describes the inter-particle interactions, and $F(X_i)$ models an internal force for each particle, such as self-propulsion.

System (1.1) is an Individual Based Model, i.e., it has an ODE for each particle coupled with others. In various applications the role of an individual can be played by atoms, social agents, bacteria in suspensions, animals in flocks, etc. The system (1.1) has two key parameters: α , the strength of interactions, and N , the number of particles. The parameter α is determined by both geometry such as interparticle distance and the *mass* of a particle (note that a model particle is just a point) as well as physics. In this paper we restrict ourselves to the case when the right hand side of (1.1) is linear in α . The magnitude of α plays an important role in analysis of the system (1.1): a small α corresponds to almost decoupled *interactions*; large α corresponds to strong interactions which is our main focus; $\alpha \sim 1$ corresponds to the classical Mean Field regime.

*The work of Leonid Berlyand and Mykhailo Potomkin was supported by DOE grant DE-FG-0208ER25862. The work of Pierre-Emmanuel Jabin was partially supported by NSF grant DMS-1312142. LB and MP wish to thank V. Rybalko for his comments and suggestions which helped to improve the manuscript.

[†]Department of Mathematics, The Pennsylvania State University, University Park, PA, USA, 16802, (berlyand@math.psu.edu).

[‡]Department of Mathematics, University of Maryland, College Park, MD 20742 USA, (pjabin@umd.edu)

[§]Department of Mathematics, The Pennsylvania State University, University Park, PA, USA, 16802, (potomkin@math.psu.edu)

Our work is motivated by experiments in bacterial suspensions [1, 2, 3, 4, 5, 6, 7]. These experiments [3, 4, 7] show the emergence of a coarse *collective scale* when the concentration of bacteria exceeds a critical value. Roughly speaking, the collective scale is the correlation length of the velocity field in a bacterial suspension. A striking universality property has been observed experimentally and numerically in [3, 4, 8] the collective scale does not change when swimming speed and concentration increase (that is more energy is injected into the system).

The motion of bacteria can be modeled by a system of the form (1.1) where the position and orientation of the i th bacteria are described by the vector $X_i(t)$. In this case the parameter α equals $NV_0(\ell/R)^2$, where V_0 is the swimming speed of a single bacterium, R is the mean distance between two bacteria, and ℓ is the characteristic size of a bacterium. The collective behavior observed in experiments [3, 4, 7] has also been qualitatively reproduced by direct numerical simulations in [8], which validates the model of the type (1.1). However, the computational cost of direct simulations of the ODE system is prohibitively high for the following reasons:

- (i) the number of bacteria N is very large (10^{10} per cm^3);
- (ii) to draw a reliable conclusion one needs to consider many realizations, which mathematically translates into random initial data;
- (iii) the main interest is in collective state corresponding to large α which leads to small time steps.

The combination of the factors (i)-(iii) makes the computational cost too high even for the most powerful particle methods such as Fast Multipole Method [9, 10, 11].

The goal of this paper is to propose a computational approach that allows one to describe numerically the collective state of this system with properties (i)-(iii). More specifically, the collective state is described by the correlation length and two-point correlation function. The objective of our study is the efficient computation of these two quantities.

The main idea is to replace the ODE system (1.1) by a PDE such that the computational cost of its solution does not grow as N goes to infinity. This idea had been used in the classical Mean Field approach which corresponds to α of the order 1 and is not valid for strong interactions, e.g., $\alpha \sim N^\gamma$ for $0 < \gamma < 1$.

This paper focuses on a PDE approach that extends beyond the Mean Field, so that it can capture correlations in a computationally efficient way such that the computational complexity *increases only slowly* with N . The main idea is to consider the BBGKY hierarchy of PDEs which consists of N equations (therefore it is even harder to solve than (1.1)) and obtain a closed system for 1-particle and 2-particle distributions by a clever truncation of the hierarchy. Then the large parameter N is present only in coefficients in a more innocuous way, *and they can be handled efficiently with high order methods*. This approach computes distribution functions and therefore it avoids computing individual realizations. Thus, it allows us to overcome the computational difficulties (i) and (ii). The contribution to the computational complexity from difficulty (iii) is *still present but much less of a problem* than (i) and (ii), because $\alpha \sim N^\gamma$ and $\gamma < 1$.

Note that a specific feature of our method is that it is efficient for random initial conditions of system (1.1) in contrast to deterministic. Indeed, a seemingly simpler case of deterministic initial data leads to a solution of the truncated BBGKY hierarchy with singular initial conditions (δ -functions), which is why the numerical cost of solving such a deterministic problem is very high. In contrast, random initial data in ODE (1.1) lead to *smooth* initial conditions in the truncated BBGKY hierarchy that

is much easier to handle numerically.

The truncation presented in this paper can be applied for various ODE systems of type (1.1), *in dimension 1 or more*. Note that different truncations of the Boltzmann hierarchy have been made before for some specific situations, which usually rely on some perturbative arguments. For example, we refer to papers [12, 13, 14] devoted to Ostwald ripening where a truncation was motivated by expansions in (small) concentration of particles.

The paper is organized as follows. We recall the Mean Field approach and discuss its limitations in Section 2. The truncation of BBGKY hierarchy is described in Section 3. Numerical simulations performed to check that the truncated PDE system is reliable are described in Section 4.

2. Random initial conditions, correlations, the Mean Field approach.

For physical reason, the initial conditions for the system (1.1) are typically random as explained below. In the classical Mean Field theory, this leads to a drastic reduction in the computational complexity: it is possible to approximate the original solution by the solution of a PDE which does not depend on N . We describe here the two classical approaches *to derive the Mean Field limit*, a first one based on the so-called empirical measure and a statistical approach which is better suited for our purpose.

The Mean Field limit is valid as long as the correlations between particles are vanishing. This phenomenon is known as *propagation of chaos*. However, our work is motivated by experimental studies of the collective state, whose key feature is the rise of *correlations* corresponding to the emergence of a *collective scale*. In this case, as we will explain below, the Mean Field approach fails.

Our approach in this paper is mostly formal. Nevertheless, we point out that the Mean Field theory described below can be made rigorous if some smoothness is assumed on K . More precisely,

$$(2.1) \quad \nabla K \in L^\infty(D), \quad K(x) \rightarrow 0 \quad \text{as } |x| \rightarrow \infty.$$

On the other hand, we believe that the numerical implementation of this approach will work well even for singular kernels (see remark 2.1).

2.1. Preliminaries. *How to choose initial conditions: Randomness and marginals.*

By assumption (2.1) and the standard Cauchy-Lipschitz theory, there exists a unique solution to (1.1) once each initial position $X_i(0)$ is chosen. However for most practical purposes, determining those initial positions can be a very delicate problem as the full information is not accessible from an experimental point of view. For $N \sim 10^{10}$, it is indeed completely unrealistic to measure the position of each particle with enough precision.

Instead, *what is accessible is some* statistical information about the positions of the particles. Hence one usually assumes that the initial position of each particle is randomly distributed. That means that the information on the initial distribution of the particles is now encoded in the N -particle distribution function at time 0, $f_N(t = 0, x_1, \dots, x_N)$. Given a subdomain $\mathcal{Q} \subset D^N$, the probability of finding the initial positions $(X_1^0, \dots, X_N^0) \in \mathcal{Q}$ is given by

$$\int_{\mathcal{Q}} f_N(t = 0, x_1, \dots, x_N) dx_1 \dots dx_N.$$

System (1.1) is deterministic but if the initial conditions are random then the randomness will be propagated defining the N -particle distribution for $t > 0$. Technically, $f_N(t, \cdot)$ is the push forward of $f_N(t = 0, \cdot)$ by the flow generated by (1.1).

From f_N one may define the k -th marginal

$$f_k(t, x_1, \dots, x_k) = \int_{D^{N-k}} f_N(t, x_1, \dots, x_k, x_{k+1}, \dots, x_N) dx_{k+1} \dots dx_N.$$

Some of marginals have a natural physical interpretation. For instance, f_1 is the 1-particle distribution function and for $\mathcal{O} \subset D$ the average number of particles in the subset \mathcal{O} is

$$\int_{\mathcal{O}} f_1(t, x) dx.$$

It is still not experimentally possible to measure f_N but it is possible to measure some marginals, especially f_1 and the 2-particle distribution function f_2 .

In the simplest case, one assumes that the particles are initially independently and identically distributed, that is

$$(2.2) \quad f_N(t=0, x_1, \dots, x_N) = \prod_{i=1}^N f^0(x_i).$$

This independence is strongly connected to the usual Mean Field limit approach as explained in subsection 2.2 (see (2.13)).

Definition of correlations. Our main goal is to understand how correlations develop in system (1.1). Those are connected to the second marginal f_2 .

We define correlation of particles' positions by

$$(2.3) \quad c = \frac{\mathbb{E}[X_1 \cdot X_2] - (\mathbb{E}[X])^2}{\mathbb{E}[X^2] - (\mathbb{E}[X])^2} = \frac{\int x_1 \cdot x_2 f_2(x_1, x_2) dx_1 dx_2 - (\int x f_1(x) dx)^2}{\int x^2 f_1(x) dx - (\int x f_1(x) dx)^2}.$$

Observe that the correlation c can only vanish if

$$f_2(x, y) = f_1(x) f_1(y),$$

that is if the particles positions are independent. Therefore, roughly speaking, the correlations in the system are determined by how far $f_2(x, y)$ is from $f_1(x) f_1(y)$.

2.2. The Mean Field approach.

Empirical measure. Assume that the $X_i(t)$ are solutions to (1.1), and define the empirical measure

$$(2.4) \quad \mu_N(t, x) = \frac{1}{N} \sum_{i=1}^N \delta(x - X_i(t)).$$

Note that if the particles are undistinguishable then there is just as much information in the empirical measure as in the position vector (X_1, \dots, X_N) . Otherwise, it only tells that there is a particle at x , but it is not clear which one.

If K is continuous, then μ_N solves the Vlasov equation in the sense of distributions

$$(2.5) \quad \partial_t f(t, x) + \nabla_x \cdot (F(x) f(t, x)) + \alpha \nabla_x \cdot \left(\int K(y-x) f(t, y) dy f(t, x) \right) = 0.$$

Consider a sequence of initial positions $\mathcal{X}_N = \{X_i(0) : i = 1, \dots, N\}$ such that the corresponding empirical measure $\mu_N(0)$ converges to some $f^0 \in \Pi(D)$ as N goes to

infinity. Here $\Pi(D)$ is the space of *probability measures*, positive measures μ on D such that $\mu(D) = 1$. Then it is natural to expect that μ_N will also converge to the corresponding solution f to (2.5) with initial data f^0 . Assuming that f^0 is smooth then it is possible to compute numerically f and hence to get a good approximation to μ_N . This is the classical Mean Field limit theory which can be made quantitative.

Those quantitative estimates require some weak distances on the space of measures. These are classically the so-called Monge-Kantorovich-Wasserstein (MKW) distances. For our purpose it is enough to understand that they *correspond to* some appropriate distance between probability measures. For the sake of completeness we define these distances in Appendix A.

Now we give the main stability estimate behind the Mean Field limit. From [15], [16], and [17], it is possible to prove that if f and g are two measure-valued solutions to (2.5), then

$$(2.6) \quad W_p(f(t, \cdot), g(t, \cdot)) \leq e^{t \alpha \|\nabla K\|_{L^\infty}} W_p(f(0, \cdot), g(0, \cdot)),$$

where $W_p(\cdot, \cdot)$ is a p -Wasserstein or MKW distance between two measures. The inequality (2.6) is a Gronwall-type inequality. Note also that the inequality (2.6) applies for any initial conditions $f(0, \cdot)$ and $g(0, \cdot)$ which are not necessarily random.

The Mean Field limit. In our context, the initial conditions are random as it was explained before. In particular, the empirical measure at time $t = 0$ is itself random.

If the initial law is chosen according to (2.2), then a large deviation for the law of large numbers applies ($\mathbb{E}\mu_N = f^0$) and ensures that, in fact, the initial measure $\mu_N(t = 0)$ is very close to f^0 . More precisely, it is proved for example in Boissard [18, Appendix A, Proposition 1.2], that if f^0 is a nonnegative measure with compact support of diameter R , then for some constant C and positive coefficients γ_1 and γ_2 , depending only on the dimension of D and R

$$(2.7) \quad \mathbb{P} \left(W_1(\mu_N(t = 0), f^0) \geq \frac{C R}{N^{\gamma_1}} \right) \leq e^{-CN^{\gamma_2}}.$$

This says that with a *probability exponentially close* to 1, $\mu_N(t = 0)$ and f^0 are polynomially close in N . Denote by f the solution to (2.5) with f^0 as an initial data. By combining the deterministic stability (2.6) with a law of large numbers in the form of (2.7) we obtain that with probability larger than $(1 - e^{-CN^{\gamma_2}})$

$$(2.8) \quad W_1(\mu_N(t, \cdot), f(t, \cdot)) \leq \frac{C R}{N^{\gamma_1}} e^{t \alpha \|\nabla K\|_{L^\infty}}.$$

Why random initial conditions make computations much easier in the Mean Field framework? Looking for a solution of the Vlasov equation (2.5) in the form of a sum of N Dirac masses like μ_N is just as complex and computationally costly as solving the original ODE system (1.1).

However looking for smooth solutions to the Vlasov equation is comparatively much faster and obviously independent of N (provided the solution is independent of N). Since the initial distribution f^0 is usually assumed to be smooth, the corresponding solution f is smooth as well and computing f numerically is thus far easier than solving (1.1) while the cost is independent of N .

The key to the reduction in the computational complexity in this Mean Field approach is that one does not solve the original ODE system (1.1) but instead one

solves the Vlasov PDE for f . The previous inequality (2.8) implies that this f will be a good approximation of the original μ_N up to a time t of order

$$(2.9) \quad \frac{\log N}{\alpha \|\nabla K\|_{L^\infty}}.$$

Note that in certain circumstances, this time can be considerably extended to become polynomial in N . This usually requires a stable equilibrium to equation (2.5), see [19] for instance.

2.3. Propagation of chaos. It is possible to interpret the Mean Field limit in terms of the propagation of chaos on the marginals. For this, we introduce the hierarchy of equations on the marginals.

First, it is well-known that f_N solves the Liouville equation:

$$(2.10) \quad \partial_t f_N + \sum_{i=1}^N \partial_{x_i} (F(x_i) f_N) + \frac{\alpha}{N} \sum_{i=1}^N \partial_{x_i} \left(\sum_{j=1}^N K(x_j - x_i) f_N \right) = 0.$$

By integrating the equation for f_N , one obtains an equation satisfied by each f_k

$$(2.11) \quad \begin{aligned} \partial_t f_k + \sum_{i=1}^k \partial_{x_i} (F(x_i) f_k) + \frac{\alpha}{N} \sum_{i=1}^k \sum_{j=1}^k \partial_{x_i} (K(x_j - x_i) f_k) \\ + \frac{\alpha(N-k)}{N} \sum_{i=1}^k \partial_{x_i} \left(\int K(y - x_i) f_{k+1}(t, x_1, \dots, x_k, y) dy \right) = 0. \end{aligned}$$

For example, the PDE for f_1 is

$$(2.12) \quad \begin{aligned} \partial_t f_1(t, x_1) + \partial_{x_1} (F(x_1) f_1(t, x_1)) + \frac{\alpha K(0)}{N} \partial_{x_1} f_1(t, x_1) \\ + \alpha \frac{N-1}{N} \partial_{x_1} \left\{ \int K(y - x_1) f_2(t, x_1, y) dy \right\} = 0. \end{aligned}$$

By taking the formal limit $N \rightarrow \infty$ in the equation (2.12) and assuming the independence condition $f_2(t, x_1, x_2) = f_1(t, x_1) f_1(t, x_2)$, we get equation (2.5).

This leads us to the crucial concept of *propagation of chaos*. Under some mild conditions on the smoothness of K , for initial positions that are close to being independent (that is (2.2) is assumed) as $N \rightarrow \infty$ we have

$$(2.13) \quad f_k(t, x_1, \dots, x_k) \rightarrow \prod_{i=1}^k f(t, x_i),$$

where $f(t, x)$ solves the Mean Field equation (2.5).

Note that for a finite N , one cannot have equality in (2.13) and, in particular, $\prod_{i=1}^N f(t, x_i)$ cannot be a solution to (2.10). Hence, for a finite but large N and for initial conditions of the form $f_N|_{t=0} = \prod_{i=1}^N f_0(x_i)$, the particles' positions are not independent but their correlation is very small, at least on the time interval when the Mean Field limit holds, *i.e.*, up to a time of order (2.9).

Beyond Mean Field. The Mean Field limit predicts that correlations are vanishing. In our context however this would imply *that* a correlation as defined in (2.3) is close to 0 which is in contradiction with the experimentally observed phenomenon which we wish to explain.

Recall that system (1.1) is scaled with a factor of $1/N$. In a given experimental setting, there is no particular reason why α should be of order 1. Instead because of the scaling, the coefficient α is typically polynomial in N . Hence the time T defined by (2.9) is also very small. After this time during which the Mean Field theory guarantees that correlations are small, those correlations may start developing and creating the collective behavior.

In general the computation of those correlations would require one to solve the full equation (2.10) on f_N . Unfortunately, f_N is a function of $N + 1$ variables and the computational cost of the numerical solution of (2.10) is much too large.

We would like to compute directly some of the marginals f_k , in particular f_1 and f_2 , since f_2 essentially captures correlations. However, the equation (2.11) on f_k is never closed unless $k = N$ as it involves f_{k+1} . This is the problem we focus on in this paper; we propose a solution consisting in truncating the hierarchy by choosing an ansatz for f_3 in terms of f_1 and f_2 .

Remark 2.1 (*On singular kernels*) As mentioned before, the rigorous justification of the classical Mean Field theory requires some smoothness on the interaction kernel, K Lipschitz. Many physical kernels are more singular, in particular in the context we are interested in, *i.e.*, the context of bacteria interacting through a fluid.

It is widely conjectured that the Mean Field theory can be extended to more singular kernels and some results are already available, see for example [20], [21], [22] or [23] in the phase space framework.

In this work, we are not concerned with rigorous justification of our results under proper assumptions on smoothness of K , however, just as in the Mean Field approach, we believe that the numerical implementation of our approach will work well for a wide class of kernels K (including singular ones).

3. Truncation of the hierarchy. In this section we first discuss the possibility of a truncation ansatz $f_3 = \mathcal{F}[f_1, f_2]$ such that the full BBGKY hierarchy becomes a system of two PDEs for marginals f_1 and f_2 only. A truncation ansatz of the form $f_3 = \mathcal{F}[f_1, f_2]$ changes the equation for f_2 :

$$\begin{aligned}
(3.1) \quad & \partial_t f_2 + \frac{\alpha K(0)}{N} \partial_{x_1} f_2 + \frac{\alpha K(0)}{N} \partial_{x_2} f_2 \\
& + \frac{\alpha}{N} \partial_{x_1} (K(x_2 - x_1) f_2) + \frac{\alpha}{N} \partial_{x_2} (K(x_1 - x_2) f_2) \\
& + \alpha \frac{N-2}{N} \partial_{x_1} \left\{ \int K(x_3 - x_1) \mathcal{F}[f_1, f_2](t, x_1, x_2, x_3) dx_3 \right\} \\
& + \alpha \frac{N-2}{N} \partial_{x_2} \left\{ \int K(x_3 - x_2) \mathcal{F}[f_1, f_2](t, x_1, x_2, x_3) dx_3 \right\} = 0.
\end{aligned}$$

For the sake of simplicity, in this section we consider the case with no self-interactions, that is $F(x) \equiv 0$. We want the solution to satisfy the following properties:

1. $f_2(x_1, x_2) = f_2(x_2, x_1)$ (*particles are identical*);
2. $\int f_2 \equiv \text{const}$, $f_1, f_2 \geq 0$ provided that initial conditions for f_1 and f_2 are positive (*mass preserving and positivity*);
3. $f_1 = \int f_2$ (*consistency*);
4. If $f_2 = f_1 \otimes f_1$, then $f_3(x_1, x_2, x_3) = \mathcal{F}[f_1, f_1 \otimes f_1] = f_1(x_1) f_1(x_2) f_1(x_3)$.

By $f_2 = f_1 \otimes f_1$ we mean the equality $f_2(x, y) = f_1(x) f_1(y)$.

Property 4 guarantees that the truncation ansatz $f_3 = \mathcal{F}[f_1, f_2]$ is compatible with the Mean Field limit as $N \rightarrow \infty$. More precisely, letting $N \rightarrow \infty$ in the equation

(3.1) for f_2 with the ansatz, the equation

$$\begin{aligned} \partial_t f_2^\infty + \partial_{x_1} \left\{ \int K(x_3 - x_1) \mathcal{F}[f_1^\infty, f_2^\infty](t, x_1, x_2, x_3) dx_3 \right\} \\ + \partial_{x_2} \left\{ \int K(x_3 - x_2) \mathcal{F}[f_1^\infty, f_2^\infty](t, x_1, x_2, x_3) dx_3 \right\} = 0 \end{aligned}$$

reduces to the Vlasov equation for the Mean Field limit, because the propagation of chaos holds: $f_2^\infty(t, x_1, x_2) = f_1^\infty(t, x_1) f_1^\infty(t, x_2)$.

We reformulate these properties as requirements on the functional \mathcal{F} and then prove that such an ansatz does not exist. Next, we present a truncation which is not based on a unique ansatz, yet the corresponding solution satisfies *the* four properties above.

Consider a representation for f_3 :

$$(3.2) \quad f_3(x_1, x_2, x_3) = \mathcal{F}[f_1, f_2](x_1, x_2, x_3),$$

where \mathcal{F} is a function (*in general* a nonlinear operator) of f_1 and f_2 . We reformulate the key properties as requirements on f_3 calculated by (3.2) for given f_1 and f_2 .

First, the symmetry with respect to arguments x_1 and x_2 is *equivalent to*:

$$(3.3) \quad \begin{aligned} f_2(x_1, x_2) = f_2(x_2, x_1) \text{ for all } x_1, x_2 \\ \Rightarrow f_3(x_1, x_2, x_3) = f_3(x_2, x_1, x_3) \text{ for all } x_1, x_2, x_3. \end{aligned}$$

Next, in order to preserve positivity of the distributions we need to impose

$$(3.4) \quad \text{for all } x_1, x_2 : (f_2(x_1, x_2) = 0 \Rightarrow f_3(x_1, x_2, x_3) = 0 \text{ for all } x_3)$$

and

$$(3.5) \quad (f_2(x_1, x_2) \geq 0 \text{ for all } x_1, x_2) \Rightarrow (f_3(x_1, x_2, x_3) \geq 0 \text{ for all } x_1, x_2, x_3).$$

The requirement (3.4) implies that there exists a function $h(x_1, x_2, x_3)$ such that $f_3(x_1, x_2, x_3) = h(x_1, x_2, x_3) f_2(x_1, x_2)$. Indeed, if $f_2(x_1, x_2) \neq 0$, then

$$h(x_1, x_2, x_3) = \frac{f_3(x_1, x_2, x_3)}{f_2(x_1, x_2)} \text{ for all } x_3.$$

If $f_2(x_1, x_2) = 0$, then h can be defined arbitrarily. By the method of characteristics, this property guarantees positivity of the solutions to the truncated system (3.1) provided that the initial data is positive.

Finally, in order to have the consistency property $f_1(x) = \int f_2(x, y) dy$ we impose

$$(3.6) \quad f_2(x_1, x_2) = \int f_3(x_1, x_3, x_2) dx_3$$

The equality (3.6) is equivalent to the statement that if we integrate the equation for $k = 2$ from the BBGKY hierarchy with respect to one of the spatial variables, say, x_2 , we get the equation for $k = 1$.

PROPOSITION 3.1. *There is no such representation (3.2) that all requirements (3.3), (3.4), (3.5) and (3.6) hold true.*

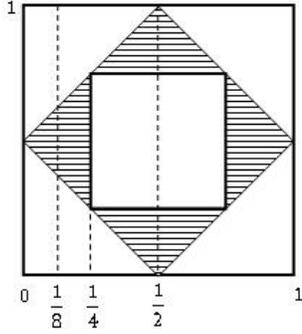


FIG. 3.1. Ω is shaded domain

Proof. The proof is by contradiction. The idea is to combine requirements (3.4) and (3.5):

$$(3.7) \quad f_2(x_1, x_2) = \int h(x_1, x_3, x_2) f_2(x_1, x_3) dx_3$$

and to find such f_2 that the LHS of (3.7) is zero, but the RHS is not zero.

Assume that (3.3), (3.4), (3.5) and (3.6) hold true. Take

$$\Omega = \left\{ (x_1, x_2) : |x_1 - \frac{1}{2}| + |x_2 - \frac{1}{2}| < \frac{1}{2} \right\} \setminus \left\{ |x_1 - \frac{1}{2}| < \frac{1}{4}, |x_2 - \frac{1}{2}| < \frac{1}{4} \right\}$$

and $f_2(x_1, x_2) = \frac{1}{|\Omega|} \chi_\Omega(x_1, x_2) = 4\chi_\Omega(x_1, x_2)$. Here χ_Ω is a characteristic function of domain Ω . Note $f_1(x) > 0$ for all $x \in (0, 1) \setminus \{\frac{1}{4}, \frac{3}{4}\}$ because of the equality $f_1(x) = \int f_2(x, y) dy$ which holds due to (3.6).

The property (3.4) implies the existence of such a function $h(x_1, x_2, x_3)$ that $f_3(x_1, x_2, x_3) = h(x_1, x_2, x_3) f_2(x_1, x_2)$. Thus, from (3.6) we obtain

$$(3.8) \quad f_2(x_1, x_2) = \int h(x_1, x_3, x_2) f_2(x_1, x_3) dx_3.$$

Let $(x_1, x_2) \notin \Omega$, then (3.8) implies that

$$(3.9) \quad 0 = \int h(x_1, x_3, x_2) f_2(x_1, x_3) dx_3 = 4 \int_{x_3: (x_1, x_3) \in \Omega} h(x_1, x_3, x_2) dx_3.$$

Thus

$$(3.10) \quad h(x_1, x_3, x_2) = 0, \text{ if } (x_1, x_2) \notin \Omega \text{ and } (x_1, x_3) \in \Omega.$$

By using the symmetry of h with respect to first two arguments we get $h(x_1, x_3, x_2) = h(x_3, x_1, x_2)$ and

$$(3.11) \quad h(x_1, x_3, x_2) \equiv 0 \text{ if } (x_2, x_3) \notin \Omega \text{ and } (x_1, x_3) \in \Omega.$$

Finally, calculate $f_1(1/8)$. On the one hand, $f_1(1/8) = \int f_2(1/8, y) dy > 0$. On the other hand,

$$(3.12) \quad f_1(x_2) = 4 \int \int_{(x_1, x_3) \in \mathcal{O}} h(x_1, x_3, x_2) \chi_\Omega(x_1, x_3) dx_3 dx_1.$$

where $\mathcal{O} = \{(x_1, x_3) : h(x_1, x_2, x_3)\chi_\Omega(x_1, x_3) \neq 0\}$. The domain \mathcal{O} depends on x_2 . We claim that \mathcal{O} is empty for $x_2 = 1/8$. Indeed,

$$\begin{aligned}\mathcal{O} &= \{h(x_1, x_3, 1/8) \neq 0 \text{ and } \chi_\Omega(x_1, x_3) \neq 0\} = [\text{definition of } \chi_\Omega] \\ &= \{(x_1, x_3) \in \Omega, h(x_1, x_3, 1/8) \neq 0\} \subset [(3.10) \text{ and } (3.11)] \\ &\subset \{(x_1, x_3) \in \Omega, (x_1, 1/8) \in \Omega, (x_3, 1/8) \in \Omega\} \\ &= \{(x_1, x_3) \in \Omega, x_1 \in (3/8, 5/8), x_3 \in (3/8, 5/8)\} = \emptyset.\end{aligned}$$

The integral in (3.12) is taken over empty set. Thus, $f_1(1/8) = 0$ and we have reached a contradiction. \square

Instead of using a unique representation ansatz for f_3 we use two different, but similar, representation ansatzes for f_3 , $f_3 = f_3^{(\text{I})}(x_1, x_2, x_3)$ and $f_3 = f_3^{(\text{II})}(x_1, x_2, x_3)$, in two different places where f_3 appears in the equation $k = 2$ such that the key properties are preserved. Namely, the equation $k = 2$ is rewritten as follows

$$\begin{aligned}(3.13) \quad &\partial_t f_2 + \frac{\alpha K(0)}{N} \partial_{x_1} f_2 + \frac{\alpha K(0)}{N} \partial_{x_2} f_2 \\ &+ \frac{\alpha}{N} \partial_{x_1} (K(x_2 - x_1) f_2) + \frac{\alpha}{N} \partial_{x_2} (K(x_1 - x_2) f_2) \\ &+ \alpha \frac{N-2}{N} \partial_{x_1} \left\{ \int K(x_3 - x_1) f_3^{(\text{I})}(t, x_1, x_2, x_3) dx_3 \right\} \\ &+ \alpha \frac{N-2}{N} \partial_{x_2} \left\{ \int K(x_3 - x_2) f_3^{(\text{II})}(t, x_1, x_2, x_3) dx_3 \right\} = 0,\end{aligned}$$

where

$$(3.14) \quad f_3^{(\text{I})}(t, x_1, x_2, x_3) = \begin{cases} \frac{f_2(t, x_1, x_2) f_2(t, x_1, x_3)}{\int f_2(t, x_1, y) dy}, & \int f_2(t, x_1, y) dy > 0, \\ 0, & \int f_2(t, x_1, y) dy = 0, \end{cases}$$

and

$$(3.15) \quad f_3^{(\text{II})}(t, x_1, x_2, x_3) = \begin{cases} \frac{f_2(t, x_1, x_2) f_2(t, x_3, x_2)}{\int f_2(t, y, x_2) dy}, & \int f_2(t, y, x_2) dy > 0, \\ 0, & \int f_2(t, y, x_2) dy = 0. \end{cases}$$

The four key properties listed above *are preserved after such truncation*:

1. Symmetry of f_2 with respect to x_1 and x_2 (provided that initially f_2 is symmetric) follows from symmetry of the equation with respect to x_1 and x_2 .
2. Conservation of mass and positivity follow from the fact that equation (3.13) can be rewritten as a standard conservation law (see (3.16) below).
3. By integrating (3.13) over, for example, x_2 , one obtains an equation for $\int f_2(x_1, x_2) dx_2$ which coincides with the equation for f_1 . By assuming uniqueness we get the consistency property: $f_1(x_1) = \int f_2(x_1, x_2) dx_2$.
4. If $f_2(x_1, x_2) = f_1(x_1) f_1(x_2)$, then $f_3^{(\text{I})} = f_3^{(\text{II})} = f_1(x_1) f_1(x_2) f_1(x_3)$. Note that in this case equation (3.13) is reduced to the Mean Field equation (2.5) in the limit $N \rightarrow \infty$.

We conclude this section by giving a physical interpretation of the introduced ansatz. To this end, we will rewrite (3.13) in a more convenient form.

The first integral term where the ansatz is applied, $\int K(x_3 - x_1) f_3^{(\text{I})}(t, x_1, x_2, x_3) dx_3$, can be rewritten as $F(t, x_1) f_2(t, x_1, x_2)$, where

$$F(t, x) = \int K(y - x) f_2(t, x, y) dy / \int f_2(t, x, y) dy$$

(assume $f_2 > 0$). Analogously, the second integral term is $F(t, x_2)f_2(t, x_1, x_2)$. Then equation (3.13) is rewritten in the form of a 2D conservation law:

$$(3.16) \quad \begin{aligned} & \partial_t f_2 + \partial_{x_1} \left(\left\{ \frac{\alpha K(0)}{N} + \frac{\alpha}{N} K(x_2 - x_1) + \alpha \frac{N-2}{N} F(t, x_1) \right\} f_2 \right) \\ & + \partial_{x_2} \left(\left\{ \frac{\alpha K(0)}{N} + \frac{\alpha}{N} K(x_1 - x_2) + \alpha \frac{N-2}{N} F(t, x_2) \right\} f_2 \right) = 0. \end{aligned}$$

For two particles X_i and X_j such that $X_i = x$, the number $(1/N)F(t, x)$ can be viewed as the average (or expected) force applied to the particle X_i by the particle X_j at time t . *In this sense, equation (3.13) can still be viewed as a mean field approximation but at a higher order.*

4. Numerical example. The goal of this section is (i) to test the truncation (3.13) on a simple 1D example and (ii) to describe *the* main difficulties which arise in the numerical *resolution* of the PDE (3.13). Here we use numerical methods which are explicit, allows for comparison with direct simulations, and *are* not necessarily the most efficient. The development of more advanced numerical methods capturing, e.g., 2D, non-smooth kernels or large times, are left for a subsequent work.

To test the truncation we compare probability distributions (marginals) f_1 and f_2 obtained by numerical solution of the truncated system (3.13) with histograms of particles satisfying the original ODE system (1.1). The histograms are built on many realizations of initial particle positions.

Specific setting. Numerical simulations are performed for the one-dimensional problem, $x \in \mathbb{R}^1$, and periodic boundary conditions with period 1. The interaction kernel K is periodic with period 1 and for $|X_j - X_i| < 1/2$ it is given by $K(X_j - X_i) = e^{-12(X_j - X_i)^2}$. Initially particles are independent:

$$(4.1) \quad f_2(0, x_1, x_2) = f_1(0, x_1)f_1(0, x_2), \text{ where } f_1(0, x_1) = .4 \sin 2\pi x_1 + 1.$$

Number of particles is $N = 100$ per one periodic cell $x \in [0, 1]$, $\alpha = 3$.

Description of numerical methods. In order to solve the PDE (3.13) we face difficulties that come from the fact that the equation is a non-local non-linear 2D conservation law. For a detailed discussion of difficulties in numerical solution of non-linear conservation laws and the way to resolve them we refer to [24]. In this example we want to simulate accurately terms of order $1/N$, since they are the source of correlations. In other words, if one erases these terms in (3.13), then the solution of equation (3.13) with initial conditions (4.1) will be of the form $f_2(t, x_1, x_2) = f_1(t, x_1)f_1(t, x_2)$, i.e., with no correlations. This motivates us to use a second order scheme and we implemented a second order scheme with flux limiters for which we have converging numerical simulations with reasonable spatial and time steps. This method is described below.

The PDE (3.13) can be rewritten as follows

$$\partial_t f_2 + \partial_{x_1} (\mathcal{A}_1 f_2) + \partial_{x_2} (\mathcal{A}_2 f_2) = 0,$$

where for $f_2(x_1, x_2) > 0$ functions \mathcal{A}_1 and \mathcal{A}_2 are given by

$$(4.2) \quad \mathcal{A}_k(t, x_1, x_2) = \frac{\alpha}{N} \sum_{i=1,2} K(x_i - x_k) + \frac{\alpha(N-2)}{N} \frac{\int K(y - x_k) f_2(t, x_k, y) dy}{\int f_2(t, x_k, y) dy}.$$

Denote by $f_{i,j}^m$ the approximation for $f_2(t, x_1, x_2)$ with $t = mdt$, $x_1 = idx$, $x_2 = jdx$, where dt and dx are time and spatial steps, respectively. For given m , i and j introduce the following finite difference approximations for $\partial_k \{\mathcal{A}_k f_2\}$, $k = 1, 2$:

$$\begin{aligned} r_{11} &:= \frac{A_{i,j}^m f_{i,j}^m - A_{i-1,j}^m f_{i-1,j}^m}{dx}, & r_{12} &:= \frac{A_{i+1,j}^m f_{i+1,j}^m - A_{i,j}^m f_{i,j}^m}{dx}, \\ r_{21} &:= \frac{A_{i,j}^m f_{i,j}^m - A_{i,j-1}^m f_{i,j-1}^m}{dx}, & r_{22} &:= \frac{A_{i,j+1}^m f_{i,j+1}^m - A_{i,j}^m f_{i,j}^m}{dx}. \end{aligned}$$

Introduce also an auxiliary function (flux limiter) $\phi(r) = \max[0, 0.5 \min(r, 1.5)]$.

The following finite difference scheme is used in the numerical solution of PDE (3.13):

$$(4.3) \quad f_{i,j}^{m+1} = f_{i,j}^m + \frac{dt}{dx} \left[\sum_{k=1,2} \left\{ r_{k1} + \phi\left(\frac{r_{k1}}{r_{k2}}\right) (r_{k2} - r_{k1}) \right\} \right].$$

In order to compute the two-particle distribution for $t > 0$ directly from the system of ODEs (1.1) we consider $R = 5 \cdot 10^5$ realizations of $N = 100$ particles initially identically distributed with probability distribution function $f(x) = 0.4 \sin 2\pi x + 1$. Denote by $X_i^{(r)}(t)$ the position of the i th particle, $i = 1, \dots, N$ in the r th realization, $r = 1, \dots, R$ at time t . For each $r = 1, \dots, R$ the positions $\{X_i^{(r)}(t)\}_{i=1, \dots, N}$, $t > 0$, are found as the solution of the ODE system (1.1) by the explicit Euler method of the first order with the time step $\Delta t = 0.01$.

We compute the following histogram which approximates the probability of that the first particle is in the interval $\Delta_j = [jh, (j+1)h)$ at time t :

$$(4.4) \quad \tilde{f}_1(t, \Delta_j) = \frac{1}{Rh} \# \left\{ X_1^{(r)}(t) \in \Delta_j, r = 1, \dots, R \right\}.$$

Here $h = 0.05$ is the size of a histogram bin.

Histogram \tilde{f}_2 which approximates the two-particle distribution can be computed as follows

$$(4.5) \quad \tilde{f}_2(t, \Delta_i, \Delta_j) = \frac{1}{Rh^2} \# \left\{ \left(X_1^{(r)}(t), X_2^{(r)}(t) \right) \in \Delta_i \times \Delta_j, r = 1, \dots, R \right\}.$$

Thus, in numerical simulations our intention is to compare $f_1 = \int f_2 dx$ and f_2 calculated by (4.3) with histograms \tilde{f}_1 and \tilde{f}_2 calculated by (4.4) and (4.5).

Simulations were performed on a machine with 3.06 Ghz Intel core CPU 8 GB of RAM. Numerical solution of (3.13) for $t = 1$ for $dx = .0025$ and $dx/dt = 200$ takes approximately 32 hours. Numerical solution of (1.1) on $R = 10^5$ realization, $t = 1$ and time step $\Delta t = 1/50$ takes approximately 83 hours. Besides the long time of computations direct simulations face another difficulty which is the large amount of data that creates technical difficulties in data movement, its analysis and visualization.

Results of numerical simulations. Plots in Fig. 4.1 show that marginal f_1 is close to histogram \tilde{f}_1 . In order to visualize comparisons between f_2 and \tilde{f}_2 we plot these functions integrated over $\mathcal{B} = \{(x_1, x_2) : 0 \leq x_1, x_2 \leq 1/2\}$:

$$\text{Qmarg} = \int_{\mathcal{B}} f_2 dx_1 dx_2, \quad \text{Qhist} = \int_{\mathcal{B}} \tilde{f}_2 dx_1 dx_2 = h^2 \sum_{i,j \leq 1/(2h)} \tilde{f}_2(t, \Delta_i, \Delta_j).$$

Plot 4.2 shows that quantities Qmarg defined by marginal f_2 and Qhist defined by histogram \tilde{f}_2 seem to be in very good agreement.

Finally, we computed correlations using marginals f_1 and f_2 . For the marginal approach (i.e., solution of (3.13)) correlations are defined as follows

$$c(t) = \int \int |f_2(t, x_1, x_2) - f_1(t, x_1)f_1(t, x_2)| dx_1 dx_2.$$

In direct simulations (i.e., solution of ODE (1.1) for many random realizations of initial conditions) correlations are defined in a similar way to the above formula with histograms in place of distributions:

$$\begin{aligned} \tilde{c}(t) &= \int \int |\tilde{f}_2(t, x_1, x_2) - \tilde{f}_1(t, x_1)\tilde{f}_1(t, x_2)| dx_1 dx_2 \\ &= h^2 \sum_{i,j} |\tilde{f}_2(t, \Delta_i, \Delta_j) - \tilde{f}_1(t, \Delta_i)\tilde{f}_1(t, \Delta_j)|. \end{aligned}$$

As it is seen on Fig. 4.3, plots for correlations computed on marginals and in direct simulations for $R = 5 \times 10^5$ have similar qualitative behavior and order of magnitude. The value of correlations is a small number and thus its computation requires high accuracy to reduce the error to an order less than that of the correlations. In direct simulations, this requires a large number of realizations which make the computations unreasonably long, in contrast to the marginal approach where the computation of correlations is much faster.

Note that correlations we observe are not large (in comparison with the maximal possible value of correlations $c_{\max} = 2$). In order to observe large correlations (e.g., ~ 0.1) we need to solve (3.13) for large times which is very costly. Moreover, *it is delicate to predict the time when correlations will reach some fixed, large value.* This question is left for subsequent works. Nevertheless, *relatively* small correlations for times of order 1 may be enough for the solution of (3.13) or the original BBGKY hierarchy to be essentially different from the one obtained by the Mean Field approach.

Convergence of numerical methods. Here we show that the numerical methods we used in this section converge.

First, consider the calculations of marginals f_1 and f_2 . Comparisons of numerical simulations for various spatial and time steps for $t = 1$, $t = 2$ and $t = 3$ are presented on Figures 4.4 and 4.5. Convergence of the numerical method in computing $\int_{\mathcal{B}} f_2 dx$ and correlations $c(t)$ is observed on plots in Figure 4.6.

Next, look at the calculations for histograms \tilde{f}_1 and \tilde{f}_2 . Plots on Figures 4.7 and 4.8 illustrate convergence of the method for histogram \tilde{f}_1 at times $t = 1$, $t = 2$ and $t = 3$, and for histogram \tilde{f}_2 summed over the set \mathcal{B} . *several time steps Δt are considered:* $\Delta t = 0.02$, $\Delta t = 0.01$, $\Delta t = 0.001$. The number of realizations, $R = 10^5$, is chosen for the width of bin $h = 0.02$. It is seen on Figures 4.7 and 4.8 that such a number of realizations R seems to be enough to have converging numerical solutions for \tilde{f}_1 and $\int_{\mathcal{B}} \tilde{f}_2$. To compute correlations $\tilde{c}(t)$ more realizations *would be* needed and plots in Figure 4.9 show that to estimate $\tilde{c}(t)$ we need *more than* $R = 5 \cdot 10^5$ realizations with $\Delta t = 0.002$.

Numerical simulations presented above show that PDE system (3.13) not only preserves the qualitative properties of the probability distribution functions (like positivity, consistency, propagation of chaos, etc.), but also may serve for the study of saturation of correlations in such many particle systems.

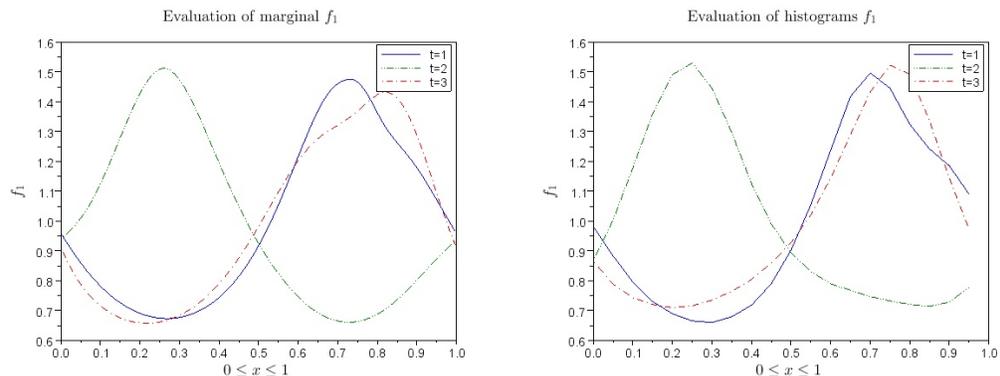


FIG. 4.1. Left: Marginal f_1 with $dx = 0.0025$ and $dt = dx/200$; Right: Histogram \tilde{f}_1 for $h = 0.05$ and $dt = 0.001$.

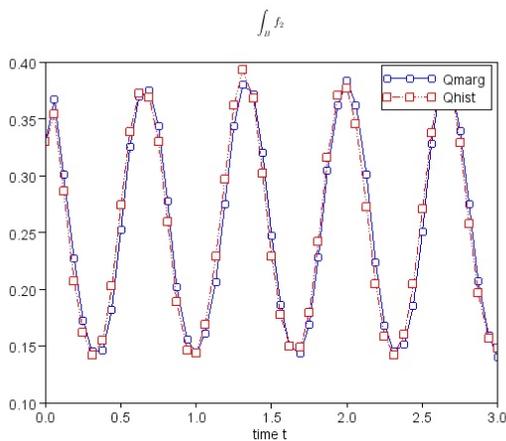


FIG. 4.2. Comparison between Qmarg and Qhist.

5. Conclusions. We developed a numerical approach which allows for study correlations and checked it in a simple setting (toy model). We believe that this approach can be successfully applied to problems in biology, physics and economics.

REFERENCES

- [1] X.-L. Wu and A. Libchaher, “Particle diffusion in a quasi-two-dimensional bacteria bath,” *Physical Review Letters*, vol. 84, p. 3017, 2000.
- [2] C. Dombrowski, L. Cisneros, S. Chatkaew, R. Goldstein, and J. Kessler, “Self-concentration and large-scale coherence in bacterial dynamics,” *Physical Review Letters*, vol. 93, p. 98103, 2004.
- [3] A. Sokolov, I. Aranson, J. Kessler, and R. Goldstein, “Concentration dependence of the collective dynamics of swimming bacteria,” *Physical Review Letters*, vol. 98, no. 15, p. 158102, 2007.
- [4] A. Sokolov and I. Aranson, “Reduction of viscosity in suspension of swimming bacteria,” *Phys. Rev. Lett.*, vol. 103, p. 148101, Sep 2009.
- [5] A. Sokolov, M. Apodaca, B. Grzybowski, and I. Aranson, “Swimming bacteria power microscopic gears,” *PNAS*, vol. 107, pp. 969–974, 2010.

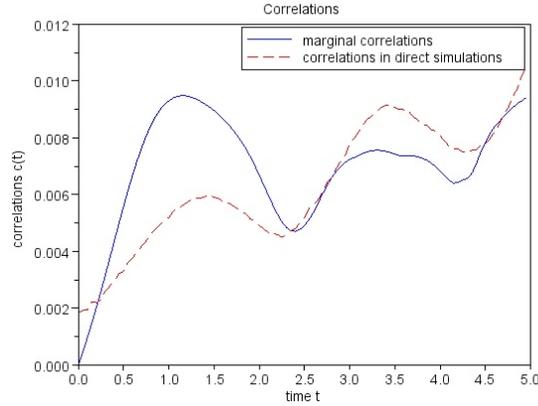


FIG. 4.3. Comparisons of correlations computed by marginal approach (3.13) and direct simulations of the system (1.1).

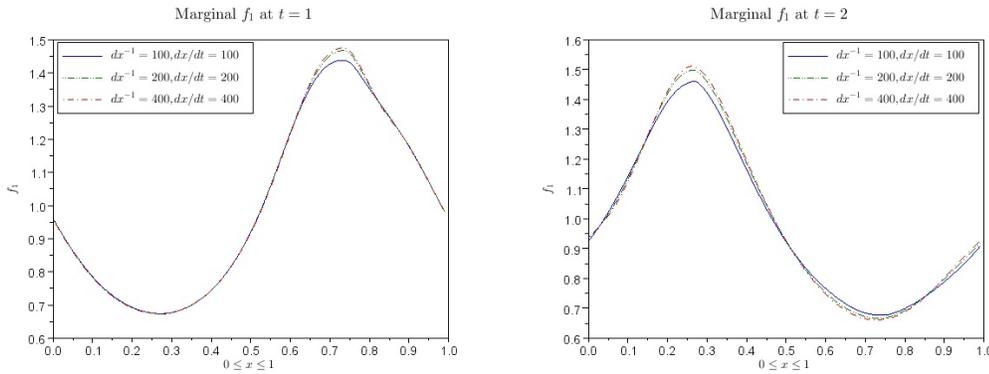


FIG. 4.4. Left: plots of marginal f_1 at $t = 1$ and Right: $t = 2$

- [6] K. Leptos, J. Guasto, J. Collub, A. Pesci, and R. Goldstein, “Dynamics of enhanced tracer diffusion in suspensions of swimming eukaryotic microorganisms,” *Physical Review Letters*, vol. 103, p. 198103, 2009.
- [7] A. Sokolov and I. Aranson, “Physical properties of collective motion in suspensions of bacteria,” *Physical Review Letters*, vol. 109, p. 248109, 2012.
- [8] S. Ryan, A. Sokolov, L. Berlyand, and I. Aranson, “Correlation properties of collective motion in bacterial suspension,” *New Journal of Physics*, 2013.
- [9] L. Greengard and V. Rokhlin, “A fast algorithm for particle simulation,” *Journal of Computational Physics*, vol. 73, pp. 325–348, 1987.
- [10] L. Greengard and V. Rokhlin, “Rapid evaluation of potential fields in three dimensions,” *Lecture Notes in Mathematics*, vol. 1360, pp. 121–141, 1988.
- [11] L. Greengard and V. Rokhlin, “On the evaluation of electrostatic interactions in molecular modeling,” *Chemica Scripta*, vol. 29A, pp. 139–144, 1989.
- [12] M. Marder, “Correlations and ostwald ripening,” *Physical Review A*, vol. 36, pp. 858–874, 1987.
- [13] A. Honig, B. Niethammer, and F. Otto, “On first-order corrections to the lsw theory i: infinite systems,” *Journal of Statistical Physics*, vol. 119, pp. 61–122, 2005.
- [14] A. Honig, B. Niethammer, and F. Otto, “On first-order corrections to the lsw theory ii: finite systems,” *Journal of Statistical Physics*, vol. 119, pp. 123–164, 2005.
- [15] W. Braun and K. Hepp, “The Vlasov dynamics and its fluctuations in the $1/N$ limit of inter-

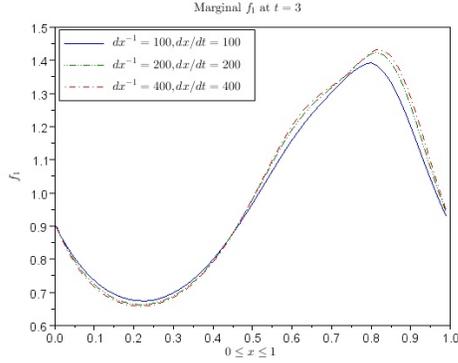


FIG. 4.5. Plots of marginal f_1 at $t = 3$ for $0 \leq x \leq 1$;

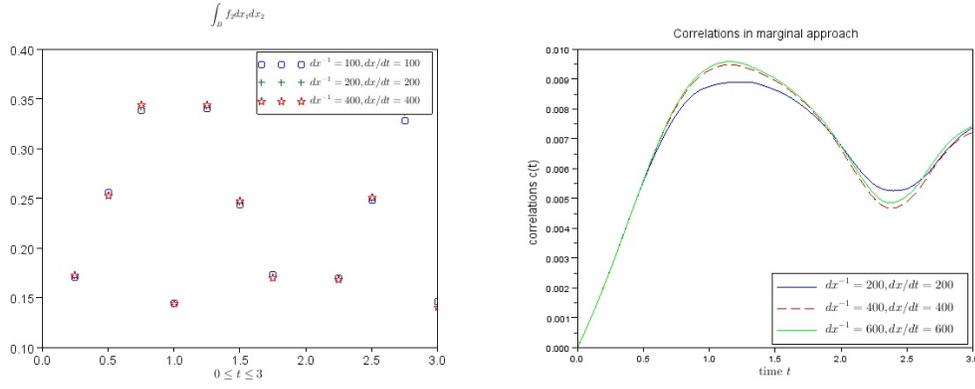


FIG. 4.6. Left: marginal f_2 integrated over B ; Right: correlations $c(t)$

- acting classical particles,” *Comm. Math. Phys.*, vol. 56, no. 2, pp. 101–113, 1977.
- [16] H. Neunzert and J. Wick, “The convergence of simulation methods in plasma physics,” in *Mathematical methods of plasmaphysics (Oberwolfach, 1979)*, vol. 20 of *Methoden Verfahren Math. Phys.*, pp. 271–286, Frankfurt: Lang, 1980.
- [17] H. Spohn, *Large scale dynamics of interacting particles*. New York: Springer Verlag, 1991.
- [18] E. Boissard, *Problèmes d’interaction discret-continu et distances de Wasserstein*. PhD thesis, Université de Toulouse III, 2011.
- [19] E. Caglioti and F. Rousset, “Long time behavior of particle systems in the mean field limit,” *Commun. Math. Sci.*, no. suppl. 1, pp. 11–19, 2007.
- [20] J. Goodman, T. Y. Hou, and J. Lowengrub, “Convergence of the point vortex method for the 2-D Euler equations,” *Comm. Pure Appl. Math.*, vol. 43, no. 3, pp. 415–430, 1990.
- [21] M. Hauray, “Wasserstein distances for vortices approximation of Euler-type equations,” *Math. Models Methods Appl. Sci.*, vol. 19, no. 8, pp. 1357–1384, 2009.
- [22] S. Schochet, “The point-vortex method for periodic weak solutions of the 2-D Euler equations,” *Comm. Pure Appl. Math.*, vol. 49, no. 9, pp. 911–965, 1996.
- [23] M. Hauray and P.-E. Jabin, “ N -particles approximation of the Vlasov equations with singular potential,” *Arch. Ration. Mech. Anal.*, vol. 183, no. 3, pp. 489–524, 2007.
- [24] R. LeVeque, *Numerical Methods for Conservation Laws*. Birkhauser, Basel, 1992.

Acknowledgments. The work of Leonid Berlyand and Mykhailo Potomkin was supported by DOE grant DE-FG- 0208ER25862. The work of Pierre-Emmanuel Jabin was partially supported by NSF grant DMS-1312142. LB and MP wish to thank V.

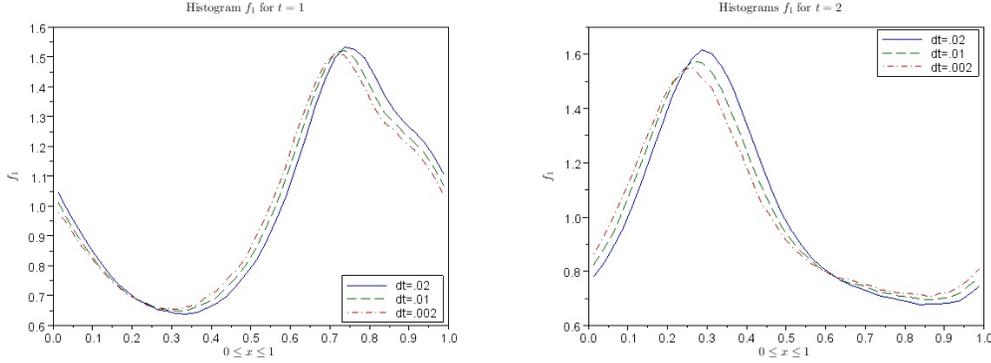


FIG. 4.7. Left: plots of histogram \tilde{f}_1 at $t = 1$; Right: plots of histogram \tilde{f}_1 at $t = 2$

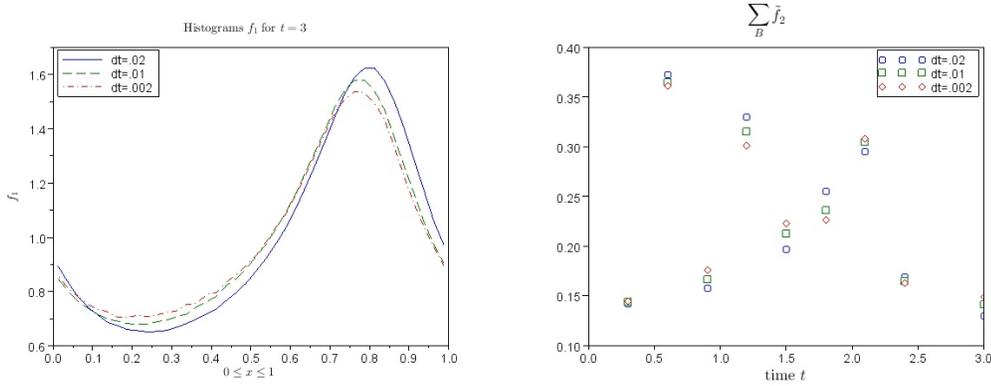


FIG. 4.8. Left: plots of histogram \tilde{f}_1 at $t = 3$; Right: plots of histogram \tilde{f}_2 summed over set \mathcal{B}

Rybalko for his comments and suggestions which helped to improve the manuscript.

Appendix A. Appendix: Wasserstein distances. Wasserstein distance or Monge-Kantorovich-Wasserstein (MKW) quantifies the difference between two given measures. Roughly speaking, a measure can be viewed as a pile of sand. The MKW distance between two such piles is an optimal work of transferring one pile into another.

Given two measures μ_1 and μ_2 in $M^1(D)$, one may define the set of transference plans between μ_1 and μ_2 as the set $\Pi(\mu_1, \mu_2)$ of measures $\pi \in M^1(D \times D)$ s.t.

$$\mu_1(x) = \int_D \pi(x, dy), \quad \mu_2(y) = \int_D \pi(dx, y).$$

The p MKW distance $W_p(\mu_1, \mu_2)$ between μ_1 and μ_2 is given by

$$W_p(\mu_1, \mu_2) = \inf_{\pi \in \Pi(\mu_1, \mu_2)} \int_{D^2} |x - y|^p \pi(dx, dy).$$

If D is the torus, then $|x - y|$ is replaced by the corresponding distance (in general in a manifold, it would be the geodesic distance).

For measures with bounded moments, the MKW distances metrize the weak-* topology. Moreover, on bounded domains the W_1 distance is essentially equivalent to

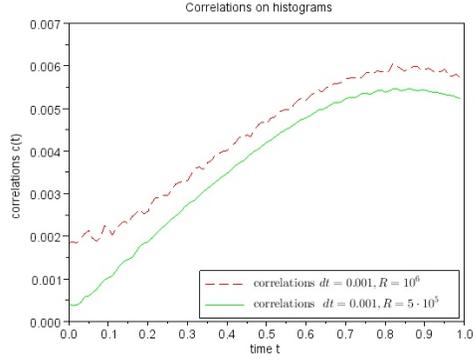


FIG. 4.9. *Correlations computed on histograms*

the negative Sobolev norm $W^{-1,1}$. The p -MKW distances play an important role for particle systems as the p -MKW distance between two empirical measures is typically comparable with the p distance between the two vectors of positions, that is

$$W_p \left(\frac{1}{N} \sum_i \delta_{x_i}, \frac{1}{N} \sum_i \delta_{y_i} \right) \sim \frac{1}{N} \sum_i |x_i - y_i|^p,$$

up to a permutation of indices on the y_i .