



Reduced Variance Random Batch Methods for Nonlocal PDEs

Lorenzo Pareschi^{1,2} · Mattia Zanella³

Received: 2 January 2024 / Accepted: 4 May 2024 / Published online: 9 May 2024
© The Author(s) 2024

Abstract

Random Batch Methods (RBM) for mean-field interacting particle systems enable the reduction of the quadratic computational cost associated with particle interactions to a near-linear cost. The essence of these algorithms lies in the random partitioning of the particle ensemble into smaller batches at each time step. The interaction of each particle within these batches is then evolved until the subsequent time step. This approach effectively decreases the computational cost by an order of magnitude while increasing the amount of fluctuations due to the random partitioning. In this work, we propose a variance reduction technique for RBM applied to nonlocal PDEs of Fokker-Planck type based on a control variate strategy. The core idea is to construct a surrogate model that can be computed on the full set of particles at a linear cost while maintaining enough correlations with the original particle dynamics. Examples from models of collective behavior in opinion spreading and swarming dynamics demonstrate the great potential of the present approach.

Keywords Random batch methods · Control variate methods · Surrogate models · Collective behavior · Nonlocal PDEs

1 Introduction

Meanfield equations are derived as reduced complexity models describing the dynamics of large systems of interacting particles [27]. Typically, the trajectories of systems of particles are governed by stochastic differential equations for the position $x_i \in \mathbb{R}^d$ and velocity $v_i \in \mathbb{R}^d$ of the i th particle and are translated in systems of stochastic differential equations of the

Dedicated to our dear friend and colleague Shi on his 60th birthday.

✉ M. Zanella
mattia.zanella@unipv.it

L. Pareschi
l.pareschi@hw.ac.uk

¹ Maxwell Institute for Mathematical Sciences & Department of Mathematics, Heriot-Watt University, Edinburgh, UK

² Department of Mathematics and Computer Science, University of Ferrara, Ferrara, Italy

³ Department of Mathematics, University of Pavia, Pavia, Italy

form

$$\begin{cases} dx_i = v_i dt, \\ dv_i = \left(\frac{1}{N} \sum_{j=1}^N P(x_i, x_j, v_i, v_j)(v_j - v_i) \right) dt + \sqrt{2D(v_i)} dW_i^t, \end{cases} \quad i = 1, \dots, N, \quad (1)$$

where $P(\cdot) \geq 0$ is a suitable interaction function. In the limit of a large number of particles N , the individual trajectories become irrelevant, and we have the opportunity to describe the average behavior of the system. In such a situation, the strength of the individual interactions becomes small, and each particle feels the effect of the force field created by all the other particles. Therefore, the statistical behavior of the system of particles is obtained by studying the evolution of the particles' density $f(x, v, t)$ which obeys a nonlocal PDE of Fokker-Planck type

$$\partial_t + v \cdot \nabla_x f = \nabla_v \cdot [\mathcal{B}[f](x, v, t) f(x, v, t)] + \Delta_v (D^2(v) f), \quad (2)$$

where the term $D(v) : \mathbb{R}^d \rightarrow \mathbb{R}$ expresses the local relevance of the diffusion and the nonlocal drift term $\mathcal{B}[f](x, v, t)$ is defined as follows

$$\mathcal{B}[f](x, v, t) = \int_{\mathbb{R}^d \times \mathbb{R}^d} P(x, y, v, w)(v - w) f(y, w, t) dw dy.$$

The derivation of the mean-field model (2) is essential to investigate mathematically the long time behavior and the formation of collective structures and self-organisation features of the multi-agent system (1). Indeed, since the trajectory of each particle is influenced by all the other particles, the computational complexity of the particle model is $O(N^2)$ and makes the microscopic approach rapidly infeasible for $N \gg 0$.

In this direction, Random Batch Methods (RBM) have gained increased popularity in the community of interacting particle systems since they are capable to efficiently reduce the computational complexity to $O(MN)$, where M is the size of the batch, at the cost of introducing an additional error in the dynamics [12, 28, 29, 32]. In the context of many-particle systems, RBM methods have been proposed in [30]. Earlier approaches based on analogous Monte Carlo strategies were developed in [1] and [7, 11] in the presence of uncertain quantities. The core of these algorithms is to randomly divide the set of particles into smaller batches at each time step, a technique closely resembling the batch methods of stochastic gradient descent algorithms. The interaction of each particle within these batches is then evolved until the subsequent time step. This approach significantly reduces the computational cost by an order of magnitude, albeit at the expense of heightened fluctuations due to the random partitioning.

In contrast with random search algorithms where the increase of fluctuations may also have a positive impact on the overall process, in RBM methods, since we aim at computing as accurate as possible the $O(N^2)$ summation, reducing the variance of the batch algorithm is of paramount importance and represents one of the main challenges. In this paper, inspired by similar ideas in the field of uncertainty quantification for kinetic and mean-field equations [15, 16, 34], we propose a control variate strategy based on a simpler surrogate model to reduce batch variance. However, compared to uncertainty quantification, where due to the static nature of stochastic samples it is easy to obtain correlations, here the samples, although initially correlated, naturally decorrelate as a consequence of their temporal evolution. For this purpose, we employ a reduced-complexity surrogate that can be calculated at the cost of $O(N)$, sharing the same asymptotic state as the original model.

We show that the present strategy permits obtaining a significant reduction in the variance of the batch, leading to a further speed-up in computational cost for a given accuracy. This is illustrated using models of collective behavior in opinion spreading and swarming dynamics.

The rest of the manuscript is organized as follows. In Sect. 2, we revise some nonlocal Fokker-Planck models that will be used as prototype examples to illustrate our approach and to test the numerical schemes. Next, Sect. 3 is devoted to presenting the reduced variance RBM methodology. Several numerical examples are then given in Sect. 4. The last section contains some concluding remarks.

2 Vlasov-Fokker-Planck Models

In this section we survey some examples of multi-agent systems whose mean-field limit corresponds to a Vlasov-Fokker-Planck model of the type (2). First, we define the classical model for transport-aggregation-diffusion. Hence, we discuss the recent extensions of this model to kinetic alignment models in \mathbb{R}^d . In the space homogeneous setting, we also survey recent examples presented in opinion dynamics that have attracted great interest in the mathematical community studying systems of interacting particles.

2.1 The Classical Vlasov-Fokker-Planck Equation

Fokker-Planck-type collision operators are commonly introduced to define reduced complexity models for collisional dynamics, see [39, 41]. The classical Vlasov-Fokker-Planck equation can be obtained from (2) with the choice

$$D^2(v) = \sigma^2 \in \mathbb{R}_+, \quad P(x, y, v, w) \equiv 1,$$

such that

$$\mathcal{B}[f](x, v, t) = v - u,$$

where

$$u = \int_{\mathbb{R}^{2d}} v f(x, v, t) dv dx, \quad \sigma^2 = \frac{1}{d} \int_{\mathbb{R}^{2d}} |v - u|^2 f(x, v, t) dv dx,$$

are the conserved global mean and temperature respectively. We may express the global Maxwellian as follows

$$f^\infty(x, v) = \left(\frac{1}{2\pi\sigma^2} \right)^{d/2} \exp \left\{ -\frac{|v - u|^2}{2\sigma^2} \right\},$$

where we assumed that the total mass is normalized to one. The study of equilibration rates on the velocity space, given a initial distribution f_0 , is a key problem in kinetic theory, see e.g. [3, 6, 39]. The trends to equilibrium are determined in terms of the Boltzmann H-functional

$$H[f](t) = \int_{\mathbb{R}^{2d}} f \log f dv dx.$$

In particular, starting from an initial distribution with finite entropy, $f(x, v, t)$ converges in relative entropy to the equilibrium solution exponentially fast.

2.2 Vlasov-Fokker-Planck Models in Collective Dynamics

Vlasov-Fokker-Planck-type equations have been used more recently to understand long-range interactions in collective phenomena. In this direction, a key role has been played by the so-called swarming dynamics in which large systems of agents tend to share the same velocity for large times, see [14]. The next example focuses on a kinetic equation for swarming dynamics. The kinetic equation can be obtained from (2) with the choice

$$P(x, y, v, w) = \frac{K(v - w)}{(\xi^2 + |x - y|^2)^\beta},$$

such that the nonlocal drift assumes the form

$$\mathcal{B}[f](x, v, t) = \int_{\mathbb{R}^{2d}} \frac{K(v - w)}{(\xi^2 + |x - y|^2)^\beta} f(y, w, t) dw dy.$$

The resulting PDE corresponds to the mean-field limit of the well-known Cucker-Smale model [8, 9, 23, 24, 38]. Several results provide conditions for the asymptotic alignment of velocities under suitable assumptions on the coefficients of the model. In particular, if $\beta < 1/2$ we have the large time collapse of the support of the emerging Maxwellian distribution and therefore asymptotic alignment of the agents' velocities [19].

The emergence of collective structures by means of space-dependent interaction forces can be also observed in multi-agent models for milling [10, 17] and synchronization [13].

2.3 Alignment Models in Consensus Dynamics

The mathematical modelling of consensus formation has seen a growing interest in recent years [21]. Classical methods of statistical mechanics served as a powerful basis to model interactions in cooperative multi-agent systems and to reduce the complexity linked to interaction forces. The study of alignment dynamics has produced a heterogeneity of models, see [2, 18, 20, 21, 25, 33, 40] and the references therein. More recently these approaches have found interesting applications in optimization and data science [26, 37].

We consider the bounded confidence model for opinion formation

$$\partial_t f(v, t) = \partial_v \left[\int_{[-1, 1]} P(|v - w|)(v - w) f(w, t) dw f(v, t) \right], \quad v \in [-1, 1], \quad (3)$$

where

$$P(|v - w|) = \chi(|v - w| \leq \delta) \quad (4)$$

being $\chi(\cdot)$ the indicator function and being $\delta > 0$ a given threshold. The large time behaviour of the model is a sum of Dirac delta distribution which express clusterization of the society towards a finite number of asymptotic opinions.

The corresponding version in presence of noise is given by

$$\partial_t f(v, t) = \partial_v \left[\int_{[-1, 1]} P(|v - w|)(v - w) f(w, t) dw f(v, t) \right] + \sigma^2 \partial_v^2 (D^2(v) f(v, t)), \quad (5)$$

where $v \in [-1, 1]$ and where $D(v) = \sqrt{1 - v^2}$. In the simplified scenario where $\delta = 2$ we get $P \equiv 1$ and we may compute the large time behaviour of the model which corresponds to the Beta distribution

$$f^\infty(v) = \frac{(1 + v)^{\frac{1+m}{\sigma^2} - 1} (1 - v)^{\frac{1-m}{\sigma^2} - 1}}{2^{\frac{2}{\sigma^2} - 1} \mathbf{B}\left(\frac{1+m}{\sigma^2}, \frac{1-m}{\sigma^2}\right)},$$

where $m = \int_{-1}^1 v f(v, 0) dv$ is the conserved mean opinion of the system. If $0 < \delta < 2$ the explicit computation of the asymptotic state is difficult to obtain. A possible strategy to get some insight on the large time trends is based on the construction of structure preserving numerical approximations [35].

3 Reduced Variance RBM

Let $f = f(x, v, t) : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ be a distribution function whose evolution is given by the Vlasov-Fokker-Planck model with nonlocal flux defined in (2) with a constant diffusion $D(v) = \sigma$. The evolution of observable quantities is obtained by rewriting the equation in weak form.

Let $\varphi(\cdot, \cdot)$ be a test function, then the Vlasov-Fokker-Planck model defined in (2) can be written in weak form as follows

$$\begin{aligned} & \frac{d}{dt} \int_{\mathbb{R}^{2d}} \varphi(x, v) f(x, v, t) dv dx + \int_{\mathbb{R}^{2d}} v \cdot \nabla_x \varphi(x, v) f(x, v, t) dv dx \\ &= \int_{\mathbb{R}^{4d}} \nabla_v \varphi(x, v) \cdot P(x, y, v, w) (w - v) f(x, v, t) f(y, w, t) dx dy dv dw + \\ & \int_{\mathbb{R}^{4d}} \Delta_v (D^2(v) \varphi(x, v)) f(x, v, t) f(y, w, t) dx dy dv dw, \end{aligned}$$

provided $f(x, v, t) \geq 0$ is such that for all $t \geq 0$

$$\begin{aligned} \mathcal{B}[f](x, v, t) + \nabla_v (D^2(v) f(x, v, t)) \Big|_{|v| \rightarrow +\infty} &= 0, \\ D^2(v) f(x, v, t) \Big|_{|v| \rightarrow +\infty} &= 0. \end{aligned}$$

Hence, by choosing $\varphi(x, v) \equiv 1$ we may observe that the total mass is conserved since

$$\frac{d}{dt} \int_{\mathbb{R}^{2d}} f(x, v, t) dv dx = 0.$$

The positivity for all $t \geq 0$ of the kinetic distribution $f(x, v, t)$ may be proven under suitable assumptions on the drift term and the diffusion function. Furthermore, for $\varphi(v) = v$ we have that the evolution of the mean velocity is given by

$$\frac{d}{dt} \int_{\mathbb{R}^{2d}} v f(x, v, t) dv dx = - \int_{\mathbb{R}^{2d}} \mathcal{B}[f](x, v, t) f(x, v, t) dv dx.$$

Therefore, the mean velocity is conserved if P is a symmetric function in both $x \in \mathbb{R}^d$ and $v \in \mathbb{R}^d$.

Given a set of sample positions and velocities $\{(x_i, v_i)\}_{i=1}^N$ at time $t \geq 0$, the kinetic distribution $f(x, v, t)$ is recovered from the empirical density function

$$f_N(x, v, t) = \frac{1}{N} \sum_{i=1}^N \delta(x - x_i(t)) \otimes \delta(v - v_i(t)), \tag{6}$$

where $\delta(\cdot)$ is the Dirac delta function and the particles evolution follows (1). Therefore, for any test function $\varphi(x, v)$, if we denote by

$$(\varphi, f)(t) = \int_{\mathbb{R}^d} \varphi(x, v) f(x, v, t) dv,$$

we have

$$(\varphi, f_N)(t) = \frac{1}{N} \sum_{i=1}^N \varphi(x_i(t), v_i(t)).$$

If $f(x, v, t)$ is a probability density function $\int_{\mathbb{R}^{2d}} f(x, v, t) dv dx$, we have

$$(\varphi, f) = \mathbb{E}_{X,V}[\varphi],$$

being $\mathbb{E}_{X,V}[\varphi]$ the expectation of the observable quantity φ . The following result holds [5]

Theorem 1 *The root-mean square error is such that for each $t \geq 0$*

$$\mathbb{E}_{X,V} [|(\varphi, f) - (\varphi, f_N)|^2]^{1/2} = \frac{\sigma_\varphi}{N^{1/2}},$$

where $\sigma_\varphi^2 = \text{Var}(\varphi)$ and

$$\text{Var}(\varphi) = \int_{\mathbb{R}^{2d}} (\varphi(x, v) - (\varphi, f))^2 f(x, v, t) dv dx.$$

Since f is a probability density, we observe that the nonlocal drift operator may be thought as an expectation with respect to the kinetic distribution $f(x, v, t)$ and in particular as the evaluation of the expected quantity of $P(x, y, v, w)(w - v)$ with respect to the pair (y, w) . Hence, we have

$$\nabla_v \cdot (\mathcal{B}[f])(x, v, t) f(x, v, t) = \nabla_v \cdot (\mathbb{E}_f[P(x, \cdot, v, \cdot)](v - \cdot)) f(x, v, t),$$

where $\mathbb{E}_f[\cdot]$ is the expectation with respect to the probability density f and is defined as

$$\mathbb{E}_f[g(x, y, v, w, t)] = \int_{\mathbb{R}^d \times \mathbb{R}^d} g(x, y, v, w, t) f(y, w, t) dy dw.$$

Therefore, the initial model in (2) is equivalently written as

$$\partial_t f + v \cdot \nabla_x f = \nabla_v \cdot [\mathbb{E}_f[P(x, \cdot, v, \cdot)](v - \cdot)] f + \sigma^2 \nabla_v f]$$

3.1 Standard RBM Method

Equation (2) can be obtained as the mean-field limit of the particle system (1). Let us consider the empirical distribution (6) and rewrite the second equation in (1) as follows

$$dv_i = \mathbb{E}_{f^N}[P(x_i, \cdot, v_i, \cdot)(\cdot - v_i)]dt + \sqrt{2\sigma^2}dW_i^t,$$

being

$$\mathbb{E}_{f^N}[P(x_i, \cdot, v_i, \cdot)(\cdot - v_i)] = \frac{1}{N} \sum_{j=1}^N P(x_i, x_j, v_i, v_j)(v_j - v_i).$$

For any $N \gg 0$ the we may observe that

$$\mathbb{E}_f[P(x, \cdot, v, \cdot)(v - \cdot)] \approx \mathbb{E}_{f^N}[P(x_i, \cdot, v_i, \cdot)(\cdot - v_i)].$$

As a result, we obtain a particle dynamics with complexity $O(N^2)$. The random batch idea is based on the use of subsets S_M of samples of size $M \ll N$ sampled uniformly from the set of indices $\{1, \dots, N\}$ to further approximate

$$\mathbb{E}_{f^N}[P(x_i, \cdot, v_i, \cdot)(\cdot - v_i)] \approx \frac{1}{M} \sum_{j \in S_M} P(x_i, x_j, v_i, v_j)(v_j - v_i), \tag{7}$$

for all $i = 1, \dots, M$. This fast summation strategy has an overall cost of $O(MN)$ instead of $O(N^2)$ and therefore substantially alleviate computational burden linked to the particle approach. In the following, we concentrate on the dynamics of the particle system given by the following system of SDEs for (x_i, v_i)

$$\begin{cases} dx_i = v_i dt, \\ dv_i = \frac{1}{M} \sum_{j \in S_M} P(x_i, x_j, v_i, v_j)(v_j - v_i)dt + \sqrt{2\sigma^2}dW_i^t \end{cases} \tag{8}$$

such that the empirical measure

$$f_N^M(x, v, t) = \frac{1}{N} \sum_{i=1}^N \delta(x - x_i(t)) \otimes \delta(v - v_i(t))$$

converges towards $f(x, v, t)$ solution to (2). Indeed, the following result can be established

Theorem 2 *The root-mean square error of the RBM method is such that for each $t \geq 0$*

$$\mathbb{E}_{X, V} \left[|(\varphi, f) - (\varphi, f_N^M)|^2 \right]^{1/2} \leq \frac{\sigma_\varphi}{\sqrt{N}} + \theta_\varphi \sqrt{\frac{1}{M} - \frac{1}{N}},$$

where

$$\sigma_\varphi^2 = \int_{\mathbb{R}^{2d}} (\varphi(x, v) - (\varphi, f))^2 f(x, v, t) dv dx$$

and

$$\theta_\varphi^2 = \int_{\mathbb{R}^{2d}} (\varphi(x, v) - (\varphi, f_N))^2 f_N(x, v, t) dv dx$$

Algorithm 1 Random Batch Method

- 1: Sample N particles from the initial distribution $f(x, v, 0)$
 - 2: Set $M \in \mathbb{N}$, $1 < M < N$
 - 3: **for** each particle i **do**
 - 4: select a random batch \mathcal{S}_M
 - 5: update (x_i, v_i) by solving a time discretization of (8)
 - 6: **end for**
 - 7: Reconstruct the kinetic density using a kernel density estimator of (6).
-

Proof For any test function φ we have

$$\begin{aligned} \mathbb{E}_{X,V} \left[\left| (\varphi, f) - (\varphi, f_N^M) \right|^2 \right] &\leq \\ &2\mathbb{E}_{X,V} \left[\left| (\varphi, f) - (\varphi, f_N) \right|^2 \right] + 2\mathbb{E}_{X,V} \left[\left| (\varphi, f_N) - (\varphi, f_N^M) \right|^2 \right]. \end{aligned}$$

Hence, we may pursue the same strategy presented in [7] to get the following estimate

$$\mathbb{E}_{X,V} \left[\left| (\varphi, f_N) - (\varphi, f_N^M) \right|^2 \right]^{1/2} \leq \theta_\varphi \sqrt{\frac{1}{M} - \frac{1}{N}},$$

from which we get thanks to Theorem 1

$$\mathbb{E}_{X,V} \left[\left| (\varphi, f) - (\varphi, f_N^M) \right|^2 \right]^{1/2} \leq \sqrt{2} \frac{\sigma_\varphi}{\sqrt{N}} + \sqrt{2} \theta_\varphi \sqrt{\frac{1}{M} - \frac{1}{N}}. \quad \square$$

From the above theorem we can conclude that the computation of the average defined in (7) approximates the original one with an accuracy $O((1/M - 1/N)^{1/2})$.

At the computational level, the RBM method, as it has been presented in [1, 11, 30], can be summarized as in Algorithm 1.

3.2 The Reduced Variance Strategy

In this section we want to define a suitable strategy to reduce the variance of the RBM method. To this end, we will develop a control variate approach which is based on the construction of simpler linear models which define a compatible asymptotic distribution with respect to the full model. We describe the method using a nonlocal PDEs of Vlasov-Fokker-Planck-type introduced in (2).

We define a new drift term as follows

$$P^\lambda(x, y, v, w)(w - v) = P(x, y, v, w)(w - v) - \lambda(\tilde{P}(x, v)(w - v) - \tilde{Q}(f)(x, v, t)), \quad (9)$$

where $\lambda \in \mathbb{R}$ and $\tilde{Q}(x, v, t)$ is a reduced complexity collision operator having the form

$$\begin{aligned} \tilde{Q}(f)(x, v, t) &= \mathbb{E}_f[\tilde{P}(x, v)(\cdot - v)] = \int_{\mathbb{R}^{2d}} \tilde{P}(x, v)(w - v) f(y, w, t) dy dw \\ &= \tilde{P}(x, v)(U - v), \end{aligned}$$

being $\tilde{P}(x, v) \geq 0$ a simplified interaction function. Hence, (9) may be rewritten as

$$\begin{aligned} P^\lambda(x, y, v, w)(w - v) &= P(x, y, v, w)(w - v) - \lambda(\tilde{P}(x, v)(w - v) - \tilde{P}(x, v)(U - v)) \\ &= P(x, y, v, w)(w - v) - \lambda\tilde{P}(x, v)(w - U). \end{aligned} \tag{10}$$

We can notice that we have

$$\mathbb{E}_f[P^\lambda(x, \cdot, v, \cdot)(\cdot - v)] = \mathbb{E}_f[P(x, \cdot, v, \cdot)(\cdot - v)] = Q(f)(x, v),$$

so that the solution of our original problem can be obtained by the evaluation of the expectation of the new term $P^\lambda(x, \cdot, v, \cdot)(\cdot - v)$. In order to determine the optimal value of the scalar value λ a common strategy relies on the minimization of the variance [15, 16, 22]. In particular, the variance of the above expression is given by

$$\mathbb{V}\text{ar}_f(P^\lambda(x, \cdot, v, \cdot)(\cdot - v)) = \int_{\mathbb{R}^{2d}} (P^\lambda(x, y, v, w)(w - v) - Q(f)(x, v))^2 f(y, w) dw dy.$$

From (9) we get

$$\begin{aligned} \mathbb{V}\text{ar}(P^\lambda(x, \cdot, v, \cdot)(\cdot - v)) &= \mathbb{V}\text{ar}(P(x, \cdot, v, \cdot)(\cdot - v)) + \lambda^2 \mathbb{V}\text{ar}(\tilde{P}(x, v)(\cdot - U)) \\ &\quad - 2\lambda \mathbb{C}\text{ov}(P(x, \cdot, v, \cdot)(\cdot - v), \tilde{P}(x, v)(\cdot - U)), \end{aligned}$$

where

$$\begin{aligned} \mathbb{C}\text{ov}_f(P(x, \cdot, v, \cdot)(\cdot - v), \tilde{P}(x, v)(\cdot - U)) &= \int_{\mathbb{R}^{2d}} (P(x, y, v, w)(w - v) - Q(f)(x, v)) \\ &\quad (P(x, v)(w - U) - \tilde{Q}(f)(x, v)) \\ &\quad f(y, w) dw dy. \end{aligned}$$

We can minimize the variance by direct differentiation with respect to λ and obtain the optimal value. The following result holds

Theorem 3 *The quantity*

$$\lambda^* = \frac{\mathbb{C}\text{ov}_f(P(x, \cdot, v, \cdot)(\cdot - v), \tilde{P}(x, v)(\cdot - U))}{\mathbb{V}\text{ar}_f(\tilde{P}(x, v)(\cdot - U))}. \tag{11}$$

minimizes the variance of $P^\lambda(x, y, v, w)(w - v)$ and is such that

$$\mathbb{V}\text{ar}(P^{\lambda^*}(x, y, v, w)(w - v)) = (1 - \rho^2) \mathbb{V}\text{ar}(P(x, y, v, w)(w - v)), \tag{12}$$

being $\rho \in [-1, 1]$ the Pearson’s correlation coefficient.

Furthermore, provided that $\tilde{P}(x, v)$ defines a consistent surrogate model, meaning that the particle system defined by a surrogate interaction is correlated with the one originated by full interaction function and $\rho \rightarrow \pm 1$, we have

$$\lim_{\rho \rightarrow \pm 1} \lambda^* = \sqrt{\frac{\mathbb{V}\text{ar}(P(x, y, v, w)(w - v))}{\mathbb{V}\text{ar}(\tilde{P}(x, v)(w - U))}}$$

Algorithm 2 Reduced Variance Random Batch Method (rvRBM)

- 1: Sample N particles from the initial distribution $f(x, v, 0)$
 - 2: Set $M \in \mathbb{N}$, $1 < M < N$
 - 3: **for** each particle i **do**
 - 4: select a random batch \mathcal{S}_M
 - 5: evaluate λ^* from (11)
 - 6: update (x_i, v_i) by solving a time discretization of (14)
 - 7: **end for**
 - 8: Reconstruct the kinetic density using a kernel density estimator of (6).
-

which leads from (12)

$$\lim_{\rho \rightarrow \pm 1} \mathbb{V}\text{ar}(P^{\lambda^*}(x, y, v, w)(w - v)) = 0.$$

On the other hand, in the absence of correlation, corresponding to the case $\rho \rightarrow 0$ we get

$$\lim_{\rho \rightarrow 0} \mathbb{V}\text{ar}(P^{\lambda^*}(x, y, v, w)(w - v)) = \mathbb{V}\text{ar}(P(x, y, v, w)(w - v)).$$

Let us assume now to evaluate in (9) the exact expectation $\tilde{P}(x, v)(U - v)$ of the surrogate linear model using N samples

$$\tilde{P}(x_i, v_i)(U - v_i) \approx \tilde{P}(x_i, v_i)(U_N - v_i)$$

where

$$U_N = \frac{1}{N} \sum_{j=1}^N v_j.$$

We can improve the standard Monte Carlo random batch estimate based on M samples (8) using

$$\begin{aligned} \frac{1}{M} \sum_{j \in \mathcal{S}_M} P^{\lambda^*}(x_i, x_j, v_i, v_j)(v_j - v_i) &= \frac{1}{M} \sum_{j \in \mathcal{S}_M} P(x_i, x_j, v_i, v_j)(v_j - v_i) \\ &\quad - \lambda_* \left(\tilde{P}(x_i, v_i)(U_M - U_N) \right). \end{aligned} \tag{13}$$

Hence, we get the following system of SDEs for (x_i, v_i)

$$\begin{cases} dx_i = v_i dt, \\ dv_i = \left[\frac{1}{M} \sum_{j \in \mathcal{S}_M} P(x_i, x_j, v_i, v_j)(v_j - v_i) - \lambda_* \left(\tilde{P}(x_i, v_i)(U_M - U_N) \right) \right] dt + \sqrt{2\sigma^2} dW_i^t, \end{cases} \tag{14}$$

where as usual $\{W_i\}$, $i = 1, \dots, N$ defines a set of independent Wiener processes and λ^* is computed as in (11). A global in time error analysis can be obtained following the approach presented in [31, 42]. After a surrogate interaction model $\tilde{P}(x_i, v_i)$ has been selected, the reduced variance strategy can be summarised as in Algorithm 2.

Remark 1 We may observe that $\mathbb{Cov}_f(P(x, y, v, w)(w - v), \tilde{P}(x, v)(w - U))$ and $\mathbb{Var}_f(\tilde{P}(x, v)(w - U))$, which define λ^* , are not known and must be estimated starting from a sample of size N . In particular, the following unbiased estimators can be defined

$$\begin{aligned} \mathbb{E}_{f^N}(\tilde{P}(x, v)(w - U)) &= \frac{1}{N} \sum_{i=1}^N \tilde{P}(x, v)(w_i - U_N), \\ \mathbb{Var}_{f^N}(\tilde{P}(x, v)(w - U)) &= \frac{1}{N-1} \sum_{i=1}^N \left(\tilde{P}(x, v)(w_i - U_N) - \mathbb{E}_{f^N}(\tilde{P}(x, v)(w - U_N)) \right)^2, \\ \mathbb{Cov}_{f^N}(P(x, y, v, w)(w - v), \tilde{P}(x, v)(w - U_N)) &= \frac{1}{N-1} \sum_{i=1}^N \left(\tilde{P}(x, v)(w_i - U_N) - \mathbb{E}_{f^N}(\tilde{P}(x, v)(w_i - U_N)) \right) \\ &\quad \left(P(x, y_i, v, w_i)(w_i - v) - \mathbb{E}_{f^N}(P(x, y_i, v, w_i)(w_i - v)) \right), \end{aligned}$$

which lead to the estimated optimal λ^* as

$$\lambda_N^* = \frac{\mathbb{Cov}_{f^N}(P(x, y, v, w)(w - v), \tilde{P}(x, v)(w - U_N))}{\mathbb{Var}_{f^N}(\tilde{P}(x, v)(w - U_N))}.$$

With reference to the computational cost, the rvRBM method requires the computation of λ_N^* , which can be done at a cost of $O(N)$ and is independent of the batch size. Therefore, the rvRBM method has the same computational cost as a standard RBM method.

Remark 2 If we consider a linear surrogate operator, i.e. $\tilde{Q} \left(\sum_{k=1}^K f_k \right) = \sum_{k=1}^K \tilde{Q}(f_k)$, where

$$\tilde{Q}(f_k)(x, v, t) = \tilde{P}(x, v)(U_k - v), \quad k = 1, 2,$$

and $U_k = \int_{\mathbb{R}^{2d}} v f_k(x, v, t) dv dx$ with $\sum_{k=1}^K U_k = U$, we can define

$$P^{\lambda_k}(x, y, v, w)(w - v) = P(x, y, v, w)(w - v) - \lambda_k \tilde{P}(x, v)(w - U_k),$$

which leads to sequence of optimal values

$$\lambda_k^* = \frac{\mathbb{Cov}_f(P(x, \cdot, v, \cdot)(\cdot - v), \tilde{P}(x, v)(\cdot - U_k))}{\mathbb{Var}_f(\tilde{P}(x, v)(\cdot - U_k))}. \tag{15}$$

Remark 3 More generally, a reduced variance approach can be obtained also in the presence of a nonlocal drift term

$$\mathcal{B}[f](x, v, t) = \int_{\mathbb{R}^d \times \mathbb{R}^d} K(x, y, v, w) f(y, w, t) dy dw = \mathbb{E}_f[K(x, y, v, w)]$$

being $K : \mathbb{R}^{4d} \rightarrow \mathbb{R}^d$, provided that it is possible to define

$$K^\lambda(x, y, v, w) = K(x, y, v, w) - \lambda \left(\tilde{K}(x, y, v, w) - \tilde{Q}(f)(x, v, t) \right) \tag{16}$$

such that

$$\mathbb{E}_f[\tilde{K}(x, y, v, w)] = \tilde{Q}(f)(x, v, t),$$

and $\tilde{K}(x, y, v, w)$ in (16) is a suitable reduced complexity interaction kernel.

4 Numerical Tests

In this section we present several numerical examples to show the performances of the introduced reduced variance RBM. In the following we will recall both deterministic and stochastic models. In all the subsequent tests, we reconstruct the particle densities through a regularization of the empirical distribution $f^N(x, v, t) = \frac{1}{N} \sum_{i=1}^N \delta(x - x_i) \otimes \delta(v - v_i)$ based on a kernel density estimator obtained through the evaluation of $\langle \varphi, f^N \rangle$, being $\varphi(x_i, v_i)$ the normal density centered in (x_i, v_i) and having fixed variance $\Sigma^2 = 10^{-5}$.

In the following numerical tests we will focus on the definition of suitable RBM and rvRBM methods pointing the reader to the nonlocal PDE models introduced in Sect. 2.

4.1 Test 1: Alignment Models

4.1.1 Test 1a: Bounded Confidence Model

We consider the bounded confidence opinion model introduced in Sect. 2.3 in the case $D \equiv 0$. We recall that it is possible to prove that (3) corresponds to the mean-field limit of the following particle system

$$\frac{d}{dt} v_i(t) = \frac{1}{N} \sum_{j=1}^N P(|v_i - v_j|)(v_j - v_i), \quad i = 1, \dots, N, \tag{17}$$

where $P(|v_i - v_j|) = \chi(|v_i - v_j| \leq \delta)$, $\delta > 0$, see e.g. [4, 36] and the references therein. Hence, the RBM method corresponds to the following reduced complexity model

$$\frac{d}{dt} v_i(t) = \frac{1}{M} \sum_{j \in \mathcal{S}_M} P(|v_i - v_j|)(v_j - v_i), \quad i = 1, \dots, N, \tag{18}$$

being \mathcal{S}_M a uniform subsample of size $M > 1$ obtained from the $\{v_i\}_{i=1}^N$. A rvRBM method can be therefore obtained by considering in (13) the interaction function $\tilde{P}(v_i)$. In particular, we will consider the cases $\tilde{P}(v_i) \equiv 1$ (denoted by case 1) and $\tilde{P}(v_i) = 1 - v_i^2$ (denoted by case 2). Therefore the rvRBM method corresponds in considering the evolution of the following particle system

$$\frac{d}{dt} v_i(t) = \frac{1}{M} \sum_{j \in \mathcal{S}_M} P(|v_i - v_j|)(v_j - v_i) - \lambda_* \tilde{P}(v_i)(u_M - u_N), \tag{19}$$

where

$$u_M = \frac{1}{M} \sum_{j \in \mathcal{S}_M} v_j, \quad u_N = \frac{1}{N} \sum_{j=1}^N v_j,$$

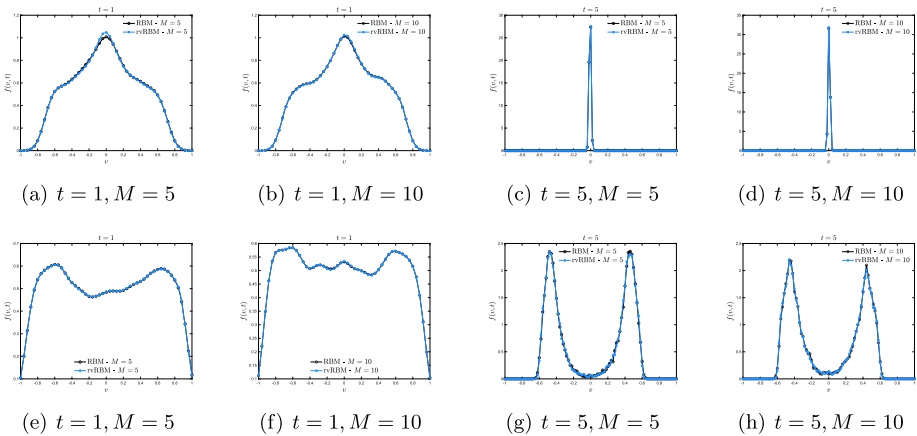


Fig. 1 Test 1a. Evolution of the densities for the deterministic BC model in (3) by means of a particle approach defined in (17) with interaction function with threshold $\delta = 1$ (top row), $\delta = 0.5$ (bottom row). We considered $N = 10^5$ particles with initial distribution defined in (20) and we compare the evolution of the approximated densities obtained with RBM and rvRBM with a subset of interacting particles of size $M = 5$ or $M = 10$. The surrogate interaction function is here considered $\tilde{P}(v_i) \equiv 1$

and $\lambda^* = \lambda^*(t)$ is computed as in (11). We remark that, since u_N is a conserved quantity of the model it is sufficient to compute u_N from the initial sample, whereas u_M is affected from the error induced by the uniform subsampling.

In Fig. 1 we report the evolution of the particle density at time $t = 1$ and $t = 5$ and by considering $M = 5$ or $M = 10$. As initial distribution we considered a sample of the uniform distribution of the interval $[-1, 1]$, i.e.

$$f(v, 0) = \begin{cases} \frac{1}{2} & v \in [-1, 1], \\ 0 & \text{elsewhere.} \end{cases} \tag{20}$$

We compare the evolution of the particle density obtained with a standard RBM together with the one obtained with rvRBM, from which we may notice good agreement in describing the transient behaviour of the distribution both in the case $\delta = 1$ (top row) and in the case $\delta = 0.5$ (bottom row). In this latter case the emerging equilibrium density displays clusterization towards the values $\pm \frac{1}{2}$.

We consider the following absolute error for the mean

$$\text{Error} = \left| \frac{1}{N} \sum_{i=1}^N v_i - m \right|, \tag{21}$$

where $\{v_i\}_{i=1}^N$ is obtained either from RBM, as in (18), or from rvRBM, as in (19). In Fig. 2 we show the evolution in time of the introduced absolute error in the case $\delta = 1$ (top row) and $\delta = 0.5$ (bottom row) by both the standard RBM and the introduced rvRBM method. We will refer to case 1 if $\tilde{P}(v_i) \equiv 1$ whereas with case 2 we will refer to $\tilde{P}(v_i) = 1 - v_i^2$. We may observe how in the case $\delta = 1$ the method rvRBM leads to a significant advantage since the emerging distribution function is highly correlated with the one characterising the reduced variance method. Indeed, as shown in Fig. 3 (left plot) the optimal λ^* is such that

$$\lambda^* \rightarrow 1, \quad t \rightarrow +\infty.$$

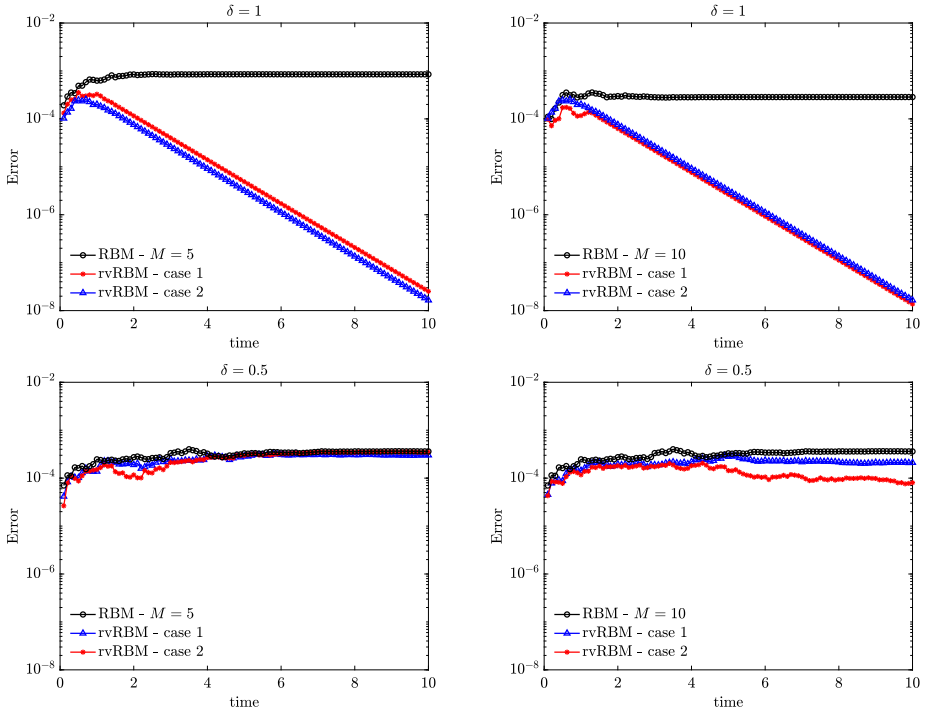


Fig. 2 Test 1a. Evolution of the absolute error defined in (21) for the bounded confidence model solved through RBM and rvRBM with $M = 5$ (left column) or $M = 10$ (right column). We considered $\delta = 1$ (top row) and $\delta = 0.5$ (bottom row)

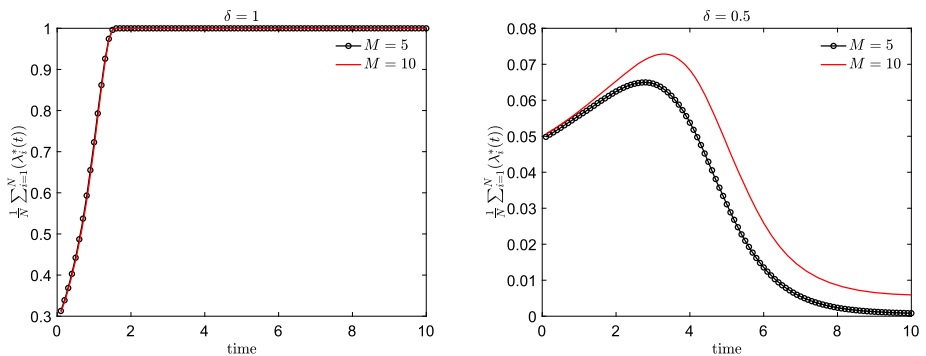


Fig. 3 Test 1a. Evolution of λ^* for the model (3) approximated through a rvRBM strategy with $M = 5, 10$ and $\delta = 1$ (left), $\delta = 0.5$ (right)

On the other hand, for the case $\delta = 0.5$ the error produced by the rvRBM strategy is comparable with the one of RBM. Indeed, the computed $\lambda^* \rightarrow 0$, since the emerging distribution has no correlation with the one of the considered for the reduced variance method.

In Fig. 4 we depict the evolution of the error produced by fixed batch $M = 5$ (left) or $M = 10$ (right) and a variable $N \in \{100, \dots, 10^4\}$. The error (21) is here computed at time

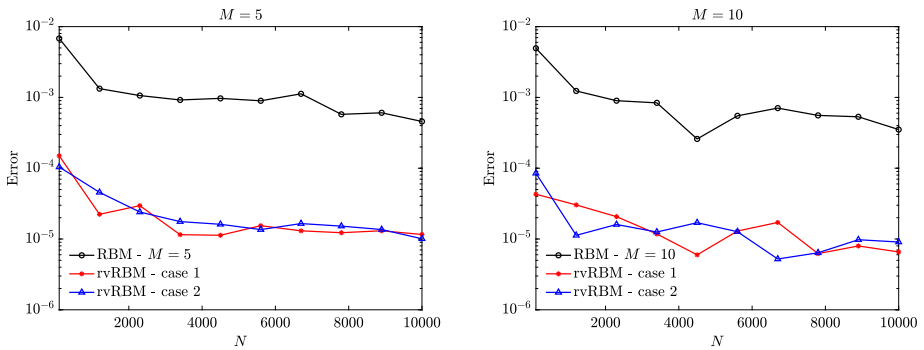


Fig. 4 Test 1a. Comparison of the errors produced by RBM and rvRBM with surrogate models obtained with $\tilde{P} \equiv 1$ (case 1) or $\tilde{P}(v_i) = 1 - v_i^2$ (case 2). We considered a fixed batch size $M = 5$ (left) and $M = 10$ (right) and variable size of the sample $N \in \{10, \dots, 10^4\}$. The error is computed from the bounded confidence model (17) with $\delta = 1$ at time $T = 5$

$T = 5$. Since the cost of the reduced variance approach is compatible to the one of a RBM method $O(MN)$, we may observe how, in the reduced variance approach, we essentially gain accuracy at the same cost of RBM.

4.1.2 Test 1b: Bounded Confidence Model with Isolated Clusters

The loss of accuracy of the rvRBM method in the case $\delta = 0.5$ is essentially due to the uncorrelated behaviour of the surrogate model with respect to the full one. Indeed, whereas the considered surrogate models would lead to the formation of global consensus, the bounded confidence model leads to the formation of local stable clusters [4]. A possible way to overcome this issue has been addressed in Remark 2. We consider a sample of the initial distribution

$$f(v, 0) = \begin{cases} 1 & v \in [\frac{1}{4}, \frac{3}{4}] \\ 1 & v \in [-\frac{3}{4}, -\frac{1}{4}]. \end{cases} \tag{22}$$

Now, for $\delta = 0.5$ no interactions are expected from the two initial clusters centered in $\pm \frac{1}{2}$. We computed the optimal λ_k^* , $k = 1, 2$, exploiting two surrogate models, the first is $\tilde{P}(v_i) \equiv 1$ (case 1) whereas for the second we considered a surrogate model defined by

$$\tilde{P}(v_i) = \left(v_i - \frac{1}{2}\right) \left(v_i + \frac{1}{2}\right)$$

In the following, we will refer to this choice as case 2. In Fig. 5 we report the evolution of the rvRBM method where we defined a sequence of optimal λ_k^* as in (15) with $k = 1, 2$ such that $U_k = \pm \frac{1}{2}$. We may observe that the evolution in time of the obtained rvRBM method is consistent with the one defined by the standard RBM method. Furthermore, from the evolution of the absolute error (21) we can observe that the method is capable to achieve much higher accuracy than RBM.

4.2 Test 2: Stochastic Opinion Formation Model

We now consider a stochastic bounded confidence model defined in Sect. 2.3 in (5), where the interaction function is defined as in (4) and that is obtained from the meanfield limit of

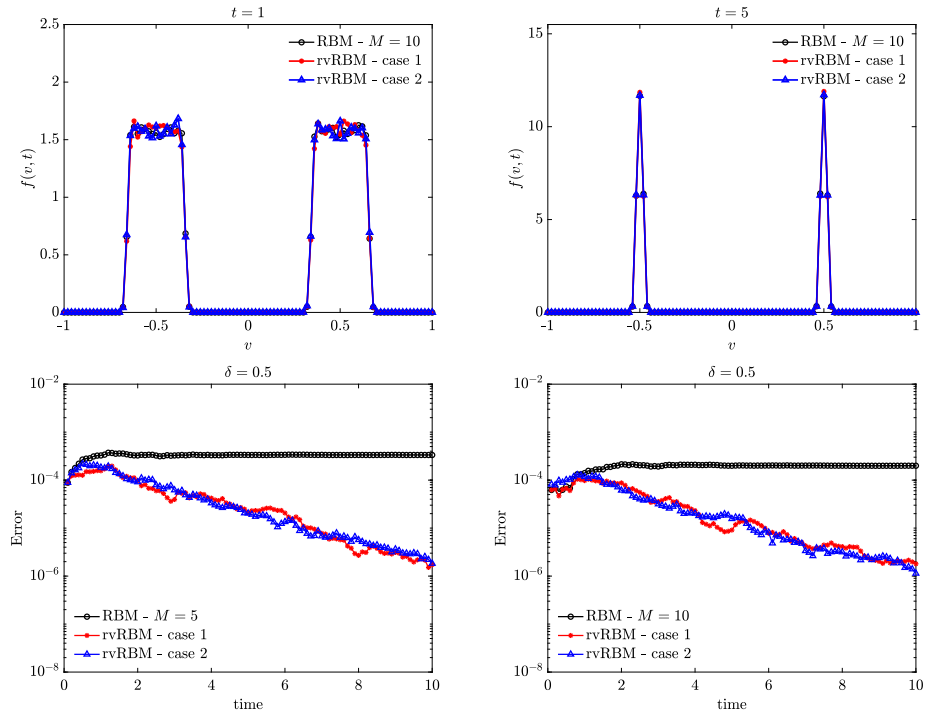


Fig. 5 Test 1b. Top row: comparison of the reconstructed distributions at time $t = 1$ and $t = 5$ for the bounded confidence model with $\delta = 0.5$ using RBM or rvRBM with $\tilde{P}(v_i) \equiv 1$ (case 1) or $\tilde{P}(v_i) = (v_i - \frac{1}{2})(v_i + \frac{1}{2})$. The batch size is $M = 10$ and the initial sample has been obtained from (22). Bottom row: evolution of the absolute errors (21) for the bounded confidence model solved through RBM and rvRBM with $M = 5$ (left) or $M = 10$ (right)

the following system of stochastic differential equations

$$dv_i(t) = \frac{1}{N} \sum_{j=1}^N P(|v_i - v_j|)(v_j - v_i)dt + \sqrt{2\sigma^2 D^2(v_i)}dW_i^t, \tag{23}$$

with $D^2(v_i) = 1 - v_i(t)$ and $\{W_i\}_{i=1}^N$ a set of independent Wiener processes. In the following, we fixed the parameter $\sigma^2 = 10^{-1}$. Therefore, the RBM approximation is obtained by considering

$$dv_i(t) = \frac{1}{M} \sum_{j \in S_M} P(|v_i - v_j|)(v_j - v_i)dt + \sqrt{2\sigma^2 D^2(v_i)}dW_i^t,$$

where S_M is a uniform subsample of size $M > 1$ from the $\{v_i\}_{i=1}^N$. Similarly, rvRBM strategy is obtained as

$$dv_i(t) = \left[\frac{1}{M} \sum_{j \in S_M} P(|v_i - v_j|)(v_j - v_i) - \lambda^* \tilde{P}(v_i)(U_N - U_M) \right] dt + \sqrt{2\sigma^2 D^2(v_i)}dW_i^t.$$

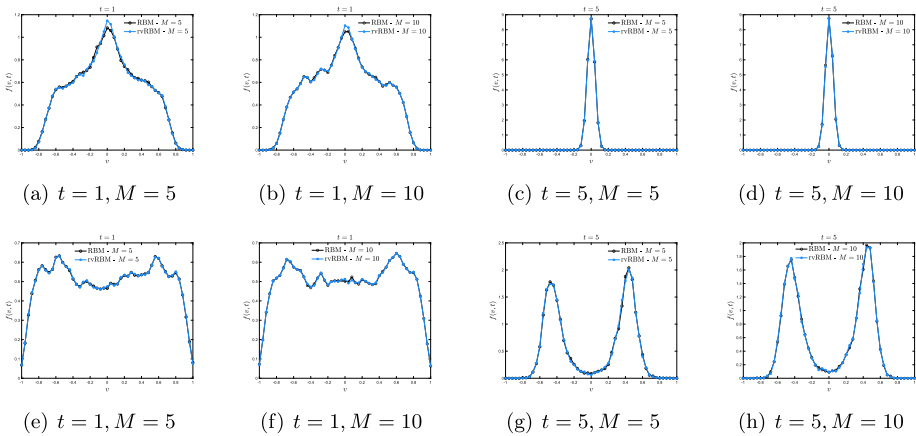


Fig. 6 Test 2. Evolution of the densities for the stochastic bounded confidence model (5) with interaction function (4) and $\delta = 1$ (top row), $\delta = 0.5$ (bottom row). We considered $N = 10^5$ particles with initial distribution (20) and we compare the evolution of the approximated densities obtained with RBM or rvRBM with a subset of interacting particles of size $M = 5$ or $M = 10$. We fixed the diffusion constant $\sigma^2 = 10^{-1}$

In the following, the methods defining RBM and rvRBM share the same set of Wiener processes of the complete model (23). The set of SDEs is solved through the Euler-Maruyama scheme.

In Fig. 6 we report the evolution of the particle density at time $t = 1$ and $t = 5$ in the stochastic setting and obtained with $M = 5, 10$. As initial distribution we considered the one introduced in (20). In each plot, we compare the evolution of the approximated density obtained with $N = 10^5$ particles and either RBM or rvRBM algorithms. In Fig. 6 we considered the case $\delta = 1$.

In Fig. 7 we show the evolution in time of the introduced absolute error in the case $\delta = 1$ (top row) obtained through the standard RBM and the introduced rvRBM method. We may observe how rvRBM leads to a substantial advantage if the emerging distributions are highly correlated, as in the case $\delta = 1$. At variance with the deterministic model, the asymptotic level of accuracy of rvRBM is influenced by the total number of particles due to the presence of the stochastic component in the dynamics. In the second row we report the error produced by a fixed batch of size $M = 5$ (left) or $M = 10$ (right) and an increasing total number of particles $N \in \{10, \dots, 10^4\}$. The error is computed at time $T = 5$. We may observe how the reduced variance approach is capable to achieve higher accuracy with the same number of particles, and therefore of cost, of the RBM method.

4.3 Test 3: Cucker-Smale Model

We consider the second order model for swarming introduced in Sect. 2.2. We recall that it is possible to prove that this model corresponds to the meanfield limit of the following particles' system

$$\begin{aligned} \frac{d}{dt}x_i &= v_i, \\ \frac{d}{dt}v_i &= \frac{1}{N} \sum_{j=1}^N P(|x_i - x_j|)(v_j - v_i), \quad i = 1, \dots, N, \end{aligned}$$

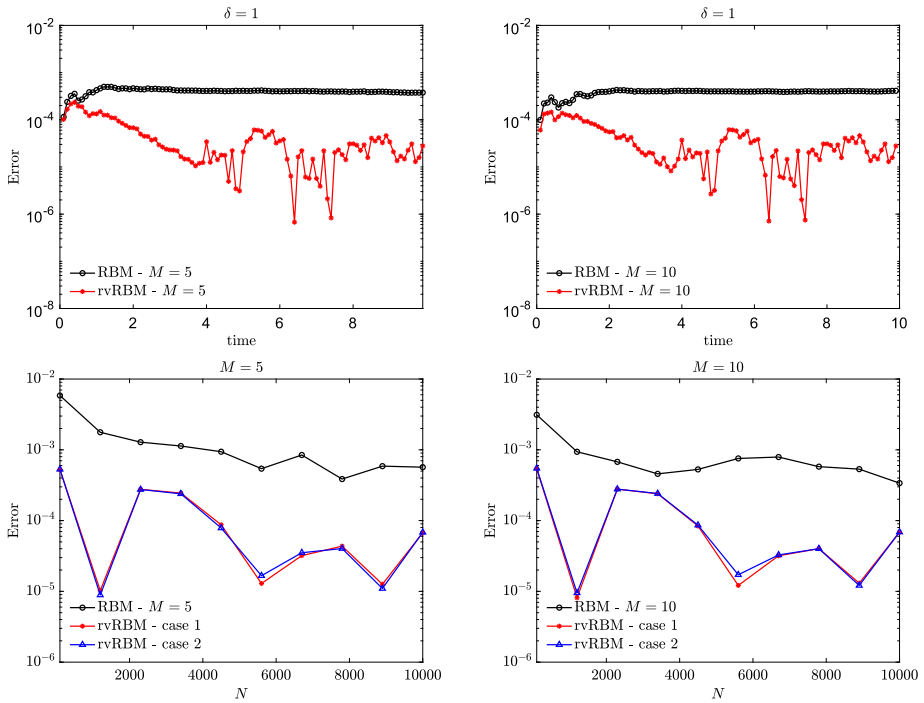


Fig. 7 Test 2. Stochastic bounded confidence model. Top row: evolution of the absolute error defined in (21) for the stochastic bounded confidence model solved through RBM and rvRBM with $M = 5$ (left) or $M = 10$ (right). Bottom row: errors of RBM and rvRBM for a fixed batch size $M = 5$ (left) or $M = 10$ (right) and variable size of the sample $N \in \{10, \dots, 10^4\}$. In all the presented tests we considered $\delta = 1$ and $N = 10^5$ particles

where $P(|x_i - x_j|) = \frac{1}{(\xi^2 + |x_i - x_j|^2)^\beta}$, we will always consider $\xi = 1$ and $\beta = 10^{-1}$, for which unconditional alignment emerges asymptotically. Hence, the RBM algorithm corresponds to consider

$$\begin{aligned} \frac{d}{dt}x_i &= v_i, \\ \frac{d}{dt}v_i &= \frac{1}{M} \sum_{j \in S_M} \frac{v_j - v_i}{(\xi^2 + |x_i - x_j|^2)^\beta}, \quad i = 1, \dots, N, \end{aligned} \tag{24}$$

where S_M is a uniform subsample of size $M > 1$ obtained from $\{(x_i, v_i)\}_{i=1}^N \in \mathbb{R}^{2d}$. In the following, we construct a rvRBM method by considering

$$\begin{aligned} \frac{d}{dt}x_i &= v_i, \\ \frac{d}{dt}v_i &= \frac{1}{M} \sum_{j \in S_M} \frac{v_j - v_i}{(\xi^2 + |x_i - x_j|^2)^\beta} - \lambda^*(x_i, v_i) \tilde{P}(x_i, v_i)(u_M - u_N), \quad i = 1, \dots, N, \end{aligned} \tag{25}$$

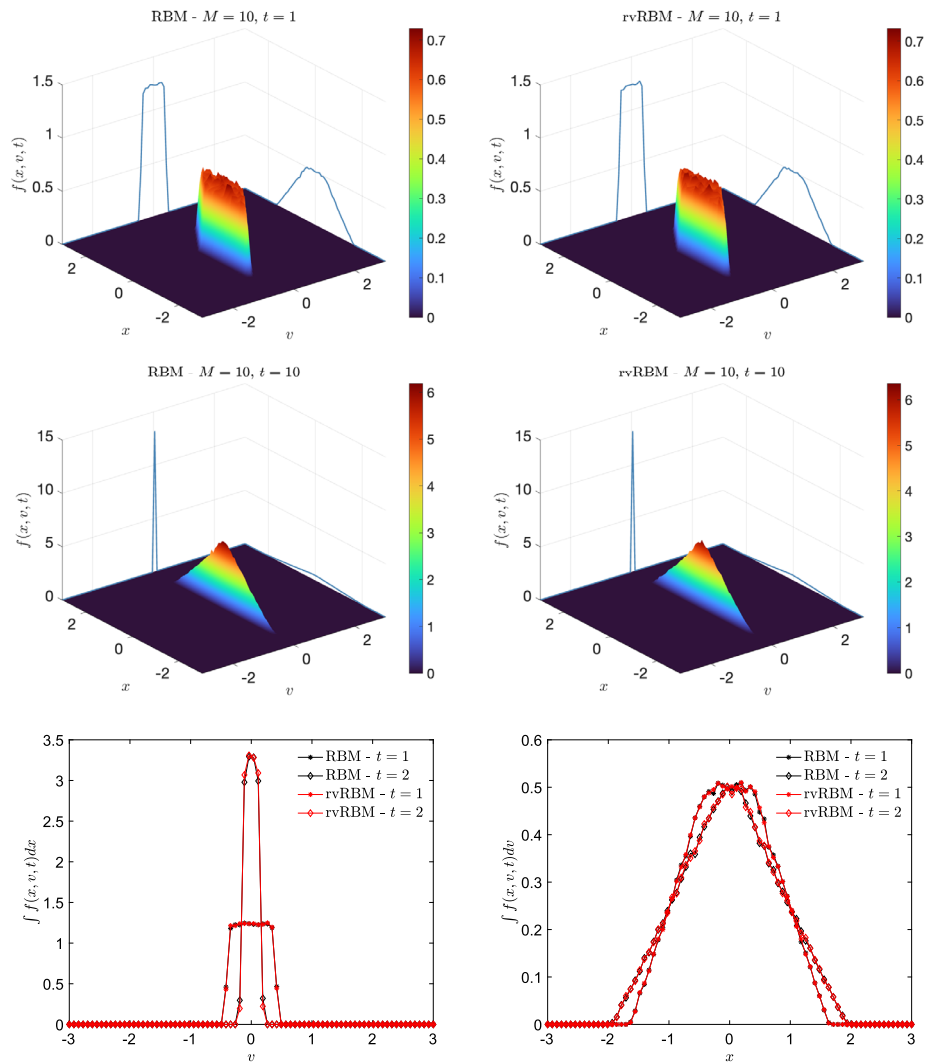


Fig. 8 **Test 3.** Evolution of the kinetic distributions obtained with RBM (left) or rvRBM (right) at time $t = 1$ (top row) and $t = 10$ (bottom row). For both methods we considered a subsample of size $M = 10$. In the bottom row we report the marginal densities obtained at time $t = 1$ and $t = 2$. The initial size of the particle system is $N = 10^5$ with initial distribution as in (26)

and $\tilde{P} \equiv 1$, being

$$u_M = \frac{1}{M} \sum_{j \in S_M} v_i, \quad u_N = \frac{1}{N} \sum_{j=1}^N v_i,$$

where $u_N(t) \equiv u_N(0)$ being $P(|x_i - v_i|)$ a symmetric interaction function. Hence, the optimal value of $\lambda(x_i, v_i)$ is computed as in (13).

In Fig. 8 we show the evolution of the reconstructed distribution functions in the phase space obtained with RBM (left panels) and rvRBM (right panels). We considered $N = 10^5$

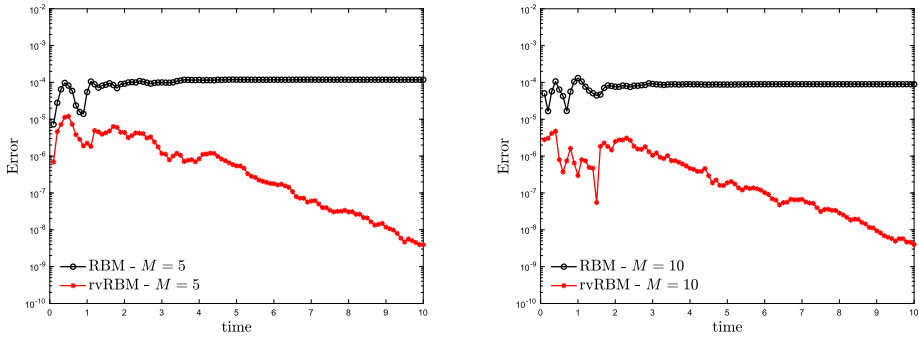
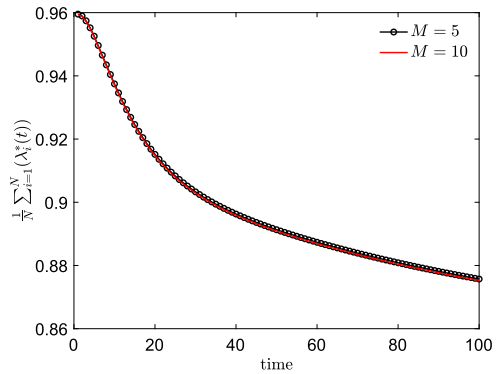


Fig. 9 Test 3. Evolution of the absolute error for the Cucker-Smale model solved through RBM as in (24) and rvRBM as in (25) with $M = 5$ (left) or $M = 10$ (right)

Fig. 10 Test 3. Evolution of the average optimal weight $\frac{1}{N} \sum_{i=1}^N \lambda_i^*(t)$ for the Cucker-Smale model approximated through rvRBM strategy with either $M = 5$ or $M = 10$



particles having uniform initial distribution in both space and velocity, i.e.

$$f_0(x, v) = \begin{cases} \frac{1}{4} & (x, v) \in [-1, 1] \times [-1, 1] \\ 0 & \text{elsewhere.} \end{cases} \tag{26}$$

Hence, we considered a RBM/rvRBM methods fixing a subsampling of size $M = 10$. We may observe how the obtained distributions consistent between the two methods also at the level of marginal distributions (bottom row).

We report in Fig. 9 the evolution of the error obtained with RBM and rvRBM. In detail, we fixed $N = 10^5$ initial particles and a subsample of size $M = 5$ (left plot) or $M = 10$ (right plot). In both cases, we can observe the advantages produced by the reduced variance strategy. In Fig. 10 we report the evolution of the average optimal λ^* computed for the Cucker-Smale model with either $M = 5$ or $M = 10$. We can observe how, due to the transport term the surrogate model lose the initial correlation with the full model.

5 Conclusions

The computational challenges associated with the $O(N^2)$ complexity of particle models for large systems with nonlocal interactions have driven the adoption of Random Batch Methods

(RBM). While these methods, using batches of size $M \ll N$, efficiently reduce the computational burden to $O(MN)$, introducing an additional error becomes an inevitable trade-off. Addressing this challenge, our paper proposes a control variate strategy inspired by ideas from uncertainty quantification. By leveraging a reduced complexity surrogate model, we successfully mitigate batch variance, achieving a substantial reduction and, consequently, an improved computational cost-to-accuracy ratio.

Of course, the effectiveness of our proposed strategy is strongly dependent on the possibility to construct a surrogate model which keeps correlations with the original model. Nevertheless, even in the worst case scenario when the surrogate model loses correlations our methodology will never deteriorate its accuracy below standard RBM methods. This is demonstrated through numerical examples, particularly in the context of models depicting collective behavior in opinion spreading and swarming dynamics. This not only showcases the practical utility of our approach but also underlines its versatility across different scenarios.

Looking ahead, our work opens directions for further research, particularly in refining control variate strategies and optimizing the implementation of Random Batch Methods in various computational contexts. For example, one could also explore the use of a multilevel strategy as an alternative to a surrogate model [22]. In this sense, the present approach marks a promising step towards bridging the gap between the microscopic complexity of interacting particle systems and the computational feasibility required for meaningful simulations, with potential implications for a broad spectrum of applications.

Acknowledgements This research has been written within the activities of the national groups GNCS (Gruppo Nazionale per il Calcolo Scientifico) and GNFM (Gruppo Nazionale di Fisica Matematica) of INdAM, Italy. MZ acknowledges the support of MUR-PRIN2020 Project No.2020JLWP23 (Integrated Mathematical Approaches to Socio-Epidemiological Dynamics). The research of LP has been supported by the Royal Society under the Wolfson Fellowship “Uncertainty quantification, data-driven simulations and learning of multiscale complex systems governed by PDEs”. The work has been partially supported by the National Centre for HPC, Big Data and Quantum Computing (CN00000013). The authors also acknowledge the support of the Banff International Research Station (BIRS) for the Focused Research Group (22frg198) “Novel perspectives in kinetic equations for emerging phenomena”, July 17–24, 2022, where part of this work was done.

Funding Statement Open access funding provided by Università degli Studi di Pavia within the CRUI-CARE Agreement. MZ acknowledges the support of MUR-PRIN2020 Project No.2020JLWP23 (Integrated Mathematical Approaches to Socio-Epidemiological Dynamics). The research of LP has been supported by the Royal Society under the Wolfson Fellowship “Uncertainty quantification, data-driven simulations and learning of multiscale complex systems governed by PDEs”. The work has been partially supported by the National Centre for HPC, Big Data and Quantum Computing (CN00000013).

Data Availability Statement The data sets generated during the current study are available from the corresponding author on reasonable request.

Declarations

Competing Interests The authors declare no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Albi, G., Pareschi, L.: Binary interaction algorithms for the simulation of flocking and swarming dynamics. *Multiscale Model. Simul.* **11**, 1–29 (2013)
2. Albi, G., Pareschi, L., Zanella, M.: Opinion dynamics over complex networks: kinetic modelling and numerical methods. *Kinet. Relat. Models* **10**(1), 1–32 (2017)
3. Barbu, V., Röckner, M.: From nonlinear Fokker-Planck equations to solutions of distribution dependent SDE. *Ann. Probab.* **48**, 1902–1920 (2020)
4. Borra, D., Lorenzi, T.: Asymptotic analysis of continuous dynamics models under bounded confidence. *Commun. Pure Appl. Anal.* **12**, 1487–1499 (2013)
5. Caflisch, R.E.: Monte Carlo and quasi-Monte Carlo methods. *Acta Numer.* **7**, 1–49 (1998)
6. Carrillo, J.A., Toscani, G.: Contractive probability metrics and asymptotic behavior of dissipative kinetic equations. *Riv. Mat. Univ. Parma* **6**(7), 75–198 (2007)
7. Carrillo, J.A., Zanella, M.: Monte Carlo gPC methods for diffusive kinetic flocking models with uncertainties. *Vietnam J. Math.* **47**(4), 931–954 (2019)
8. Carrillo, J.A., Fornasier, M., Rosado, J., Toscani, G.: Asymptotic flocking dynamics for the kinetic Cucker-Smale model. *SIAM J. Math. Anal.* **42**(1), 218–236 (2010)
9. Carrillo, J.A., Fornasier, M., Toscani, G., Vecil, F.: Particle, kinetic, and hydrodynamic models of swarming. In: *Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences*. Birkhäuser, Boston (2010)
10. Carrillo, J.A., Klar, A., Roth, A.: Single to double mill small noise transition via semi-Lagrangian finite volume methods. *Commun. Math. Sci.* **14**, 1111–1136 (2016)
11. Carrillo, J.A., Pareschi, L., Zanella, M.: Particle based gPC methods for mean-field models of swarming with uncertainty. *Commun. Comput. Phys.* **25**, 508–531 (2019)
12. Carrillo, J.A., Jin, S., Tang, Y.: Random batch particle methods for the homogeneous Landau equation. *Commun. Comput. Phys.* **31**, 997–1019 (2022)
13. Choi, Y.-P., Ha, S.-Y., Yun, S.-B.: Complete synchronization of Kuramoto oscillators with finite inertia. *Physica D* **240**, 32–44 (2011)
14. Cucker, F., Smale, S.: Emergent behavior in flocks. *IEEE Trans. Autom. Control* **52**, 852–862 (2007)
15. Dimarco, G., Pareschi, L.: Multi-scale control variate methods for uncertainty quantification in kinetic equations. *J. Comput. Phys.* **388**, 63–89 (2019)
16. Dimarco, G., Pareschi, L.: Multiscale variance reduction methods based on multiple control variates for kinetic equations with uncertainties. *Multiscale Model. Simul.* **18**(1), 351–382 (2020)
17. D’Orsogna, M.R., Chuang, Y.L., Bertozzi, A.L., Chayes, L.: Self-propelled particles with soft-core interactions: patterns, stability and collapse. *Phys. Rev. Lett.* **96**(10), 104302 (2006)
18. Düring, B., Markowich, P., Pieschmann, J.F., Wolfram, M.-T.: Boltzmann and Fokker-Planck equations modelling opinion formation in the presence of strong leaders. *Proc. R. Soc. A* **465**, 3687–3708 (2009)
19. Figalli, A., Kang, M.-J.: A rigorous derivation from the kinetic Cucker-Smale model to the pressureless Euler system with nonlocal alignment. *Anal. PDE* **12**(3), 843–866 (2019)
20. Furioli, G., Pulvirenti, A., Terraneo, E., Toscani, G.: Fokker-Planck equations in the modeling of socio-economic phenomena. *Math. Models Methods Appl. Sci.* **27**(01), 115–158 (2017)
21. Garnier, J., Papanicolaou, G., Yang, T.-W.: Consensus convergence with stochastic effects. *Vietnam J. Math.* **45**(1–2), 51–75 (2017)
22. Giles, M.B.: Multilevel Monte Carlo methods. *Acta Numer.* **24**, 259–328 (2015)
23. Ha, S.-Y., Liu, J.-G.: A simple proof of the Cucker-Smale flocking dynamics and mean-field limit. *Commun. Math. Sci.* **7**, 297–325 (2009)
24. Ha, S.-Y., Tadmor, E.: From particle to kinetic and hydrodynamic descriptions of flocking. *Kinet. Relat. Models* **1**(3), 415–435 (2008)
25. Hegselmann, R., Krause, U.: Opinion dynamics and bounded confidence: models, analysis and simulation. *J. Artif. Soc. Soc. Simul.* **5**(3) (2022)
26. Herty, M., Pareschi, L., Visconti, G.: Mean field models for large data-clustering problems. *Netw. Heterog. Media* **15**(3), 463–487 (2020)
27. Jabin, P.E.: A review of the mean field limits for Vlasov equations. *Kinet. Relat. Models* **7**, 661–711 (2020)
28. Jin, S., Li, L.: On the mean field limit of the Random Batch Method for interacting particle systems. *Sci. China Math.* **65**, 169–202 (2022)
29. Jin, S., Li, L.: Random batch methods for classical and quantum interacting particle systems and statistical samplings. In: Bellomo, N., Carrillo, J.A., Tadmor, E. (eds.) *Active Particles*, vol. 3. Modeling and Simulation in Science, Engineering and Technology. Birkhäuser, Cham (2022)
30. Jin, S., Li, L., Liu, J.-G.: Random Batch Methods (RBM) for interacting particle systems. *J. Comput. Phys.* **400**, 108877 (2020)

31. Jin, S., Li, L., Ye, X., Zhou, Z.: Ergodicity and long-time behavior of the Random Batch Method for interacting particle systems. *Math. Models Methods Appl. Sci.* **33**(01), 67–102 (2023)
32. Ko, D., Zuazua, E.: Model predictive control with random batch methods for a guiding problem. *Math. Models Methods Appl. Sci.* **31**(8), 1569–1592 (2021)
33. Motsch, S., Tadmor, E.: Heterophilious dynamics enhances consensus. *SIAM Rev.* **56**(4), 577–621 (2014)
34. Pareschi, L.: An introduction to uncertainty quantification for kinetic equations and related problems. In: Albi, G., Merino-Aceituno, S., Nota, A., Zanella, M. (eds.) *Trails in Kinetic Theory*. SEMA SIMAI Springer Series, vol. 25, pp. 141–181 (2021)
35. Pareschi, L., Zanella, M.: Structure preserving schemes for nonlinear Fokker-Planck equations and applications. *J. Sci. Comput.* **74**(3), 1575–1600 (2018)
36. Pareschi, L., Toscani, G., Tosin, A., Zanella, M.: Hydrodynamic models of preference formation in multi-agent societies. *J. Nonlinear Sci.* **29**(6), 2761–2796 (2019)
37. Pinnau, R., Totzeck, C., Tse, O., Martin, S.: A consensus-based model for global optimization and its mean-field limit. *Math. Models Methods Appl. Sci.* **27**(1), 183–204 (2017)
38. Shu, R., Tadmor, E.: Anticipation breeds alignment. *Arch. Ration. Mech. Anal.* **240**, 203–241 (2021)
39. Toscani, G.: Entropy production and the rate of convergence to equilibrium for the Fokker-Planck equation. *Q. Appl. Math.* **57**, 521–541 (1999)
40. Toscani, G.: Kinetic models of opinion formation. *Commun. Math. Sci.* **4**(3), 481–496 (2006)
41. Villani, C.: A review of mathematical topics in collisional kinetic theory. In: *Handbook of Mathematical Fluid Dynamics*, vol. 1, pp. 3–8 (2002)
42. Ye, X., Zhou, Z.: Error analysis of time-discrete random batch method for interacting particle systems and associated mean-field limits. *IMA J. Numer. Anal.* drad043 (2023)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.