



Continuum Approximations to Systems of Correlated Interacting Particles

Leonid Berlyand¹ · Robert Creese¹ · Pierre-Emmanuel Jabin² · Mykhailo Potomkin¹

Received: 30 April 2018 / Accepted: 5 December 2018 / Published online: 15 December 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

We consider a system of interacting particles with random initial conditions. Continuum approximations of the system, based on truncations of the BBGKY hierarchy, are described and simulated for various initial distributions and types of interaction. Specifically, we compare the mean field approximation (MFA), the Kirkwood superposition approximation (KSA), and a recently developed truncation of the BBGKY hierarchy (the truncation approximation—TA). We show that KSA and TA perform more accurately than MFA in capturing approximate distributions (histograms) obtained from Monte Carlo simulations. Furthermore, TA is more numerically stable and less computationally expensive than KSA.

Keywords Many particle system · Mean field approximation · Closure of BBGKY hierarchy

1 Introduction

Systems of interacting particles are ubiquitous and can be found in many problems of physics, chemistry, biology, economics, and social science. A wide class of such systems can be presented as follows. Consider N interacting particles described by a coupled system of N ODEs:

$$dX_i(t) = S(X_i)dt + \sqrt{2D} dW_i(t) + \frac{1}{N} \sum_{j=1}^N u(X_i, X_j)dt, \quad \text{for } i = 1, \dots, N. \quad (1)$$

Here $X_i(t)$ is the position of the i th particle at time t , S is either self-propulsion, internal frequency (as in Kuramoto model, see Sect. 3.3), or a conservative force field (e.g., gravity), $\{W_i\}_{i=1}^N$ are N independent Wiener processes and $u(x, y)$ is an interaction force between two particles at positions x and y . Oftentimes the force is represented by a function of the directed distance between particles, so u can be written as follows

✉ Robert Creese
rec5208@psu.edu

¹ Department of Mathematics, The Pennsylvania State University, University Park, PA, USA

² Department of Mathematics, University of Maryland, College Park, MD 20742, USA

$$u(X_i, X_j) = \hat{u}(X_j - X_i). \quad (2)$$

System (1) has to be supplied with initial conditions. For a large number of particles N , finding the initial position for each particle is not practical. Instead, it is reasonable to assume that initially the positions $X_1, \dots, X_N$ are random, independent and identically distributed (i.i.d.). Thus, instead of determining a massive tuple of N initial conditions, a single continuous probability distribution function is introduced.

Note that system (1) represents first order dynamics in which the net force is proportional to velocity, i.e. $F_i \sim V_i = \dot{X}_i$, as opposed to second order dynamics, usually obtained from Newton's Law, in which the net force is proportional to acceleration $F_i \sim a_i = \ddot{X}_i$. First order dynamics are commonly used for models such as those of ants marching [29], bacteria swimming [38,39], hierarchies in pigeons [32], opinion dynamics [31], point vortices [16], etc.

To solve (1) with random initial conditions means to find a joint probability distribution function (or N -particle pdf):

$$f_N(t, x_1, x_2, \dots, x_N).$$

Then the probability of finding a tuple $(X_1, \dots, X_N)$ in a given domain Ω at time t is

$$\int_{\Omega} f_N(t, x_1, \dots, x_N) dx_1 \dots dx_N.$$

The function f_N can be found as a solution of the Liouville equation [44]. However finding a function of N arguments, such as f_N , numerically means computing an N -dimensional array which is prohibitively computationally expensive even for moderately large N . Therefore a simplification for the problem for f_N is required.

A classical approach is the mean field approximation (MFA) [2,20,44] which relies on the assumption that initially uncorrelated particles remain uncorrelated as time evolves. Then the joint probability distribution function f_N is determined by a function of a two variables, $f_1(t, x)$:

$$f_N(t, x_1, x_2, \dots, x_N) \approx \prod_{i=1}^N f_1(t, x_i). \quad (3)$$

One can substitute (3) into the Liouville equation for f_N to obtain a partial differential equation (PDE) for f_1 (McKean–Vlasov equation). In the limit $N \rightarrow \infty$ (the mean field limit), formula (3) holds exactly [6,13]. The function f_1 has the meaning of a one-particle pdf. Alternatively, in the limit $N \rightarrow \infty$ one can describe the set of all particles as a continuum with density $f_1(t, x)$. Though MFA is useful in many applications, it is generally not as accurate for moderate or small N [28]. MFA is also not applicable when the impact of correlations (which are neglected in MFA) is investigated. An example is collective behavior in bacterial suspensions [42,43]: density of bacteria—or equivalently one-particle pdf (since the number of bacteria is $N = 10^{10}$ per cm^3)—remains uniform, while correlation length increases, so that the two-particle pdf changes due to emergence of correlations.

One set of approaches to account for correlations is based on using closures of the Bogoliubov–Born–Green–Kirkwood–Yvon (BBGKY) hierarchy. This hierarchy is the system of N PDEs: one for the one-particle pdf f_1 , one for the two-particle pdf f_2 , ..., and one for the N -particle pdf f_N . The PDE for f_k ($k = 1, \dots, N - 1$) in the BBGKY hierarchy is obtained by integration of the Liouville equation with respect to $x_{k+1}, \dots, x_N$. The equation for f_N is the Liouville equation itself. Solving the BBGKY hierarchy is equivalent to solving

the Liouville equation which is computationally prohibitive as explained above. On the other hand, the PDEs in the BBGKY hierarchy are coupled as follows: the PDE for f_k depends on f_{k+1} . Therefore, one can obtain a closed system for $f_1, \dots, f_k$ by introducing a closure approximation for f_{k+1} in terms of $f_1, \dots, f_k$. For example, MFA can be considered as a closure of the BBGKY hierarchy at level $k = 1$ using the closure approximation:

$$f_2(t, x_1, x_2) = f_1(t, x_1)f_1(t, x_2). \quad (4)$$

The closure approximation (4) means that MFA relies on the assumption that correlations in the system of interacting particles are negligible. *To account for correlations one needs a closure approximation at least at level $k = 2$.*

The Kirkwood superposition approximation (KSA), developed in [23], is the most widely used closure of the BBGKY hierarchy at level $k = 2$ and was applied, for example, in gas dynamics [25], simple liquids [14] and recently employed in biology [3,28]. Following the general idea of closure approximations of the BBGKY hierarchy described above, in KSA a single ansatz for f_3 in terms of f_1 and f_2 is substituted in the equation for f_2 . This ansatz is presented in Sect. 2 and may be formulated in words as follows: the probability of finding the particle triple in a given configuration equals to the probability of finding each pair independently from the third particle [10]. Though KSA is a phenomenological ansatz, formal justification and further improvements are available [9,36]. However, we note that, up to our best knowledge, there is no rigorous asymptotic approach to derive a closure of BBGKY hierarchy that takes into account correlations.

Recently, a closure at level $k = 2$, alternative to KSA, has been introduced in [5]. The main difference between KSA and the closure from [5]—referred below to as the truncation approximation (TA)—is that instead of a single ansatz for f_3 , TA introduces an individual representation for each of the two terms in the equation for f_2 where f_3 appears. The choices in TA are made so that key properties of pdfs f_1 and f_2 are preserved (the properties are listed in Sect. 2). It was also proven that there is no such single representation ansatz for f_3 that preserves the key properties. Moreover TA is less computationally expensive than KSA.

Note that in order to construct a correction to MFA in a way alternative to finding correlations, one can take into account fluctuations. This is typically done by so called Large Deviations or Concentration Estimates which in simple terms are upper bounds for probabilities that a given realization of system (1) is far from the mean field limit; see [7,8,11,12,15] and review [21]. In this work we aim to characterize the collective or correlated state (for example, via the correlation length determined by f_2), which is not a result of random fluctuations from the mean field limit.

In this paper we consider system (1) with various types of interactions u . We compare the closures obtained from MFA, KSA and TA with each other and with Monte Carlo Simulations of (1). We show that TA is at least as accurate as KSA (when comparing to Monte Carlo simulations). Moreover, we observe that TA is less computationally expensive and more numerically stable than KSA. Finally for each type of interaction considered in this paper we describe the effect of correlations by comparing MFA, which neglects correlations, with other methods. Here we consider not very large N for the following two reasons. First, one- and two-particle histograms $\hat{f}_1$ and $\hat{f}_2$ obtained from Monte Carlo simulations do not require excessive computations for such N . Note that $\hat{f}_1$ and $\hat{f}_2$ converge to true f_1 and f_2 (that is, solutions of the original not truncated BBGKY hierarchy) as sample size, the number of realizations, grows to infinity. The second reason to choose N not large is to have an observable impact of correlations (correlations vanish as $N \rightarrow \infty$).

The paper is organized as follows. In Sect. 2 we formulate the main problem and review and discuss the application of the BBGKY hierarchy and its truncations, such as MFA, KSA and TA. Results of numerical simulations are presented in Sect. 3 and discussed in Sect. 4.

2 Description of Continuum Approximations

In this section we review the continuum approximations that are used in this work. First, we describe the BBGKY hierarchy and then discuss its closures such as MFA, KSA, and TA. Next, we compare the three approximations noting some key differences as well as similarities between them. Numerical integration of the corresponding PDEs is presented in Sect. 3.

BBGKY Hierarchy Consider system (1) with u satisfying (2). The one-particle pdf $f_1(t, x_1)$ solves the following evolution equation

$$\partial_t f_1(t, x_1) + \nabla_{x_1} \cdot ((S(x_1) + \mathcal{F}(t, x_1)) f_1(t, x_1)) = D \Delta_{x_1} f_1(t, x_1), \quad (5)$$

where $\mathcal{F}$ is the conditional expectation of force exerted on the first particle which occupies the position $X_1(t) = x_1$ by all other particles:

$$\mathcal{F}(t, x_1) = \mathbb{E} \left\{ \frac{1}{N} \sum_{j=2}^N u(X_1(t), X_j(t)) \middle| X_1(t) = x_1 \right\}. \quad (6)$$

Equation (5) is an advection–diffusion equation for f_1 or, alternatively, it can be derived from the Liouville equation by direct integration with respect to all variables except t and x_1 .

An explicit formula for $\mathcal{F}$ in terms f_1 and f_2 follows from the definition of conditional expectation:

$$\mathcal{F}(t, x_1) = \frac{N-1}{N} \int u(x_1, y) \frac{f_2(t, x_1, x_2)}{f_1(t, x_1)} dy. \quad (7)$$

In view of formula (7), we note that Eq. (5) depends on the two-particle pdf $f_2(t, x_1, x_2)$. In order to find f_2 we need to consider an equation for f_2 , analogous to (5) for f_1 :

$$\begin{aligned} & \partial_t f_2(t, x_1, x_2) + \nabla_{x_1} \cdot (\mathcal{F}_1 f_2(t, x_1, x_2)) + \nabla_{x_2} \cdot (\mathcal{F}_2 f_2(t, x_1, x_2)) \\ & + \nabla_{x_1} \cdot (S(x_1) f_2(t, x_1, x_2)) + \nabla_{x_2} \cdot (S(x_2) f_2(t, x_1, x_2)) \\ & = D(\Delta_{x_1} f_2(t, x_1, x_2) + \Delta_{x_2} f_2(t, x_1, x_2)), \end{aligned} \quad (8)$$

where $\mathcal{F}_i(t, x_1, x_2)$ ($i = 1, 2$) are the conditional expectation of force exerted on the i th particle by other particles given that $X_1(t) = x_1$ and $X_2(t) = x_2$:

$$\mathcal{F}_i(t, x_1, x_2) = \mathbb{E} \left\{ \frac{1}{N} \sum_{j \neq i}^N u(X_i(t), X_j(t)) \middle| \begin{matrix} X_1(t) = x_1 \\ X_2(t) = x_2 \end{matrix} \right\}. \quad (9)$$

Using that all particles are identical and substituting conditions $X_1(t) = x_1$ and $X_2(t) = x_2$ into the sum in the right hand side of (9), we simplify the formula for $\mathcal{F}_i$

$$\mathcal{F}_1(t, x_1, x_2) = \frac{1}{N} u(x_1, x_2) + \frac{N-2}{N} \int \frac{u(x_1, y) f_3(t, x_1, x_2, y)}{f_2(t, x_1, x_2)} dy, \quad (10)$$

$$\mathcal{F}_2(t, x_1, x_2) = \frac{1}{N} u(x_2, x_1) + \frac{N-2}{N} \int \frac{u(x_2, y) f_3(t, x_1, x_2, y)}{f_2(t, x_1, x_2)} dy. \quad (11)$$

It is clear from (10), (11) that to solve (8) one needs f_3 , the three particle pdf. One can write the equation for f_3 similar to (5) and (8), and this equation will depend on f_4 . We can continue in this manner to obtain a system of N coupled partial differential equations for $f_1, f_2, \dots, f_N$. The resulting system is the BBGKY hierarchy described in Sect. 1. This system is prohibitively computationally expensive to solve. Instead we look at various truncations of the BBGKY hierarchy which are computationally feasible and do not rely on the assumption that correlations are negligible unlike MFA which is a truncation in the equation for f_1 (at level $k = 1$). Specifically, we focus on truncations in the equations for f_1 and f_2 (at level $k = 2$). One can consider truncations at higher levels but the computational expense increases greatly with increasing the level of a truncation. As a result, we only consider truncations in the equations for f_1 and f_2 .

Mean field approximation MFA is a truncation of the BBGKY hierarchy at the equation for f_1 using the assumption

$$f_2(t, x_1, x_2) = f_1(t, x_1) f_1(t, x_2). \quad (12)$$

Substituting this assumption into (5) results in the following PDE

$$\begin{aligned} \partial_t f_1(t, x_1) + \frac{(N-1)}{N} \nabla_{x_1} \cdot \left(\int u(x_1, y) f_1(t, y) dy f_1(t, x_1) \right) + \nabla_{x_1} \cdot (S(x_1) f_1(t, x_1)) \\ = D \Delta_{x_1} f_1(t, x_1). \end{aligned} \quad (13)$$

Notice that the assumption (12) is equivalent to particles being uncorrelated. In other words, MFA does not take into account the effects of correlations. Taking the limit as $N \rightarrow \infty$ results in the coefficient $\frac{(N-1)}{N}$ being dropped and yields the McKean–Vlasov equation,

$$\begin{aligned} \partial_t f_1(t, x_1) + \nabla_{x_1} \cdot \left(\int u(x_1, y) f_1(t, y) dy f_1(t, x_1) \right) + \nabla_{x_1} \cdot (S(x_1) f_1(t, x_1)) \\ = D \Delta_{x_1} f_1(t, x_1). \end{aligned} \quad (14)$$

It was shown in [13] that Eq (14) is well-posed for smooth and bounded $u(x, y)$.

The McKean–Vlasov equation (14) can also be understood as follows. Write the BBGKY hierarchy for $N = \infty$, that is, the hierarchy is an infinite system of coupled PDEs for $f_1, f_2, \dots$. Assume in addition that all particles are initially independent:

$$f_n(0, x_1, \dots, x_n) = \prod_{i=1}^n f_1(0, x_i), \quad n \geq 1. \quad (15)$$

Then one can show that

$$f_n(t, x_1, \dots, x_n) = \prod_{i=1}^n f_1(t, x_i) \text{ for all } t \geq 0 \text{ and } n \geq 1. \quad (16)$$

This is so called *propagation of chaos*: if particles are initially independent (no correlations, “chaotic”), then they stay independent as time evolves. The notion of propagation of chaos was introduced in [22] and was rigorously justified in the limit $N \rightarrow \infty$ in the framework of Newtonian systems, that is, for the second order in time analogue of (1), for both stochastic [26] and deterministic [6] systems with Lipschitz interaction force F , whereas the first order dynamics with stochasticity was considered in [27, 34]. For further review and extensions of these results we refer to [44, 45].

Therefore, MFA assumption holds exactly in the limit $N \rightarrow \infty$ and it implies that correlations in the system (1) are negligible. Since we are interested in capturing how correlations affect the evolution of f_1 , we must go beyond MFA.

Kirkwood superposition approximation KSA is a truncation of the BBGKY hierarchy at the equation for f_2 . KSA is based on the following representation ansatz for f_3 in terms of f_1 and f_2 :

$$f_3(t, x_1, x_2, x_3) = \frac{f_2(t, x_1, x_2) f_2(t, x_2, x_3) f_2(t, x_1, x_3)}{f_1(t, x_1) f_1(t, x_2) f_1(t, x_3)}. \quad (17)$$

Substitute this approximation into (8) and obtain the following equation for f_2 ,

$$\begin{aligned} & \partial_t f_2(t, x_1, x_2) + \frac{1}{N} \nabla_{x_1} \cdot (u(x_1, x_2) f_2(t, x_1, x_2)) + \frac{1}{N} \nabla_{x_2} \cdot (u(x_2, x_1) f_2(t, x_1, x_2)) \\ & + \frac{N-2}{N} \nabla_{x_1} \cdot \int u(x_1, y) \frac{f_2(t, x_1, x_2) f_2(t, x_1, y) f_2(t, x_2, y)}{f_1(t, x_1) f_1(t, x_2) f_1(t, y)} dy \\ & + \frac{N-2}{N} \nabla_{x_2} \cdot \int u(x_2, y) \frac{f_2(t, x_1, x_2) f_2(t, x_1, y) f_2(t, x_2, y)}{f_1(t, x_1) f_1(t, x_2) f_1(t, y)} dy \\ & + \nabla_{x_1} \cdot (S(x_1) f_2(t, x_1, x_2)) + \nabla_{x_2} \cdot (S(x_2) f_2(t, x_1, x_2)) \\ & = D(\Delta_{x_1} f_2(t, x_1, x_2) + \Delta_{x_2} f_2(t, x_1, x_2)). \end{aligned} \quad (18)$$

Note that KSA representation ansatz (17) can be formally derived from the maximization of a truncated entropy functional [41]. The derivation in [41] is based on the observation that the equilibrium N -particle pdf f_N minimizes the Helmholtz free energy

$$\begin{aligned} & \frac{1}{N} \iiint_{\mathbb{R}^N} \sum_{i,j} \hat{U}(x_j - x_i) f_N(x_1, \dots, x_N) dx_1 \dots dx_N \\ & + D \iiint_{\mathbb{R}^N} f_N(x_1, \dots, x_N) \ln(f_N(x_1, \dots, x_N)) dx_1 \dots dx_N, \end{aligned} \quad (19)$$

subject to normalization constraint $\iiint_{\mathbb{R}^N} f_N dx_1 \dots dx_N = 1$. Here $-\hat{U}$ is the potential of interaction force $\hat{u}$, that is $\hat{u} = -\hat{U}'$. Since the goal is to find an expression for f_N in terms of f_2 , the next step in [9] and [41] is to assume that f_2 is given and to replace the normalization constraint by:

$$f_2(x_i, x_j) = \iiint f_N(x_1, \dots, x_N) \prod_{k \neq i, j} dx_k, \quad i, j = 1, \dots, N, \quad i \neq j. \quad (20)$$

Then the first term in (19) can be rewritten in terms of f_2 only and thus does not influence the minimization subject to (20), so one can replace (19) by the maximization of the entropy functional:

$$H(f_N) = - \iiint f_N \ln(f_N) dx_1 \dots dx_N \quad \text{subject to (20)}.$$

The idea of the truncated entropy maximization principle in [41] is to maximize $H(f_3)$ instead of $H(f_N)$. This substitution of f_N by f_3 introduces an approximation error but leads to the KSA representation of f_3 in terms of f_2 ; for further details we refer to [41]. This method can be applied to find similar approximations for f_n , $n > 3$, however the numerical cost of solving the associated PDEs becomes prohibitive.

Truncation approximation TA is obtained from the following observation. Consider $i = 1$ and rewrite the first term in the sum (9) by substituting conditions $X_1(t) = x_1$ and $X_2(t) = x_2$:

$$\mathcal{F}_1(t, x_1, x_2) = \frac{1}{N}u(x_1, x_2) + \mathbb{E}\left\{\frac{1}{N}\sum_{j \neq 1,2}^N u(X_1(t), X_j(t)) \left\| \begin{matrix} X_1(t) = x_1 \\ X_2(t) = x_2 \end{matrix} \right\| \right\} \quad (21)$$

Next observe that the sum in (21) does not have a term depending on $X_2(t)$ and thus it is natural to assume that the dependence of the expected value in (21) on the condition $X_2(t) = x_2$ is weak and therefore can be ignored. This observation leads to the following approximation for $\mathcal{F}_1$:

$$\mathcal{F}_1(t, x_1, x_2) = \frac{1}{N}u(x_1, x_2) + \mathbb{E}\left\{\frac{1}{N}\sum_{j \neq 1,2}^N u(X_1(t), X_j(t)) \left\| X_1(t) = x_1 \right\| \right\}. \quad (22)$$

A similar approximation can be written for $\mathcal{F}_2$:

$$\mathcal{F}_2(t, x_1, x_2) = \frac{1}{N}u(x_2, x_1) + \mathbb{E}\left\{\frac{1}{N}\sum_{j \neq 1,2}^N u(X_2(t), X_j(t)) \left\| X_2(t) = x_2 \right\| \right\}. \quad (23)$$

Next we use the definition of conditional probability to rewrite (22) and (23):

$$\mathcal{F}_1(t, x_1, x_2) = \frac{1}{N}u(x_1, x_2) + \frac{N-2}{N} \int u(x_1, y) \frac{f_2(t, x_1, y)}{f_1(t, x_1)} dy, \quad (24)$$

$$\mathcal{F}_2(t, x_1, x_2) = \frac{1}{N}u(x_2, x_1) + \frac{N-2}{N} \int u(x_2, y) \frac{f_2(t, x_2, y)}{f_1(t, x_2)} dy. \quad (25)$$

Substituting (24), (25) into (8) yields the following PDE for f_2 , without f_3 :

$$\begin{aligned} & \partial_t f_2(t, x_1, x_2) + \frac{1}{N} \nabla_{x_1} \cdot (u(x_1, x_2) f_2(t, x_1, x_2)) + \frac{1}{N} \nabla_{x_2} \cdot (u(x_2, x_1) f_2(t, x_1, x_2)) \\ & + \frac{N-2}{N} \nabla_{x_1} \cdot \int u(x_1, y) \frac{f_2(t, x_1, x_2) f_2(t, x_1, y)}{f_1(t, x_1)} dy \\ & + \frac{N-2}{N} \nabla_{x_2} \cdot \int u(x_2, y) \frac{f_2(t, x_1, x_2) f_2(t, x_2, y)}{f_1(t, x_2)} dy \\ & + \nabla_{x_1} \cdot (S(x_1) f_2(t, x_1, x_2)) + \nabla_{x_2} \cdot (S(x_2) f_2(t, x_1, x_2)) \\ & = D(\Delta_{x_1} f_2(t, x_1, x_2) + \Delta_{x_2} f_2(t, x_1, x_2)). \end{aligned} \quad (26)$$

Solutions f_1 and f_2 of the system (5)–(26) satisfy the following key properties of probability distribution functions [5]:

1. f_2 is symmetric with respect to x_1 and x_2 :

$$f_2(t, x_1, x_2) = f_2(t, x_2, x_1). \quad (27)$$

2. f_2 conserves its mass and positivity as time evolves:

$$\int f_2(t, x_1, x_2) dx_1 dx_2 = \int f_2(0, x_1, x_2) dx_1 dx_2, \quad (28)$$

and

$$f_1(t, x_1) \geq 0, f_2(t, x_1, x_2) \geq 0 \text{ if } f_1(0, x_1) \geq 0, f_2(0, x_1, x_2) \geq 0. \quad (29)$$

3. f_1 and f_2 are consistent:

$$f_1(t, x_1) = \int f_2(t, x_1, x_2) dx_2. \quad (30)$$

4. Propagation of chaos: $f_2(t, x_1, x_2) = f_1(t, x_1)f_1(t, x_2)$ where f_1 solves the McKean–Vlasov equation (14) is a solution of (26) in the limit $N \rightarrow \infty$.

It was also shown in [5] that no single representation for f_3 is able to satisfy all four of these properties. Namely, consider a representation for f_3 of the following form:

$$f_3(x_1, x_2, x_3) = \mathcal{F}(f_1(x_1), f_1(x_2), f_1(x_3), f_2(x_1, x_2), f_2(x_2, x_3), f_2(x_1, x_3)), \quad (31)$$

where $\mathcal{F} : \mathbb{R}^6 \rightarrow \mathbb{R}$ is a nonlinear function. Then solutions f_1 and f_2 of system (5), (8) with f_3 replaced by (31) (f_3 enters (8) through $\mathcal{F}_i$; see (10) and (11)) cannot satisfy all the four above listed properties. For example, solutions of KSA, which is derived from a single representation (17), do not satisfy the property of consistency (30). The fact that solutions of TA satisfy (30) implies that we can substitute (30) into (26) to obtain a closed form equation for f_2 .

Comparison between approximations Here we focus on the comparison between TA and KSA since we are most interested in the effect of correlations which MFA neglects. First we present heuristics on how TA and KSA can be derived in a simple way. Consider a triplet of particles 1, 2, and 3 with positions at x_1 , x_2 , and x_3 , respectively. Assume that one studies how particles 2 and 3 affect particle 1. If correlations in the system are not low, then we need to take into account all correlations including the correlation between particles 2 and 3. On the other hand, if overall correlations are not high, then one would expect that the contribution from correlation between particles 2 and 3 only appear at a lower order for particle 1, compared to correlations between particles 1 and 2 as well as 1 and 3. Therefore, take as an approximation assumption that particles 2 and 3 are almost independent:

$$1 \approx \frac{f_2(t, x_2, x_3)}{f_1(t, x_2)f_1(t, x_3)}. \quad (32)$$

Furthermore, using Bayes' Theorem and again independence of particles 2 and 3 one obtains that

$$\begin{aligned} f_3(t, x_1, x_2, x_3) &= f_3(t, x_3|x_1, x_2)f_2(t, x_1, x_2) \\ &\approx f_2(t, x_3|x_1)f_2(t, x_1, x_2) \\ &= \frac{f_2(t, x_1, x_2)f_2(t, x_1, x_3)}{f_1(t, x_1)}. \end{aligned} \quad (33)$$

Here $f_3(t, x_3|x_1, x_2)$ and $f_2(t, x_3|x_1)$ denote conditional pdfs. The formula (33) can serve as an approximation for f_3 with the specific assumption that particles 2 and 3 are almost independent. To extend the formula to the case when a pair from the three particles (not specifically particles 2 and 3) is almost independent, multiply (33) by (32). By doing this we get a representation for f_3 which is symmetric with respect to x_1 , x_2 , and x_3 , and it exactly coincides with KSA representation (17). However, multiplication by (32) introduces an additional approximation error. Instead, TA uses exactly (33) in the equation for f_2 where f_3 appears in $\mathcal{F}_1$, and (33) with the assumption that particles 1 and 3 are almost independent in the term where f_3 appears in $\mathcal{F}_2$. From these observations it follows that TA is more accurate than KSA, and both KSA and TA are more accurate than MFA since they are derived from less restrictive assumptions.

Finally we compare computational complexity between solving KSA and TA. To this end, note that equations for f_2 in KSA and TA only have a difference in the integral terms. For example, the first integral term in these equations looks as follows:

$$\text{KSA: } \int u(x_1, y) \frac{f_2(t, x_1, y) f_2(t, x_2, y)}{f_1(t, y)} dy \frac{f_2(t, x_1, x_2)}{f_1(t, x_1) f_1(t, x_2)}, \quad (34)$$

$$\text{TA: } \int u(x_1, y) f_2(t, x_1, y) dy \frac{f_2(t, x_1, x_2)}{f_1(t, x_1)}. \quad (35)$$

We see that the integral for KSA involves variables x_1 , x_2 , and y , whereas there are only x_1 and y for TA. This allows for the reduction of computational complexity for TA as compared to KSA since at each time step the following integral can be pre-computed:

$$c(x) := \int u(x, y) f_2(t, x, y) dy, \quad (36)$$

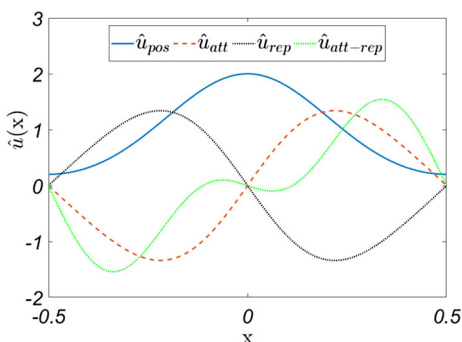
and used in both integral terms of the TA equation for f_2 . Thus, TA is less computationally expensive than KSA.

3 Results of Numerical Simulations

In this section we compare numerical solutions of the continuum approximations MFA, KSA, and TA with direct simulations for various examples of interaction forces $u(X_i, X_j) = \hat{u}(X_j - X_i)$. Throughout this section, by direct simulations we mean Monte Carlo simulations of the individual based system (1). First, we present our results for smooth interaction forces including positive, attraction, repulsion and attraction–repulsion interactions. Next, we consider these continuum approximations for the Morse interaction force and the Kuramoto model. All the interaction forces are introduced below.

In all cases, we consider dynamics of the system of interacting particles for $0 < t < T$ with one-dimensional positions $X_i(t)$ and periodic boundary conditions in $0 \leq x \leq 1$. For the direct simulations of system (1) we use the Euler-Maruyama scheme in order to capture the stochastic term. The time step is $\Delta t = 5 \cdot 10^{-3}$ and the number of realizations is 10^6 . To compute solutions for the continuum approximations a finite difference scheme was used with the spatial and time steps $\Delta x = 10^{-2}$ and $\Delta t = 10^{-5}$, respectively. We note here that KSA has a specific drawback, it does not satisfy the consistency relation between f_1 and f_2 , that is $f_1 \neq \int f_2$. In order to find f_1 for KSA we use Eq. (5). The choice of the number of particles N and magnitude of diffusion D was made so the difference between the continuum approximations is visible and one can draw a conclusion on how the approximation captures properties of the system. For large N , approximations are nearly indistinguishable from each other as well as from the direct simulations, which is consistent with the mean field limit. Thus, we used small N , which in addition allowed us to have a reasonable computational time for the direct simulations, since they require many realizations. The computational cost of each realization depends linearly in N if a more powerful particle method, such as the Fast Multipole Method [17–19], is applied. Since we considered small N the benefit of the Fast Multipole Method would be negligible and a direct explicit method was used. We compare probability distribution functions f_1 and f_2 obtained from the continuum approximations with histograms of the positions of particles and pairs of particles obtained from the direct simulations. Here our focus is on qualitative comparison, such as description of peak formation or convergence to uniform distributions, rather than on quantitative comparison such as L^p errors since they are not informative about the effects of correlations.

Fig. 1 Plot of the smooth interaction forces given by (37)–(40). Note that the sign of $x \cdot \hat{u}(x)$ determines attraction and repulsion, with a positive sign implying attraction and a negative sign signifying repulsion



3.1 Smooth Interaction Forces

We consider cases of positive, attracting and repulsive interaction forces as well as the one which combines short range repulsion and long range attraction. These forces are defined by (2) with $\hat{u}$ given by:

$$\hat{u}_{\text{pos}}(x) = 2e^{-10x^2}, \quad (37)$$

$$\hat{u}_{\text{att}}(x) = 10xe^{-10x^2}, \quad (38)$$

$$\hat{u}_{\text{rep}}(x) = -10xe^{-10x^2}, \quad (39)$$

$$\hat{u}_{\text{att-rep}}(x) = -100x(0.1^2 - x^2)e^{-10x^2}. \quad (40)$$

Note that all these interaction forces are smooth functions (Fig. 1). In particular, they are continuous at 0, unlike, for example, the Morse interaction force, considered in the next subsection. Initial conditions are chosen as follows:

$$f_1(0, x) = 1.0 + 0.4 \sin(2\pi x) \quad 0 \leq x \leq 1, \quad (41)$$

$$f_2(0, x, y) = f_1(0, x)f_1(0, y) \quad 0 \leq x, y \leq 1. \quad (42)$$

In other words, we consider initial one-particle distribution function f_1 as a perturbation of the uniform distribution $f_1 \equiv 1$, and the condition (42) means that the particles are initially independent. Throughout this subsection we set $N = 10$ and $D = 0.005$.

Positive interaction force $\hat{u}_{\text{pos}}$ given by (37) This force acts so that particles exert forces on each other in the positive direction only, that is, particles in front pull particles behind and those behind push those in front. One way to think of this system is unidirectional swimming, for example fishes swimming in a narrow channel. The fish in front will lower the resistance for those behind causing them to swim faster, while the fish behind will push the water around them forward, helping those in front to move faster. Cannibalistic locusts oriented in the same direction in a one dimensional tunnel would also follow this type of interactions, a locust will chase the locusts in front trying to eat them, while running away from those trying to eat it from behind [4,37].

Direct simulations for this system showed that it exhibits interesting qualitative behavior for large times. Namely, particles interacting via the positive force $\hat{u}_{\text{pos}}$ tend to form a cluster which moves with speed close to $\hat{u}_{\text{pos}}(0)$. Nevertheless, the one-particle probability distribution function f_1 converges to a uniform value for large times, that is $f_1 \approx 1$ as $t \rightarrow \infty$. This does not contradict to cluster formation, since as $t \rightarrow \infty$ the many particle system is in

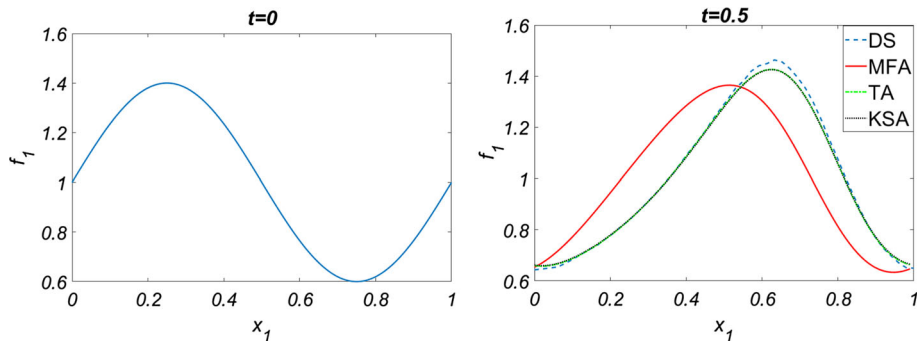


Fig. 2 Left figure: the initial distribution at $t = 0$ for all approximations. Right figure: $f_1(0.5, x)$ from the various approximations to (1) with the positive interaction force $\hat{u}_{\text{pos}}(x) = 2e^{-10x^2}$, $N = 10$, and initial conditions (41)

a highly correlated regime, therefore f_2 concentrates around the diagonal $x_1 = x_2$ and f_1 is essentially the probability distribution function of the cluster location.

In direct simulations we observe that the peak of $f_1(t, x)$ moves to the right and slightly grows as time increases. Motion to the right is because all the particles' velocities in this case are positive, that is, $\dot{X}_i(t) > 0$ (if diffusion is disregarded). The growth of the peak is due to the particles clustering.

In continuum approximations the solution f_1 for both KSA and TA moves to the right at almost the same pace as for the direct simulations and also captures the growth of the peak while MFA moves slower and is unable to capture the growth of the peak, see Fig. 2. This is due to particles clustering and that particles move faster when they are part of a cluster. Since these effects come from correlations, the methods such as TA and KSA, which take into account correlations, capture the speed and the peak growth better than the MFA. In Fig. 3, one can see that like in the direct simulations the two-particle distribution function f_2 computed by TA and KSA has a single non-round peak (the yellow spot), while f_2 in MFA has smaller wider and round peak.

Attracting interaction force $\hat{u}_{\text{att}}$ given by (38) This force results in particles approaching to one another and as time evolves the particles tend to concentrate at a single location determined by initial conditions. As in the case of positive interaction force, for $N < \infty$ one should distinguish between the concentration of many interacting particles and one-particle probability distribution function f_1 . While the particles tend to cluster at a single location, the one-particle probability distribution function does not become a δ -function. Moreover, if initially the distribution f_1 is close to uniform, it stays nearly uniform for all $t > 0$, even though the particles will almost surely form a point cluster. This is because particles tend to concentrate but the point of concentration is random and almost uniformly distributed. On the other hand, for fixed initial f_1 , if N increases, then the one-particle distribution function f_1 eventually (i.e., as $t \rightarrow \infty$) exhibits larger peaks (unless it is initially uniform), and in the limit $N \rightarrow \infty$ it becomes a δ -function. This is consistent with the mean field limit: as $N \rightarrow \infty$ correlations vanish, and the notion of one-particle probability distribution function f_1 coincides with the particles concentration. We also note that the two-particle distribution function $f_2(t, x_1, x_2)$ for all N concentrates along the diagonal $x_1 = x_2$ as $t \rightarrow \infty$.

All three approximations capture the growth of the peak of f_1 at $t = 0.15$, see Fig. 4 (left). The MFA overestimates the peak but KSA and TA both capture it well with TA being

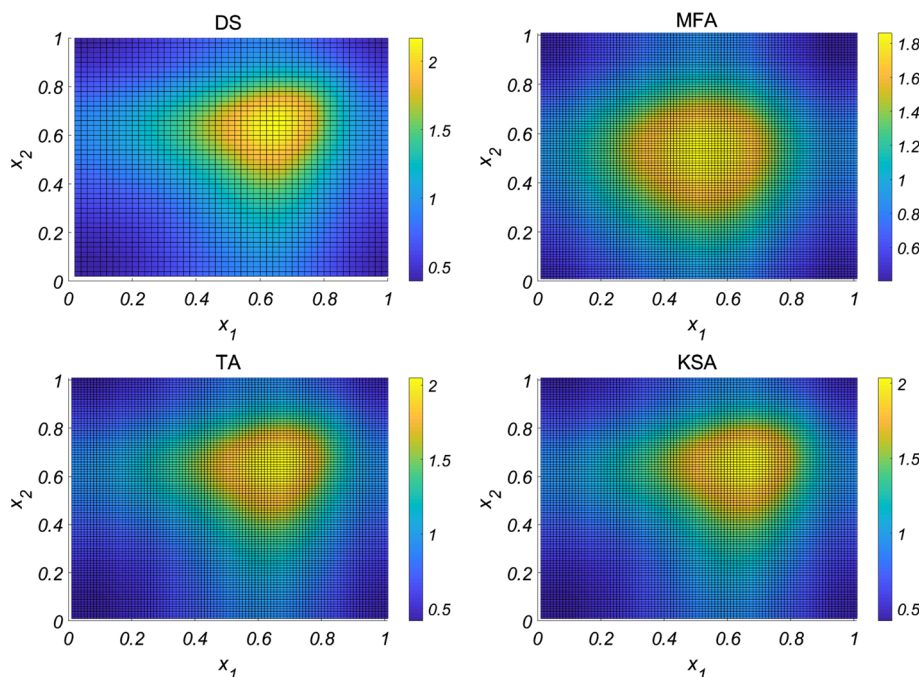


Fig. 3 Approximations of f_2 at $t = 0.5$ with the positive interaction force $\hat{u}_{\text{pos}}$. Top left: direct simulations, top right: mean field, bottom left: truncation approximation, bottom right: Kirkwood superposition approximation

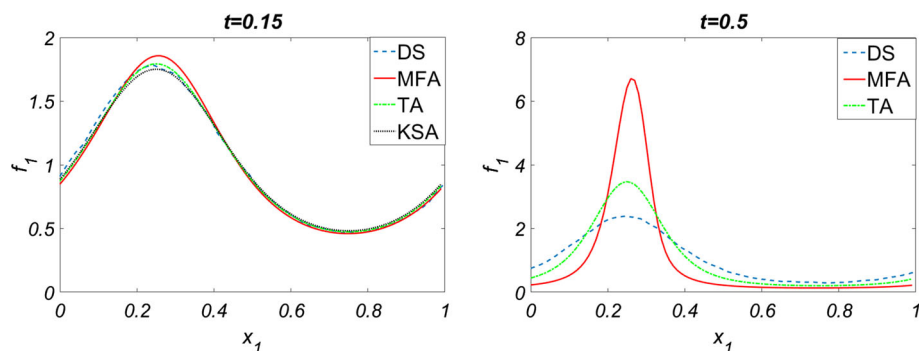


Fig. 4 Approximations of f_1 with the attraction interaction force $\hat{u}_{\text{att}}(x)$, $N = 10$ and initial conditions (41). Left: $t = 0.15$, right: $t = 0.5$

slightly more accurate. Both KSA and TA, unlike MFA, capture large values of f_2 near the diagonal, $x_1 = x_2$ (see Fig. 5: the peak represented by the yellow spot is elongated along the diagonal for KSA and TA, while for MFA it is round). Recall that concentration of f_2 near the diagonal $x_1 = x_2$ means that any two particles are located close to each other. All approximations underestimate the maximum value of f_2 obtained from the direct simulations, with KSA being the closest. At $t = 0.5$ MFA greatly overestimates the growth of the peak of f_1 , TA also overestimates the growth of the peak but is much closer to the direct simulations than MFA, see Fig. 4 (right). Moreover, among the two truncations at

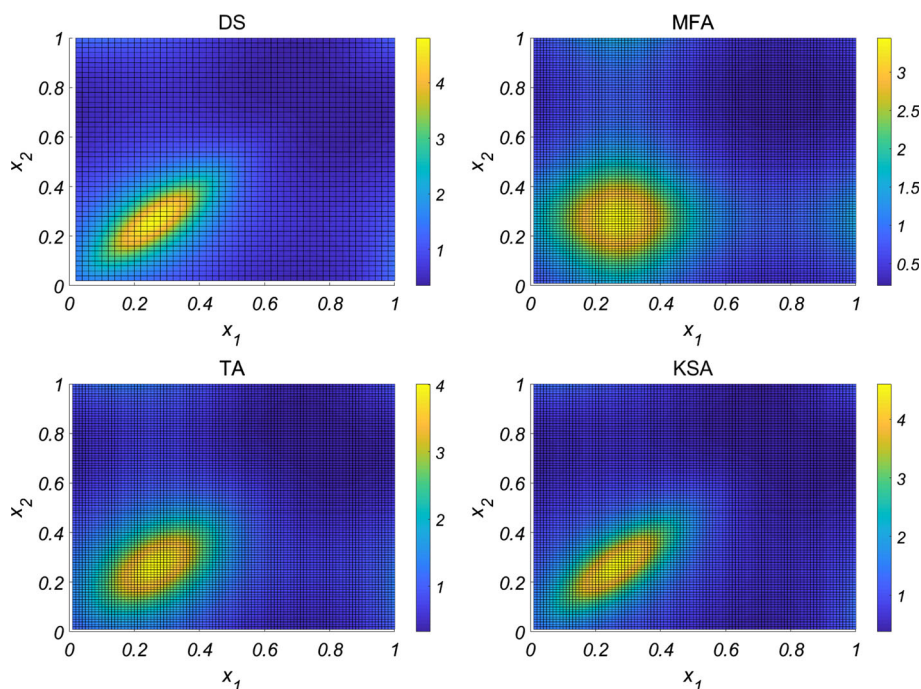


Fig. 5 Approximations of f_2 at $t = 0.15$ with the attraction interaction force $\hat{u}_{\text{att}}$. Top left: direct simulations, top right: mean field approximation, bottom left: truncation approximation, bottom right: Kirkwood superposition approximation

level $k = 2$ considered in this work, TA was capable of producing results with the explicit numerical scheme, while the numerical simulations for KSA became unstable and are not presented. For f_2 TA approximates values obtained from direct simulations near the diagonal better than MFA, see Fig. 6. The maximal value of f_2 in direct simulations is also closer to TA than MFA.

Repulsion interaction force $\hat{u}_{\text{rep}}$ given by (39) This force results in particles pushing away from one another. As time evolves, particles form a lattice with even spacing, where the final locations are determined by initial conditions. The one-particle probability distribution function f_1 becomes uniform as $t \rightarrow \infty$. Repulsion between particles leads to that values of the two-particle probability distribution function f_2 near the diagonal $x_1 = x_2$ decrease with time, and f_2 concentrates at lines $|x_1 - x_2| = \frac{k}{N}$, $k = 1, \dots, N$ as $t \rightarrow \infty$ (these lines are parallel to the diagonal $x_1 = x_2$ but the diagonal is not one of these lines).

All three approximations capture tendency of f_1 to become uniform, see Fig. 7. The KSA and the TA, unlike the MFA, capture that the values of f_2 along the diagonal $x_1 = x_2$ decrease in time, see Fig. 8.

Interaction force with repulsion at short range and attraction at long range $\hat{u}_{\text{att-rep}}$ given by (40) This interaction force results in particles pushing away from one another when the distance between the particles is less than 0.1 and attracting otherwise. Interaction forces which are repulsive at short range and attracting at long range are very common in physics. For example, in order to describe forces between atoms, a variety of such interaction functions introduced via potentials is used, among them the Morse, the Yukawa, and the Lennard–Jones

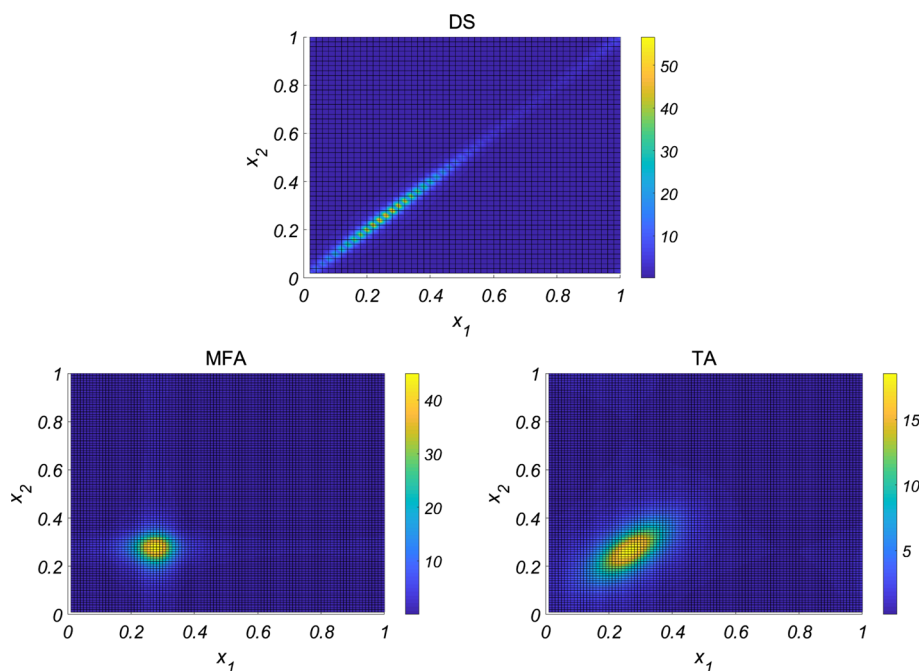
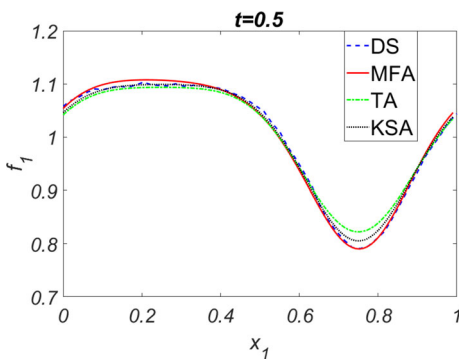


Fig. 6 Approximations of f_2 at $t = 0.5$ with the attraction interaction force $\hat{u}_{\text{att}}$. Top: direct simulations, bottom left: mean field, bottom right: truncation approximation

Fig. 7 Approximations of $f_1(0.5, x_1)$ for the repulsion interaction force $\hat{u}_{\text{rep}}$ and initial conditions (41)



potentials. The Morse potential will be considered in Sect. 3.2. The main difference between $\hat{u}_{\text{att-rep}}$ and these potential forces that $\hat{u}_{\text{att-rep}}$ is smooth at 0 which leads to that the repulsion part of $\hat{u}_{\text{att-rep}}$ is weaker than for the potential forces and hence it can not be considered as a good choice for modeling if interactions between particles at short range are steric, that is, particles have a finite size and do not penetrate each other. However, since equations for f_1 and f_2 contain derivatives of terms with $\hat{u}$, smooth interaction forces are the most convenient for numerical simulations among all short-range-repelling/long-range-attracting interaction forces.

Results of numerical simulations for all approximations of one-particle probability distribution function f_1 with interaction force $\hat{u}_{\text{att-rep}}$ are depicted in Fig. 9. MFA overestimates the peak of f_1 obtained from direct simulations. Note that this observation is similar to the

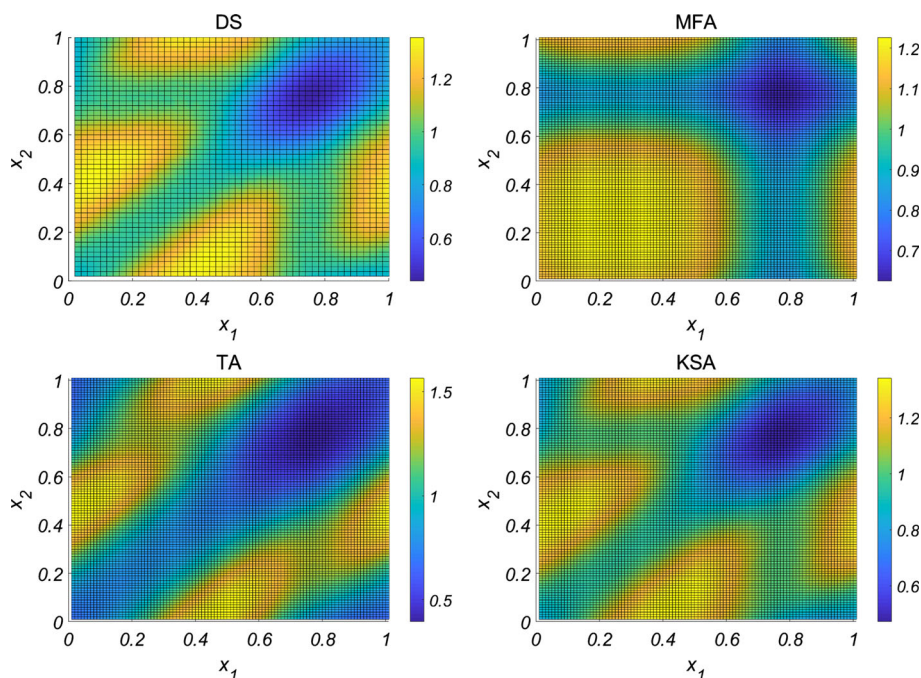
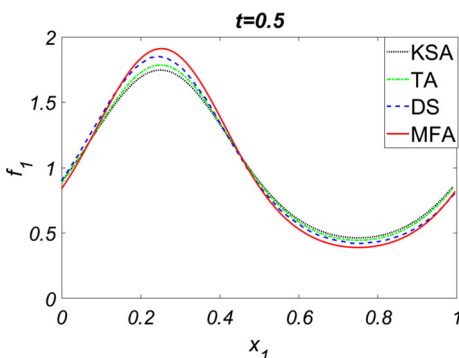


Fig. 8 Approximations of f_2 at $t = 0.5$ for $\hat{u}_{\text{rep}}$. Top left: direct simulations, top right: mean field approximation, bottom left: truncation approximation, bottom right: Kirkwood superposition approximation

Fig. 9 Approximations of $f_1(0.5, x)$ for $\hat{u}_{\text{att-rep}}$ with $N = 10$ where initial conditions are given by (41)



one for the attraction interaction force, see Fig. 4; this is because the attraction component of interactions dominates due to the specific form of $\hat{u}_{\text{att-rep}}$, see Fig. 1. However, TA and KSA underestimate the peak; this is presumably because these approximations overestimate the repulsion at the peak as in Fig. 7. TA, unlike MFA and KSA, captures that f_2 tends to decrease at the diagonal $x_1 = x_2$, see Fig. 10.

3.2 Morse Interaction Force

The Morse interaction force was originally introduced in physics and chemistry to model inter-atomic forces, see e.g. [30,40], and was further used in other disciplines such as for

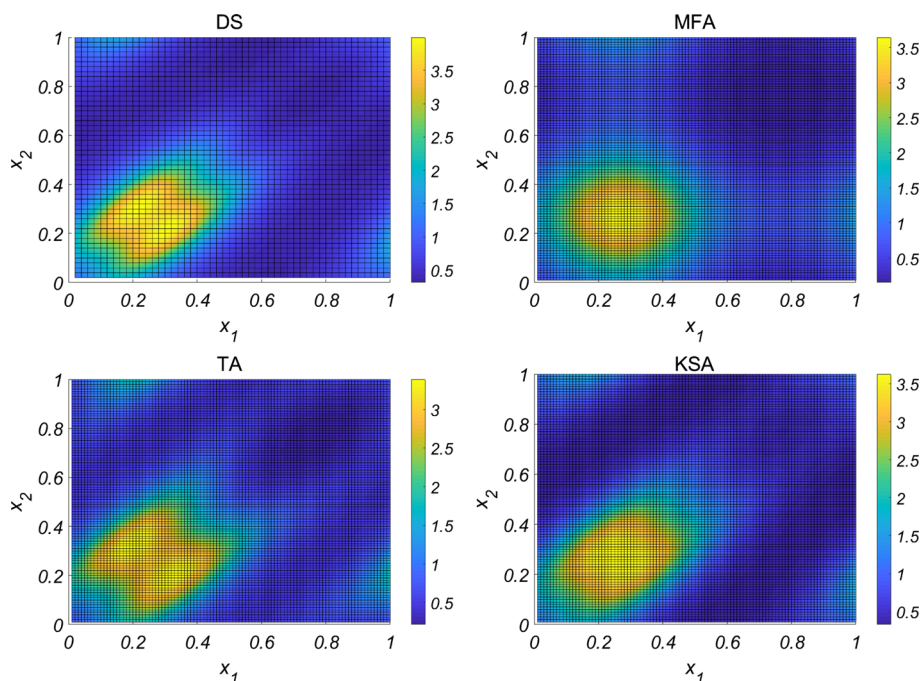


Fig. 10 Approximations of f_2 at $t = 0.5$ for $\hat{u}_{\text{att-rep}}$. Top left: direct simulations, top right: mean field, bottom left: truncation approximation, bottom right: Kirkwood superposition approximation

example mathematical biology, see [28,33]. As in the case of $u_{\text{att-rep}}$ from Sect. 3.1, particles which interact through the Morse interaction force repel each other if they are close and attract otherwise. On the other hand, unlike $u_{\text{att-rep}}$, the repulsion of the Morse interaction force does not vanish as particles approach each other. The growth of repulsion as inter-particle distance goes to zero is relevant if for instance the repulsion component serves to model flexible volume constraints (that is, particles push each other away if they share the same place). Specifically, the system of many particles interacting through the Morse interaction force is

$$dX_i = \sum_{j \neq i} \hat{u}(X_j - X_i) dt + \sqrt{2D} dW_t, \text{ where} \quad (43)$$

$$\hat{u}_M(x) = \begin{cases} 120 \left[e^{-2(|x|-r_e)} - e^{-(|x|-r_e)} \right] \frac{x}{|x|}, & |x| \leq c, \\ 0, & |x| > c. \end{cases} \quad (44)$$

The Morse interaction force is defined in (44), see Fig. 11. The parameter $r_e = 0.1$ is the equilibrium distance, that is, the distance at which the force vanishes, and c is the radius of truncation of the Morse force or in other words c is the range of interactions.

For the numerical simulations of system (43), (44) the following initial conditions were chosen

$$f_1(t = 0, x) = \frac{5}{6} (\tanh(30(x - 0.2)) + \tanh(30(0.8 - x))). \quad (45)$$

We note here that we chose to present results of numerical simulations for initial conditions (45) instead of (41) from Sect. 3.1, since the system (43), (44) for latter conditions showed

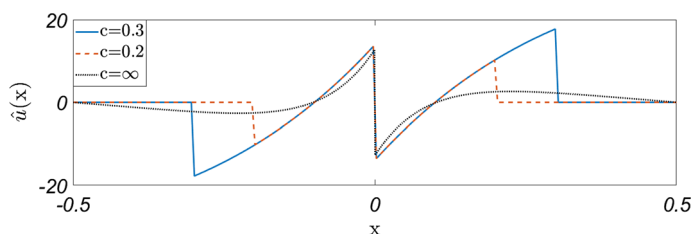
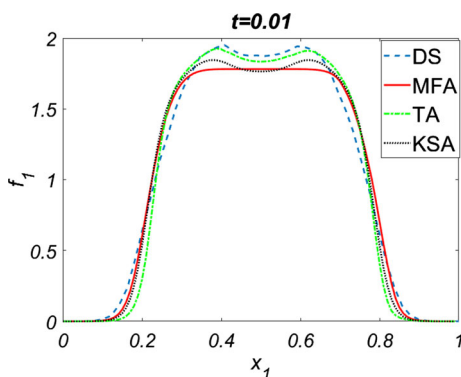


Fig. 11 The Morse interaction force with truncations at $c = 0.2$ and $c = 0.3$. Note that the sign of $x \cdot \hat{u}(x)$ determines attraction and repulsion, with a positive sign implying attraction and a negative sign signifying repulsion. Dashed line depicts the plot for the Morse force for $c = \infty$ (rescaled for better visibility)

Fig. 12 Approximations of $f_1(0.01, x)$ with the Morse interaction force, $c = 0.2$, $D = 0.045$, and initial conditions (45)



very slow dynamics of the probability distribution function f_1 . Visible changes in f_1 with initial condition (45) as time evolves are due to large gradients of f_1 at $x \approx 0.2$ and $x \approx 0.8$.

First we consider the system (43), (44) with $N = 5$, $D = 0.045$ and $c = 0.2$. Two distinct peaks in the plot of f_1 obtained by direct simulations are observed, see Fig. 12. Both TA and KSA capture these peaks, and TA approximates the peaks more accurately. MFA does not exhibit any peaks. In capturing f_2 , both TA and KSA capture the low values along the diagonal lines, $x_1 = x_2$ and $x_1 \approx x_2 \pm 0.2$, whereas MFA does not, see Fig. 13. Additionally, TA captures the maximum values of f_2 better than KSA as well as the narrowness of the peaks' width.

Next consider the system (43), (44) with a larger range of interactions, specifically, $c = 0.3$ with all other parameters remaining the same: $N = 5$ and $D = 0.045$. Note that increasing the range of interactions effectively increases the strength of attraction between particles in the system. In this case the one-particle probability distribution function f_1 has a single peak at center, see Fig. 14. TA and KSA both capture the peak, while MFA does not. Comparing the approximations of f_2 depicted in Fig. 15, we see that both TA and KSA capture low values along the diagonal lines, $x_1 = x_2$ and $x_1 \approx x_2 \pm 0.3$, whereas MFA does not.

3.3 Kuramoto Interaction Force

Among all models used for the description of synchronization phenomena, the Kuramoto model is the most popular one and it was successfully used in various branches of science such as chemistry, physics, neural science, biology and even social science [1,24,35]. In general, this model considers N oscillators so that each oscillator has the phase $X_i(t)$ at time

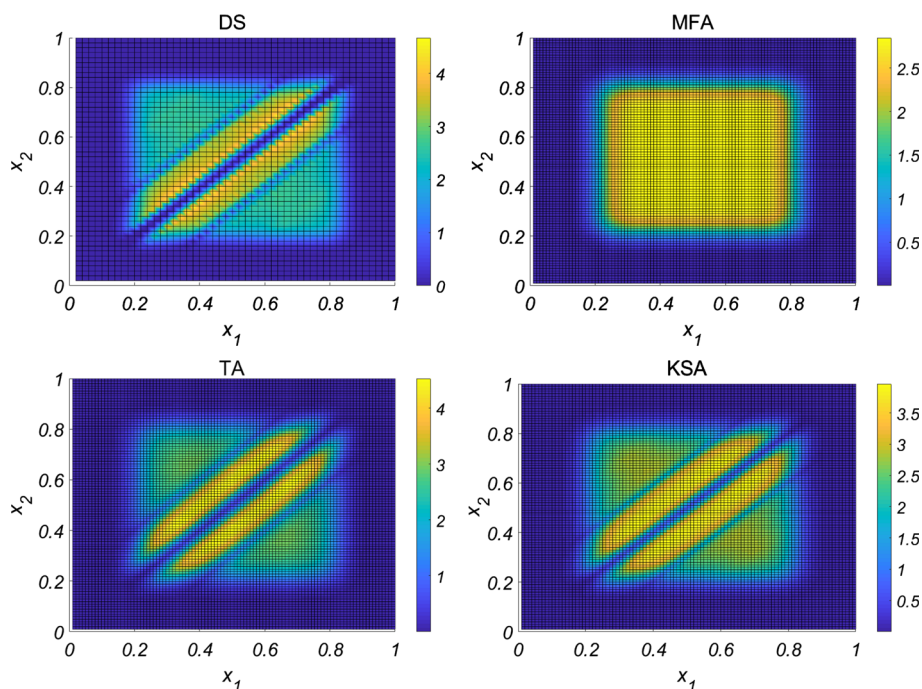
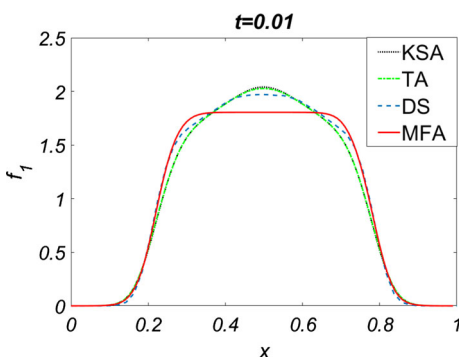


Fig. 13 Approximations of f_2 at $t = 0.01$ with the Morse interaction force and $c = 0.2$. Top left: direct simulations, top right: mean field approximation, bottom left: truncation approximation, bottom right: Kirkwood superposition approximation

Fig. 14 Approximations of $f_1(0.01, x)$ for the system (1) with the Morse interaction force, $c = 0.3$, $D = 0.045$, and initial conditions (45)



t , and the oscillators are coupled by the following attracting interaction force (which we call here the Kuramoto interaction force):

$$\hat{u}_K(x) = K \sin(2\pi x). \quad (46)$$

The sub-index K in the left hand side of (46) stands for “Kuramoto” and the parameter $K = 2.0$ in the right hand side of (46) is the strength of interactions.

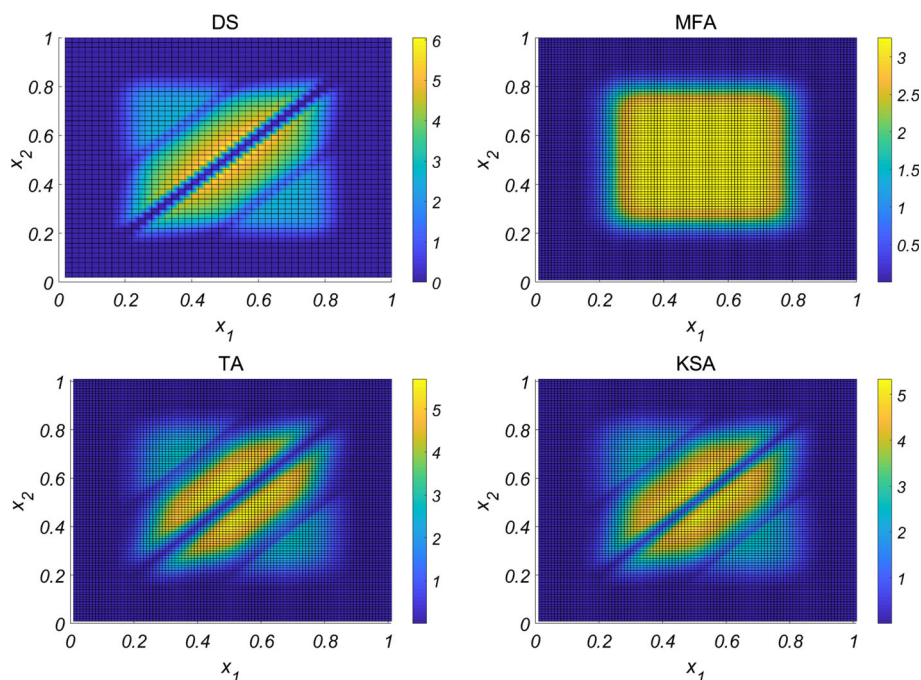


Fig. 15 Approximations of f_2 at $t = 0.01$ with the Morse interaction force and $c = 0.3$. Top left: direct simulations, top right: mean field approximation, bottom left: truncation approximation, bottom right: Kirkwood superposition approximation

A distinguishing feature of the Kuramoto model is that in addition to pairwise interactions, each oscillator also has a given intrinsic frequency w_i . The resulting individual based system is

$$dX_i(t) = w_i dt + \frac{1}{N} \sum_{j=1}^N \hat{u}_K(X_i - X_j) dt + \sqrt{2D} dW_i(t), \quad \text{for } i = 1, \dots, N. \quad (47)$$

The interactions are purely attractive, therefore the particles tend to occupy a single location at each moment of time moving with the same frequency/velocity. However, if values of frequencies w_i are high, then they dominate the attractive interactions and in this case the one-particle probability distribution function becomes uniform.

In numerical simulations we choose $N = 5$ and $D = 0.045$. The frequencies w_i are random variables, independently and identically distributed with the uniform distribution on $(-1, 1)$.

From Fig. 16 we see that both TA and MFA exhibit a peak in f_1 and TA is more accurate than MFA. TA also approximates direct simulations more accurately than MFA. Specifically, from Fig. 17 we see that while the MFA distribution converges towards uniform, both direct simulations and TA, which takes into account correlations, concentrate on diagonal $x_1 = x_2$. It suggests that even when MFA does not predict synchronization of oscillators, they may still exhibit collective motion (roughly speaking, a weak synchronization) since correlations in a system of a finite number of interacting particles inevitably appear. TA, capturing correlations, may potentially serve for analysis of such “weakly” synchronized states.

Fig. 16 Approximations of $f_1(0.025, x)$ for the system (47) with the Kuramoto interaction force and initial conditions given by (41)

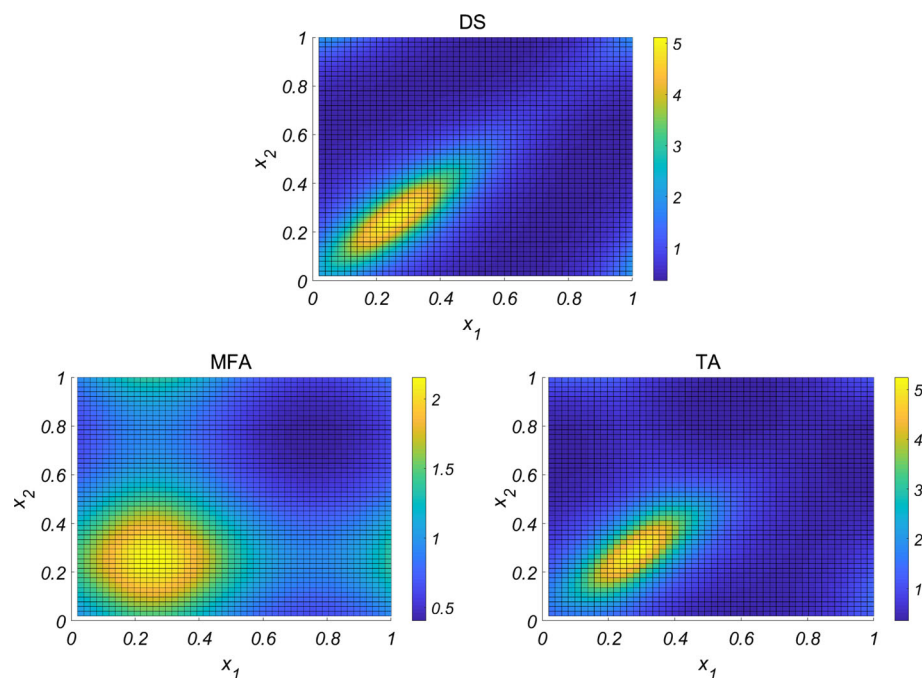
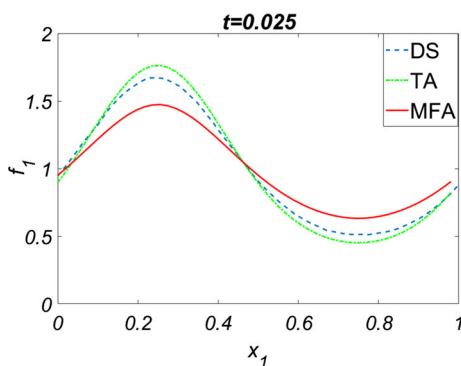


Fig. 17 Approximations of f_2 at $t = 0.025$ with the Kuramoto interaction force. Top: direct simulations, bottom left: mean field approximation, bottom right: truncation approximation

KSA was not used here since the introduction of intrinsic frequencies significantly increases computational complexity of KSA.

4 Conclusions

In this paper three continuum approximations of a system of interacting particles—MFA, Kirkwood superposition approximation, and truncation approximation—were tested and compared to direct simulations of the individual based system for various types of interactions. It was shown that in all the considered cases TA and KSA performed noticeably

better than MFA. When comparing TA and KSA, TA performed significantly better for all tested interactions except the repulsive interaction. For attractive interactions TA was stable while KSA was not. The major advantage TA had over KSA was the computational complexity. Due to the form of integral terms TA has significantly shorter computational time. This advantage becomes more important as the dimension of a problem is increased. In comparison to direct simulations, continuum approximations are faster as they do not depend on the number of particles N and do not require many realizations to take into account randomness of initial particles' locations.

Acknowledgements PEJ was partially supported by NSF Grant 1614537, and NSF Grant RNMS (Ki-Net) 1107444. LB and MP were supported by NSF DMREF Grant DMS-1628411.

References

1. Acebrón, J.A., Bonilla, L.L., Vicente, C.J., Ritort, F., Spigler, R.: The Kuramoto model: a simple paradigm for synchronization phenomena. *Rev. Mod. Phys.* **77**(1), 137 (2005)
2. Aldous, D.J.: Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists. *Bernoulli* **5**(1), 3–48 (1999)
3. Baker, R.E., Simpson, M.J.: Correcting mean-field approximations for birth-death-movement processes. *Phys. Rev. E* **82**, 041905 (2010)
4. Bazazi, S., Buhl, J., Hale, J.J., Anstey, M.L., Sword, G.A., Simpson, S.J., Couzin, I.D.: Collective motion and cannibalism in locust migratory bands. *Curr. Biol.* **18**(10), 735–739 (2008)
5. Berlyand, L., Jabin, P.-E., Potomkin, M.: Complexity reduction in many particle systems with random initial data. *SIAM/ASA J. Uncertain. Quantif.* **4**(1), 446–474 (2016)
6. Braun, W., Hepp, K.: The Vlasov dynamics and its fluctuations in the $1/N$ limit of interacting classical particles. *Commun. Math. Phys.* **56**(2), 101–113 (1977)
7. Bolley, F., Guillin, A., Villani, C.: Quantitative concentration inequalities for empirical measures on non-compact spaces. *Probab. Theory Relat. Fields* **137**(3–4), 541–593 (2007)
8. Budhiraja, A., Dupuis, P., Fischer, M.: Large deviation properties of weakly interacting processes via weak convergence methods. *Ann. Probab.* **40**(1), 74–102 (2012)
9. Bugaenko, N.N., Gorban', A.N., Karlin, I.V.: Universal expansion of three-particle distribution function. *Theor. Math. Phys.* **88**(3), 977–985 (1991)
10. Cole, G.H.A.: Classical fluids and the superposition approximation. *Rep. Prog. Phys.* **31**(2), 419 (1968)
11. Dawson, D.A., Gärtner, J.: Large deviations from the McKean–Vlasov limit for weakly interacting diffusions. *Stochastics* **20**(4), 247–308 (1987)
12. Dawson, D.A., Gärtner, J.: Multilevel large deviations and interacting diffusions. *Probab. Theory Relat. Fields* **98**, 423–487 (1994)
13. Dobrushin, R.L.: Vlasov equations. *Funct. Anal. Appl.* **13**(2), 115–123 (1979)
14. Egelstaff, P.A., Page, D.I., Heard, C.R.T.: Experimental study of the triplet correlation function for simple liquids. *J. Phys. C* **4**(12), 14–53 (1971)
15. Gärtner, J.: On the McKean–Vlasov limit for interacting diffusions. *Math. Nachr.* **137**, 197–248 (1988)
16. Goodman, J., Hou, T.Y.: New stability estimates for the 2-D vortex method. *Commun. Pure Appl. Math.* **44**(8–9), 1015–1031 (1991)
17. Greengard, L., Rokhlin, V.: A fast algorithm for particle simulation. *J. Comput. Phys.* **73**, 325–348 (1987)
18. Greengard, L., Rokhlin, V.: Rapid Evaluation of Potential Fields in Three Dimensions, in *Vortex Methods*, Lecture Notes in Mathematics, pp. 121–141. Springer, New York (1988)
19. Greengard, L., Rokhlin, V.: On the evaluation of electrostatic interactions in molecular modeling. *Chem. Scr.* **29A**, 139–144 (1989)
20. Jabin, P.-E.: A review of the mean field limits for Vlasov equations. *Kinet. Relat. Models* **7**(4), 661–711 (2014)
21. Jabin, P.E., Wang, Z.: Mean Field Limit for Stochastic Particle Systems. *Active Particles*. Birkhäuser, Cham (2017)
22. Kac, M.: Foundations of kinetic theory. In: *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955*, pp. 3, 171–197. University of California Press, Berkeley and Los Angeles (1956)
23. Kirkwood, J.G.: Statistical mechanics of fluid mixtures. *J. Chem. Phys.* **3**, 300–313 (1935)

24. Kuramoto, Y.: Self-entrainment of a population of coupled non-linear oscillators. In: International Symposium on Mathematical Problems in Theoretical Physics, pp. 420–422. Springer, Berlin (1975)
25. Leutheusser, E.: Dynamics of a classical hard-sphere gas I. Formal theory. *J. Phys. C* **15**(13), 2827–2843 (1982)
26. McKean Jr., H.P.: Propagation of Chaos for a Class of Non-linear Parabolic Equations. Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic University), pp. 41–57. Air Force Office of Scientific Research, Arlington (1967)
27. Méléard, S., Roelly-Coppoletta, S.: A propagation of chaos result for a system of particles with moderate interaction. *Stoch. Process. Appl.* **26**, 317–332 (1987)
28. Middleton, A., Fleck, C., Grima, R.: A continuum approximation to an off-lattice individual-cell based model of cell migration and adhesion. *J. Theor. Biol.* **359**, 220–232 (2014)
29. Morale, D., Capasso, V., Oelschläger, K.: An interacting particle system modeling aggregation behavior: from individuals to populations. *J. Math. Biol.* **50**(1), 49–66 (2005)
30. Morse, P.M.: Diatomic molecules according to the wave mechanics. II. Vibrational levels. *Phys. Rev.* **34**(1), 57 (1929)
31. Motsch, S., Tadmor, E.: Heterophilious dynamics enhances consensus. *SIAM Rev.* **56**(4), 577–621 (2014)
32. Nagy, M., Vásárhelyi, G., Pettit, B., Roberts-Mariani, I., Vicsek, T., Biro, D.: Context-dependent hierarchies in pigeons. *Proc. Natl. Acad. Sci. USA* **110**, 13049–13054 (2013)
33. Newmann, T.J.: Modeling multicellular systems using subcellular elements. *Math. Biosci. Eng.* **2**(3), 611–622 (2005)
34. Oelschläger, K.: A martingale approach to the law of large numbers for weakly interacting stochastic processes. *Ann. Probab.* **12**, 458–479 (1984)
35. Pikovsky, A., Rosenblum, M., Kurths, J.: Synchronization: A Universal Concept in Nonlinear Sciences, vol. 12. Cambridge University Press, Cambridge (2003)
36. Rice, S.A., Lekner, J.: On the equation of state of the rigid-sphere fluid. *J. Chem. Phys.* **42**(10), 3559–3565 (1965)
37. Romanczuk, P., Couzin, I.D., Schimansky-Geier, L.: Collective motion due to individual escape and pursuit response. *Phys. Rev. Lett.* **102**(1), 010602 (2009)
38. Ryan, S.D., Haines, B.M., Berlyand, B.M., Ziebert, L., Aranson, F.: Viscosity of bacterial suspensions: hydrodynamic interactions and self-induced noise. *Phys. Rev. E* **83**, 050904(R) (2011)
39. Ryan, S.D., Berlyand, L., Haines, B.M., Karpeev, D.: A kinetic model for semidilute bacterial suspensions. *SIAM Multiscale Model. Simul.* **11**(4), 1176–1196 (2013)
40. Schiff, L.I.: Quantum Mechanics, vol. 3, pp. 61–62. McGraw-Hill, New York (1968)
41. Singer, A.: Maximum entropy formulation of the Kirkwood superposition approximation. *J. Chem. Phys.* **121**, 36–57 (2004)
42. Sokolov, A., Aranson, I.S.: Physical properties of collective motion in suspensions of bacteria. *Phys. Rev. Lett.* **109**, 248109 (2012)
43. Sokolov, A., Aranson, I.S., Kessler, J.O., Goldstein, R.E.: Concentration dependence of the collective dynamics of swimming bacteria. *Phys. Rev. Lett.* **98**, 158102 (2007)
44. Spohn, H.: Large Scale Dynamics of Interacting Particles. Springer, Berlin (1991)
45. Sznitman, A.-S.: Topics in propagation of chaos, in *École d'Été de Probabilités de Saint-Flour XIX - 1989. Lecture Notes in Mathematics*, pp. 165–251. Berlin, Springer (1991)