On the Mean-Field Limit for the Vlasov–Poisson–Fokker–Planck System

Hui Huang[1] · Jian-Guo Liu[2] · Peter Pickl[3,4]

Abstract

We rigorously justify the mean-field limit of an N-particle system subject to Brownian motions and interacting through the Newtonian potential in $\mathbb{R}^3$. Our result leads to a derivation of the Vlasov–Poisson–Fokker–Planck (VPFP) equations from the regularized microscopic N-particle system. More precisely, we show that the maximal distance between the exact microscopic trajectories and the mean-field trajectories is bounded by $N^{-\frac{1}{3}+\varepsilon}$ ($\frac{1}{63} \leq \varepsilon < \frac{1}{36}$) with a blob size of $N^{-\delta}$ ($\frac{1}{3} \leq \delta < \frac{19}{54} - \frac{2\varepsilon}{3}$) up to a probability of $1 - N^{-\alpha}$ for any $\alpha > 0$. Moreover, we prove the convergence rate between the empirical measure associated to the regularized particle system and the solution of the VPFP equations. The technical novelty of this paper is that our estimates rely on the randomness coming from the initial data and from the Brownian motions.

1 Introduction

Systems of interacting particles are quite common in physics and biosciences, and they are usually formulated according to first principles (such as Newton's second law). For instance,

---

✉ Peter Pickl
  pickl@math.lmu.de; peter.pickl@duke.edu

  Hui Huang
  hui.huang@tum.de

  Jian-Guo Liu
  jliu@phy.duke.edu

[1] Department of Mathematics, Technische Universität München, München, Germany

[2] Departments of Physics and Mathematics, Duke University, Durham, NC, USA

[3] Mathematisches Institut, Universität München, München, Germany

[4] Duke Kunshan University, Kunshan, Jiangsu, China

🖄 Springer

particles can represent galaxies in cosmological models [1], molecules in a fluid [33], or ions and electrons in plasmas [61]. Such particle systems are also relevant as models for the collective behavior of certain animals like birds, fish, insects, and even micro-organisms (such as cells or bacteria) [5,13,48].

Such systems are often described by time dependent trajectory on the N-body configuration space. Depending on the equation of motion, in particular if it is first or second order, the configuration space is given by the collection of all particle positions or all positions and velocities. The latter is used to translate the second order equation into a first order system. The trajectory is then given by the integral curves of the respective flow.

Since numeric and analytic treatment of many body systems is often complicated, effective descriptions which describe collective degrees of freedom of the system are sought for. These effective equations describe collective degrees of freedom of the system, for example temperature, pressure or, as it is the topic of the present manuscript, the density.

In this paper, we are interested in the second order case, in particular the classical Newtonian dynamics of N indistinguishable particles interacting through pair interaction forces and subject to Brownian noise. Denote by $x_i \in \mathbb{R}^3$ and $v_i \in \mathbb{R}^3$ the position and velocity of particle i. The evolution of the system is given by the following stochastic differential equations (SDEs),

$$dx_i = v_i dt, \quad dv_i = \frac{1}{N-1} \sum_{j \neq i}^{N} k(x_i - x_j) dt + \sqrt{2\sigma} dB_i, \quad i = 1, \ldots, N, \qquad (1)$$

where $k(x)$ models the pairwise interaction between the individuals, and $\{B_i\}_{i=1}^{N}$ are independent realizations of Brownian motions which count for extrinsic random perturbations such as random collisions against the background. In the presence of friction, model (1) is known as the interacting Ornstein–Uhlenbeck model in the probability or statistical mechanics community. In particular, we refer readers to [49,58] by Olla, Varadhan and Tremoulet for the scaling limit of the Ornstein–Uhlenbeck system. In this manuscript, we take the interaction kernel to be the Coulombian kernel

$$k(x) = a \frac{x}{|x|^3}, \qquad (2)$$

for some real number a. The case $a > 0$ corresponds, for example, to the electrostatic (repulsive) interaction of charged particles in a plasma, while the case $a < 0$ describes the attraction between massive particles subject to gravitation. We refer readers to [37,61] for the original modelings.

Since the number N of particles is large, it is extremely complicated to investigate the microscopic particle system (1) directly. Fortunately, it can be studied through macroscopic descriptions of the system based on the probability density for the particles on phase space. These macroscopic descriptions are usually expressed as continuous partial differential equations (PDEs). The analysis of the scaling limit of the interacting particle system to the macroscopic continuum model is usually called the *mean-field limit*. For the second order particle system (1), it is expected to be approximated by the following Vlasov–Poisson–Fokker-Planck (VPFP) equations

$$\begin{cases} \partial_t f(x, v, t) + v \cdot \nabla_x f(x, v, t) + k * \rho(x, t) \cdot \nabla_v f(x, v, t) = \sigma \Delta_v f(x, v, t), \\ f(x, v, 0) = f_0(x, v), \end{cases} \qquad (3)$$

where $f(x, v, t) : (x, v, t) \in \mathbb{R}^3 \times \mathbb{R}^3 \times [0, \infty) \to \mathbb{R}^+$ is the probability density function in the phase space (x, v) at time t, and

$$\rho(x, t) = \int_{\mathbb{R}^3} f(x, v, t) dv, \tag{4}$$

is the charge density introduced by $f(x, v, t)$. We denote by $E(x, t) := k * \rho(x, t)$ the Coulombian or gravitational force field.

The intent of this research is to show the mean-field limit of the particle system (1) towards the Vlasov–Poisson–Fokker-Planck equations (3). In particular, we quantify how close these descriptions are for a given N. Where $\sigma = 0$ (there is no randomness coming from the noise), mean-field limit results for interacting particle systems with globally Lipschitz forces have been obtained by Braun and Hepp [9] and Dobrushin [16]. Bolley et al. [5] presented an extension of the classical theory to the particle system with only locally Lipschitz interacting force. Such case concerning kernels $k \in W_{loc}^{1,\infty}$ were also used in the context of neuroscience [6,57]. The last few years have seen great progress in mean-field limits for singular forces by treating them with an N-dependent cut-off. In particular, Hauray and Jabin [32] discussed mildly singular force kernels satisfying $|k(x)| \leq \frac{C}{|x|^\alpha}$ with $\alpha < d - 1$ in dimensions $d \geq 3$. For $1 < \alpha < d - 1$, they performed the mean-field limit for typical initial data, where they chose the cut-off to be $N^{-\frac{1}{2d}}$. For $\alpha < 1$, they proved molecular chaos without cut-off. Unfortunately, their method failed precisely at the Coulomb threshold when $\alpha = d - 1$. More recently, Boers and Pickl [4] proposed a novel method for deriving mean-field equations with interaction forces scaling like $\frac{1}{|x|^{3\lambda-1}}$ ($5/6 < \lambda < 1$), and they were able to obtain a cut-off as small as $N^{-\frac{1}{d}}$. Furthermore, Lazarovici and Pickl [39] extended the method in [4] to include the Coulomb singularity and they obtained a microscopic derivation of the Vlasov–Poisson equations with a cut-off of $N^{-\delta}$ ($0 < \delta < \frac{1}{d}$). More recently, the cut-off parameter was reduced to as small as $N^{-\frac{7}{18}}$ in [23] by using the second order nature of the dynamics. Where $\sigma > 0$, the random particle method for approximating the VPFP system with the Coulombian kernel was studied in [26], where the initial data was chose on a mesh and the cut-off parameter could be $N^{-\delta}$ ($0 < \delta < \frac{1}{d}$). Most recently, Carrillo et al. [12] also investigated the singular VPFP system but with the i.i.d. initial data, and obtained the propagation of chaos through a cut-off of $N^{-\delta}$ ($0 < \delta < \frac{1}{d}$), which was a generalization of [39]. We also note that Jabin and Wang [34] rigorously justified the mean-field limit and propagation of chaos for the Vlasov systems with L^∞ forces and vanishing viscosity ($\sigma_N \to 0$ as $N \to \infty$) by using a relative entropy method. They were also able to give quantitative estimates for stochastic systems with integration kernel in the negative Sobolev space $W^{-1,\infty}$ [36]. Moreover, [54] introduced a relative entropy method at the the level of the empirical measure based on the energy of the system, which lead to mean-field limit for quite singular interactions of Riesz potential form without a cut-off. Lastly, for a general overview of this topic we refer readers to [13,31,35,56].

When the interacting kernel k is singular, it poses problems for both theory and numerical simulations. An easy remedy is to regularize the force with an N-dependent cut-off parameter and get k^N. The delicate question is how to choose this cut-off. On the one hand, the larger the cut-off is, the smoother k^N will be, and the easier it will be to show the convergence. However, the regularized system is not a good approximation of the actual system. On the other hand, the smaller the cut-off is, the closer k^N is to the real k, thus the less information will be lost through the cut-off. Consequently, the necessary balance between accuracy (small cut-off) and regularity (large cut-off) is crucial. The analysis we reviewed above tried to justify that. In this manuscript, we set $\sigma > 0$. Compared with the recent work [12], the main

technical innovation of this paper is that we fully use the randomness coming from the initial conditions *and* the Brownian motions to significantly improve the cut-off. Note that in [12] the size of cut-off can be very close to but larger than $N^{-\frac{1}{d}}$. However we manage to reduce the cut-off size to be smaller than $N^{-\frac{1}{d}}$ (see Remark 1.4), which is a sort of average minimal distance between N particles in dimension d. This manuscript significantly improves the ideas presented in [10]. There the potential is split up into a more singular and less singular part. The less singular part is controlled in the usual manner while the mixing coming from the Brownian motion is used to estimate the more singular part. The technical innovation in the present paper is that the possible number of particles subject to the singular part of the interaction can be bounded due to the fact that the support of the singular part is small using a Law of Large Numbers argument. Again using the Law of Large Numbers based on the randomness coming from the Brownian motion, we show that the leading order of the singular part of the interaction can be replaced by its expectation value. This step is a key point of the present manuscript. The replacement by the expectation value, i.e. the integration of the force against the probability density, gives the regularization of the singular part and gives a significant improvement of our estimates. This is carried out in Lemma 3.3, the proof of which can be found in Sect. 5. [10] and the present paper are, to our knowledge, so far the only results where the mixing from the Brownian motion has been used in the derivation of a mean-field limit for an interacting many-body system.

As a companion of (1), some also consider the first order stochastic system

$$dx_i = \frac{1}{N-1} \sum_{j \neq i}^{N} k(x_i - x_j)dt + \sqrt{2\sigma}\, dB_i, \quad i = 1, \cdots, N. \tag{5}$$

As before, one can expect that as the number of the particles N goes to infinity we can get the continuous description of the dynamics as the following nonlinear PDE

$$\partial_t f(x, t) + \nabla \cdot (f(k * f)) = \sigma \Delta_x f, \tag{6}$$

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

The particle system (5) has many important applications. One of the best known classical applications is in fluid dynamics with the Biot-Savart kernel

$$k(x) = \frac{1}{2\pi} \left(\frac{-x_2}{|x|^2}, \frac{x_1}{|x|^2} \right). \tag{7}$$

It can be treated by the well-known vortex method introduced by Chorin in 1973 [14]. The convergence of the vortex method for two and three dimensional inviscid ($\sigma = 0$) incompressible fluid flows was first proved by Hald et al. [24,25], Beale and Majda [2,3]. When the effect of viscosity is involved ($\sigma > 0$), the vortex method is replaced by the so called random vortex method by adding a Brownian motion to every vortex. The convergence analysis of the random vortex method for the Navier–Stokes equations was given by [22,46, 47,51] in the 1980s. For more recent results we refer to [17,19,36,54]. Another well-known application of the system (5) is to choose the interaction to be the Poisson kernel

$$k(x) = -C_d \frac{x}{|x|^d}, \quad d \geq 2, \tag{8}$$

where $C_d > 0$ and k is set to be attractive. Now the system (5) coincides with the particle models to approximate the classical Keller–Segel (KS) equation for chemotaxis [38,52]. We mainly refer to [10,18,28–30,43,44] for the mean-field limit of the KS system. Concerning the size of the cut-off, more specifically, [43] chose the cut-off to be $(\ln N)^{-\frac{1}{d}}$, which was

significantly improved in [28], where the cut-off size could be as small as $N^{-\frac{1}{d(d+1)}} \log(N)$. In [10,29], the cut-off size was almost optimal, coming fairly close to $N^{-\frac{1}{d}}$. Many techniques used in this manuscript are adapted from these papers. For the Poisson-Nernst-Planck equation (k is set to be repulsive), [41,43] proved the mean-field limit without a cut-off.

The rest of the introduction will be split into three parts: We start with introducing the microscopic random particle system in Sect. 1.1. Then we present some results on the existence of the macroscopic mean-field VPFP equations in Sect. 1.2. Lastly, our main theorem will be stated in Sect. 1.3, where we prove the closeness of the approximation of solutions to VPFP equations by the microscopic system.

1.1 Microscopic Random Particle System

We are interested in the time evolution of a system of N-interacting Newtonian particles with noise in the $N \to \infty$ limit. The motion of the system studied in this paper is described by trajectories on phase space, i.e. a time dependent $\Phi_t : \mathbb{R} \to \mathbb{R}^{6N}$. We use the notation

$$\Phi_t := (X_t, V_t) := \left(x_1^t, \ldots, x_N^t, v_1^t, \ldots v_N^t \right), \tag{9}$$

where x_j^t stands for the position of the jth particle at time t and v_j^t stands for the velocity of the jth particle at time t. The system is a Newtonian system with a noise term coupled to the velocity, whose evolution is governed by a system of SDEs of the type

$$\begin{cases} dx_i^t = v_i^t dt, \quad i = 1, \ldots, N, \\ dv_i^t = \dfrac{1}{N-1} \displaystyle\sum_{j \neq i}^{N} k(x_i^t - x_j^t) dt + \sqrt{2\sigma} dB_i^t, \end{cases} \tag{10}$$

where k is the Coulomb kernel (2) modeling interaction between particles and B_i^t are independent realizations of Brownian motions.

We regularize the kernel k by a blob function $0 \leq \psi(x) \in C^2(\mathbb{R}^3)$, supp $\psi(x) \subseteq B(0, 1)$ and $\int_{\mathbb{R}^3} \psi(x) dx = 1$. Let $\psi_\delta^N = N^{3\delta} \psi(N^\delta x)$, then the Coulomb kernel with regularization has the form

$$k^N(x) = k * \psi_\delta^N. \tag{11}$$

Thus one has the regularized microscopic N-particle system for $i = 1, 2 \ldots, N$

$$\begin{cases} dx_i^t = v_i^t dt, \\ dv_i^t = \dfrac{1}{N-1} \displaystyle\sum_{i \neq j}^{N} k^N(x_i^t - x_j^t) dt + \sqrt{2\sigma} dB_i^t. \end{cases} \tag{12}$$

Here the initial condition Φ_0 of the system is independently, identically distributed (i.i.d.) with the common probability density given by f_0. And the corresponding regularized VPFP equations are

$$\begin{cases} \partial_t f^N(x, v, t) + v \cdot \nabla_x f^N(x, v, t) + k^N * \rho^N(x, t) \cdot \nabla_v f^N(x, v, t) = \sigma \Delta_v f^N(x, v, t), \\ \rho^N(x, t) = \displaystyle\int_{\mathbb{R}^3} f^N(x, v, t) dv, \\ f^N(x, v, 0) = f_0(x, v). \end{cases} \tag{13}$$

1.2 Existence of Classical Solutions to the Vlasov–Poisson–Fokker-Planck System

The existence of weak and classical solutions to VPFP equations (3) and related systems has been very well studied. Degond [15] first showed the existence of a global-in-time smooth solution for the Vlasov–Fokker–Planck equations in one and two space dimensions in the electrostatic case. Later on, Bouchut [7,8] extended the result to three dimensions when the electric field was coupled through a Poisson equation, and the results were given in both the electrostatic and gravitational case. Also, Victory and O'Dwyer [59] showed existence of classical solutions for VPFP equations when the spacial dimension was less than or equal to two, and local existence for all other dimensions. Then Bouchut in [7] proved the global existence of classical solutions for the VPFP system (3) in dimension $d = 3$. His proof relied on the techniques introduced by Lions and Perthame [42] concerning the existence to the Vlasov–Poisson system in three dimensions. The long time behavior of the VPFP system was studied by Ono and Strauss [50], Carpio [11] and Carrillo et al. [55].

The existence results in [7,59] are most appropriate for this work. We summarize them in the following theorem, which was also used in [26, Theorem 2.1].

Theorem 1.1 (Classical solutions of the VPFP equations) *Let the initial data $0 \leq f_0(x, v)$ satisfies the following properties:*

(a) $f_0 \in W^{1,1} \cap W^{1,\infty}(\mathbb{R}^6)$;
(b) there exists a $m_0 > 6$, such that

$$(1 + |v|^2)^{\frac{m_0}{2}} f_0 \in W^{1,\infty}(\mathbb{R}^6). \tag{14}$$

Then for any $T > 0$, the VPFP equations (3) admits a unique classical solution on $[0, T]$.

Remark 1.1 The proof of the above theorem given in [7,59] indicates that the map

$$t \to E(\cdot, t) := k * \rho(\cdot, t) \tag{15}$$

is a continuous map from $[0, T]$ to $W^{1,\infty}(\mathbb{R}^3)$. This implies that initial smooth data remains smooth for all time intervals $[0, T]$. So if we assume the initial data satisfies the following for any $k \geq 1$

(a) $f_0 \in W^{k,1} \cap W^{k,\infty}(\mathbb{R}^6)$;
(b) there exists a $m_0 > 6$, such that

$$(1 + |v|^2)^{\frac{m_0}{2}} f_0 \in W^{k,1} \cap W^{k,\infty}(\mathbb{R}^6). \tag{16}$$

Then the unique classical solution f maintains the regularity on $[0, T]$ for any $k \geq 1$:

$$\max_{0 \leq t \leq T} \|(1 + |v|^2)^{\frac{m_0}{2}} f_t\|_{W^{k,1} \cap W^{k,\infty}(\mathbb{R}^6)} < \infty. \tag{17}$$

The present paper also needs the uniform-in-time L^∞ bound of the charge density ρ:

$$\max_{0 \leq t \leq T} \|\rho(\cdot, t)\|_{W^{1,\infty}(\mathbb{R}^3)} < \infty, \tag{18}$$

which was obtained in [53] by means of the stochastic characteristic method under the assumption the f_0 was compactly supported in velocity. We also note that [12] provided a proof of the local-in-time L^∞ bound for ρ by employing Feynman–Kac's formula and assuming the initial data has polynomial decay.

In this paper, we assume that the initial data f_0 satisfies the following assumption:

Assumption 1.1 The initial data $0 \leq f_0(x, v)$ satisfies

1. $f_0 \in W^{1,1} \cap W^{1,\infty}(\mathbb{R}^6)$;
2. there exists a $m_0 > 6$, such that

$$(1 + |v|^2)^{\frac{m_0}{2}} f_0 \in W^{1,1} \cap W^{1,\infty}(\mathbb{R}^6); \tag{19}$$

3. $f_0(x, v) = 0$ when $|v| > Q_v$.

The above assumption makes sure that we have the regularity needed for this article: for any $T > 0$,

$$\max_{t \in [0,T]} \|\rho(\cdot, t)\|_{W^{1,\infty}(\mathbb{R}^3)} + \max_{0 \leq t \leq T} \|(1 + |v|^2)^{\frac{m_0}{2}} f_t\|_{W^{1,1} \cap W^{1,\infty}(\mathbb{R}^6)} \leq C_{f_0}, \tag{20}$$

where C_{f_0} depends only on $\|f_0\|_{W^{1,1} \cap W^{1,\infty}(\mathbb{R}^6)}$, $\|(1+|v|^2)^{\frac{m_0}{2}} f_0\|_{W^{1,1} \cap W^{1,\infty}(\mathbb{R}^6)}$ and Q_v. Note that the charge density ρ satisfies

$$\partial_t \rho(x, t) + \int_{\mathbb{R}^3} v \cdot \nabla_x f(x, v, t) dv = 0 . \tag{21}$$

Thus we have

$$\max_{t \in [0,T]} \|\partial_t \rho(\cdot, t)\|_{L^\infty(\mathbb{R}^3)}$$

$$\leq \max_{t \in [0,T]} \left\| \int_{\mathbb{R}^3} |v| |\nabla_x f(x, v, t)| dv \right\|_{L^\infty(\mathbb{R}^3)}$$

$$\leq \max_{t \in [0,T]} \|(1 + |v|^2)^{\frac{m_0}{2}} f_t\|_{W^{1,\infty}(\mathbb{R}^6)} \int_{\mathbb{R}^3} \frac{|v|}{(1 + |v|^2)^{\frac{m_0}{2}}} dv \leq C_{f_0} . \tag{22}$$

We also note that equivalently one can estimate a bound for f^N and ρ^N uniformly in N.

Remark 1.2 The assumption that f_0 is compactly supported in the velocity variable is not required for the existence of the VPFP system. However it is used to get the L^∞ bound of the charge density ρ (see in [53]) and also in the proof of Lemma 3.1 (see in (84)).

Remark 1.3 All our estimates below are also possible in the presence of sufficiently smooth external fields. Due to the fluctuation-dissipation principle it is more natural to add an external, velocity-dependent friction force to the system.

1.3 Statement of the Main Results

Our objective is to derive the macroscopic mean-field PDE (3) from the regularized microscopic particle system (12). We will do this by using probabilistic methods as in [10,28,29,39]. More precisely, we shall prove the convergence rate between the solution of VPFP equations (3) and the empirical measure associated to the regularized particle system Φ_t satisfying (12). We assume that the initial condition Φ_0 of the system is independently, identically distributed (i.i.d.) with the common probability density given by f_0.

Given the solution f^N to the mean-field Eq. (13), we first construct an auxiliary trajectory Ψ_t from (13). Then we prove the closeness between Φ_t and Ψ_t. For the auxiliary trajectory

$$\Psi_t := \left(\overline{X}_t, \overline{V}_t\right) = \left(\overline{x}_1^t, \ldots, \overline{x}_N^t, \overline{v}_1^t, \ldots \overline{v}_N^t\right), \tag{23}$$

we shall consider again a Newtonian system with noise, however, this time not subject to the pair interaction but under the influence of the external mean field $k^N * \rho^N(x, t)$

$$\begin{cases} d\overline{x}_i^t = \overline{v}_i^t dt, \quad i = 1, \ldots, N, \\ d\overline{v}_i^t = \int_{\mathbb{R}^3} k^N(\overline{x}_i^t - x)\rho^N(x, t)dxdt + \sqrt{2\sigma}dB_i^t. \end{cases} \tag{24}$$

Here we let Ψ_t have the same initial condition as Φ_t (i.i.d. with the common density f_0). Since the particles are just N identical copies of evolution, the independence is conserved. Therefore the Ψ_t are distributed i.i.d. according to the common probability density f^N. We remark that the VPFP equation (13) is Kolmogorov's forward equation for any solution of (24), and in particular their probability density f^N solves (13). This i.i.d. property will play a crucial role below, where we shall use the concentration inequality (see in Lemma 2.5) on some functions depending on Ψ_t.

Our main result states that the N-particle trajectory Φ_t starting from Φ_0 (i.i.d. with the common density f_0) remains close to the mean-field trajectory Ψ_t with the same initial configuration $\Phi_0 = \Psi_0$ during any finite time $[0, T]$. More precisely, we prove that the measure of the set where the maximal distance $\max_{t \in [0, T]} \|\Phi_t - \Psi_t\|_\infty$ on $[0, T]$ exceeds $N^{-\lambda_2}$ decreases exponentially as the number N of particles grows to infinity. Here the distance $\|\Phi_t - \Psi_t\|_\infty$ is measured by

$$\|\Phi_t - \Psi_t\|_\infty := \sqrt{\log(N)}\|X_t - \overline{X}_t\|_\infty + \|V_t - \overline{V}_t\|_\infty. \tag{25}$$

Theorem 1.2 *For any $T > 0$, assume that trajectories $\Phi_t = (X_t, V_t)$, $\Psi_t = (\overline{X}_t, \overline{V}_t)$ satisfy (12) and (24) respectively with the initial data $\Phi_0 = \Psi_0$, which is i.i.d. sharing the common density f_0 that satisfies Assumption 1.1. Then for any $\alpha > 0$ and $0 < \lambda_2 < \frac{1}{3}$, there exists some $0 < \lambda_1 < \frac{\lambda_2}{3}$ and a $N_0 \in \mathbb{N}$ which both depend only on α, T and C_{f_0}, such that for $N \geq N_0$, the following estimate holds with the cut-off index $\delta \in \left[\frac{1}{3}, \min\left\{\frac{\lambda_1 + 3\lambda_2 + 1}{6}, \frac{1 - \lambda_2}{2}\right\}\right)$*

$$\mathbb{P}\left(\max_{t \in [0, T]} \|\Phi_t - \Psi_t\|_\infty \leq N^{-\lambda_2}\right) \geq 1 - N^{-\alpha},$$

where $\|\Phi_t - \Psi_t\|_\infty$ is defined in (25).

Remark 1.4 *In particular, for any $\frac{1}{63} \leq \varepsilon < \frac{1}{36}$, choosing $\lambda_2 = \frac{1}{3} - \varepsilon$ and $\lambda_1 = \frac{1}{9} - \varepsilon$, we have a convergence rate $N^{-\frac{1}{3} + \varepsilon}$ with a cut-off size of $N^{-\delta}$ ($\frac{1}{3} \leq \delta < \frac{19}{54} - \frac{2\varepsilon}{3}$). In other words, the cut-off parameter δ can be chosen very close to $\frac{19}{54}$ and in particular larger than $\frac{1}{3}$, which is a significant improvement over previous results in the literature.*

Strategy of the proof

Before we explain our strategy in more details, we would like to explain the novelty of this paper, namely the use of the Brownian motion. Of course Brownian motion helps for the solution theory of nonlinear PDE. Existence, both local and global, as well as uniqueness are typically easier to prove or hold with more generality, given that the PDE includes a term which smooths the density. In this case the smoothing comes from the heat kernel, which is due to the Brownian motion on the microscopic level. This paper, however, focuses on the derivation of the effective description from the microscopic dynamics. We cite the respective results on the existence of classical solutions with the properties we need, which is done in Theorem 1.1. For the proof we refer to the literature.

However, in addition to stronger results on the solution theory, there is also a direct advantage for the proof of validity of the effective description from the microscopic dynamics,

which plays an important part in the technicalities of this paper. The argument is, on the heuristic level, implemented in the following way: All our proof is based on controlling the growth of the difference between the trajectory of the many body system and the trajectory of the mean field system, which starts with the same initial configuration but assumes that the particles follow the flow generated from the mean-field interaction. This distance is measured in maximum norm, so it gives a control for the deviation of each single particle from its mean-field trajectory. We use Gronwall's inequality, i.e. assuming that this distance has a certain value, we control the growth of this distance. Under this assumption we are blind to many details of our system. Having information on the distance for each particle, for example, one does not get a good control of the clustering of particles. This is because the density in our system is high, and the worst case scenario can still lead to large clusters assuming that the estimate on the distance of the true positions of the particles from the mean-field trajectory is much larger than the intermediate distance of particles. Without using the effect of Brownian motion in our proof we would now have to assume the worst case, i.e. that the particles form the largest possible clusters. Of course, given the Brownian motion, this is not very probable. Therefore the idea is to show that the worst clusters have only small probability using a law of large numbers argument based on the randomness we get from the Brownian motion.

Note that for other arguments the randomness coming from the initial condition is helpful. Thus we will distinguish between the two in the proof below. In more detail: the strategy is to obtain a Gronwall-type inequality for $\max_{t \in [0,T]} \|\Phi_t - \Psi_t\|_\infty$. Notice that

$$\frac{d(V_t - \overline{V}_t)}{dt} = K^N(X_t) - \overline{K}^N(\overline{X}_t),$$

where $K^N(X_t)$ and $\overline{K}^N(\overline{X}_t)$ are defined as

$$(K^N(X_t))_i := \frac{1}{N-1} \sum_{j \neq i} k^N(x_i^t - x_j^t); \quad (\overline{K}^N(\overline{X}_t))_i := \int_{\mathbb{R}^3} k^N(\overline{x}_i^t - x) \rho^N(x,t) dx \tag{26}$$

One can compute

$$\frac{d\|\Phi_t - \Psi_t\|_\infty}{dt} \leq \sqrt{\log(N)} \left\| V_t - \overline{V}_t \right\|_\infty + \left\| K^N(X_t) - \overline{K}^N(\overline{X}_t) \right\|_\infty$$
$$\leq \sqrt{\log(N)} \left\| V_t - \overline{V}_t \right\|_\infty + \left\| K^N(X_t) - K^N(\overline{X}_t) \right\|_\infty + \left\| K^N(\overline{X}_t) - \overline{K}^N(\overline{X}_t) \right\|_\infty. \tag{27}$$

If the force k^N is Lipschitz continuous with a Lipschitz constant independent of N, the desired convergence follows easily [9,16]. However the force considered here becomes singular as $N \to \infty$, hence it does not satisfy a uniform Lipschitz bound.

The first term in (27) is already a sufficient bound in view of Gronwall's Lemma.

• By the Law of Large Numbers, carried out in detail for our purpose here in Proposition 3.1, we show for any $T > 0$

$$\max_{t \in [0,T]} \left\| K^N(\overline{X}_t) - \overline{K}^N(\overline{X}_t) \right\|_\infty \leq CN^{2\delta-1} \log(N), \tag{28}$$

where for convenience we abused the notation $a \preceq b$ to denote $a \leq b$ except for an event with probability approaching zero. This direct error estimate can be seen as a consistency of the two evolution in high probability.

• In Proposition 3.2 and Lemma 3.4, we show that the propagation of errors, coming from the second term in (27), is stable. This stability is important to be able to close the Gronwall

argument. We show that for any $T > 0$

$$\|K^N(X_t) - K^N(\overline{X}_t)\|_\infty \preceq C \log(N) \left\|X_t - \overline{X}_t\right\|_\infty + C \log^2(N) N^{-\lambda_3} , \text{ for all } t \in [0, T] \quad (29)$$

holds under the condition that

$$\max_{t \in [0,T]} \|\Phi_t - \Psi_t\|_\infty \preceq N^{-\lambda_2} . \quad (30)$$

Here it is crucial to ensure the constant λ_3 satisfies $2\delta - 1 \leq -\lambda_3 < -\lambda_2 < 0$.

To get this improvement of the cutoff parameter compared to previous results in the literature, we make use of the mixing caused by the Brownian motion. Therefore we split potential $K^N := K_1^N + K_2^N$, where K_2^N is chosen to have a wider cut-off of order $N^{-\lambda_2} > N^{-\delta}$. The less singular part K_2^N is controlled in the usual manner [4,28,29,39] (see in estimate (116)):

$$\|K_2^N(X_t) - K_2^N(\overline{X}_t)\|_\infty \preceq C \log(N) \left\|X_t - \overline{X}_t\right\|_\infty . \quad (31)$$

Thus we are left with the force K_1^N. We shall first estimate the number of particles that will be present in the support of K_1^N. Since the latter is small, this number will always be very small compared to N.

Under the condition (30) we can not track the particles of the Newtonian time evolution with an accuracy larger than $N^{-\lambda_2}$. Thus – without using the Brownian motion – we have to assume the worst case scenario, which is all particles giving the maximal possible solution to the force, i.e. sitting close to the edge of the cutoff region, and all forces summing up, i.e. all particles sitting on top of each other.

But the Brownian motion in our system will lead to mixing. For a short time interval the effect of mixing will be much larger than the effect of the correlations coming from the pair interaction, and we can make use of the independence of the Brownian motions. This mixing, which happens on a larger spacial scale than the range of the potential, causes the particles to be distributed roughly equally over the support of the interaction resulting in a cancellation of the leading order of K_1^N. It follows for the more singular part K_1^N that there exists some constant λ_3 satisfying $2\delta - 1 \leq -\lambda_3 < -\lambda_2 < 0$ such that

$$\|K_1^N(X_t) - K_1^N(\overline{X}_t)\|_\infty \preceq C \log^2(N) N^{-\lambda_3} . \quad (32)$$

This is mainly carried out in Sect. 5 (the proof of Lemma 3.3).

- Combining consistency (28) and stability (29), we conclude that

$$\frac{d\|\Phi_t - \Psi_t\|_\infty}{dt} \preceq C \log(N) \left\|X_t - \overline{X}_t\right\|_\infty + C \log^2(N) N^{-\lambda_3} , \text{ for all } t \in (0, T] , \quad (33)$$

where $-\lambda_3 < -\lambda_2 < 0$. Gronwall's inequality implies

$$\max_{t \in [0,T]} \|\Phi_t - \Psi_t\|_\infty \preceq e^{C\sqrt{\log(N)}T} C \log^2(N) N^{-\lambda_3} T \leq \frac{1}{2} N^{-\lambda_2} < N^{-\lambda_2} , \quad (34)$$

which verifies the condition (30). Hence it completes our proof.

To quantify the convergence of probability measures, we give a brief introduction on the topology of the p-Wasserstein space. In the context of kinetic equations, it was first introduced by Dobrushin [16]. Consider the following probability space with finite p-th moment:

$$\mathcal{P}_p(\mathbb{R}^d) = \left\{\mu | \, \mu \text{ is a probability measure on } \mathbb{R}^d \text{ and } \int_{\mathbb{R}^d} |x|^p d\mu(x) < +\infty\right\}. \quad (35)$$

We denote the Monge-Kantorovich-Wasserstein distance in $\mathcal{P}_p(\mathbb{R}^d)$ as follows

$$W_p^p(\mu, \nu) = \inf_{\pi \in \Lambda(\mu, \nu)} \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p d\pi(x, y) \right\} = \inf_{X \sim \mu, Y \sim \nu} \left\{ \mathbb{E}[|X - Y|^p] \right\}, \quad (36)$$

where $\Lambda(\mu, \nu)$ is the set of joint probability measures on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals μ and ν respectively and (X, Y) are all possible couples of random variables with μ and ν as respective laws. For rotational simplicity, the notation for a probability measure and its probability density is often abused. So if μ, ν have densities ρ_1, ρ_2 respectively, we also denote the distance as $W_p^p(\rho_1, \rho_2)$. For further details, we refer the reader to the book of Villani [60].

Following the same argument as [39, Corollary 4.3], Theorem 1.2 implies molecular chaos in the following sense:

Corollary 1.1 *For any $T > 0$, let $F_0^N := \otimes^N f_0$ and F_t^N be the N-particle distribution evolving with the microscopic flow* (12) *starting from F_0^N. Then the k-particle marginal*

$$^{(k)}F_t^N(z_1, \cdots, z_k) := \int F_t^N(Z) dz_{k+1} \cdots dz_N$$

converges weakly to $\otimes^k f_t$ as $N \to \infty$ for all $k \in N$, where f_t is the unique solution of the VPFP equations (3) *with $f_t|_{t=0} = f_0$. More precisely, under the assumptions of Theorem 1.2, for any $\alpha > 0$, there exists some constants $C > 0$ and $N_0 > 0$ depending only on α, T and C_{f_0}, such that for $N \geq N_0$, the following estimate holds*

$$W_1 \left(^{(k)}F_t^N, \otimes^k f_t \right) \leq k \exp \left(TC\sqrt{\log(N)} \right) N^{-\lambda_2}, \quad \forall \, 0 \leq t \leq T,$$

where λ_2 is used in Theorem 1.2.

Another result from Theorem 1.2 is the derivation of the macroscopic mean-field VPFP equations (3) from the microscopic random particle system (12). We define the empirical measure associated to the microscopic N-particle systems (12) and (24) respectively as

$$\mu_\Phi(t) := \frac{1}{N} \sum_{i=1}^{N} \delta(x - x_i^t)\delta(v - v_i^t), \quad \mu_\Psi(t) := \frac{1}{N} \sum_{i=1}^{N} \delta(x - \overline{x}_i^t)\delta(v - \overline{v}_i^t). \quad (37)$$

The following theorem shows that under additional moment control assumptions on f_0, the empirical measure $\mu_\Phi(t)$ converges to the solution of VPFP equations (3) in W_p distance with high probability.

Theorem 1.3 *Under the same assumption as in Theorem 1.2, let f_t be the unique solution to the VPFP equations* (3) *with the initial data satisfying Assumption 1.1 and $\mu_\Phi(t)$ be the empirical measure defined in* (37) *with Φ_t being the particle flow solving* (12). *Let $p \in [1, \infty)$ and assume that there exists $m > 2p$ such that $\iint_{\mathbb{R}^6} |x|^m f_0(x, v) dx dv < +\infty$. Then for any $T > 0$ and $\kappa < \min\{\frac{1}{6}, \frac{1}{2p}, \delta\}$, there exists a constant C_1 depending only on T and C_{f_0} and constants C_2, C_3 depending only on m, p, κ, such that for all $N \geq e^{\left(\frac{C_1}{1-3\lambda_2}\right)^2}$ it holds that*

$$\mathbb{P}\left(\max_{t \in [0,T]} W_p(f_t, \mu_\Phi(t)) \leq N^{-\kappa+1-3\lambda_2} + N^{-\lambda_2} \right)$$

$$\geq 1 - C_2 \left(e^{-C_3 N^{1-\max\{6,2p\}\kappa}} + N^{1-\frac{m}{2p}} \right). \quad (38)$$

where δ and λ_2 are used in Theorem 1.2.

This theorem provides a derivation of the VPFP equations from an interacting N-particle system, bridging the gap between the microscopic descriptions in terms of agent based models and macroscopic or hydrodynamic descriptions for the particle probability density.

2 Preliminaries

In this section we collect the technical lemmas that are used in the proofs of the main theorems. Throughout this manuscript, generic constants will be denoted generically by C (independent of N), even if they are different from line to line. We use $\|\cdot\|_p$ for the L^p ($1 \leq p \leq \infty$) norm of a function. Moreover if $v = (v_1, \ldots, v_N)$ is a vector, then $\|v\|_\infty := \max_{i=1,\ldots,N} |v_i|$.

2.1 Local Lipschitz Bound

First let us recall some estimates of the regularized kernel k^N defined in (11):

Lemma 2.1 (*Regularity of k^N*)

(i) $k^N(0) = 0$, $k^N(x) = k(x)$, for any $|x| \geq N^{-\delta}$ and $|k^N(x)| \leq |k(x)|$, for any $x \in \mathbb{R}^3$;
(ii) $|\partial^\beta k^N(x)| \leq C N^{(2+|\beta|)\delta}$, for any $x \in \mathbb{R}^3$;
(iii) $\|k^N\|_2 \leq C N^{\frac{\delta}{2}}$.

The estimate (*i*) has been proved in [63, Lemma 2.1] and (*ii*) follows from [2, Lemma 5.1]. As for (*iii*), it is a direct result of Young's inequality.

Next we define a cut-off function ℓ^N, which will provide the local Lipschitz bound for k^N.

Definition 2.1 Let

$$\ell^N(x) = \begin{cases} \dfrac{6^3}{|x|^3}, & \text{if } |x| \geq 6N^{-\delta}, \\ N^{3\delta}, & \text{else}, \end{cases} \tag{39}$$

and $L^N : \mathbb{R}^{3N} \to \mathbb{R}^N$ be defined by $(L^N(X_t))_i := \frac{1}{N-1} \sum_{i \neq j} \ell^N(x_i^t - x_j^t)$. Furthermore, we define $\overline{L}^N(\overline{X}_t)$ by $(\overline{L}^N(\overline{X}_t))_i := \int_{\mathbb{R}^3} \ell^N(\overline{x}_i^t - x)\rho^N(x,t)dx$.

We summarize our first observation of k^N and ℓ^N in the following lemma:

Lemma 2.2 *There is a constant $C > 0$ independent of N such that for all $x, y \in \mathbb{R}^3$ with $|x - y| \leq N^{-\lambda_2} \gg N^{-\delta}$ ($\lambda_2 < \delta$) the following holds:*

$$\frac{|\nabla k^N(x)|}{\ell^N(y)} \leq C N^{3(\delta-\lambda_2)},$$

where k^N is the regularization of the Coulomb kernel (2) and ℓ^N satisfies Definition 2.1.

Proof Let us first consider the case $|y| < 2N^{-\lambda_2}$. It follows from the bound from Lemma 2.1 and the decrease of ℓ^N that

$$\frac{|\nabla k^N(x)|}{\ell^N(y)} \leq \frac{N^{3\delta}}{\ell^N(2N^{-\lambda_2})} = C N^{3(\delta-\lambda_2)}, \tag{40}$$

where we used $2N^{-\lambda_2} > 6N^{-\delta}$, thus $\ell^N(2N^{-\lambda_2}) = 27N^{3\lambda_2}$.

Next we consider the case $|y| \geq 2N^{-\lambda_2}$. It follows that $|x| \geq N^{-\lambda_2}$ and thus by Lemma 2.1 (i)

$$\frac{|\nabla k^N(x)|}{\ell^N(y)} \leq \frac{C|x|^{-3}}{|y|^{-3}} \leq C \frac{(|y| - N^{-\lambda_2})^{-3}}{|y|^{-3}} \leq C, \tag{41}$$

where in the last step we used $|x| \geq (|y| - N^{-\lambda_2}) \geq \frac{|y|}{2}$ for $|y| \geq 2N^{-\lambda_2}$. Collecting (40) and (41) finishes the proof. $\qquad\square$

Recall the notations

$$(K^N(X_t))_i := \frac{1}{N-1} \sum_{j \neq i} k^N(x_i^t - x_j^t), \quad (K^N(\overline{X}_t))_i := \frac{1}{N-1} \sum_{j \neq i} k^N(\overline{x}_i^t - \overline{x}_j^t), \tag{42}$$

and we have the local Lipschitz continuity of K^N:

Lemma 2.3 *If $\|X_t - \overline{X}_t\|_\infty \leq 2N^{-\delta}$, then it holds that*

$$\left\| K^N(X_t) - K^N(\overline{X}_t) \right\|_\infty \leq C \|L^N(\overline{X}_t)\|_\infty \|X_t - \overline{X}_t\|_\infty, \tag{43}$$

for some $C > 0$ independent of N.

Proof For any $\xi \in \mathbb{R}^3$ with $|\xi| < 4N^{-\delta}$, we claim that

$$|k^N(x + \xi) - k^N(x)| \leq C\ell^N(x)|\xi|, \tag{44}$$

where $\ell^N(x)$ is defined in (39). Indeed, for $|x| < 6N^{-\delta}$, estimate (44) holds due to the fact that $\|\nabla k^N\|_\infty \leq N^{3\delta}$. For $|x| \geq 6N^{-\delta}$, there exists $s \in [0, 1]$ such that

$$|k^N(x + \xi) - k^N(x)| \leq |\nabla k^N(x + s\xi)||\xi|,$$

where

$$|\nabla k^N(x + s\xi)| \leq C|x + s\xi|^{-3}.$$

The right hand side of the above expression takes its largest value when $s = 1$ and

$$|x + s\xi|^{-3} \leq |x(1 - \frac{|\xi|}{|x|})|^{-3}.$$

Since $|\xi| < 4N^{-\delta}$ and $|x| \geq 6N^{-\delta}$, it follows that $\frac{|\xi|}{|x|} < \frac{2}{3}$. Therefore, we get

$$|k^N(x + \xi) - k^N(x)| \leq C\left(\frac{3}{|x|}\right)^2 |\xi| \leq C\frac{|\xi|}{|x|^3}.$$

Applying claim (44) one has

$$|(K^N(X_t))_i - (K^N(\overline{X}_t))_i| \leq \frac{1}{N-1} \sum_{j \neq i}^N |k^N(x_i^t - x_j^t) - k^N(\overline{x}_i^t - \overline{x}_j^t)|$$

$$\leq \frac{1}{N-1} \sum_{j \neq i}^N C\ell^N(\overline{x}_i^t - \overline{x}_j^t)|x_i^t - x_j^t - \overline{x}_i^t + \overline{x}_j^t|$$

$$\leq C(L^N(\overline{X}_t))_i \|X_t - \overline{X}_t\|_\infty \leq C\|L^N(\overline{X}_t)\|_\infty \|X_t - \overline{X}_t\|_\infty, \tag{45}$$

which leads to (43). □

The following observations of k^N and ℓ^N turn out to be very helpful in the sequel:

Lemma 2.4 *Let* $\ell^N(x)$ *be defined in Definition* 2.1 *and* $\rho \in W^{1,1} \cap W^{1,\infty}(\mathbb{R}^3)$. *Then there exists a constant* $C > 0$ *independent of* N *such that*

$$\|\ell^N * \rho\|_\infty \leq C \log(N)(\|\rho\|_1 + \|\rho\|_\infty), \quad \|(\ell^N)^2 * \rho\|_\infty \leq CN^{(3\delta)}(\|\rho\|_1 + \|\rho\|_\infty); \tag{46}$$

and

$$\|k^N * \rho\|_\infty \leq C(\|\rho\|_1 + \|\rho\|_\infty), \quad \|\nabla k^N * \rho\|_\infty \leq C(\|\nabla\rho\|_1 + \|\nabla\rho\|_\infty). \tag{47}$$

Proof We only prove one of the estimates above, since all the estimates can be obtained through the same procedure. One can estimate

$$\|\ell^N * \rho\|_\infty = \left\| \int_{\mathbb{R}^3} \ell^N(x-y)\rho(y)dy \right\|_\infty$$

$$\leq \left\| \int_{|x-y|<6N^{-\delta}} \ell^N(x-y)\rho(y)dy \right\|_\infty$$

$$+ \left\| \int_{6N^{-\delta}\leq|x-y|\leq 1} \ell^N(x-y)\rho(y)dy \right\|_\infty$$

$$+ \left\| \int_{1\leq|x-y|} \ell^N(x-y)\rho(y)dy \right\|_\infty. \tag{48}$$

We estimate the first term

$$\left\| \int_{|x-y|<6N^{-\delta}} \ell^N(x-y)\rho(y)dy \right\|_\infty \leq \|\rho\|_\infty \|\ell^N\|_\infty |B(6N^{-\delta})|$$

$$\leq \frac{4\pi}{3}(6N^{-\delta})^3 N^{3\delta}\|\rho\|_\infty \leq C\|\rho\|_\infty, \tag{49}$$

where $B(r)$ denotes the ball with radius r in $\mathbb{R}^3$. The second term is bounded by

$$\left\| \int_{6N^{-\delta}\leq|x-y|\leq 1} \ell^N(x-y)\rho(y)dy \right\|_\infty \leq \|\rho\|_\infty \int_{6N^{-\delta}\leq|y|\leq 1} \frac{C}{|y|^3}dy \leq C\log(N)\|\rho\|_\infty. \tag{50}$$

It is easy to compute the last term

$$\left\| \int_{1\leq|x-y|} \ell^N(x-y)\rho(y)dy \right\|_\infty \leq C\|\rho\|_1. \tag{51}$$

Collecting estimates (49), (50) and (51), one has

$$\|\ell^N * \rho\|_\infty \leq C\|\rho\|_\infty + C\log(N)\|\rho\|_\infty + C\|\rho\|_1 \leq C\log(N)(\|\rho\|_\infty + \|\rho\|_1). \tag{52}$$

□

2.2 Law of Large Numbers

Also, we need the following concentration inequality to provide us the probability bounds of random variables:

Lemma 2.5 *Let $Z_1, \ldots, Z_N$ be i.i.d. random variables with $\mathbb{E}[Z_i] = 0$, $\mathbb{E}[Z_i^2] \leq g(N)$ and $|Z_i| \leq C\sqrt{Ng(N)}$. Then for any $\alpha > 0$, the sample mean $\bar{Z} = \frac{1}{N}\sum_{i=1}^{N} Z_i$ satisfies*

$$\mathbb{P}\left(|\bar{Z}| \geq \frac{C_\alpha \sqrt{g(N)}\log(N)}{\sqrt{N}} \right) \leq N^{-\alpha}, \tag{53}$$

where C_α depends only on C and α.

The proof can be seen in [22, Lemma 1], which is a direct result of Taylor's expansion and Markov's inequality.

Recall the notation

$$(\overline{K}^N(\overline{X}_t))_i := \int_{\mathbb{R}^3} k^N(\overline{x}_i^t - x)\rho^N(x, t)dx. \tag{54}$$

We can introduce the following version of the Law of Large Numbers:

Lemma 2.6 *At any fixed time $t \in [0, T]$, suppose that $\overline{X}_t$ satisfies the mean-field dynamics (24), K^N and $\overline{K}^N$ are defined in (42) and (54) respectively, L^N and $\overline{L}^N$ are introduced in Definition 2.1. For any $\alpha > 0$ and $\frac{1}{3} \leq \delta < 1$, there exist a constant $C_{1,\alpha} > 0$ depending only on α, T and C_{f_0} such that*

$$\mathbb{P}\left(\left\| K^N(\overline{X}_t) - \overline{K}^N(\overline{X}_t) \right\|_\infty \geq C_{1,\alpha} N^{2\delta-1}\log(N) \right) \leq N^{-\alpha}, \tag{55}$$

and

$$\mathbb{P}\left(\left\| L^N(\overline{X}_t) - \overline{L}^N(\overline{X}_t) \right\|_\infty \geq C_{1,\alpha} N^{3\delta-1}\log(N) \right) \leq N^{-\alpha}. \tag{56}$$

Proof We can prove this lemma by using Lemma 2.5.

Due to the exchangeability of the particles, we are ready to bound

$$(K^N(\overline{X}_t))_1 - (\overline{K}^N(\overline{X}_t))_1 = \frac{1}{N-1}\sum_{j=2}^{N} k^N(\overline{x}_1^t - \overline{x}_j^t) - \int_{\mathbb{R}^3} k^N(\overline{x}_1^t - x)\rho^N(x, t)dx$$

$$= \frac{1}{N-1}\sum_{j=2}^{N} Z_j, \tag{57}$$

where

$$Z_j := k^N(\overline{x}_1^t - \overline{x}_j^t) - \int_{\mathbb{R}^3} k^N(\overline{x}_1^t - x)\rho^N(x, t)dx.$$

Since $\overline{x}_1^t$ and $\overline{x}_j^t$ are independent when $j \neq 1$ and $k^N(0) = 0$, let us consider $\overline{x}_1^t$ as given and denote $\mathbb{E}'[\cdot] = \mathbb{E}[\cdot | \overline{x}_1^t]$. It is easy to show that $\mathbb{E}'[Z_j] = 0$ since

$$\mathbb{E}'\left[k^N(\overline{x}_1^t - \overline{x}_j^t) \right] = \iint_{\mathbb{R}^6} k^N(\overline{x}_1^t - x) f^N(x, v, t)dxdv$$

$$= \int_{\mathbb{R}^3} k^N(\overline{x}_1^t - x)\rho^N(x, t)dx. \tag{58}$$

To use Lemma 2.5, we need a bound for the variance

$$\mathbb{E}'[|Z_j|^2] = \mathbb{E}'\left[\left|k^N(\overline{x}_1^t - \overline{x}_j^t) - \int_{\mathbb{R}^3} k^N(\overline{x}_1^t - x)\rho^N(x,t)dx\right|^2\right]. \tag{59}$$

Since it follows from Lemma 2.4 that

$$\int_{\mathbb{R}^3} k^N(\overline{x}_1^t - x)\rho^N(x,t)dx \leq C(\|\rho^N\|_1 + \|\rho^N\|_\infty), \tag{60}$$

it suffices to bound

$$\mathbb{E}'\left[k^N(\overline{x}_1^t - \overline{x}_j^t)\right] = \int_{\mathbb{R}^3} k^N(\overline{x}_1^t - x)\rho^N(x,t)dx \leq C(\|\rho^N\|_1 + \|\rho^N\|_\infty) \leq C(T, C_{f_0}), \tag{61}$$

and

$$\mathbb{E}'\left[k^N(\overline{x}_1^t - \overline{x}_j^t)^2\right] = \int_{\mathbb{R}^3} k^N(\overline{x}_1^t - x)^2\rho^N(x,t)dx \leq \|\rho^N\|_\infty \|k^N\|_2^2 \leq C(T, C_{f_0})N^\delta, \tag{62}$$

where we have used $\|k^N\|_2 \leq CN^{\frac{\delta}{2}}$ in Lemma 2.1 (iii). Hence one has

$$\mathbb{E}'[|Z_j|^2] \leq CN^\delta. \tag{63}$$

So the hypotheses of Lemma 2.5 are satisfied with $g(N) = CN^{4\delta-1}$. In addition, it follows from (ii) in Lemma 2.1 that $|Z_j| \leq CN^{2\delta} \leq C\sqrt{Ng(N)}$. Hence, using Lemma 2.5, we have the probability bound

$$\mathbb{P}\left(\left|(K^N(\overline{X}_t))_1 - (\overline{K}^N(\overline{X}_t))_1\right| \geq C(\alpha, T, C_{f_0})N^{2\delta-1}\log(N)\right) \leq N^{-\alpha}. \tag{64}$$

Similarly, the same bound also holds for all other indexes $i = 2, \ldots, N$, which leads to

$$\mathbb{P}\left(\left\|K^N(\overline{X}_t) - \overline{K}^N(\overline{X}_t)\right\|_\infty \geq C(\alpha, T, C_{f_0})N^{2\delta-1}\log(N)\right) \leq N^{1-\alpha}. \tag{65}$$

Let $C_{1,\alpha}$ be the constant $C(\alpha, T, C_{f_0})$ in (65), then we conclude (55).

To prove (56), we follow the same procedure as above

$$(L^N(\overline{X}_t))_1 - (\overline{L}^N(\overline{X}_t))_1 = \frac{1}{N-1}\sum_{j=2}^N \ell^N(\overline{x}_1^t - \overline{x}_j^t) - \int_{\mathbb{R}^3} \ell^N(\overline{x}_1^t - x)\rho^N(x,t)dx$$

$$= \frac{1}{N-1}\sum_{j=2}^N Z_j, \tag{66}$$

where

$$Z_j = \ell^N(\overline{x}_1^t - \overline{x}_j^t) - \int_{\mathbb{R}^3} \ell^N(\overline{x}_1^t - x)\rho^N(x,t)dx.$$

It is easy to show that $\mathbb{E}'[Z_j] = 0$. To use Lemma 2.5, we need a bound for the variance. One computes that

$$\mathbb{E}'\left[\ell^N(\overline{x}_1^t - \overline{x}_j^t)\right] = \int_{\mathbb{R}^3} \ell^N(\overline{x}_1^t - x)\rho^N(x,t)dx \leq C\log(N)(\|\rho\|_1 + \|\rho\|_\infty)$$

$$\leq C(T, C_{f_0})\log(N), \tag{67}$$

and

$$\mathbb{E}'[\ell^N(\overline{x}_1^t - \overline{x}_j^t)^2] = \int_{\mathbb{R}^3} \ell^N(\overline{x}_1^t - x)^2 \rho^N(x, t)dx$$
$$\leq CN^{3\delta}(\|\rho\|_1 + \|\rho\|_\infty) \leq C(T, C_{f_0})N^{3\delta}, \tag{68}$$

where we have used the estimates of ℓ^N in Lemma 2.4. Hence one has

$$\mathbb{E}'[|Z_j|^2] \leq CN^{3\delta}. \tag{69}$$

So the hypotheses of Lemma 2.5 are satisfied with $g(N) = CN^{6\delta-1}$. In addition, it follows from Definition 2.1 that $|Z_j| \leq CN^{3\delta} \leq C\sqrt{Ng(N)}$. Hence, we have the probability bound

$$\mathbb{P}\left(\left|(L^N(\overline{X}_t))_1 - (\overline{L}^N(\overline{X}_t))_1\right| \geq C(\alpha, T, C_{f_0})N^{3\delta-1}\log(N)\right) \leq N^{-\alpha}, \tag{70}$$

by Lemma 2.5, which leads to

$$\mathbb{P}\left(\left\|L^N(\overline{X}_t) - \overline{L}^N(\overline{X}_t)\right\|_\infty \geq C(\alpha, T, C_{f_0})N^{3\delta-1}\log(N)\right) \leq N^{1-\alpha}. \tag{71}$$

Thus, (56) follows from (71). □

3 Proof of Theorem 1.2

We do the proof by following the idea in [28,29], which is that consistency and stability imply convergence. This at least in principle corresponds to the Lax's equivalence theorem of proving the convergence of a numerical algorithm, which is that stability and consistency of an algorithm imply its convergence.

3.1 Consistency

In order to obtain the consistency error for the entire time interval, we divide $[0, T]$ into $M+1$ subintervals with length $\Delta\tau = N^{-\frac{\gamma}{3}}$ for some $\gamma > 4$ and $\tau_k = n\Delta\tau$, $k = 0, \ldots, M+1$. The choice of γ will be clear from the discussion below. Here the choice of $\Delta\tau$ is only for the purpose of proving consistency and it can be sufficiently small. Note that it is different from Δt in the proof of stability in the next subsection.

First, we establish the following lemma on the traveling distance of $\overline{X}_t$ in a short time interval $[\tau_k, \tau_{k+1}]$:

Lemma 3.1 *Assume that $(\overline{X}_t, \overline{V}_t)$ satisfies the mean-field dynamics (24). For $\gamma > 4$ it holds*

$$\mathbb{P}\left(\max_k \max_{t \in [\tau_k, \tau_{k+1}]} \left\|\overline{X}_t - \overline{X}_{\tau_k}\right\|_\infty \geq C_B N^{-\frac{\gamma-1}{3}}\right) \leq C_B N^{\frac{\gamma-1}{3}} \exp(-C_B N^{\frac{2}{3}}), \tag{72}$$

where C_B depends only on T and C_{f_0}.

Proof Notice that for $t \in [\tau_k, \tau_{k+1}]$

$$\overline{X}_t - \overline{X}_{\tau_k} = \int_{\tau_k}^t \overline{V}_s ds = \int_{\tau_k}^t \int_{\tau_k}^s \overline{K}^N(\overline{X}_\tau)d\tau ds + \sqrt{2\sigma}\int_{\tau_k}^t (B(s) - B(\tau_k))ds + \int_{\tau_k}^t \overline{V}_{\tau_k} ds$$
$$=: I_1^k(t) + I_2^k(t) + I_3^k(t), \tag{73}$$

where

$$\overline{V}_{\tau_k} = V_0 + \int_0^{\tau_k} \overline{K}^N(\overline{X}_s)ds + \sqrt{2\sigma}\,B(\tau_k). \tag{74}$$

The estimate of $I_1^k(t)$ follows from Lemma 2.4

$$\int_{\tau_k}^t \int_{\tau_k}^s \overline{K}^N(\overline{X}_\tau)d\tau ds \le (\Delta t)^2 \|\overline{K}^N\|_\infty \le CN^{-\frac{2\gamma}{3}}. \tag{75}$$

So we have

$$\max_k \max_{t\in[\tau_k,\tau_{k+1}]} \|I_1^k(t)\|_\infty \le CN^{-\frac{2\gamma}{3}}. \tag{76}$$

To estimate $I_2^k(t)$, recall a basic property of Brownian motion [20, Chap. 1.2]:

$$\mathbb{P}\left(\max_{t\le s\le t+\Delta t} \|B(s) - B(t)\|_\infty \ge b\right) \le C_1(\sqrt{\Delta t}/b)\exp(-C_2 b^2/\Delta t), \tag{77}$$

which leads to

$$\mathbb{P}\left(\max_{t\in[\tau_k,\tau_{k+1}]} \|B(t) - B(\tau_k)\|_\infty \ge N^{-\frac{1}{3}}\right) \le C_1 N^{-\frac{\gamma-2}{6}}\exp(-C_2 N^{\frac{\gamma-2}{3}}), \tag{78}$$

where we choose $b = N^{-\frac{1}{3}}$.

Since $\max_{t\in[\tau_k,\tau_{k+1}]} \|I_2^k(t)\|_\infty \le \Delta t\sqrt{2\sigma}\max_{t\in[\tau_k,\tau_{k+1}]} \|B(t) - B(\tau_k)\|_\infty$, it follows from (78) that

$$\mathbb{P}\left(\max_{t\in[\tau_k,\tau_{k+1}]} \|I_2^k(t)\|_\infty \ge CN^{-\frac{\gamma+1}{3}}\right) \le C_1 N^{-\frac{\gamma-2}{6}}\exp(-C_2 N^{\frac{\gamma-2}{3}}), \tag{79}$$

which leads to

$$\mathbb{P}\left(\max_k \max_{t\in[\tau_k,\tau_{k+1}]} \|I_2^k(t)\|_\infty \ge CN^{-\frac{\gamma+1}{3}}\right) \le C_1 N^{\frac{\gamma+2}{6}}\exp(-C_2 N^{\frac{\gamma-2}{3}}), \tag{80}$$

where we used the fact that $n \le \frac{T}{\Delta t} = TN^{\frac{\gamma}{3}}$.

Lastly, we prove the estimate of $I_3^k(t)$. It is obvious that

$$\int_0^{\tau_k} \overline{K}^N(\overline{X}_s)ds \le n\Delta t\|\overline{K}^N\|_\infty \le CT, \tag{81}$$

and it follows from (77) that

$$\mathbb{P}(\|B(\tau_k)\|_\infty \ge N^{\frac{1}{3}}) \le C_1 N^{-\frac{1}{3}}\sqrt{T}\exp(-C_2 N^{\frac{2}{3}}/T). \tag{82}$$

Moreover, it follows from the assumption in Theorem 1.1 *b*) the distribution $f_0^v(v)$ of V_0 has a compact support:

$$f_0^v(v) = \int_{\mathbb{R}^3} f_0(x,v)dx = 0, \text{ when } |v| > Q_v. \tag{83}$$

Then one has

$$\mathbb{P}(\|V_0\|_\infty \ge N^{\frac{1}{3}}) = \int_{|v|\ge N^{\frac{1}{3}}} f_0^v(v)dv = 0, \text{ when } N > Q_v^3. \tag{84}$$

It follows from (73) that

$$
\max_{t\in[\tau_k,\tau_{k+1}]}\|I_3^k(t)\|_\infty = \int_{\tau_k}^t \|\overline{V}_{\tau_k}\|_\infty ds \le N^{-\frac{\gamma}{3}}\left(\|V_0\|_\infty + \sqrt{2\sigma}\|B(\tau_k)\|_\infty + \int_0^{\tau_k}\overline{K}^N(\overline{X}_s)ds\right)
$$
$$
\le N^{-\frac{\gamma}{3}}(\|V_0\|_\infty + \sqrt{2\sigma}\|B(\tau_k)\|_\infty) + CN^{-\frac{\gamma}{3}}, \tag{85}
$$

then it yields

$$
\mathbb{P}\left(\max_{t\in[\tau_k,\tau_{k+1}]}\left\|I_3^k(t)\right\|_\infty \ge 3N^{-\frac{\gamma-1}{3}}\right)
$$
$$
\le \mathbb{P}\left(N^{-\frac{\gamma}{3}}\|V_0\|_\infty \ge N^{-\frac{\gamma-1}{3}}\right) + \mathbb{P}\left(\sqrt{2\sigma}N^{-\frac{\gamma}{3}}\|B(\tau_k)\|_\infty \ge N^{-\frac{\gamma-1}{3}}\right)
$$
$$
+ \mathbb{P}\left(CN^{-\frac{\gamma}{3}} \ge N^{-\frac{\gamma-1}{3}}\right)
$$
$$
\le 0 + CN^{-\frac{1}{3}}\exp(-CN^{\frac{2}{3}}) + 0 \le CN^{-\frac{1}{3}}\exp(-CN^{\frac{2}{3}}), \tag{86}
$$

which leads to

$$
\mathbb{P}\left(\max_k \max_{t\in[\tau_k,\tau_{k+1}]}\left\|I_3^k(t)\right\|_\infty \ge 3N^{-\frac{\gamma-1}{3}}\right) \le CN^{\frac{\gamma-1}{3}}\exp(-CN^{\frac{2}{3}}). \tag{87}
$$

Then it follows from (76), (80) and (87) that

$$
\mathbb{P}\left(\max_k \max_{t\in[\tau_k,\tau_{k+1}]}\|\overline{X}_t - \overline{X}_{\tau_k}\|_\infty \ge CN^{-\frac{\gamma-1}{3}}\right)
$$
$$
\le C_1 N^{\frac{\gamma+2}{6}}\exp(-C_2 N^{\frac{\gamma-2}{3}}) + CN^{\frac{\gamma-1}{3}}\exp(-CN^{\frac{2}{3}}) \le CN^{\frac{\gamma-1}{3}}\exp(-CN^{\frac{2}{3}}),
$$

for $\gamma > 4$, which completes the proof of (72). $\qquad\square$

Now we can prove the consistency error for the entire time interval $[0,T]$.

Proposition 3.1 *(Consistency) For any $T > 0$, let $(\overline{X}_t, \overline{V}_t)$ satisfy the mean-field dynamics (24) with initial density $f_0(x,v)$, K^N and $\overline{K}^N$ be defined in (42) and (54) respectively. For any $\alpha > 0$ and $\frac{1}{3} \le \delta < 1$, there exist a constant $C_{2,\alpha} > 0$ depending only on α, T and C_{f_0} such that*

$$
\mathbb{P}\left(\max_{t\in[0,T]}\left\|K^N(\overline{X}_t) - \overline{K}^N(\overline{X}_t)\right\|_\infty \ge C_{2,\alpha}N^{2\delta-1}\log(N)\right) \le N^{-\alpha}, \tag{88}
$$

and

$$
\mathbb{P}\left(\max_{t\in[0,T]}\left\|L^N(\overline{X}_t) - \overline{L}^N(\overline{X}_t)\right\|_\infty \ge C_{2,\alpha}N^{3\delta-1}\log(N)\right) \le N^{-\alpha}. \tag{89}
$$

Proof Denote the events:

$$
\overline{\mathcal{H}} := \left\{\max_k \max_{t\in[\tau_k,\tau_{k+1}]}\|\overline{X}_t - \overline{X}_{\tau_k}\|_\infty \le C_B N^{-\frac{\gamma-1}{3}}\right\}, \tag{90}
$$

and

$$
\mathcal{C}_{\tau_k} := \left\{\left\|K^N(\overline{X}_{\tau_k}) - \overline{K}^N(\overline{X}_{\tau_k})\right\|_\infty \le C_{1,\alpha}N^{2\delta-1}\log(N)\right\}, \tag{91}
$$

where C_B and $C_{1,\alpha}$ are used in Lemma 2.6 and Lemma 3.1 respectively. According to Lemma 2.6 and Lemma 3.1, one has

$$\mathbb{P}(\mathcal{C}_{\tau_k}^c) \le N^{-\alpha}, \quad \mathbb{P}(\overline{\mathcal{H}}^c) \le C_B N^{\frac{\gamma-1}{3}} \exp(-C_B N^{\frac{2}{3}}), \tag{92}$$

for any $\alpha > 0$ and $\gamma > 4$.

Furthermore, we denote

$$\mathcal{B}_{\tau_k} := \left\{ \left\| L^N(\overline{X}_{\tau_k}) - \overline{L}^N(\overline{X}_{\tau_k}) \right\|_\infty \le C_{1,\alpha} N^{3\delta-1} \log(N) \right\}, \tag{93}$$

then one has

$$\mathbb{P}(\mathcal{B}_{\tau_k}^c) \le N^{-\alpha}, \tag{94}$$

by Lemma 2.6. Also, under the event $\mathcal{B}_{\tau_k}$, it holds that

$$\|L^N(\overline{X}_{\tau_k})\|_\infty \le \|\overline{L}^N(\overline{X}_{\tau_k})\|_\infty + C_{1,\alpha} N^{3\delta-1} \log(N) \le C(\alpha, T, C_{f_0}) N^{3\delta-1} \log(N), \tag{95}$$

where we have used $\|\overline{L}^N(\overline{X}_{\tau_k})\|_\infty \le C \log(N)$ from Lemma 2.4.

For all $t \in [\tau_k, \tau_{k+1}]$, under the event $\mathcal{B}_{\tau_k} \cap \mathcal{C}_{\tau_k} \cap \overline{\mathcal{H}}$, we obtain

$$\begin{aligned}
&\left\| K^N(\overline{X}_t) - \overline{K}^N(\overline{X}_t) \right\|_\infty \\
&\le \left\| K^N(\overline{X}_t) - K^N(\overline{X}_{\tau_k}) \right\|_\infty + \left\| K^N(\overline{X}_{\tau_k}) - \overline{K}^N(\overline{X}_{\tau_k}) \right\|_\infty + \left\| \overline{K}^N(\overline{X}_{\tau_k}) - \overline{K}^N(\overline{X}_t) \right\|_\infty \\
&\le C\|L^N(\overline{X}_{\tau_k})\|_\infty \|\overline{X}_t - \overline{X}_{\tau_k}\|_\infty + C_{1,\alpha} N^{2\delta-1} \log(N) + C \left\| \overline{X}_t - \overline{X}_{\tau_k} \right\|_\infty + C N^{-\frac{\gamma}{3}} \\
&\le C(\alpha, T, C_{f_0}) N^{3\delta-1} \log(N) N^{-\frac{\gamma-1}{3}} + C_{1,\alpha} N^{2\delta-1} \log(N) \\
&\le C(\alpha, T, C_{f_0}) N^{2\delta-1} \log(N),
\end{aligned} \tag{96}$$

due to the fact that $3\delta + 1 < 4 < \gamma$. In the second inequality we have used the local Lipschitz bound of K^N

$$\left\| K^N(X_t) - K^N(\overline{X}_{\tau_k}) \right\|_\infty \le C\|L^N(\overline{X}_{\tau_k})\|_\infty \|X_t - \overline{X}_{\tau_k}\|_\infty, \tag{97}$$

under the event $\overline{\mathcal{H}}$ (see in Lemma 2.3). To bound the third term $\left\| \overline{K}^N(\overline{X}_{\tau_k}) - \overline{K}^N(\overline{X}_t) \right\|_\infty$, we used the uniform control of $\max_{\tau_k \le t \le \tau_{k+1}} \|\partial_t \rho^N\|_{L^\infty(\mathbb{R}^3)}$ in (22). Indeed,

$$\begin{aligned}
&\|k^N * \rho_t(X_t) - k^N * \rho_{\tau_k}(X_{\tau_k})\|_\infty \\
&\le \|k^N * \rho_t(X_t) - k^N * \rho_t(X_{\tau_k})\|_\infty + \|k^N * \rho_t(X_{\tau_k}) - k^N * \rho_{\tau_k}(X_{\tau_k})\|_\infty \\
&\le C \left\| \overline{X}_t - \overline{X}_{\tau_k} \right\|_\infty + C\Delta\tau \le C \left\| \overline{X}_t - \overline{X}_{\tau_k} \right\|_\infty + C N^{-\frac{\gamma}{3}}.
\end{aligned} \tag{98}$$

In the third inequality we have used (90) and (95). This yields that

$$\max_{t \in [0,T]} \left\| K^N(\overline{X}_t) - \overline{K}^N(\overline{X}_t) \right\|_\infty \le C(\alpha, T, C_{f_0}) N^{2\delta-1} \log(N), \tag{99}$$

holds under the event $\bigcap_{k=0}^M \mathcal{B}_{\tau_k} \cap \mathcal{C}_{\tau_k} \cap \overline{\mathcal{H}}$. Therefore it follows from (92) and (94) that

$$\mathbb{P}\left(\max_{t \in [0,T]} \left\| K^N(\overline{X}_t) - \overline{K}^N(\overline{X}_t) \right\|_\infty \ge C(\alpha, T, C_{f_0}) N^{2\delta-1} \log(N) \right)$$

$$\leq \sum_{k=0}^{M} P(\mathcal{B}_{\tau_k}^c) + \sum_{k=0}^{M} P(\mathcal{C}_{\tau_k}^c) + P(\overline{\mathcal{H}}^c)$$

$$\leq TN^{-\frac{3\alpha-\gamma}{3}} + TN^{-\frac{3\alpha-\gamma}{3}} + C_B N^{\frac{\gamma-1}{3}} \exp(-C_B N^{\frac{2}{3}}) \leq N^{-\alpha'}. \tag{100}$$

Denote $C_{2,\alpha'}$ to be the constant $C(\alpha, T, C_{f_0})$ in (100). Since $\alpha > 0$ is arbitrary and so is α', (88) holds true. The proof of (89) can be done similarly. $\qquad\square$

3.2 Stability

In this subsection we obtain the stability result.

Definition 3.1 Let $\mathcal{A}_T$ be the event given by

$$\mathcal{A}_T := \left\{ \max_{t \in [0,T]} \sqrt{\log(N)} \left\| X_t - \overline{X}_t \right\|_\infty + \left\| V_t - \overline{V}_t \right\|_\infty \leq N^{-\lambda_2} \right\}. \tag{101}$$

Proposition 3.2 (Stability) For any $T > 0$, assume that the trajectories $\Phi_t = (X_t, V_t)$, $\Psi_t = (\overline{X}_t, \overline{V}_t)$ satisfy (12) and (24) respectively with the initial data $\Phi_0 = \Psi_0$ which is i.i.d. sharing the common density f_0 satisfying Assumption 1.1. Let K^N be introduced in (42). For any $0 < \lambda_2 < \frac{1}{3}$, $0 < \lambda_1 < \frac{\lambda_2}{3}$ and $\frac{1}{3} \leq \delta < 1$, we denote the event:

$$\mathcal{S}_T(\Lambda) := \left\{ \|K^N(X_t) - K^N(\overline{X}_t)\|_\infty \leq \Lambda \log(N) \left\| X_t - \overline{X}_t \right\|_\infty \right.$$

$$\left. + \Lambda \log^2(N)(N^{6\delta-1-\lambda_1-4\lambda_2} + N^{3\lambda_1-2\lambda_2} + N^{2\delta-1}), \ \forall\, t \in [0,T] \right\}. \tag{102}$$

Then for any $\alpha > 0$, there exists some $C_{3,\alpha} > 0$ and a $N_0 \in \mathbb{N}$ depending only on α, T and C_{f_0} such that

$$\mathbb{P}\left(\mathcal{A}_T \cap \mathcal{S}_T^c(C_{3,\alpha}) \right) \leq N^{-\alpha}, \tag{103}$$

for all $N \geq N_0$.

Remark 3.1 This proposition is one of the crucial statements in our paper, which later leads to Lemma 3.4. Proving propagation of chaos for systems like the one we consider under the assumptions of Lipschitz-continuous forces is standard, as explained in the introduction. The forces we consider are more singular. However our techniques allow us to show that the Lipschitz condition encoded in the definition of $\mathcal{S}$ holds typically, i.e. with probability close to one. In this sense, Proposition 3.2 is only helpful if we find an argument that $\mathcal{A}_T$ holds. But as long as we have good estimates on the difference of the forces and thus the growth of $\max_{t \in [0,T]} \sqrt{\log(N)} \left\| X_t - \overline{X}_t \right\|_\infty + \left\| V_t - \overline{V}_t \right\|_\infty$, we are in fact able to control $\mathcal{A}_T$ (see argument in (154)).

Proof Let $\alpha > 0$. First, we write $\mathcal{S}_T(\Lambda)$ as the intersection of non-overlapping sets $\{\mathcal{S}_n(\Lambda)\}_{n=0}^{M'}$, where

$$\mathcal{S}_n(\Lambda) := \left\{ \|K^N(X_t) - K^N(\overline{X}_t)\|_\infty \leq \Lambda \log(N) \left\| X_t - \overline{X}_t \right\|_\infty \right.$$

$$\left. + \Lambda \log^2(N)(N^{6\delta-1-\lambda_1-4\lambda_2} + N^{3\lambda_1-2\lambda_2} + N^{2\delta-1}), \ \forall\, t \in [t_n, t_{n+1}], \ 0 \leq n \leq M' \right\}, \tag{104}$$

with $\Delta t := t_{n+1} - t_n = N^{-\lambda_1}$, then $\mathcal{S}_T(\Lambda) = \bigcap_{n=0}^{M'} \mathcal{S}_n(\Lambda)$. Note that here the choice of Δt is for the purpose of proving stability and it is different from $\Delta \tau$ in the proof of consistency.

To prove this proposition, we split the interaction force k^N into $k^N = k_1^N + k_2^N$, where k_2^N is the result of choosing a wider cut-off of order $N^{-\lambda_2} > N^{-\delta}$ in the force kernel k and

$$k_1^N := k^N - k_2^N, \quad k_2^N = k * \psi_{\lambda_2}^N, \tag{105}$$

which means that for k_2^N and ℓ_2^N we choose $\delta = \lambda_2$ in (11) and (39) respectively.

Following the approach in [10], we introduce the following auxiliary trajectory

$$\begin{cases} d\widetilde{x}_i^t = \widetilde{v}_i^t dt, \\ d\widetilde{v}_i^t = \displaystyle\int_{\mathbb{R}^3} k^N(\widetilde{x}_i^t - x)\rho^N(x, t) dx dt + \sqrt{2\sigma} dB_i^t \ . \end{cases} \tag{106}$$

We consider the above auxiliary trajectory with two different initial phases. For any $1 \leq n \leq M'$ and $t \in [t_n, t_{n+1}]$, we consider the auxiliary trajectory starting from the initial phase

$$(\widetilde{x}_i^{t_{n-1}}, \widetilde{v}_i^{t_{n-1}}) = (x_i^{t_{n-1}}, v_i^{t_{n-1}}), \tag{107}$$

where $(x_i^{t_{n-1}}, v_i^{t_{n-1}})$ satisfies (12) at time t_{n-1}. However when $n = 1$, i.e. $t \in [0, t_1]$, the initial phase of the auxiliary trajectory is chosen to be $(\widetilde{x}_i^0, \widetilde{v}_i^0) = (x_i^0, v_i^0)$, which has the distribution f_0. Moreover in the latter case the distribution of $(\widetilde{x}_i^t, \widetilde{v}_i^t)$ is exactly f_t^N, which solves the regularized VPFP equations (13) with the initial data f_0.

For later reference let us estimate the difference $\|\overline{X}_t - \widetilde{X}_t\|_\infty$ and $\|\overline{V}_t - \widetilde{V}_t\|_\infty$. Using the equations of these trajectories, we have for $t \in [t_n, t_{n+1}]$,

$$\frac{d}{dt}\|\overline{X}_t - \widetilde{X}_t\|_\infty = \|\overline{V}_t - \widetilde{V}_t\|_\infty, \tag{108}$$

and

$$\begin{aligned} \frac{d}{dt}\|\overline{V}_t - \widetilde{V}_t\|_\infty &= \|\overline{K}^N(\overline{X}_t) - \overline{K}^N(\widetilde{X}_t)\|_\infty \\ &\leq \max_{1 \leq j \leq N} |k^N * \rho^N(\cdot, t)(\overline{x}_j) - k^N * \rho^N(\cdot, t)(\widetilde{x}_j)| \\ &\leq \max_{1 \leq j \leq N} |\overline{x}_j - \widetilde{x}_j| \|\nabla k^N * \rho^N(\cdot, t)\|_\infty \\ &\leq C(\|\nabla \rho^N\|_1 + \|\nabla \rho^N\|_\infty)\|\overline{X}_t - \widetilde{X}_t\|_\infty \\ &\leq C\|\overline{X}_t - \widetilde{X}_t\|_\infty \ , \end{aligned} \tag{109}$$

where C depends only on T and C_{f_0}. Summarizing, we get

$$\frac{d}{dt}\left(\|\overline{X}_t - \widetilde{X}_t\|_\infty + \|\overline{V}_t - \widetilde{V}_t\|_\infty\right) \leq C\left(\|\overline{X}_t - \widetilde{X}_t\|_\infty + \|\overline{V}_t - \widetilde{V}_t\|_\infty\right)$$

Using Gronwall's inequality it follows that

$$\begin{aligned} \max_{t_n \leq t \leq t_{n+1}}\left(\|\overline{X}_t - \widetilde{X}_t\|_\infty + \|\overline{V}_t - \widetilde{V}_t\|_\infty\right) &\leq \exp(C\Delta t)(\|\overline{X}_{t_n} - X_{t_n}\|_\infty + \|\overline{V}_{t_n} - V_{t_n}\|_\infty) \\ &\leq \exp(CN^{-\lambda_1})N^{-\lambda_2} \leq CN^{-\lambda_2}, \end{aligned} \tag{110}$$

under the event $\mathcal{A}_T$ defined in (101).

Then for any $t \in [t_n, t_{n+1}]$, one splits the error

$$\|K^N(X_t) - K^N(\overline{X}_t)\|_\infty$$

$$\leq \|K_2^N(X_t) - K_2^N(\overline{X}_t)\|_\infty + \|K_1^N(X_t) - K_1^N(\widetilde{X}_t)\|_\infty + \|K_1^N(\widetilde{X}_t) - K_1^N(\overline{X}_t)\|_\infty$$
$$=: \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3. \tag{111}$$

First, let us compute $\mathcal{I}_1$:

$$\|K_2^N(X_t) - K_2^N(\overline{X}_t)\|_\infty \leq C\|L_2^N(\overline{X}_t)\|_\infty \|X_t - \overline{X}_t\|_\infty, \tag{112}$$

where we have used the local Lipschitz bound of K_2^N under the event $\mathcal{A}_T$ (see in Lemma 2.3). Furthermore, we denote

$$\mathcal{B}_2 := \left\{ \max_{t \in [0,T]} \left\| L_2^N(\overline{X}_t) - \overline{L}_2^N(\overline{X}_t) \right\|_\infty \leq C_{2,\alpha} N^{3\lambda_2 - 1} \log(N) \right\}. \tag{113}$$

Since Proposition 3.1 also holds for the case $\lambda_2 < \frac{1}{3}$, one has

$$\mathbb{P}(\mathcal{B}_2^c) \leq N^{-\alpha}. \tag{114}$$

Under the event $\mathcal{B}_2$, it holds that

$$\|L_2^N(\overline{X}_t)\|_\infty \leq \|\overline{L}_2^N(\overline{X}_t)\|_\infty + C_{2,\alpha} N^{3\lambda_2 - 1} \log(N) \leq C \log(N), \tag{115}$$

since $\lambda_2 < \frac{1}{3}$, where $\|\overline{L}_2^N(\overline{X}_t)\|_\infty \leq C \log(N)$ follows from Lemma 2.4. Hence, one has

$$\mathcal{I}_1 \leq \|K_2^N(X_t) - K_2^N(\overline{X}_t)\|_\infty \leq C \log(N) \|X_t - \overline{X}_t\|_\infty, \quad \forall\, t \in [t_n, t_{n+1}], \tag{116}$$

under event $\mathcal{A}_T \cap \mathcal{B}_2$.

To estimate $\mathcal{I}_2$, notice that by triangle inequality and (110) one has

$$\|X_t - \widetilde{X}_t\|_\infty \leq \int_{t_n}^t \|V_s - \widetilde{V}_s\|_\infty ds \leq \int_{t_n}^t \|V_s - \overline{V}_s\|_\infty + \|\overline{V}_s - \widetilde{V}_s\|_\infty ds \tag{117}$$

$$\leq \Delta t \max_{s \in [t_n, t]} \left(\|V_s - \overline{V}_s\|_\infty + \|\overline{V}_s - \widetilde{V}_s\|_\infty \right) \tag{118}$$

$$\leq C N^{-\lambda_1 - \lambda_2}, \tag{119}$$

under the event $\mathcal{A}_T$, which leads to

$$\|K_1^N(X_t) - K_1^N(\widetilde{X}_t)\|_\infty \leq (\|\nabla K_1^N(X_t)\|_\infty + \|\nabla K_1^N(\widetilde{X}_t)\|_\infty)\|X_t - \widetilde{X}_t\|_\infty$$
$$\leq C N^{3(\delta - \lambda_2)}\|L^N(\overline{X}_t)\|_\infty \|X_t - \widetilde{X}_t\|_\infty \leq C N^{3\delta - \lambda_1 - 4\lambda_2}\|L^N(\overline{X}_t)\|_\infty. \tag{120}$$

Here the bound $\frac{\|\nabla K_1^N(X_t)\|_\infty}{\|L^N(\overline{X}_t)\|_\infty} \leq C N^{3(\delta - \lambda_2)}$ uses Lemma 2.2 since

$$\|X_t - \overline{X}_t\|_\infty \leq N^{-\lambda_2} \gg N^{-\delta}. \tag{121}$$

And a similar estimate leads to $\frac{\|\nabla K_1^N(\widetilde{X}_t)\|_\infty}{\|L^N(\overline{X}_t)\|_\infty} \leq C N^{3(\delta - \lambda_2)}$.

We denote the event

$$\mathcal{B}_3 := \left\{ \max_{t \in [0,T]} \left\| L^N(\overline{X}_t) - \overline{L}^N(\overline{X}_t) \right\|_\infty \leq C_{2,\alpha} N^{3\delta - 1} \log(N) \right\}. \tag{122}$$

It has been proved in Proposition 3.1 that

$$\mathbb{P}(\mathcal{B}_3^c) \leq N^{-\alpha}. \tag{123}$$

Then under the event $\mathcal{B}_3$ it follows that

$$\|L^N(\overline{X}_t)\|_\infty \leq \|\overline{L}^N(\overline{X}_t)\|_\infty + C_{2,\alpha} N^{3\delta - 1} \log(N) \leq C N^{3\delta - 1} \log(N), \tag{124}$$

since $\|\overline{L}^N(\overline{X}_t)\|_\infty \le C \log(N)$ and $\frac{1}{3} \le \delta < 1$. Thus, we have

$$\mathcal{I}_2 = \|K_1^N(X_t) - K_1^N(\widetilde{X}_t)\|_\infty \le C N^{6\delta - 1 - \lambda_1 - 4\lambda_2} \log(N), \quad \forall \, t \in [t_n, t_{n+1}], \quad (125)$$

under the event $\mathcal{A}_T \cap \mathcal{B}_3$.

The estimate of $\mathcal{I}_3$ is a result of Lemma 3.2. Indeed, we denote the event

$$\mathcal{G}_n := \left\{ \max_{t \in [t_n, t_{n+1}]} \left\| K_1^N(\widetilde{X}_t) - K_1^N(\overline{X}_t) \right\|_\infty \le C_{4,\alpha} N^{2\delta - 1} \log(N) \right.$$
$$\left. + C_{4,\alpha} \log^2(N) N^{3\lambda_1 - \lambda_2} \|k_1^N\|_1 \right\}, \quad (126)$$

so by Lemma 3.2 one has that for any $0 \le n \le M'$

$$\mathbb{P}\left(\mathcal{A}_T \cap \mathcal{G}_n^c \right) \le N^{-\alpha}. \quad (127)$$

Furthermore, it holds that

$$\mathcal{I}_3 = \|K_1^N(\widetilde{X}_t) - K_1^N(\overline{X}_t)\|_\infty \le C_{4,\alpha} N^{2\delta - 1} \log(N) + C_{4,\alpha} \log^2(N) N^{3\lambda_1 - 2\lambda_2}$$
$$\le C(\alpha, T, C_{f_0}) \log^2(N)(N^{3\lambda_1 - 2\lambda_2} + N^{2\delta - 1}), \quad \forall \, t \in [t_n, t_{n+1}], \quad (128)$$

under the event $\mathcal{G}_n$, where we have used the fact that $\|k_1^N\|_1 \le C N^{-\lambda_2}$. Indeed, it is easy to compute that

$$\|k_1^N\|_1 = \|k^N - k_2^N\|_1 \le C \int_{0 \le |x| \le N^{-\lambda_2}} \frac{1}{|x|^2} dx \le C N^{-\lambda_2}. \quad (129)$$

Collecting (116), (125) and (128) yields that

$$\|K^N(X_t) - K^N(\overline{X}_t)\|_\infty$$
$$\le C \log(N) \left\| X_t - \overline{X}_t \right\|_\infty + C N^{6\delta - 1 - \lambda_1 - 4\lambda_2} \log(N) + C \log^2(N)(N^{3\lambda_1 - 2\lambda_2} + N^{2\delta - 1})$$
$$\le C \log(N) \left\| X_t - \overline{X}_t \right\|_\infty + C \log^2(N)(N^{6\delta - 1 - \lambda_1 - 4\lambda_2} + N^{3\lambda_1 - 2\lambda_2} + N^{2\delta - 1}),$$
$$\forall t \in [t_n, t_{n+1}],$$

under the event $\mathcal{B}_2 \cap \mathcal{B}_3 \cap \mathcal{A}_T \cap \mathcal{G}_n$, where C depends on α, T and C_{f_0}. To distinguish it from other constants we will denote this C by $C_{3,\alpha}$. This implies $\mathcal{B}_2 \cap \mathcal{B}_3 \cap \mathcal{A}_T \cap \mathcal{G}_n \subseteq \mathcal{S}_n(C_{3,\alpha})$, which yields that

$$\mathcal{B}_2 \cap \mathcal{B}_3 \cap \mathcal{A}_T \cap \left(\bigcap_{n=0}^{M'} \mathcal{G}_n \right) \subseteq \left(\bigcap_{n=0}^{M'} \mathcal{S}_n(C_{3,\alpha}) \right) = \mathcal{S}_T(C_{3,\alpha}). \quad (130)$$

It follows that

$$\mathbb{P}\left(\mathcal{A}_T \cap \mathcal{S}_T^c(C_{3,\alpha}) \right) \le \mathbb{P}\left(\mathcal{B}_2^c \right) + \mathbb{P}\left(\mathcal{B}_3^c \right) + \sum_{n=0}^{M'} \mathbb{P}\left(\mathcal{A}_T \cap \mathcal{G}_n^c \right)$$
$$\le (M' + 3) N^{-\alpha} \le 2T N^{\lambda_1 - \alpha} \le N^{-\alpha'},$$

where we used the estimates in (114), (123) and (127). Here α is arbitrary and so is α'. □

Lemma 3.2 *Consider two trajectories $(\widetilde{X}_t, \widetilde{V}_t)$, $(\overline{X}_t, \overline{V}_t)$ on $t \in [t_n, t_{n+1}]$ satisfying (106)-(107) and (24) respectively. When $1 \le n \le M'$, the two different initial phases are chosen to be $(X_{t_{n-1}}, V_{t_{n-1}})$ and $(\overline{X}_{t_{n-1}}, \overline{X}_{t_{n-1}})$ at time $t = t_{n-1}$, and when $n = 0$ the two different initial phases are chosen to be (X_0, V_0) and $(\overline{X}_0, \overline{V}_0)$ at time $t = 0$. Then for any $\alpha > 0$,*

there exists a $C_{4,\alpha} > 0$ depending only on α, T and C_{f_0} such that for N sufficiently large it holds that

$$\mathbb{P}\left(\mathcal{A}_T \cap \left\{ \max_{t \in [t_n, t_{n+1}]} \left\| K_1^N(\widetilde{X}_t) - K_1^N(\overline{X}_t) \right\|_\infty \geq C_{4,\alpha} N^{2\delta - 1} \log(N) \right. \right.$$
$$\left. \left. + C_{4,\alpha} \log^2(N) N^{3\lambda_1 - \lambda_2} \|k_1^N\|_1 \right\} \right) \leq N^{-\alpha}, \tag{131}$$

where we require $t_{n+1} - t_n = N^{-\lambda_1}$ with $0 < \lambda_1 < \frac{\lambda_2}{3}$ and $0 < \lambda_2 < \frac{1}{3}$. Here

$$(K_1^N(\widetilde{X}_t))_i = \frac{1}{N-1} \sum_{j \neq i}^N k_1^N(\widetilde{x}_i^t - \widetilde{x}_j^t), \quad (K_1^N(\overline{X}_t))_i$$

$$= \frac{1}{N-1} \sum_{j \neq i}^N k_1^N(\overline{x}_i^t - \overline{x}_j^t), \quad t \in [t_n, t_{n+1}],$$

where k_1^N is defined in (105).

Lemma 3.2 is used in the proof of Proposition 3.2. It follows from the following estimate of the term in (131) at any fixed time $t \in [t_n, t_{n+1}]$, a statement which will later be generalized to hold for the maximum of $\max t \in [t_n, t_{n+1}]$.

Lemma 3.3 *Under the same assumptions as in Lemma 3.2, for any $\alpha > 0$, there exists $C_{5,\alpha} > 0$ depending only on α, T and C_{f_0} such that for N sufficiently large it holds that for any fixed time $t \in [t_n, t_{n+1}]$*

$$\mathbb{P}\left(\mathcal{A}_T \cap \left\{ \left\| K_1^N(\widetilde{X}_t) - K_1^N(\overline{X}_t) \right\|_\infty \geq C_{5,\alpha} N^{2\delta - 1} \log(N) \right. \right.$$
$$\left. \left. + C_{5,\alpha} \log^2(N) N^{3\lambda_1 - \lambda_2} \|k_1^N\|_1 \right\} \right) \leq N^{-\alpha}. \tag{132}$$

The proof of Lemma 3.3 is carried out in Sect. 5. The novel technique in the proof used the fact that k_1^N has a support with the radius $N^{-\lambda_2}$ (small). This means that in order to contribute to the interaction, $\widetilde{x}_j^t$ (or $\overline{x}_j^t$) has to get close enough (less than $N^{-\lambda_2}$) to $\widetilde{x}_i^t$ (or $\overline{x}_i^t$). Due to the effect of Brownian motion we get mixing of the positions of the particles over the whole support of k_1^N. Using a Law of Large Numbers argument one can show that the leading order of the interaction can in good approximation be replaced by the respective expectation value. Due to symmetry of k_1^N this expectation value is zero. Significant fluctuations of the interaction k_1^N have very small probability.

The proof of Lemma 3.2 We follow the similar procedure as in Proposition 3.1. We divide $[t_n, t_{n+1}]$ into $M + 1$ subintervals with length $\Delta \tau = N^{-\frac{\gamma}{3}}$ for some $\gamma > 4$ and $\tau_k = k \Delta \tau$, $k = 0, \ldots, M+1$. Recall the event $\overline{\mathcal{H}}$ as in (90) and denote the event

$$\widetilde{\mathcal{H}} := \left\{ \max_k \max_{t \in [\tau_k, \tau_{k+1}]} \left\| \widetilde{X}_t - \widetilde{X}_{\tau_k} \right\|_\infty \leq C_B N^{-\frac{\gamma - 1}{3}} \right\}. \tag{133}$$

It follows from Lemma 3.1 that

$$\mathbb{P}(\overline{\mathcal{H}}^c), \mathbb{P}(\widetilde{\mathcal{H}}^c) \leq C_B N^{\frac{\gamma - 1}{3}} \exp(-C_B N^{\frac{2}{3}}), \tag{134}$$

for any $\gamma > 4$. Furthermore we denote the event

$$
\begin{aligned}
\mathcal{G}_{\tau_k} := \Big\{ & \left\| K_1^N(\widetilde{X}_{\tau_k}) - K_1^N(\overline{X}_{\tau_k}) \right\|_\infty \leq C_{5,\alpha} N^{2\delta - 1} \log(N) \\
& + C_{5,\alpha} \log^2(N) N^{3\lambda_1 - \lambda_2} \| k_1^N \|_1 \Big\}
\end{aligned}
\tag{135}
$$

in (132), then it follow from Lemma 3.3 that

$$
\mathbb{P}(\mathcal{G}_{\tau_k}^c) \leq N^{-\alpha}.
\tag{136}
$$

For all $t \in [\tau_k, \tau_{k+1}]$, under the event $\mathcal{G}_{\tau_k} \cap \overline{\mathcal{H}} \cap \widetilde{\mathcal{H}}$, we obtain

$$
\begin{aligned}
& \left\| K_1^N(\widetilde{X}_t) - K_1^N(\overline{X}_t) \right\|_\infty \\
& \leq \left\| K_1^N(\widetilde{X}_t) - K_1^N(\widetilde{X}_{\tau_k}) \right\|_\infty + \left\| K_1^N(\widetilde{X}_{\tau_k}) - K_1^N(\overline{X}_{\tau_k}) \right\|_\infty + \left\| K_1^N(\overline{X}_{\tau_k}) - K_1^N(\overline{X}_t) \right\|_\infty \\
& \leq \| \nabla K_1^N \|_\infty \left(\| \widetilde{X}_t - \widetilde{X}_{\tau_k} \|_\infty + \| \overline{X}_t - \overline{X}_{\tau_k} \|_\infty \right) + \left\| K_1^N(\widetilde{X}_{\tau_k}) - K_1^N(\overline{X}_{\tau_k}) \right\|_\infty \\
& \leq C N^{3\delta - \frac{\gamma - 1}{3}} + C_{5,\alpha} N^{2\delta - 1} \log(N) + C_{5,\alpha} \log^2(N) N^{3\lambda_1 - \lambda_2} \| k_1^N \|_1 \\
& \leq C(\alpha, T, C_{f_0}) N^{2\delta - 1} \log(N) + C(\alpha, T, C_{f_0}) \log^2(N) N^{3\lambda_1 - \lambda_2} \| k_1^N \|_1
\end{aligned}
$$

when $\gamma > 4$ is sufficiently large. This yields that under the event $\bigcap_{k=0}^M \mathcal{G}_{\tau_k} \cap \overline{\mathcal{H}} \cap \widetilde{\mathcal{H}}$ it holds that

$$
\begin{aligned}
\max_{t \in [t_n, t_{n+1}]} & \left\| K_1^N(\widetilde{X}_t) - K_1^N(\overline{X}_t) \right\|_\infty \\
& \leq C(\alpha, T, C_{f_0}) N^{2\delta - 1} \log(N) + C(\alpha, T, C_{f_0}) \log^2(N) N^{3\lambda_1 - \lambda_2} \| k_1^N \|_1.
\end{aligned}
$$

Therefore it follows from (134) and (136) that

$$
\begin{aligned}
\mathbb{P}\Big(\max_{t \in [t_n, t_{n+1}]} & \left\| K_1^N(\widetilde{X}_t) - K_1^N(\overline{X}_t) \right\|_\infty \geq C(\alpha, T, C_{f_0}) N^{2\delta - 1} \log(N) \\
& + C(\alpha, T, C_{f_0}) \log^2(N) N^{3\lambda_1 - \lambda_2} \| k_1^N \|_1 \Big) \leq \sum_{k=0}^M \mathbb{P}(\mathcal{G}_{\tau_k}^c) + \mathbb{P}(\overline{\mathcal{H}}^c) + \mathbb{P}(\widetilde{\mathcal{H}}^c) \\
& \leq T N^{-\frac{3\alpha - \gamma}{3}} + 2 C_B N^{\frac{\gamma - 1}{3}} \exp(-C_B N^{\frac{2}{3}}) \leq N^{-\alpha'}.
\end{aligned}
\tag{137}
$$

Denote $C_{4,\alpha'}$ to be the constant $C(\alpha, T, C_{f_0})$ in (137). Since $\alpha > 0$ is arbitrary and so is α', (131) holds true. This completes the proof of Lemma 3.2. $\qquad\square$

3.3 Convergence and the Proof of Theorem 1.2

In this section, we achieve the convergence by using the consistency from Proposition 3.1 and the stability from Proposition 3.2. To do this, we first prove the following lemma.

Lemma 3.4 *Under the same assumptions as in Proposition* 3.2, *for any $\alpha > 0$ assume that*

$$
\mathbb{P}(\mathcal{A}_T^c) \leq N^{-\alpha},
\tag{138}
$$

and let $C_{3,\alpha}$ be the constant given in Proposition 3.2. *Then it holds that*

$$\mathbb{P}(\mathcal{S}_T^c(C_{3,\alpha})) \leq 2N^{-\alpha} \tag{139}$$

for N large enough.

Proof We apply Proposition 3.2 and obtain

$$\mathbb{P}(\mathcal{S}_T^c(C_{3,\alpha})) = \mathbb{P}(\mathcal{A}_T \cap \mathcal{S}_T^c(C_{3,\alpha})) + \mathbb{P}(\mathcal{A}_T^c \cap \mathcal{S}_T^c(C_{3,\alpha})) \leq N^{-\alpha} + \mathbb{P}(\mathcal{A}_T^c) \leq 2N^{-\alpha}$$

for *N* large enough. □

We now return to the proof of Theorem 1.2.

Proof of Theorem 1.2 We shall also use the quantity $e(t)$ defined as

$$e(t) := \|\Phi_t - \Psi_t\|_\infty = \sqrt{\log(N)} \|X_t - \overline{X}_t\|_\infty + \|V_t - \overline{V}_t\|_\infty. \tag{140}$$

Using the fact that $\frac{d\|x\|_\infty}{dt} \leq \|\frac{dx}{dt}\|_\infty$, one has for all $t \in (0, T]$

$$\begin{aligned}
\frac{de(t)}{dt} &\leq \sqrt{\log(N)} \|V_t - \overline{V}_t\|_\infty + \|K^N(X_t) - \overline{K}^N(\overline{X}_t)\|_\infty \\
&\leq \sqrt{\log(N)} \|V_t - \overline{V}_t\|_\infty + \|K^N(X_t) - K^N(\overline{X}_t)\|_\infty + \|K^N(\overline{X}_t) - \overline{K}^N(\overline{X}_t)\|_\infty.
\end{aligned} \tag{141}$$

Recall the event $\mathcal{S}_T(C_{3,\alpha})$ defined in (102)

$$\begin{aligned}
\mathcal{S}_T(C_{3,\alpha}) := \Big\{ &\|K^N(X_t) - K^N(\overline{X}_t)\|_\infty \leq \Lambda \log(N) \|X_t - \overline{X}_t\|_\infty \\
&+ C_{3,\alpha} \log^2(N)(N^{6\delta-1-\lambda_1-4\lambda_2} + N^{3\lambda_1-2\lambda_2} + N^{2\delta-1}), \ \forall \, t \in [0, T] \Big\},
\end{aligned} \tag{142}$$

and denote

$$\mathcal{C}_T := \left\{ \max_{t \in [0,T]} \|K^N(\overline{X}_t) - \overline{K}^N(\overline{X}_t)\|_\infty \leq C_{2,\alpha} N^{2\delta-1} \log(N) \right\}. \tag{143}$$

Then under the event $\mathcal{S}_T(C_{3,\alpha}) \cap \mathcal{C}_T$ it follows that

$$\begin{aligned}
\frac{de(t)}{dt} \leq &\sqrt{\log(N)} \|V_t - \overline{V}_t\|_\infty \\
&+ C_{3,\alpha} \log(N) \|X_t - \overline{X}_t\|_\infty + C_{3,\alpha} \log^2(N)(N^{6\delta-1-\lambda_1-4\lambda_2} + N^{3\lambda_1-2\lambda_2} + N^{2\delta-1}) \\
&+ C_{2,\alpha} N^{2\delta-1} \log(N) \\
\leq &\, C(\alpha, T, C_{f_0}) \sqrt{\log(N)} e(t) \\
&+ C(\alpha, T, C_{f_0}) \log^2(N)(N^{6\delta-1-\lambda_1-4\lambda_2} + N^{3\lambda_1-2\lambda_2} + N^{2\delta-1}) \\
\leq &\, C(\alpha, T, C_{f_0}) \sqrt{\log(N)} e(t) + C(\alpha, T, C_{f_0}) \log^2(N) N^{-\lambda_3},
\end{aligned} \tag{144}$$

where in the second inequality we used the fact that

$$\sqrt{\log(N)} \|V_t - \overline{V}_t\|_\infty + \log(N) \|X_t - \overline{X}_t\|_\infty = \sqrt{\log(N)} e(t). \tag{145}$$

Here we denote

$$-\lambda_3 := \max\{6\delta - 1 - \lambda_1 - 4\lambda_2, 3\lambda_1 - 2\lambda_2, 2\delta - 1\}. \tag{146}$$

Notice that for

$$0 < \lambda_2 < 1/3; \ 0 < \lambda_1 < \frac{\lambda_2}{3}; \ \frac{1}{3} \leq \delta < \min\left\{\frac{\lambda_1 + 3\lambda_2 + 1}{6}, \frac{1 - \lambda_2}{2}\right\}, \quad (147)$$

one has $-\lambda_3 < -\lambda_2$. In other words, we obtain that for $-\lambda_3 < -\lambda_2$, there exists some $N_0 \in \mathbb{N}$ such that for all $N \geq N_0$ it holds that

$$\frac{de(t)}{dt} \leq C(\alpha, T, C_{f_0})\sqrt{\log(N)}e(t) + C(\alpha, T, C_{f_0})\log^2(N)N^{-\lambda_3}, \text{ for all } t \in (0, T], \tag{148}$$

which leads to

$$\max_{0 \leq t \leq T} e(t) \leq e^{C\sqrt{\log(N)}T} C \log^2(N) N^{-\lambda_3} T, \tag{149}$$

by Gronwall's Lemma. Hence we have

$$\mathbb{P}\left(\max_{0 \leq t \leq T} e(t) \geq e^{C\sqrt{\log(N)}T} C \log^2(N) N^{-\lambda_3} T\right) \leq \mathbb{P}\left((\mathcal{S}_T(C_{3,\alpha}) \cap \mathcal{C}_T)^c\right)$$
$$\leq \mathbb{P}(\mathcal{C}_T^c) + \mathbb{P}(\mathcal{S}_T^c(C_{3,\alpha})). \tag{150}$$

According to Lemma 3.4, assuming that for any $\alpha > 0$

$$\mathbb{P}(\mathcal{A}_T^c) \leq N^{-\alpha}, \tag{151}$$

where $\mathcal{A}_T = \left\{\max_{t \in [0,T]} e(t) \leq N^{-\lambda_2}\right\}$ implies that

$$\mathbb{P}(\mathcal{S}_T^c(C_{3,\alpha})) \leq 2N^{-\alpha}. \tag{152}$$

This combining the consistency in Proposition 3.1, i.e. $\mathbb{P}\left(\mathcal{C}_T^c\right) \leq N^{-\alpha}$, (150) implies that

$$\mathbb{P}\left(\max_{0 \leq t \leq T} e(t) \geq e^{C\sqrt{\log(N)}T} C \log^2(N) N^{-\lambda_3} T\right) \leq \mathbb{P}(\mathcal{C}_T^c) + \mathbb{P}(\mathcal{S}_T^c(C_{3,\alpha})) \leq 3N^{-\alpha}. \tag{153}$$

We are left to verify the assumption (151). Indeed since $-\lambda_3 < -\lambda_2$, and $e^{C\sqrt{\log(N)}T}$, $\log^2(N)$ are asymptotically bounded by any positive power of N, we can find a $N_0 \in \mathbb{N}$ depending only on C and T such that for any $N \geq N_0$

$$e^{C\sqrt{\log(N)}T} C \log^2(N) N^{-\lambda_3} T \leq \frac{1}{2} N^{-\lambda_2} < N^{-\lambda_2}. \tag{154}$$

Hence it holds that

$$\mathbb{P}\left(\max_{t \in [0,T]} e(t) \geq \frac{1}{2} N^{-\lambda_2}\right) \leq \mathbb{P}\left(\max_{0 \leq t \leq T} e(t) \geq e^{C\sqrt{\log(N)}T} C \log^2(N) N^{-\lambda_3} T\right)$$
$$\leq 3N^{-\alpha} \leq N^{-\alpha'},$$

where, similar as above, for sufficiently large N the exponents α and α' can be any positive real number. This means that $\max_{t \in [0,T]} e(t)$ can hardly reach $N^{-\lambda_2}$, which verifies the assumption (151). Summarizing, we have proven that for any $\alpha > 0$, there exists some $N_0 \in \mathbb{N}$ such that

$$\mathbb{P}\left(\max_{0 \leq t \leq T} e(t) \leq e^{C\sqrt{\log(N)}T} C \log^2(N) N^{-\lambda_3} T\right) \geq 1 - N^{-\alpha}$$

for all $N \geq N_0$, where C depends only on α, T, and C_{f_0}. This leads to Theorem 1.2. $\qquad\square$

4 Proof of Theorem 1.3

In order to prove the error estimate between f_t and $\mu_\Phi(t)$, let us split the error into three parts

$$W_p(f_t, \mu_\Phi(t)) \leq W_p(f_t, f_t^N) + W_p(f_t^N, \mu_\Psi(t)) + W_p(\mu_\Psi(t), \mu_\Phi(t)). \tag{155}$$

The Theorem 1.3 is proven once we obtain the respective error estimates of those three parts.

Proof of Theorem 1.3 • *The first term* $W_p(f_t, f_t^N)$. The convergence of this term is a deterministic result: solutions of the regularized VPFP equations (13) approximate solutions of the original VPFP equations (3) as the width of the cut-off goes to zero. It follows from [12, Lemma 3.2] that

$$\max_{t \in [0,T]} W_p(f_t, f_t^N) \leq N^{-\delta} e^{C_1 \sqrt{\log(N)}}, \tag{156}$$

where $p \in [1, \infty)$, $N > 3$ and C_1 depends only on T and C_{f_0}. The proof is inspired by the method of Leoper [45]. Note that here we can't follow the method in [39] directly since the support of f^N and f are not compact in our present case.

• *The second term* $W_p(f_t^N, \mu_\Psi(t))$. This term concerns the sampling of the mean-field dynamics by discrete particle trajectories. The convergence rate has been proved in [39, Corollary 9.4] by using the concentration estimate of Fournier and Guillin [18]. We summarize the result as follows: let $p \in [1, \infty)$, $\kappa < \min\{\delta, \frac{1}{6}, \frac{1}{2p}\}$ and $N > 3$. Assume that there exists $m > 2p$ such that

$$\iint_{\mathbb{R}^6} (|x|^m + |v|^m) f_0(x, v) dx dv < +\infty.$$

Then there exist constants C_2 and C_3 such that it holds

$$\mathbb{P}\left(\max_{t \in [0,T]} W_p(f_t^N, \mu_\Psi(t)) \leq \sqrt{\log(N)} N^{-\kappa} e^{C_2 \sqrt{\log(N)}} \right)$$
$$\geq 1 - C_3 \left(e^{-C_4 N^{1-\max\{6,2p\}\kappa}} + N^{1-\frac{m}{2p}} \right). \tag{157}$$

• *The third term* $W_p(\mu_\Psi(t), \mu_\Phi(t))$. The convergence of this term is a direct result of Theorem 1.2. Indeed, it follows from [39, Lemma 5.2] that for all $p \in [0, \infty]$

$$\max_{t \in [0,T]} W_p(\mu_\Psi(t), \mu_\Phi(t)) \leq \max_{t \in [0,T]} \|\Psi(t) - \Phi(t)\|_\infty. \tag{158}$$

Then we choose $\alpha = \frac{m}{2p} - 1$ in Theorem 1.2 so that

$$\mathbb{P}\left(\max_{t \in [0,T]} W_p(\mu_\Psi(t), \mu_\Phi(t)) \leq N^{-\lambda_2} \right) \geq 1 - N^{1-\frac{m}{2p}}. \tag{159}$$

•*Convergence of* $W_p(f_t, \mu_\Phi(t))$. Collecting estimates (156), (157) and (159) and choosing $\kappa < \min\{\delta, \frac{1}{6}, \frac{1}{2p}\}$, it follows that

$$\mathbb{P}\left(\max_{t \in [0,T]} W_p(f_t, \mu_\Phi(t)) \leq (1 + \sqrt{\log(N)}) N^{-\kappa} e^{C_5 \sqrt{\log(N)}} + N^{-\lambda_2} \right)$$
$$\geq 1 - C_6 \left(e^{-C_7 N^{1-\max\{6,2p\}\kappa}} + N^{1-\frac{m}{2p}} \right), \tag{160}$$

where C_5 depends only on T and C_{f_0}, and C_6, C_7 depend only on m, p, κ. We can simplify this result by demanding $N \geq e^{\left(\frac{2C_5}{1-3\lambda_2}\right)^2}$, which yields $N^{1-3\lambda_2} \geq (1+\sqrt{\log(N)})e^{C_5\sqrt{\log(N)}}$. Hence we conclude that

$$\mathbb{P}\left(\max_{t\in[0,T]} W_p(f_t, \mu_\Phi(t)) \leq N^{-\kappa+1-3\lambda_2} + N^{-\lambda_2}\right)$$
$$\geq 1 - C_6\left(e^{-C_7 N^{1-\max\{6,2p\}\kappa}} + N^{1-\frac{m}{2p}}\right). \qquad (161)$$

$\square$

5 The Proof of Lemma 3.3

In this section, we present the proof of Lemma 3.3, which provides the distance between $K_1^N(\widetilde{X}_t)$ and $K_1^N(\overline{X}_t)$ ($t \in [t_n, t_{n+1}]$), where $(\widetilde{X}_t, \widetilde{V}_t)$, $(\overline{X}_t, \overline{V}_t)$ satisfying (106)–(107) and (24) respectively with two different initial phases $(X_{t_{n-1}}, V_{t_{n-1}})$ and $(\overline{X}_{t_{n-1}}, \overline{X}_{t_{n-1}})$ at time $t = t_{n-1}$ when $1 \leq n \leq M'$, or (X_0, V_0) and $(\overline{X}_0, \overline{V}_0)$ at time $t = 0$ when $n = 0$. To do this, we introduce the following stochastic process: For time $0 \leq s \leq t$ and $a := (a_x, a_v) \in \mathbb{R}^{6N}$, let $Z_{t,s}^{a,N} := (Z_{x,t,s}^{a,N}, Z_{v,t,s}^{a,N})$ be the process starting at time s at the position (a_x, a_v) and evolving from time s up to time t according to the mean-field force $\overline{K}^N$:

$$\begin{cases} dZ_{x,t,s}^{a,i,N} = Z_{v,t,s}^{a,i,N}dt, \quad t > s, \\ dZ_{v,t,s}^{a,i,N} = \int_{\mathbb{R}^3} k^N(Z_{x,t,s}^{a,i,N} - x)\rho^N(x,t)dx + \sqrt{2\sigma}dB_i^t, \quad i = 1, \ldots, N, \end{cases} \qquad (162)$$

and

$$(Z_{x,s,s}^{a,i,N}, Z_{v,s,s}^{a,i,N}) = (a_x^i, a_v^i), \quad \text{at } t = s. \qquad (163)$$

Note that here $(Z_{x,t,s}^{a,i,N}, Z_{v,t,s}^{a,i,N})$, $i = 1, \ldots, N$ are independent. Furthermore $(Z_{x,t,s}^{a,N}, Z_{v,t,s}^{a,N})$ has the strong Feller property (see [21] Definition (A)), implying in particular that it has a transition probability density $u_{t,s}^{a,N}$ which is given by the product $u_{t,s}^{a,N} := \prod_{i=1}^N u_{t,s}^{a,i,N}$. Hence each term $u_{t,s}^{a,i,N}$ is the transition probability density of $(Z_{x,t,s}^{a,i,N}, Z_{v,t,s}^{a,i,N})$ and is also the solution to the linearized equation for $t >$:

$$\partial_t u_{t,s}^{a,i,N} + v \cdot \nabla_x u_{t,s}^{a,i,N} + k^N * \rho^N \cdot \nabla_v u_{t,s}^{a,i,N} = \Delta_v u_{t,s}^{a,i,N}, \quad u_{s,s}^{a,i,N} = \delta_{a_i}, \quad (164)$$

where $\rho^N = \int_{\mathbb{R}^3} f^N(t,x,v)dv$, and f^N solves the regularized VPFP equations (13) with initial condition f_0.

Consider now the process $Z_{t,s}^{a,N}$ and $Z_{t,s}^{b,N}$ for two different starting points $a, b \in \mathbb{R}^{6N}$. It is intuitively clear that the probability density $u_{t,s}^{a,i,N}$ and $u_{t,s}^{b,i,N}$ are just a shift of each other. The next lemma gives an estimate for the distance between any two densities in terms of the distance between the starting points a and b and the elapsed time $t - s$. The proof is carried out in Appendix A.

Lemma 5.1 *There exists a positive constant C depending only on C_{f_0} and T such that for each $N \in \mathbb{N}$, any starting points $a, b \in \mathbb{R}^{6N}$ and any time $0 < t \leq T$, the following estimates for the transition probability densities $u_{t,s}^{a,i,N}$ resp. $u_{t,s}^{b,i,N}$ of the processes $Z_{t,s}^{a,i,N}$ resp. $Z_{t,s}^{b,i,N}$ given by (162) hold for $t - s < \min\{1, T - s\}$:*

(i) $\|u_{t,s}^{a,i,N}\|_{\infty,1} \leq C\left((t-s)^{-\frac{9}{2}} + 1\right)$,

(ii) $\|u_{t,s}^{a,i,N} - u_{t,s}^{b,i,N}\|_{\infty,1} \le C|a - b| \left((t - s)^{-6} + 1\right)$.

The norm $\|\cdot\|_{p,q}$ *denotes the p-norm in the x and q-norm in the v-variable, i.e. for any* $f : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$

$$\|f\|_{p,q} := \left(\int_{\mathbb{R}^3} \left(\int_{\mathbb{R}^3} |f(x, v)|^q \, dv\right)^{p/q} dx\right)^{1/p}. \tag{165}$$

To this end one assumes $\Delta t = t_{n+1} - t_n = N^{-\lambda_1}$. Next we define for $t \in [t_n, t_{n+1}]$ the random sets

$$M_{t_n}^t := \left\{2 \le j \le N : \left|x_1^{t_n} - x_j^{t_n} + (t - t_n)(v_1^{t_n} - v_j^{t_n})\right| \le N^{-\lambda_2} + \log(N)\Delta t^{\frac{3}{2}}\right\} \tag{166}$$

and

$$\overline{M}_{t_n}^t := \left\{2 \le j \le N : \left|\overline{x}_1^{t_n} - \overline{x}_j^{t_n} + (t - t_n)(\overline{v}_1^{t_n} - \overline{v}_j^{t_n})\right| \le 3N^{-\lambda_2} + \log(N)\Delta t^{\frac{3}{2}}\right\}. \tag{167}$$

Here $M_{t_n}^t$ is at time t_n the set of indices of those particles $x_j^{t_n}$ which are in the ball of radius $N^{-\lambda_2} + \log(N)\Delta t^{\frac{3}{2}}$ around $x_1^{t_n} + (t - t_n)(v_1^{t_n} - v_j^{t_n})$, and $\overline{M}_{t_n}^t$ is an intermediate set introduced to help to control $M_{t_n}^t$. Note that under the event $\mathcal{A}_T$, we have $M_{t_n}^t \subseteq \overline{M}_{t_n}^t$.

We also define random sets for $t \in [t_n, t_{n+1}]$

$$\mathcal{S}_{t_n}^t = \left\{\text{card}\,(M_{t_n}^t) < 2C_*N \left(3N^{-\lambda_2} + \log(N)\Delta t^{\frac{3}{2}}\right)^2\right\}, \tag{168}$$

and

$$\overline{\mathcal{S}}_{t_n}^t = \left\{\text{card}\,(\overline{M}_{t_n}^t) < 2C_*N \left(3N^{-\lambda_2} + \log(N)\Delta t^{\frac{3}{2}}\right)^2\right\}, \tag{169}$$

where C_* will be defined later. Here $\mathcal{S}_{t_n}^t$ indicates the event where the number of particles inside the set $M_{t_n}^t$ is smaller than $2C_*N \left(3N^{-\lambda_2} + \log(N)\Delta t^{\frac{3}{2}}\right)^2$, and the event $\overline{\mathcal{S}}_{t_n}^t$ is introduced to help estimate $\mathbb{P}(\mathcal{S}_{t_n}^t)$.

Our next lemma provides the probability estimate of the event where particle $\widetilde{x}_j^t$ (or $\overline{x}_j^t$) is close to $\widetilde{x}_1^t$ (or $\overline{x}_1^t$) (distance smaller than $N^{-\lambda_2}$) during a short time interval $t - t_n$, which contributes to the interaction of k_1^N defined in (105), since the support of k_1^N has radius $N^{-\lambda_2}$.

Lemma 5.2 *Let* $(\widetilde{x}_j^t, \widetilde{v}_j^t)$ *satisfy* (106) *and* (107) *on* $t \in [t_n, t_{n+1}]$ *and the random set* $M_{t_n}^t$ *satisfy* (166), *then for any* $\alpha > 0$, *there exists some constant* $N_0 > 0$ *depending only on* α, T *and* C_{f_0} *such that for all* $N \ge N_0$ *it holds*

$$\mathbb{P}\left(\min_{t \in [t_n, t_{n+1}]} \max_{j \in (M_{t_n}^t)^c} \left\{\left|\widetilde{x}_1^t - \widetilde{x}_j^t\right|\right\} < N^{-\lambda_2}\right) \le N^{-\alpha},$$

$$\mathbb{P}\left(\min_{t \in [t_n, t_{n+1}]} \max_{j \in (\overline{M}_{t_n}^t)^c} \left\{\left|\overline{x}_1^t - \overline{x}_j^t\right|\right\} < N^{-\lambda_2}\right) \le N^{-\alpha}.$$

This means that for some particle index j *outside* $M_{t_n}^t$, $\widetilde{x}_j^t$ *for some* $t \in [t_n, t_{n+1}]$ *such that* $\left|\widetilde{x}_1^t - \widetilde{x}_j^t\right| < N^{-\lambda_2}$ *(i.e.* $\widetilde{x}_j^t$ *contributes to the interaction of* k_1^N*) with probability less than* $N^{-\alpha}$. *Here* $\mathbb{P}$ *is understood to be taken on the initial condition* $\widetilde{x}_j^{t_n}$.

Proof Let $(1, j)$ be fixed and $a_1^t := (a_{1,x}^t, a_{1,v}^t)$, $b_j^t := (b_{j,x}^t, b_{j,v}^t) \in \mathbb{R}^6$ satisfy the stochastic differential equations

$$da_{1,x}^t = a_{1,v}^t dt, \ da_{1,v}^t = \sqrt{2\sigma} dB_1^t; \quad db_{j,x}^t = b_{j,v}^t dt, \ db_{j,v}^t = \sqrt{2\sigma} dB_j^t, \quad t_n < t \leq t_{n+1},$$

with the initial data $a_1^{t_n} = 0$ and $b_j^{t_n} = 0$. Here B_j^t is the same as in (106). It follows from the evolution Eq. (106) that

$$d\left(\widetilde{x}_1^t - a_{1,x}^t\right) = \left(\widetilde{v}_1^t - a_{1,v}^t\right) dt \text{ and } d\left(\widetilde{v}_1^t - a_{1,v}^t\right) = \overline{k}^N(\widetilde{x}_1^t)dt, \quad t_n < t \leq t_{n+1},$$

and

$$d\left(\widetilde{x}_j^t - b_{j,x}^t\right) = \left(\widetilde{v}_j^t - b_{j,v}^t\right) dt \text{ and } d\left(\widetilde{v}_j^t - b_{j,v}^t\right) = \overline{k}^N(\widetilde{x}_j^t)dt, \quad t_n < t \leq t_{n+1},$$

where

$$\overline{k}^N := k^N * \rho^N,$$

which is bounded by $\|\overline{k}^N\|_\infty \leq C(\|\rho^N\|_1 + \|\rho^N\|_\infty)$ according to Lemma 2.4. Integrating twice we get for any $s \geq t_n$

$$\left(\widetilde{v}_j^s - b_{j,v}^s\right) = \widetilde{v}_j^{t_n} + \int_{t_n}^s \overline{k}^N(\widetilde{x}_j^\tau)d\tau$$

and

$$\left(\widetilde{x}_j^t - b_{j,x}^t\right) = \widetilde{x}_j^{t_n} + \int_{t_n}^t \left(\widetilde{v}_j^{t_n} + \int_{t_n}^s \overline{k}^N(\widetilde{x}_j^\tau)d\tau\right) ds .$$

And by the same argument one has

$$\widetilde{x}_1^t - \widetilde{x}_j^t - (a_{1,x}^t - b_{j,x}^t) = \widetilde{x}_1^{t_n} - \widetilde{x}_j^{t_n} + \int_{t_n}^t \left(\widetilde{v}_1^{t_n} - \widetilde{v}_j^{t_n} + \int_{t_n}^s \overline{k}^N(\widetilde{x}_1^\tau)d\tau - \int_{t_n}^s \overline{k}^N(\widetilde{x}_j^\tau)d\tau\right) ds .$$

Since $\overline{k}^N(\widetilde{x}_j^\tau)$ is bounded by $\|\overline{k}^N\|_\infty \leq C(\|\rho^N\|_1 + \|\rho^N\|_\infty)$ according to Lemma 2.4., it follows that there is a constant $0 < C < \infty$ depending only on $\|\overline{k}^N\|_\infty$ such that

$$\left|\widetilde{x}_1^t - \widetilde{x}_j^t\right| \geq \left|\widetilde{x}_1^{t_n} - \widetilde{x}_j^{t_n} + (t - t_n)(\widetilde{v}_1^{t_n} - \widetilde{v}_j^{t_n})\right| - \left|a_{1,x}^t - b_{j,x}^t\right| - C\Delta t^2, \text{ for all } t \in [t_n, t_{n+1}]. \tag{170}$$

For $j \in (M_{t_n}^t)^c$ for some $t \in [t_n, t_{n+1}]$, i.e.

$$\left|\widetilde{x}_1^{t_n} - \widetilde{x}_j^{t_n} + (t - t_n)(\widetilde{v}_1^{t_n} - \widetilde{v}_j^{t_n})\right| \geq N^{-\lambda_2} + \log(N)\Delta t^{\frac{3}{2}} \tag{171}$$

together with $\min\limits_{t \in [t_n, t_{n+1}]} \max\limits_{j \in (M_{t_n}^t)^c} \left\{\left|\widetilde{x}_1^t - \widetilde{x}_j^t\right|\right\} < N^{-\lambda_2}$, (170) and (171) imply

$$\max\limits_{t \in [t_n, t_{n+1}]} \min\limits_{j \in (M_{t_n}^t)^c} \left\{\left|a_{1,x}^t - b_{j,x}^t\right|\right\} > (\ln N - C)\Delta t^{\frac{3}{2}}.$$

Hence

$$
\mathbb{P}\left(\min_{t \in [t_n, t_{n+1}]} \max_{j \in (M_{t_n}^t)^c} \left\{ \left| \widetilde{x}_1^t - \widetilde{x}_j^t \right| \right\} < N^{-\lambda_2} \right)
$$

$$
\leq \mathbb{P}\left(\max_{t \in [t_n, t_{n+1}]} \min_{j \in (M_{t_n}^t)^c} \left\{ \left| a_{1,x}^t - b_{j,x}^t \right| \right\} > (\ln N - C)\Delta t^{\frac{3}{2}} \right)
$$

$$
\leq \mathbb{P}\left(\max_{t \in [t_n, t_{n+1}]} \min_{j \in (M_{t_n}^t)^c} \left\{ \left| a_{1,v}^t - b_{j,v}^t \right| \right\} > (\ln N - C)\Delta t^{\frac{1}{2}} \right), \tag{172}
$$

where we used $a_x^t = \int_{t_n}^t a_v^s ds$ and $b_x^t = \int_{t_n}^t b_v^s ds$ in the second inequality. In the same way we can argue that

$$
\mathbb{P}\left(\min_{t \in [t_n, t_{n+1}]} \max_{j \in (\overline{M}_{t_n}^t)^c} \left\{ \left| \overline{x}_1^t - \overline{x}_j^t \right| \right\} < N^{-\lambda_2} \right)
$$

$$
\leq \mathbb{P}\left(\max_{t \in [t_n, t_{n+1}]} \min_{j \in (\overline{M}_{t_n}^t)^c} \left\{ \left| a_{1,v}^t - b_{j,v}^t \right| \right\} > (\ln N - C)\Delta t^{\frac{1}{2}} \right). \tag{173}
$$

Due to independence the difference $c_{j,v}^t = (c_{j,1}^t, c_{j,2}^t, c_{j,3}^t) = a_{1,v}^t - b_{j,v}^t$ is itself a Wiener process [62] since

$$
dc_{j,v}^t = d(a_{1,v}^t - b_{j,v}^t) = d(B_1^t - B_j^t). \tag{174}
$$

Splitting up this Wiener process into its three spacial components we get

$$
\mathbb{P}\left(\max_{t \in [t_n, t_{n+1}]} \min_{j \in (M_{t_n}^t)^c} \left\{ \left| a_{1,v}^t - b_{j,v}^t \right| \right\} > (\ln N - C)\Delta t^{\frac{1}{2}} \right)
$$

$$
\leq 3\mathbb{P}\left(\max_{t \in [t_n, t_{n+1}]} \min_{j \in (M_{t_n}^t)^c} \left\{ \left| c_{j,1}^t \right| \right\} > (\ln N - C)\Delta t^{\frac{1}{2}} \right)
$$

$$
\leq 6\mathbb{P}\left(\max_{t \in [t_n, t_{n+1}]} \min_{j \in (M_{t_n}^t)^c} \left\{ c_{j,1}^t \right\} > (\ln N - C)\Delta t^{\frac{1}{2}} \right)
$$

$$
= 12\mathbb{P}\left(\min_{j \in (M_{t_n}^t)^c} \left\{ c_{j,1}^{t_{n+1}} \right\} > (\ln N - C)\Delta t^{\frac{1}{2}} \right). \tag{175}
$$

where in the last equality we used the reflection principle based on the Markov property [40].

Recall that the time evolution of $a_{1,v}^t$ and $b_{j,v}^t$ are standard Brownian motions, i.e. the density is a Gaussian with standard deviation $\sigma_t = \sigma(t - t_n)^{\frac{1}{2}}$. Due to the independence of $a_{1,v}^t$ and $b_{j,v}^t$, $c_{j,1}^t$ is also normal distributed with the standard deviation of order $(t - t_n)^{\frac{1}{2}}$. Hence for N sufficiently large, following from (175), it holds that

$$
\mathbb{P}\left(\max_{t \in [t_n, t_{n+1}]} \min_{j \in (M_{t_n}^t)^c} \left\{ \left| a_{1,v}^t - b_{j,v}^t \right| \right\} > (\ln N - C)\Delta t^{\frac{1}{2}} \right) \leq N^{-\alpha}
$$

and

$$\mathbb{P}\left(\max_{t \in [t_n, t_{n+1}]} \min_{j \in (\overline{M}_{t_n}^t)^c} \left\{ \left| a_{1,v}^t - b_{j,v}^t \right| \right\} > (\ln N - C)\Delta t^{\frac{1}{2}} \right) \le N^{-\alpha}$$

With (172) and (173) the lemma follows. □

Now we have all the estimates needed for the proof of Lemma 3.3.

Proof of Lemma 3.3 We show that under the event $\mathcal{A}_T$ defined in (101), for any $\alpha > 0$ there exists a C_α depending only on α, T and C_{f_0} such that at any fixed time $t \in [t_n, t_{n+1}]$

$$\mathbb{P}\left(\left| \frac{1}{N-1} \sum_{j \neq 1}^{N} \left(k_1^N(\widetilde{x}_1^t - \widetilde{x}_j^t) - k_1^N(\overline{x}_1^t - \overline{x}_j^t) \right) \right| \right.$$
$$\left. \ge C_\alpha N^{2\delta - 1} \log(N) + C_\alpha \log^2(N) N^{3\lambda_1 - \lambda_2} \|k_1^N\|_1 \right) \le N^{-\alpha}. \tag{176}$$

This is done under the event $\mathcal{A}_T$ in three steps:

(1) We prove that for any $t \in [t_n, t_{n+1}]$ the number of particles inside $M_{t_n}^t$ is larger than

$$M_* := 2C_* N \left(3N^{-\lambda_2} + \log(N)\Delta t^{3/2} \right)^2 \tag{177}$$

with probability less than $N^{-\alpha}$. Note that M_* is used as a bound in the definition of (168) and (169). For any $t \in [t_n, t_{n+1}]$ we prove that

$$\mathbb{P}\left(\text{card}\,(M_{t_n}^t) > M_* \right) = \mathbb{P}((\mathcal{S}_{t_n}^t)^c) \le \mathbb{P}((\overline{\mathcal{S}}_{t_n}^t)^c) \le N^{-\alpha}. \tag{178}$$

(2) We prove that at any fixed time $t \in [t_n, t_{n+1}]$, particles outside $M_{t_n}^t$ contribute to the interaction of k_1^N with probability less than $N^{-\alpha}$, namely

$$\mathbb{P}\left(\left| \frac{1}{N-1} \sum_{j \in (M_{t_n}^t)^c} \left(k_1^N(\widetilde{x}_1^t - \widetilde{x}_j^t) - k_1^N(\overline{x}_1^t - \overline{x}_j^t) \right) \right| > 0 \right) \le N^{-\alpha}. \tag{179}$$

(3) According to step (2) above, at any fixed time $t \in [t_n, t_{n+1}]$, particles outside $M_{t_n}^t$ do not contribute to the interaction of k_1^N with high probability, so we only consider particles that are inside $M_{t_n}^t$. And we know already from step (1) above that the number of particles inside $M_{t_n}^t$ is larger than M_*, with low probability. To prove (176), we only need to prove

$$\mathbb{P}\left(\mathcal{X}(M_{t_n}^t) \cap \left\{ \text{card}\,(M_{t_n}^t) \le M_* \right\} \right) \le N^{-\alpha} \tag{180}$$

at any fixed time $t \in [t_n, t_{n+1}]$, where the event $\mathcal{X}(M_{t_n}^t)$ is defined by

$$\mathcal{X}(M_{t_n}^t) := \left\{ \left| \frac{1}{N-1} \sum_{j \in M_{t_n}^t} \left(k_1^N(\widetilde{x}_1^t - \widetilde{x}_j^t) - k_1^N(\overline{x}_1^t - \overline{x}_j^t) \right) \right| \right.$$
$$\left. \ge C_\alpha N^{2\delta - 1} \log(N) + C_\alpha \log^2(N) N^{3\lambda_1 - \lambda_2} \|k_1^N\|_1 \right\}. \tag{181}$$

• *Step 1:* To prove the first part of (178), note that on the event $\mathcal{A}_T$ defined in (101) and assuming that $t \in [t_n, t_{n+1}]$

$$\left| x_1^{t_n} - x_j^{t_n} + (t - t_n)(v_1^{t_n} - v_j^{t_n}) \right| \le N^{-\lambda_2} + \log(N)\Delta t^{\frac{3}{2}}$$
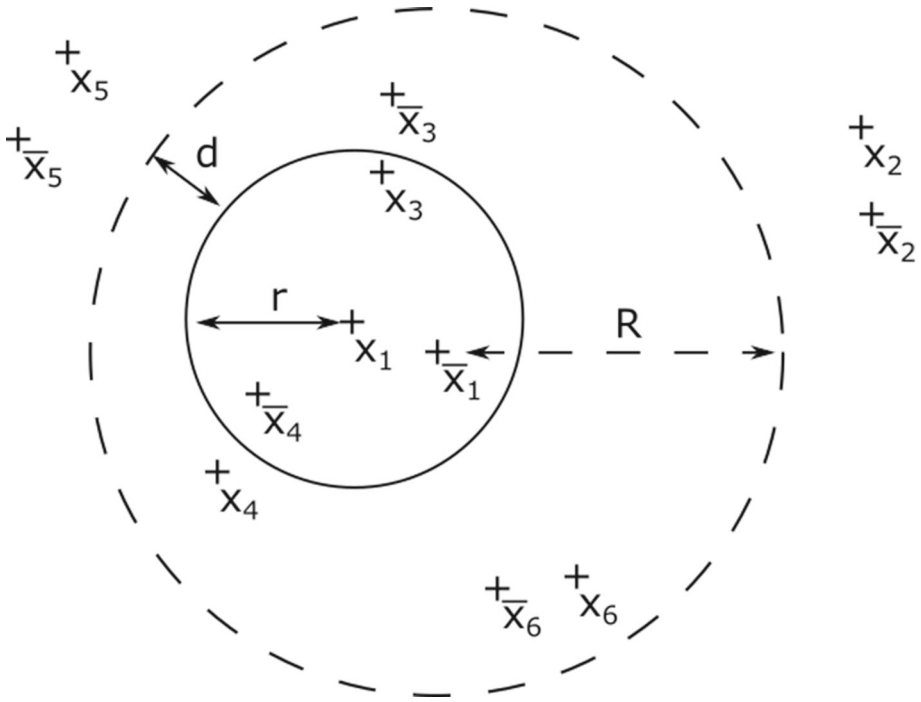
Fig. 1 Illustration of the sets $M_{t_n}^t$ and $\overline{M}_{t_n}$ under the assumption that $\mathcal{A}_T$ holds: the set $M_{t_n}^t$ contains all indices of particles with respect to X which are in the ball of radius $r = N^{-\lambda_2} + \log(N)(\Delta t)^{3/2}$ around x_1. In the figure this is the ball with solid lines and $M_{t_n}^t = \{1, 3\}$. The set $\overline{M}_{t_n}^t$ contains all indices of particles with respect to $\overline{X}$ which are in the ball of radius $R = 3N^{-\lambda_2} + \log(N)(\Delta t)^{3/2}$ around $\overline{x}_1$. In the figure this is the ball with dashed lines and $\overline{M}_{t_n}^t = \{1, 3, 4, 6\}$. Since on the set $\mathcal{A}_T$ the distance d of the particles x_1 and $\overline{x}_1$ cannot be larger than $N^{-\lambda_2}$, it follows that, given that the event $\mathcal{A}_T$ holds, a particle $\overline{x}_j$ is in the solid ball only if the particle x_j is in the ball with dashed lines, i.e. with radius $R = 3N^{-\lambda_2} + \log(N)(\Delta t)^{3/2}$ around x_1 (see for example particles x_3 and $\overline{x}_3$). Thus $M_{t_n}^t \subseteq \overline{M}_{t_n}^t$. Controlling $M_{t_n}^t$ by $\overline{M}_{t_n}$ will be helpful to estimate the number of particles inside these sets. The $\overline{x}_j$ are distributed independently, and the probability of finding any of these $\overline{x}_j$ inside the solid ball is small due to the small volume of the ball. This helps to estimate the number of particles in the set $\overline{M}_{t_n}$ (see *Step* 1). Particles outside the ball, i.e. indices not in $\overline{M}_{t_n}$ do not contribute to the interaction k_1. This comes from the fact that in order to get a sufficiently small distance for x_1 to interact, they have to travel a long distance during the short time interval $(t - t_n)$: the distance $\log(N)(\Delta t)^{3/2}$ (recall that the support of k_1 has radius $N^{-\lambda_2}$). Due to the Brownian motion, this is possible, of course, but the probability to travel that far will be smaller than any polynomial in N. This argument is worked out in *Step* 2. The main contribution thus comes from *Step* 3. Knowing that the number of particles in M_{t_n} is quite small helps to estimate this term

implies

$$\left| \overline{x}_1^{t_n} - \overline{x}_j^{t_n} + (t - t_n)(\overline{v}_1^{t_n} - \overline{v}_j^{t_n}) \right| \leq 3N^{-\lambda_2} + \log(N)\Delta t^{\frac{3}{2}} .$$

Hence $M_{t_n}^t \subseteq \overline{M}_{t_n}^t$ and thus for any $R > 0$, card $(\overline{M}_{t_n}^t) < R$ implies that card $(M_{t_n}^t) \leq$ card $(\overline{M}_{t_n}^t) < R$, consequently $\mathcal{S}_{t_n}^t \supseteq \overline{\mathcal{S}}_{t_n}^t$, i.e. $(\mathcal{S}_{t_n}^t)^c \subseteq (\overline{\mathcal{S}}_{t_n}^t)^c$.

The second part of (178) is trivial. For the third part we use the independence of the $\overline{x}$-particles. Note that the law of $(\overline{x}_{t_n}^j, \overline{v}_{t_n}^j)$ has a density $f^N(x, v, t_n)$. For any $j \in \{2, \ldots, N\}$

the probability to find $j \in \overline{M}^t_{t_n}$ for any $t \in [t_n, t_{n+1}]$ is given by

$$\mathbb{P}\left(j \in \overline{M}^t_{t_n}\right) = \int_{\mathbb{R}^3} \int_{B_R(\Xi^t)} f^N(x, v, t_n) dx dv, \tag{182}$$

where the center Ξ^t of the ball is given by $\Xi^t = \overline{x}^{t_n}_1 + (t - t_n)(\overline{v}^{t_n}_1 - v)$, and the radius of the ball is given by $R = 3N^{-\lambda_2} + \log(N)\Delta t^{3/2}$.

Define

$$g^N(x, v, s) := f^N(x - vs, v, t_n) \tag{183}$$

which then satisfies the following transport equation

$$\begin{cases} \partial_s g^N(x, v, s) + v \cdot \nabla_x g^N(x, v, s) = 0, \ 0 < s \leq \Delta t, \\ g^N(x, v, 0) = f^N(x, v, t_n). \end{cases} \tag{184}$$

Then one has

$$\int_{B_R(\Xi^t)} f^N(x, v, t_n) dx = \int_{B_R(\Xi^t_0)} g^N(x, v, t - t_n) dx, \tag{185}$$

where the center Ξ^t_0 of the ball is given by $\Xi^t_0 = \overline{x}^{t_n}_1 + (t - t_n)\overline{v}^{t_n}_1$, in particular the integration area is independent of v. It follows that the probability of finding $j \in \overline{M}^t_{t_n}$ for any $t \in [t_n, t_{n+1}]$ is equivalent to

$$\mathbb{P}\left(j \in \overline{M}^t_{t_n}\right) = \int_{\mathbb{R}^3} \int_{B_R(\Xi^t_0)} g^N(x, v, t - t_n) dx dv. \tag{186}$$

Next, we compute for $0 < s \leq \Delta t$

$$\begin{aligned} \bar{\rho}^N(x, s) &:= \int_{\mathbb{R}^3} g^N(x, v, s) dv = \int_{|v| \leq r(s)} g^N(x, v, s) dv + \int_{|v| > r(s)} g^N(x, v, s) dv \\ &\leq C_1 \|g^N(\cdot, \cdot, s)\|_\infty r(s)^3 + \frac{1}{r(s)^6} \int_{|v| > r(s)} |v|^6 g^N(x, v, s) dv \\ &= 2 C_1^{\frac{2}{3}} \|g^N(\cdot, \cdot, s)\|_\infty^{\frac{2}{3}} \left(\int_{|v| > r(s)} |v|^6 g^N(x, v, s) dv\right)^{\frac{1}{3}}, \end{aligned} \tag{187}$$

where we have chosen

$$r(s) = \left(\frac{\int_{|v| > r(s)} |v|^6 g^N(x, v, s) dv}{C_1 \|g^N(\cdot, \cdot, s)\|_\infty}\right)^{\frac{1}{9}}. \tag{188}$$

It follows that

$$\begin{aligned} \int_{\mathbb{R}^3} |\bar{\rho}^N(x, s)|^3 dx &\leq 8 C_1^2 \|g^N(\cdot, \cdot, s)\|_\infty^2 \iint_{\mathbb{R}^6} |v|^6 g^N(x, v, s) dx dv \\ &= 8 C_1^2 \|f^N(x - vs, v, t_n)\|_\infty^2 \iint_{\mathbb{R}^6} |v|^6 f^N(x - vs, v, t_n) dx dv \\ &\leq C\left(\|f^N(\cdot, \cdot, t_n)\|_{L^\infty(\mathbb{R}^6)}, \ \||v|^6 f^N(\cdot, \cdot, s)\|_{L^1(\mathbb{R}^6)}\right), \end{aligned} \tag{189}$$

which leads to

$$\max_{s \in [0, \Delta t]} \|\bar{\rho}^N(\cdot, s)\|_3 \leq C_2, \tag{190}$$

because of (20), where C_2 depends only on T, and C_{f_0}. It follows from (186) that

$$
\begin{aligned}
\mathbb{P}\left(j \in \overline{M}_{t_n}^t\right) = \int_{B_R(\Xi_0^t)} \bar{\rho}^N(x, t - t_n) dx &\leq \|\bar{\rho}^N\|_3 |B_R(\Xi_0^t)|^{\frac{2}{3}} \\
&\leq C_2(\frac{4}{3}\pi)^{\frac{2}{3}}\left(3N^{-\lambda_2} + \log(N)\Delta t^{\frac{3}{2}}\right)^2 \\
&= C_*\left(3N^{-\lambda_2} + \log(N)\Delta t^{3/2}\right)^2 =: p,
\end{aligned} \tag{191}
$$

where we define $C_* := C_2(\frac{4}{3}\pi)^{\frac{2}{3}}$, which depends only on T and C_{f_0}.

The probability of finding k particles inside the set $\overline{M}_{t_n}^t$ is thus bounded from above by the binomial probability mass function with parameter p at position k, i.e. for any natural number $0 \leq A \leq N$ and any $t \in [t_n, t_{n+1}]$

$$
\mathbb{P}\left(\operatorname{card}(\overline{M}_{t_n}^t) \geq A\right) \leq \sum_{j=A}^{N}\binom{N}{j} p^j (1-p)^{N-j}.
$$

Binomially distributed random variables have mean Np and standard deviation $\sqrt{Np(1-p)} < \sqrt{Np}$, and the probability to find more than $Np + a\sqrt{Np}$ particles in the set $\overline{M}_{t_n}^t$ is exponentially small in a, i.e. there is a sufficiently large N for any $\alpha > 0$ and any $t \in [t_n, t_{n+1}]$ such that

$$
\mathbb{P}\left(\operatorname{card}(\overline{M}_{t_n}^t) \geq Np + a\sqrt{Np}\right) \leq a^{-\alpha}.
$$

This is because of the central limit theory and so the binomial distribution can be seen as a normal distribution when N is sufficiently large. Since $p \geq CN^{-3\lambda_2}$, we get that $\sqrt{Np} > CN^{\frac{1}{2}(1-3\lambda_2)}$ ($\lambda_2 < 1/3$). Hence the probability of finding more than $2Np = Np + \sqrt{Np}\sqrt{Np}$ (i.e. $a = \sqrt{Np} > CN^{\frac{1}{2}(1-3\lambda_2)}$) particles is the set $\overline{M}_{t_n}^t$ is smaller than any polynomial in N, i.e. there is a C_α for any $\alpha > 0$ and any $t \in [t_n, t_{n+1}]$ such that

$$
\mathbb{P}((\overline{S}_{t_n}^t)^c) = \mathbb{P}\left(\operatorname{card}(\overline{M}_{t_n}^t) \geq 2Np\right) \leq N^{-\alpha}.
$$

• *Step 2:* For (179) it is sufficient to show that for any $\alpha > 0$ there is a sufficiently large N such that for some $j \in (M_{t_n}^t)^c$

$$
\mathbb{P}\left(\max_{t \in [t_n, t_{n+1}]}\left|\frac{1}{N-1}\left(k_1^N(\tilde{x}_1^t - \tilde{x}_j^t) - k_1^N(\overline{x}_1^t - \overline{x}_j^t)\right)\right| > 0\right) \leq N^{-\alpha}.
$$

The total probability we have to control in (179) is at maximum the N-fold value of this. The key to prove that is Lemma 5.2. To have an interaction $k_1^N(\tilde{x}_1^t - \tilde{x}_j^t) \neq 0$ for all $t \in [t_n, t_{n+1}]$ the distance between particle 1 and particle j has to be reduced to a value smaller than $N^{-\lambda_2}$. Due to the Brownian motion, this is possible, but suppressed. Due to the fast decay of the Gaussian it is very unlikely that $k_1^N(\tilde{x}_1^t - \tilde{x}_j^t) \neq 0$. The probability is smaller than any polynomial in N (see Lemma 5.2).The same holds true for $k_1^N(\overline{x}_1^t - \overline{x}_j^t)$.

In more detail: due to the cut-off $N^{-\lambda_2}$ we introduced for k_1^N

$$
\mathbb{P}\left(\max_{t \in [t_n, t_{n+1}]}\left|\frac{1}{N-1}\sum_{j \in (M_{t_n}^t)^c}\left(k_1^N(\tilde{x}_1^t - \tilde{x}_j^t) - k_1^N(\overline{x}_1^t - \overline{x}_j^t)\right)\right| > 0\right)
$$

$$\leq \mathbb{P}\left(\max_{t\in[t_n,t_{n+1}]} \left| \frac{1}{N-1} \sum_{j\in(M^t_{t_n})^c} k_1^N(\widetilde{x}_1^t - \widetilde{x}_j^t) \right| > 0 \right)$$

$$+ \mathbb{P}\left(\max_{t\in[t_n,t_{n+1}]} \left| \frac{1}{N-1} \sum_{j\in(M^t_{t_n})^c} k_1^N(\overline{x}_1^t - \overline{x}_j^t) \right| > 0 \right)$$

$$\leq N\mathbb{P}\left(\min_{t\in[t_n,t_{n+1}]} \max_{j\in(M^t_{t_n})^c} \left\{ \left| \widetilde{x}_1^t - \widetilde{x}_j^t \right| \right\} < N^{-\lambda_2} \right)$$

$$+ N\mathbb{P}\left(\min_{t\in[t_n,t_{n+1}]} \max_{j\in(\overline{M}^t_{t_n})^c} \left\{ \left| \overline{x}_1^t - \widetilde{x}_j^t \right| \right\} < N^{-\lambda_2} \right),$$

where we used the fact that $(\overline{M}^t_{t_n})^c \subseteq (M^t_{t_n})^c$ in the last inequality. With Lemma 5.2 we get the bound for (179).

- *Step 3:* To get (180) we prove that for any natural number

$$0 \leq M \leq M_* = 2C_*N\left(3N^{-\lambda_2} + \log(N)\Delta t^{3/2} \right)^2$$

one has

$$\mathbb{P}\left(\mathcal{X}(M^t_{t_n}) \cap \{ \operatorname{card}(M^t_{t_n}) = M \} \right) \leq N^{-\alpha}, \tag{192}$$

where the event $\mathcal{X}(M^t_{t_n})$ is defined in (181). This can be recast without relabeling j as

$$\mathbb{P}\left(\left| \frac{1}{N-1} \sum_{j=1}^M \left(k_1^N(\widetilde{x}_1^t - \widetilde{x}_j^t) - k_1^N(\overline{x}_1^t - \overline{x}_j^t) \right) \right| \right.$$
$$\left. \geq C_\alpha N^{2\delta-1}\log(N) + C_\alpha \log^2(N)N^{3\lambda_1-\lambda_2}\|k_1^N\|_1 \right) \leq N^{-\alpha}. \tag{193}$$

Lemma 5.3 *Let $Z_1, \cdots, Z_M$ be independent random variables with $\mathbb{E}[|Z_i|] \leq CM^{-2}$ and $|Z_i| \leq C$ for any $i \in \{1, \cdots, M\}$. Then for any $\alpha > 0$, it holds that*

$$\mathbb{P}\left(\sum_{i=1}^M |Z_i| \geq C_\alpha \ln(M) \right) \leq M^{-\alpha}, \tag{194}$$

where C_α depends only on C and α.

Proof We first split the random variables $Z_i = Z_i^a + Z_i^b$ such that Z_i^a and Z_i^b are sequences of independent random variables with

$$\mathbb{P}(|Z_i^a| > 0) = M^{-1} \text{ and } |Z_i^b| \leq CM^{-1}.$$

This can be achieved by defining

$$Z_i^a(\omega) = \begin{cases} Z_i(\omega) & \text{if } Z_i(\omega) > \gamma, \\ 0 & \text{else}. \end{cases}$$

and $Z_i^b = Z_i - Z_i^a$. Here we choose γ such that $\mathbb{P}(|Z_i^a| > 0) = M^{-1}$. Applying Markov's inequality, one computes

$$M^{-1} = \mathbb{P}(|Z_i^a| > 0) = \mathbb{P}(Z_i > \gamma) \le \mathbb{P}\left(|Z_i| > \frac{\gamma}{\mathbb{E}[|Z_i|]}\mathbb{E}[|Z_i|]\right) \le \frac{\mathbb{E}[|Z_i|]}{\gamma} \le C\frac{M^{-2}}{\gamma}. \tag{195}$$

This implies that $\gamma \le CM^{-1}$.

For the sum of Z_i^b we get the trivial bound

$$\sum_{j=1}^{M} |Z_i^b| \le CM^{-1}M = C \,.$$

Thus the lemma follows if we can show that

$$\mathbb{P}\left(\sum_{i=1}^{M} |Z_i^a| \ge C_\alpha \ln(M)\right) \le N^{-\alpha}, \tag{196}$$

where C_α has been changed.

Let

$$X_i(\omega) = \begin{cases} 0 & \text{if } Z_i^a(\omega) = 0, \\ 1 & \text{else} \,. \end{cases}$$

Since $|Z_i| \le C$, one has

$$\sum_{i=1}^{M} |Z_i^a| = \sum_{i=1}^{M} X_i|Z_i| \le C\sum_{i=1}^{M} X_i. \tag{197}$$

Then it follows that

$$\sum_{i=1}^{M} |Z_i^a| \ge C_\alpha \ln(M) \implies \sum_{j=1}^{M} X_i \ge \frac{C_\alpha}{C} \ln(M).$$

Noticing that X_i are i.i.d. Bernoulli random variables with $\mathbb{P}(X_i = 1) = \mathbb{P}(|Z_i^a| > 0) = M^{-1}$, we get

$$\mathbb{P}\left(\sum_{i=1}^{M} |Z_i^a| \ge C_\alpha \ln(M)\right) \le \mathbb{P}\left(\sum_{j=1}^{M} X_i \ge \frac{C_\alpha}{C} \ln(M)\right)$$

$$= \sum_{j=a}^{M} \frac{M!}{j!(M-j)!}M^{-j}(1 - M^{-1})^{M-j} \le \sum_{j=a}^{M} \frac{M^j}{j!}M^{-j} \le \frac{2}{a!},$$

where $a = \frac{C_\alpha}{C} \ln(M)$. Notice the decay property of the factorial

$$(\ln(M))! \ge \left(\frac{\ln(M)}{2}\right)^{\frac{\ln(M)}{2}} = \exp\left(\frac{\ln(M)}{2} \ln\left(\frac{\ln(M)}{2}\right)\right) = M^{\frac{1}{2}\ln\left(\frac{\ln(M)}{2}\right)}. \tag{198}$$

Thus one chooses M large enough and concludes (196), which proves the lemma. $\qquad\square$

Using the lemma above, now we proceed to prove (193). Define

$$Z_j := N^{-2\delta} \left(k_1^N(\widetilde{x}_1^t - \widetilde{x}_j^t) - k_1^N(\overline{x}_1^t - \overline{x}_j^t) \right).$$

It follows that $|Z_j|$ is bounded and

$$\mathbb{E}(|Z_j|) \le C N^{-2\delta} \|k_1^N\|_1 \begin{cases} \|u_{t,t_{n-1}}^{a,N}\|_{\infty,1}, & \text{for } 1 \le n \le M' \\ \|f_t^N\|_{\infty,1}, & \text{for } n = 1 \end{cases} \le C N^{\frac{9}{2}\lambda_1 - 2\delta - \lambda_2},$$

where we use the fact $\|u_{t,t_{n-1}}^{a,N}\|_{\infty,1} \le C N^{\frac{9}{2}\lambda_1}$ ($\Delta t \le t - t_{n-1} \le 2\Delta t$) from (i) in Lemma 5.1, $\|f_t^N\|_{\infty,1} \le C_{f_0}$ and $\|k_1^N\|_1 \le N^{-\lambda_2}$ from (129). Using Lemma 5.3 with $M = N^{\delta + \frac{\lambda_2}{2} - \frac{9}{4}\lambda_1}$ one obtains

$$\mathbb{P}\left(\frac{1}{N-1} \sum_{j=1}^{N^{\delta + \frac{\lambda_2}{2} - \frac{9}{4}\lambda_1}} \left| \left(k_1^N(\widetilde{x}_1^t - \widetilde{x}_j^t) - k_1^N(\overline{x}_1^t - \overline{x}_j^t) \right) \right| \ge C_\alpha N^{2\delta - 1} \ln(N) \right) \le N^{-\alpha}, \tag{199}$$

which leads to (193) for $M = N^{\delta + \frac{\lambda_2}{2} - \frac{9}{4}\lambda_1}$. It is obvious that

$$\sum_{i=1}^{M} |Z_i| \le \sum_{i=1}^{N^{\delta + \frac{\lambda_2}{2} - \frac{9}{4}\lambda_1}} |Z_i|, \tag{200}$$

for any $M \le N^{\delta + \frac{\lambda_2}{2} - \frac{9}{4}\lambda_1}$. Thus one concludes (193) holds for the case $M \le N^{\delta + \frac{\lambda_2}{2} - \frac{9}{4}\lambda_1}$.

For the remaining M we note that

$$2C_* N \left(3N^{-\lambda_2} + \log(N)N^{-\frac{3}{2}\lambda_1} \right)^2 \le 4C_* N \log^2(N) N^{-3\lambda_1}, \tag{201}$$

due to the fact that $0 < \lambda_1 < \frac{2}{3}\lambda_2$. Thus we are left to prove (193) for the case

$$N^{\delta + \frac{\lambda_2}{2} - \frac{9}{4}\lambda_1} < M \le 4C_* N \log^2(N) N^{-3\lambda_1}. \tag{202}$$

This can be done by Lemma 2.5, which we repeat below for easier reference:

Lemma 2.5 Let $Z_1, \ldots, Z_M$ be $i.i.d.$ random variables with $\mathbb{E}[Z_i] = 0$, $\mathbb{E}[Z_i^2] \le g(M)$ and $|Z_i| \le C\sqrt{Mg(M)}$. Then for any $\alpha > 0$, the sample mean $\overline{Z} = \frac{1}{M}\sum_{i=1}^{M} Z_i$ satisfies

$$\mathbb{P}\left(|\overline{Z}| \ge \frac{C_\alpha \sqrt{g(M)} \log(M)}{\sqrt{M}} \right) \le M^{-\alpha}, \tag{203}$$

where C_α depends only on C and α.

For any fixed $t \in [t_n, t_{n+1}]$ we choose $Z_j^t := \frac{M}{N-1} k_1^N(\widetilde{x}_1^t - \widetilde{x}_j^t) - \frac{M}{N-1} \mathbb{E}[k_1^N(\widetilde{x}_1^t - \widetilde{x}_j^t)]$ and $g(M) := CMN^{4\delta-2}$, where $N^{\delta + \frac{\lambda_2}{2} - \frac{9}{4}\lambda_1} < M \le 4C_* N \log^2(N) N^{-3\lambda_1}$. Then following the same argument as in (63), the condition

$$\mathbb{E}[(Z_j^t)^2] \le C \frac{M^2}{(N-1)^2} N^\delta \begin{cases} \|u_{t,t_{n-1}}^{a,N}\|_{\infty,1}, & \text{for } 1 \le n \le M' \\ \|f_t^N\|_{\infty,1}, & \text{for } n = 1 \end{cases} \le C M^2 N^{\delta-2} N^{\frac{9}{2}\lambda_1} \le g(M),$$

is satisfied. We can also deduce that

$$|Z_j^t| \leq C \frac{M}{N-1} N^{2\delta} \leq \sqrt{M(CMN^{4\delta-2})} = \sqrt{Mg(M)}.$$

Applying Lemma 2.5 we obtain at any fixed time $t \in [t_n, t_{n+1}]$

$$\mathbb{P}\left(\left| \frac{1}{N-1} \sum_{j=1}^{M} \left(k_1^N(\widetilde{x}_1^t - \widetilde{x}_j^t) - \mathbb{E}[k_1^N(\widetilde{x}_1^t - \widetilde{x}_j^t)] \right) \right| \geq C_\alpha N^{2\delta-1} \log(N) \right) \leq N^{-\alpha},$$

(204)

and similarly

$$\mathbb{P}\left(\left| \frac{1}{N-1} \sum_{j=1}^{M} \left(k_1^N(\overline{x}_1^t - \overline{x}_j^t) - \mathbb{E}[k_1^N(\overline{x}_1^t - \overline{x}_j^t)] \right) \right| \geq C_\alpha N^{2\delta-1} \log(N) \right) \leq N^{-\alpha}.$$

(205)

It is left to control the difference

$$\left| \frac{1}{N-1} \sum_{j=1}^{M} \left(\mathbb{E}[k_1^N(\widetilde{x}_1^t - \widetilde{x}_j^t)] - \mathbb{E}[k_1^N(\overline{x}_1^t - \overline{x}_j^t)] \right) \right|,$$

where M satisfies (202). This can be done by using Lemma 5.1. For any $t \in [t_n, t_{n+1}]$, when $1 \leq n \leq M'$ we write $a = (\widetilde{X}_{t_{n-1}}, \widetilde{V}_{t_{n-1}}) = (X_{t_{n-1}}, V_{t_{n-1}})$ and $b = (\overline{X}_{t_{n-1}}, \overline{V}_{t_{n-1}})$. Then it follows that

$$\left| \frac{1}{N-1} \sum_{j=1}^{M} \left(\mathbb{E}[k_1^N(\widetilde{x}_1^t - \widetilde{x}_j^t)] - \mathbb{E}[k_1^N(\overline{x}_1^t - \overline{x}_j^t)] \right) \right|$$

$$= \frac{1}{N-1} \left| \sum_{j=1}^{M} \int k_1^N(x_1 - x_j) \left(u_{t,t_{n-1}}^{a,1,N}(x_1, v_1) u_{t,t_{n-1}}^{a,j,N}(x_j, v_j) \right. \right.$$

$$\left. \left. - u_{t,t_{n-1}}^{b,1,N}(x_1, v_1) u_{t,t_{n-1}}^{b,j,N}(x_j, v_j) \right) dx_1 dv_1 dx_j dv_j \right|$$

$$\leq \frac{1}{N-1} \sum_{j=1}^{M} \left| \int k_1^N(x_1 - x_j) u_{t,t_{n-1}}^{a,1,N}(x_1, v_1) \left(u_{t,t_{n-1}}^{a,j,N}(x_j, v_j) - u_{t,t_{n-1}}^{b,j,N}(x_j, v_j) \right) dx_1 dv_1 dx_j dv_j \right|$$

$$+ \frac{1}{N-1} \sum_{j=1}^{M} \left| \int k_1^N(x_1 - x_j) u_{t,t_{n-1}}^{b,j,N}(x_1, v_1) \left(u_{t,t_{n-1}}^{a,1,N}(x_j, v_j) - u_{t,t_{n-1}}^{b,1,N}(x_j, v_j) \right) dx_1 dv_1 dx_j dv_j \right|$$

$$\leq \frac{1}{N-1} \sum_{j=1}^{M} \left(\|u_{t,t_{n-1}}^{a,j,N} - u_{t,t_{n-1}}^{b,j,N}\|_{\infty,1} \|k_1^N * \rho_{t,t_{n-1}}^{a,1,N}\|_1 + \|u_{t,t_{n-1}}^{a,1,N} - u_{t,t_{n-1}}^{b,1,N}\|_{\infty,1} \|k_1^N * \rho_{t,t_{n-1}}^{b,j,N}\|_1 \right)$$

$$\leq \frac{1}{N-1} \sum_{j=1}^{M} C(t-s)^{-6} |a_i - b_i| (\|k_1^N\|_1 \|\rho_{t,t_{n-1}}^{a,1,N}\|_1 + \|k_1^N\|_1 \|\rho_{t,t_{n-1}}^{b,j,N}\|_1)$$

$$\leq C \log^2(N) N^{3\lambda_1} |a_i - b_i| \|k_1^N\|_1 \leq C \log^2(N) N^{3\lambda_1 - \lambda_2} \|k_1^N\|_1,$$

(206)

where $\rho_{t,t_{n-1}}^{a,1,N}(x_1) = \int_{\mathbb{R}^3} u_{t,t_{n-1}}^{a,1,N}(x_1, v_1) dv_1$. Here we have used the fact that when $1 \leq n \leq M'$

$$\|u_{t,t_{n-1}}^{a,j,N} - u_{t,t_{n-1}}^{b,j,N}\|_{\infty,1} \leq C|a_i - b_i|((t-t_{n-1})^{-6} + 1) \leq C|a_i - b_i| N^{6\lambda_1}$$

by Lemma 5.1 since $N^{-\lambda_1} \leq t - t_{n-1} \leq 2N^{-\lambda_1}$. When $n = 1$, since $a = (\widetilde{X}_0, \widetilde{V}_0) = (X_0, V_0) = (\overline{X}_0, \overline{V}_0) = b$, one has

$$\left| \frac{1}{N-1} \sum_{j=1}^{M} \left(\mathbb{E}[k_1^N (\widetilde{x}_1^t - \widetilde{x}_j^t)] - \mathbb{E}[k_1^N (\overline{x}_1^t - \overline{x}_j^t)] \right) \right| = 0.$$

Collecting (204), (205) and (206) we get (193) for M satisfying (202), which finishes the proof of (193) for any M. Hence we conclude (180).

•*Step 4:* Now we prove (176). To see this, we split the summation $\sum_{j \neq 1}^{N}$ into two parts: the part where $j \in M_{t_n}^t$ and the part where $j \in (M_{t_n}^t)^c$

$$\mathbb{P}\left(\left| \frac{1}{N-1} \sum_{j \neq 1}^{N} \left(k_1^N (\widetilde{x}_1^t - \widetilde{x}_j^t) - k_1^N (\overline{x}_1^t - \overline{x}_j^t) \right) \right| \geq 2C_\alpha N^{2\delta - 1} \log(N) \right.$$
$$\left. + 2C_\alpha \log^2(N) N^{3\lambda_1 - \lambda_2} \|k_1^N\|_1 \right) \leq \mathbb{P}\left(\mathcal{X}(M_{t_n}^t) \right) + \mathbb{P}\left(\mathcal{X}((M_{t_n}^t)^c) \right), \qquad (207)$$

where $\mathcal{X}(M_{t_n}^t)$ is defined in (181) and

$$\mathcal{X}((M_{t_n}^t)^c) := \left\{ \left| \frac{1}{N-1} \sum_{j \in (M_{t_n}^t)^c} \left(k_1^N (\widetilde{x}_1^t - \widetilde{x}_j^t) - k_1^N (\overline{x}_1^t - \overline{x}_j^t) \right) \right| \right.$$
$$\left. \geq C_\alpha N^{2\delta - 1} \log(N) + C_\alpha \log^2(N) N^{3\lambda_1 - \lambda_2} \|k_1^N\|_1 \right\}. \qquad (208)$$

For the part in the event $\mathcal{X}((M_{t_n}^t)^c)$ where $j \in (M_{t_n}^t)^c$, it follows from (179) that

$$\mathbb{P}\left(\mathcal{X}((M_{t_n}^t)^c) \right) \leq N^{-\alpha}. \qquad (209)$$

Thus we have

$$\mathbb{P}\left(\left| \frac{1}{N-1} \sum_{j \neq 1}^{N} \left(k_1^N (\widetilde{x}_1^t - \widetilde{x}_j^t) - k_1^N (\overline{x}_1^t - \overline{x}_j^t) \right) \right| \geq 2C_\alpha N^{2\delta - 1} \log(N) \right.$$
$$\left. + 2C_\alpha \log^2(N) N^{3\lambda_1 - \lambda_2} \|k_1^N\|_1 \right) \leq \mathbb{P}\left(\mathcal{X}(M_{t_n}^t) \right) + N^{-\alpha}. \qquad (210)$$

Next we split the summation $\sum_{j \in M_{t_n}^t}$ in the event $\mathcal{X}(M_{t_n}^t)$ (181) into two cases: the case where $\text{card}\,(M_{t_n}^t) \leq M_*$ and the case where $\text{card}\,(M_{t_n}^t) > M_*$. Here M_* is defined in (177).

$$\mathbb{P}\left(\mathcal{X}(M_{t_n}^t) \right) \leq \mathbb{P}\left(\mathcal{X}(M_{t_n}^t) \cap \left\{ \text{card}\,(M_{t_n}^t) \leq M_* \right\} \right) + \mathbb{P}\left(\mathcal{X}(M_{t_n}^t) \cap \left\{ \text{card}\,(M_{t_n}^t) > M_* \right\} \right)$$
$$\leq \mathbb{P}\left(\mathcal{X}(M_{t_n}^t) \cap \left\{ \text{card}\,(M_{t_n}^t) > M_* \right\} \right) + N^{-\alpha}, \qquad (211)$$

where in the last inequality we used (180).

According to (178), for any $t \in [t_n, t_{n+1}]$ one has

$$\mathbb{P}\left(\text{card}\,(M_{t_n}^t) > M_* \right) \leq N^{-\alpha}, \qquad (212)$$

which leads to

$$\mathbb{P}\left(\mathcal{X}(M_{t_n}^t) \cap \left\{ \text{card}\,(M_{t_n}^t) > M_* \right\} \right) \leq \mathbb{P}\left(\text{card}\,(M_{t_n}^t) > M_* \right) \leq N^{-\alpha}. \qquad (213)$$

Therefore it follows from (211) that

$$\mathbb{P}\left(\mathcal{X}(M_{t_n}^t)\right) \leq 2N^{-\alpha}. \tag{214}$$

Together with (210), it implies

$$\mathbb{P}\Bigg(\left|\frac{1}{N-1}\sum_{j\neq 1}^{N}\left(k_1^N(\widetilde{x}_1^t - \widetilde{x}_j^t) - k_1^N(\overline{x}_1^t - \overline{x}_j^t)\right)\right| \geq 2C_\alpha N^{2\delta-1}\log(N)$$

$$+ 2C_\alpha \log^2(N)N^{3\lambda_1-\lambda_2}\|k_1^N\|_1\Bigg) \leq 3N^{-\alpha}. \tag{215}$$

Finally, since the particles are exchangeable, the same result holds for changing $(\widetilde{x}_1^t, \overline{x}_1^t)$ in (215) into $(\widetilde{x}_i^t, \overline{x}_i^t)$, $i = 2, \ldots, N$, which completes the proof of Lemma 3.3. □

Appendix

A Proof of Lemma 5.1

First, let us consider the fundamental solution $G(x, v, t)$ of the equation

$$\partial_t G + v \cdot \nabla_x G = \Delta_v G, \quad G\,|_{t=0} = \delta(x)\delta(v), \tag{216}$$

which can be calculated explicitly as

$$G(x, v, t) = C\frac{1}{t^6}\exp\left(-\frac{|v|^2}{4t} - \frac{3|x - tv/2|^2}{t^3}\right), \tag{217}$$

where C is a normalization constant. The following lemma states some estimates of the fundamental solution.

Lemma A.1 *Let $G(x, v, t)$ be defined in (217) and $p \in [1, \infty]$. There exists a C_p such that for any $j \in \mathbb{N}_0$ the following holds*

$$\left\||x|^j \nabla_v G\right\|_{p,1} \leq C_p t^{\frac{-10p+3jp+9}{2p}}, \quad \left\||x|^j \nabla_x G\right\|_{p,1} \leq C_p t^{\frac{-12p+3jp+9}{2p}}, \tag{218}$$

and

$$\left\|G\left(\cdot - \frac{1}{2}(a_i - b_i)\right) - G\left(\cdot - \frac{1}{2}(b_i - a_i)\right)\right\|_{p,1} \leq C_p|a_i - b_i|\left(t^{\frac{-12p+9}{2p}} + t^{\frac{-10p+9}{2p}}\right), \tag{219}$$

as well as

$$\left\| |\cdot| \left(G\left(\cdot - \frac{1}{2}(a_i - b_i) \right) - G\left(\cdot - \frac{1}{2}(b_i - a_i) \right) \right) \right\|_{p,1} \leq C_p |a_i - b_i| \left(t^{\frac{-7p+9}{2p}} + t^{\frac{-9p+9}{2p}} \right).$$

(220)

The norm $\| \cdot \|_{p,q}$ *denotes the p-norm in the x and q-norm in the v-variable, i.e. for any* $f : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$

$$\|f\|_{p,q} := \left(\int_{\mathbb{R}^3} \left(\int_{\mathbb{R}^3} |f(x,v)|^q dv \right)^{p/q} dx \right)^{1/p}.$$

(221)

Proof It is easy to compute that

$$G = C \frac{1}{t^6} \exp\left(-\frac{3|x|^2}{4t^3} \right) \exp\left(-\frac{|v - \frac{3x}{2t}|^2}{t} \right)$$

(222)

and

$$\nabla_v G = C \frac{1}{t^6} \exp\left(-\frac{3|x|^2}{4t^3} \right) \exp\left(-\frac{|v - \frac{3x}{2t}|^2}{t} \right) \left(-\frac{2v}{t} + \frac{3x}{t^2} \right).$$

(223)

Now we can do the calculation of $\int_{\mathbb{R}^3} |G| dv$ and $\int_{\mathbb{R}^3} |\nabla_v G| dv$:

$$\int_{\mathbb{R}^3} |G| dv = C \frac{1}{t^6} \exp\left(-\frac{3|x|^2}{4t^3} \right) \int_{\mathbb{R}^3} \exp\left(-\frac{|v - \frac{3x}{2t}|^2}{t} \right) dv$$

$$\leq C \frac{1}{t^{9/2}} \exp\left(-\frac{3|x|^2}{4t^3} \right),$$

(224)

and

$$\int_{\mathbb{R}^3} |\nabla_v G| dv = C \frac{1}{t^6} \exp\left(-\frac{3|x|^2}{4t^3} \right) \int_{\mathbb{R}^3} \exp\left(-\frac{|v - \frac{3x}{2t}|^2}{t} \right) \left(-\frac{2v}{t} + \frac{3x}{t^2} \right) dv$$

$$\leq C \frac{1}{t^5} \exp\left(-\frac{3|x|^2}{4t^3} \right) \int_{\mathbb{R}^3} u \exp\left(-u^2 \right) du \leq C \frac{1}{t^5} \exp\left(-\frac{3|x|^2}{4t^3} \right), \quad (225)$$

respectively. As a direct result from (224) and (225), one has

$$\left\| |\cdot|^j G \right\|_{\infty,1} \leq C \frac{1}{t^{(9-3j)/2}} \qquad \left\| |\cdot|^j \nabla_v G \right\|_{\infty,1} \leq C \frac{1}{t^{5-3j/2}}.$$

(226)

For $1 \leq p < \infty$

$$\left\| |\cdot|^j G \right\|_{p,1} \leq C t^{-9/2} \left(\int_{\mathbb{R}^3} |x|^{pj} \exp\left(-\frac{3px^2}{4t^3} \right) dx \right)^{\frac{1}{p}}$$

$$\leq C_p t^{-\frac{9}{2} + \frac{9+3pj}{2p}} \left(\int_{\mathbb{R}^3} |y|^j \exp\left(-y^2 \right) dy \right)^{\frac{1}{p}} \leq C_p t^{\frac{-9p+3jp+9}{2p}}, \quad (227)$$

and

$$\left\| |\cdot|^j \nabla_v G \right\|_{p,1} \leq C t^{-5} \left(\int_{\mathbb{R}^3} |x|^j \exp\left(-\frac{3px^2}{4t^3} \right) dx \right)^{\frac{1}{p}}$$

$$\le C_p t^{-5+\frac{9+3jp}{2p}} \left(\int_{\mathbb{R}^3} |y|^j \exp\left(-y^2\right) dy \right)^{\frac{1}{p}} \le C_p t^{\frac{-10p+3jp+9}{2p}}. \tag{228}$$

We also have

$$\nabla_x G = \frac{1}{t^6} \exp\left(-\frac{3|x|^2}{4t^3}\right) \exp\left(-\frac{|v-\frac{3x}{2t}|^2}{t}\right) \left(-\frac{6x}{t^2} + \frac{3v}{t}\right), \tag{229}$$

which leads to

$$\int_{\mathbb{R}^3} |\nabla_x G| dv = C \frac{1}{t^6} \exp\left(-\frac{3|x|^2}{4t^3}\right) \int_{\mathbb{R}^3} \exp\left(-\frac{|v-\frac{3x}{2t}|^2}{t}\right) \left(-\frac{6x}{t^3} + \frac{3v}{t^2}\right) dv$$

$$\le C \frac{1}{t^8} \exp\left(-\frac{3|x|^2}{4t^3}\right) \int_{\mathbb{R}^3} \exp\left(-\frac{|v-\frac{3x}{2t}|^2}{t}\right) \left(-\frac{2x}{t} + v\right) dv$$

$$\le C t^{-8+\frac{3}{2}} \exp\left(-\frac{3|x|^2}{4t^3}\right) \int_{\mathbb{R}^3} (\sqrt{t}u - \frac{x}{2t}) \exp\left(-u^2\right) du$$

$$\le C \frac{1}{t^6} \exp\left(-\frac{3|x|^2}{4t^3}\right) + C \frac{1}{t^6} \frac{x}{t^{\frac{3}{2}}} \exp\left(-\frac{3|x|^2}{4t^3}\right). \tag{230}$$

It follows from the above that

$$\left\| |\cdot|^j \nabla_x G \right\|_{\infty,1} \le C \frac{1}{t^{6-3j/2}}. \tag{231}$$

For $1 \le p < \infty$

$$\left\| |\cdot|^j \nabla_x G \right\|_{p,1}$$

$$\le C \frac{1}{t^6} \left(\left(\int_{\mathbb{R}^3} |x|^j \exp\left(-\frac{3px^2}{4t^3}\right) dx \right)^{\frac{1}{p}} + \left(\int_{\mathbb{R}^3} |x|^j \exp\left(-\frac{3p|x|^2}{4t^3}\right) \left(\frac{x}{t^{\frac{3}{2}}}\right)^p dx \right)^{\frac{1}{p}} \right)$$

$$\le C_p t^{\frac{-12p+3jp+9}{2p}} \left(\left(\int_{\mathbb{R}^3} |y|^j \exp\left(-y^2\right) dy \right)^{\frac{1}{p}} + \left(\int_{\mathbb{R}^3} |y|^j \exp\left(-py^2\right) |y|^p dy \right)^{\frac{1}{p}} \right)$$

$$\le C_p t^{\frac{-12p+3jp+9}{2p}}, \tag{232}$$

which concludes the proof of (218).

As a direct result of (218) we can prove (219). Indeed,

$$\left| G\left(\cdot -\frac{1}{2}(a_i - b_i)\right) - G\left(\cdot -\frac{1}{2}(b_i - a_i)\right) \right|$$

$$\le |a_i - b_i| \int_0^1 \left| \nabla G\left(\cdot -\frac{1}{2}(b_i - a_i) + s(b_i - a_i)\right) \right| ds$$

$$\le |a_i - b_i| \int_0^1 \left| \nabla_v G\left(\cdot -\frac{1}{2}(b_i - a_i) + s(b_i - a_i)\right) \right| ds$$

$$+ |a_i - b_i| \int_0^1 \left| \nabla_x G\left(\cdot -\frac{1}{2}(b_i - a_i) + s(b_i - a_i)\right) \right| ds, \tag{233}$$

which leads to

$$\left\| \left(G\left(\cdot -\frac{1}{2}(a_i - b_i)\right) - G\left(\cdot -\frac{1}{2}(b_i - a_i)\right) \right) \right\|_{p,1}$$

$$\leq C|a_i - b_i| \left(\|\nabla_v G\|_{p,1} + \|\nabla_x G\|_{p,1} \right) \leq C_p |a_i - b_i| \left(t^{\frac{-12p+9}{2p}} + t^{\frac{-10p+9}{2p}} \right). \quad (234)$$

Next we prove (220):

$$|\cdot| \left(G \left(\cdot - \frac{1}{2}(a_i - b_i) \right) - G \left(\cdot - \frac{1}{2}(b_i - a_i) \right) \right)$$

$$\leq \left(|\cdot - \frac{1}{2}(a_i - b_i)| G \left(\cdot - \frac{1}{2}(a_i - b_i) \right) - |\cdot - \frac{1}{2}(b_i - a_i)| G \left(\cdot - \frac{1}{2}(b_i - a_i) \right) \right)$$

$$+ \frac{1}{2}|a_i - b_i| \left(G \left(\cdot - \frac{1}{2}(a_i - b_i) \right) + G \left(\cdot - \frac{1}{2}(b_i - a_i) \right) \right). \quad (235)$$

In view of (227), the $(p, 1)$-norm of the terms in the last line have the right bound. With the other term we proceed as above, using the function $H = |\cdot| G$:

$$\left(|\cdot - \frac{1}{2}(a_i - b_i)| G \left(\cdot - \frac{1}{2}(a_i - b_i) \right) - |\cdot - \frac{1}{2}(b_i - a_i)| G \left(\cdot - \frac{1}{2}(b_i - a_i) \right) \right)$$

$$\leq |a_i - b_i| \int_0^1 \left| \nabla H \left(\cdot - \frac{1}{2}(b_i - a_i) + s(b_i - a_i) \right) \right| ds$$

$$\leq |a_i - b_i| \int_0^1 \left| \nabla_v H \left(\cdot - \frac{1}{2}(b_i - a_i) + s(b_i - a_i) \right) \right| ds$$

$$+ |a_i - b_i| \int_0^1 \left| \nabla_x H \left(\cdot - \frac{1}{2}(b_i - a_i) + s(b_i - a_i) \right) \right| ds. \quad (236)$$

It follows from our estimates in (218) that

$$\left\| |\cdot - \frac{1}{2}(a_i - b_i)| G \left(\cdot - \frac{1}{2}(a_i - b_i) \right) - |\cdot - \frac{1}{2}(b_i - a_i)| G \left(\cdot - \frac{1}{2}(b_i - a_i) \right) \right\|_{p,1} \quad (237)$$

$$\leq C|a_i - b_i| \left(\| |\cdot| \nabla_v G\|_{p,1} + \| |\cdot| \nabla_x G\|_{p,1} + \|G\|_{p,1} \right)$$

$$\leq C_p |a_i - b_i| \left(t^{\frac{-7p+9}{2p}} + t^{\frac{-9p+9}{2p}} \right), \quad (238)$$

which leads to (220). $\qquad \square$

Proof of Lemma 5.1 The proof of the estimates follows the ideas of [10, Lemma 2]. However, the evolution equation for the present system is more difficult to handle, and in particular, the spacial overlap is suppressed for short periods of time since we have a noise term in the momentum variable only. Both estimates can be proved in the same way. We just give the proof for the more difficult part (ii), which can be easily adapted for part (i). Without loss of generality we set $s = 0$ and $t < 1$. What we need to show then is

$$\|u_{t,s}^{a,i,N} - u_{t,s}^{b,i,N}\|_{\infty,1} \leq C|a_i - b_i| \left((t - s)^{-6} + 1 \right)$$

holds for all $i = 1, \ldots, N$.

Note that the force $\bar{k}_t^N(x) := k^N * \rho_t^N$ we consider is globally Lipschitz and L^∞ because of (47), thus there exists a $C > 0$ independent of N such that

$$\max_{0 \leq t \leq T; x, y \in \mathbb{R}^3} \frac{|\bar{k}_t(x) - \bar{k}_t(y)|}{|x - y|} \leq C. \quad (239)$$

Let c_t be the trajectory on phase space following the Newtonian equations of motion with respect to the force $\bar{k}_t^N$, starting with $\frac{1}{2}(a_i + b_i)$ at time 0, i.e.

$$c_t = (x_t^c, v_t^c), \qquad \frac{d}{dt}x_t^c = v_t^c, \qquad \frac{d}{dt}v_t^c = \bar{k}_t^N(x_t^c), \qquad c_0 = \frac{1}{2}(a_i + b_i).$$

We use the trajectory c to change the frame of inertia that we use to look at $u_{t,s}^{d,i,N}$ for $d \in \{a, b\}$, i.e. we define for any $t > 0$ the density $w_{t,0}^{a,i,N}$ on phase space by

$$w_{t,0}^{a,i,N}((x, v)) := u_{t,0}^{a,i,N}((x, v) + c_t). \tag{240}$$

From the evolution equation of $u_{t,s}^{d,i,N}$ for $d \in \{a, b\}$ and c_t one gets directly

$$\frac{\partial}{\partial t}w_{t,0}^{d,i,N}(x, v) := \Delta_v w_{t,0}^{d,i,N}(x, v)$$
$$- \nabla_x w_{t,0}^{d,i,N} \cdot v - \nabla_v w_{t,0}^{d,i,N} \cdot \left(\bar{k}_t^N(x + x_t^c) - \bar{k}_t^N(x_t^c)\right), \tag{241}$$

with $w_{0,0}^{a,i,N} = \delta\left(\cdot - \left(\frac{1}{2}(a_i - b_i)\right)\right)$ and $w_{0,0}^{b,i,N} = \delta\left(\cdot - \left(\frac{1}{2}(b_i - a_i)\right)\right)$.

Since w is built from u by translation we have for any $1 \le p \le \infty$

$$\|u_{t,0}^{a,i,N} - u_{t,0}^{b,i,N}\|_{p,1} = \|w_{t,0}^{a,i,N} - w_{t,0}^{b,i,N}\|_{p,1}. \tag{242}$$

Before proceeding we would like to explain the advantage of looking at w instead of u first on a heuristic level. The difficulties arise when dealing with short periods of time. There the $u^d, d \in \{a, b\}$ are roughly given by a Gaussian around the center at $\frac{1}{2}(a+b)$, respectively the w^d are roughly given by a Gaussian around the center at 0. Here the force term of w – which is zero at $x = 0$ – suppresses the last term of (241). Thus w will be very close to the heat-kernel G_t of our time evolution.

Using (241) and the properties of the heat kernel we get

$$w_{t,0}^{a,i,N} = G_t * \delta\left(\cdot - \left(\frac{1}{2}(a_i - b_i)\right)\right) - \int_0^t G_{t-s} * \left(\nabla_v w_{s,0}^{a,i,N} \cdot \left(\bar{k}_s(\cdot + x_s^c) - \bar{k}_s(x_s^c)\right)\right) ds$$
$$= G_t\left(\cdot - \frac{1}{2}(a_i - b_i)\right) - \int_0^t \nabla_v G_{t-s} * \left(w_{s,0}^{a,i,N}\left(\bar{k}_s(\cdot + x_s^c) - \bar{k}_s(x_s^c)\right)\right) ds, \tag{243}$$

and

$$w_{t,0}^{b,i,N} = G_t\left(\cdot - \frac{1}{2}(b_i - a_i)\right) - \int_0^t \nabla_v G_{t-s} * \left(w_{s,0}^{b,i,N}\left(\bar{k}_s(\cdot + x_s^c) - \bar{k}_s(x_s^c)\right)\right) ds,$$

thus

$$w_{t,0}^{a,i,N} - w_{t,0}^{b,i,N} = \left(G_t\left(\cdot - \frac{1}{2}(a_i - b_i)\right) - G_t\left(\cdot - \frac{1}{2}(b_i - a_i)\right)\right)$$
$$- \int_0^t \nabla_v G_{t-s} * \left(\left(w_{s,0}^{a,i,N} - w_{s,0}^{b,i,N}\right)\left(\bar{k}_s(\cdot + x_s^c) - \bar{k}_s(x_s^c)\right)\right) ds. \tag{244}$$

Defining $\eta_{t,0}^N : \mathbb{R}^6 \to \mathbb{R}_0^+$ by $\eta_{t,0}^N(x, v) := |(x, v)| \left|w_{t,0}^{a,i,N} - w_{t,0}^{b,i,N}\right|$ and using (239), we can find a constant C such that

$$\eta_{t,0}^N \le |\cdot| \left|G_t\left(\cdot - \frac{1}{2}(a_i - b_i)\right) - G_t\left(\cdot - \frac{1}{2}(b_i - a_i)\right)\right|$$

$$+ C \left| \int_0^t \nabla_v G_{t-s} * \eta_{s,0}^N ds \right|. \tag{245}$$

Using the properties of the heat kernel (218), (220) and Young's inequality in (245), we get

$$\left\| \eta_{t,0}^N \right\|_{1,1} \leq C|a_i - b_i| + C \int_0^t (t-s)^{-\frac{1}{2}} \left\| \eta_{s,0}^N \right\|_{1,1} ds. \tag{246}$$

Applying a generalized Gronwall's inequality with weak singularities [27, Lemma 7.1.1] leads to

$$\|\eta_{t,0}^N\|_{1,1} \leq C|a_i - b_i| \qquad \text{uniform in } t \in [0, T]. \tag{247}$$

Further (245) gives for any $1 \leq p \leq \infty$ and $t \in [0, T]$

$$\left\| \eta_{t,0}^N \right\|_{p,1} \leq \left\| |\cdot| G_t \left(\cdot - \frac{1}{2} (a_i - b_i) \right) - G_t \left(\cdot - \frac{1}{2} (b_i - a_i) \right) \right\|_{p,1}$$
$$+ C \int_0^{t/2} \left\| \nabla_v G_{t-s} * \eta_{s,0}^N \right\|_{p,1} ds + C \int_{t/2}^t \left\| \nabla_v G_{t-s} * \eta_{s,0}^N \right\|_{p,1} ds. \tag{248}$$

Using Young's inequality we get for $1 + p^{-1} = \frac{9}{10} + q^{-1}$ and $t \in [0, T]$,

$$\left\| \eta_{t,0}^N \right\|_{p,1} \leq C|a_i - b_i| t^{\frac{-9p+9}{2p}} + C \int_0^{t/2} \left\| \nabla_v G_{t-s} \right\|_{p,1} \left\| \eta_{s,0}^N \right\|_{1,1} ds$$
$$+ C \int_{t/2}^t \left\| \nabla_v G_{t-s} \right\|_{10/9,1} \left\| \eta_{s,0}^N \right\|_{q,1} ds.$$

Due to (218), one has $\|\nabla_v G_{t-s}\|_{10/9,1} \leq C(t-s)^{-19/20}$. This and (247) give

$$\left\| \eta_{t,0}^N \right\|_{p,1} \leq C|a_i - b_i| t^{\frac{-9p+9}{2p}} + C|a_i - b_i| \int_0^{t/2} \left\| \nabla_v G_{t-s} \right\|_{p,1} ds + C \max_{t/2 \leq s \leq t} \left\| \eta_{s,0}^N \right\|_{q,1}. \tag{249}$$

We use this formula starting at $p_1 = 1$ and setting $p_{k+1} = \frac{10 p_k}{10 - p_k}$. Therefore, starting with our estimate for $\left\| \eta_{t,0}^{a,i,N} \right\|_{1,1}$ (see (247)) we can then iteratively estimate the L^p norms of $\eta_{t,0}^N$ for higher exponents, i.e.

$$\left\| \eta_{t,0}^N \right\|_{p_{k+1},1} \leq C|a_i - b_i| t^{\frac{-9p_{k+1}+9}{2p_{k+1}}} + C|a_i - b_i| \int_0^{t/2} \left\| \nabla_v G_{t-s} \right\|_{p_{k+1},1} ds + C \max_{t/2 \leq s \leq t} \left\| \eta_{s,0}^N \right\|_{p_k,1}. \tag{250}$$

The exponent $p_{k+1} = \infty$ is attained after $k = 10$ steps. It follows that

$$\left\| \eta_{t,0}^N \right\|_{\infty,1} \leq C|a_i - b_i| (t^{\frac{-9}{2}} + 1). \tag{251}$$

Having good control of $\|\eta_{t,0}^N\|_{\infty,1}$ we can now estimate $w_{t,0}^{a,i,N} - w_{t,0}^{b,i,N}$ using (244):

$$\left\| w_{t,0}^{a,i,N} - w_{t,0}^{b,i,N} \right\|_{\infty,1} \leq \left\| G_t \left(\cdot - \frac{1}{2} (a_i - b_i) \right) - G_t \left(\cdot - \frac{1}{2} (b_i - a_i) \right) \right\|_{\infty,1}$$
$$+ \int_0^t \left\| \nabla_v G_{t-s} * \left(\left(w_{s,0}^{a,i,N} - w_{s,0}^{b,i,N} \right) \left(\bar{k}_s \left(\cdot + x_s^c \right) - \bar{k}_s \left(x_s^c \right) \right) \right) \right\|_{\infty,1} ds$$

$$\leq C|a_i - b_i|t^{-6} + C\int_0^{t/2} \|\nabla_v G_{t-s}\|_{\infty,1} \left\|\eta_{s,0}^N\right\|_{1,1} ds$$

$$+ \int_{t/2}^t \|\nabla_v G_{t-s}\|_{1,1} \left\|\eta_{s,0}^N\right\|_{\infty,1} ds$$

$$\leq C|a_i - b_i|t^{-6} + C|a_i - b_i|\int_0^{t/2} (t-s)^{-5} ds + C|a_i - b_i|\int_{t/2}^t (t-s)^{-\frac{1}{2}}(s^{-\frac{9}{2}} + 1) ds$$

$$\leq C|a_i - b_i|\left(t^{-6} + t^{-4} + t^{\frac{1}{2}}\right) \leq C|a_i - b_i|(t^{-6} + 1). \tag{252}$$

With (242) statement (ii) of the lemma follows. $\qquad\square$

References

1. Aarseth, S.J.: Gravitational N-body simulations: tools and algorithms. Cambridge University Press, Cambridge (2003)
2. Beale, J.T., Majda, A.: Convergence in three dimensions. Vortex methods. I. Math. Comput. **39**(159), 1–27 (1982)
3. Beale, J.T., Majda, A.: Higher order accuracy in two and and three dimensions. Vortex methods. II. Math. Comput. **39**(159), 29–52 (1982)
4. Boers, N., Pickl, P.: On mean field limits for dynamical systems. J. Stat. Phys. **164**(1), 1–16 (2016)
5. Bolley, F., Canizo, J.A., Carrillo, J.A.: Stochastic mean-field limit: non-Lipschitz forces and swarming. Math. Models Methods Appl. Sci. **21**(11), 2179–2210 (2011)
6. Bossy, M., Faugeras, O., Talay, D.: Clarification and complement to mean-field description and propagation of chaos in networks of Hodgkin-Huxley and Fitzhugh–Nagumo neurons. J. Math. Neurosci. (JMN) **5**(1), 19 (2015)
7. Bouchut, F.: Existence and uniqueness of a global smooth solution for the Vlasov–Poisson–Fokker–Planck system in three dimensions. J. Funct. Anal. **111**(1), 239–258 (1993)
8. Bouchut, F.: Smoothing effect for the non-linear Vlasov–Poisson–Fokker–Planck system. J. Differ. Equ. **122**(2), 225–238 (1995)
9. Braun, W., Hepp, K.: The Vlasov dynamics and its fluctuations in the $\frac{1}{N}$ limit of interacting classical particles. Commun. Math. Phys. **56**(2), 101–113 (1977)
10. Cañizares-García, A., Pickl, P.: Microscopic derivation of the Keller–Segel equation in the sub-critical regime. arXiv preprint arXiv:1703.04376 (2017)
11. Carpio, A.: Long-time behaviour for solutions of the Vlasov—Poisson-Fokker–Planck equation. Math. Methods Appl. Sci. **21**(11), 985–1014 (1998)
12. Carrillo, J.A., Choi, Y.-P., Salem, S.: Propagation of chaos for the Vlasov—Poisson-Fokker–Planck equation with a polynomial cut-off. Commun. Contemp. Math. **21**(04), 1850039 (2019)
13. Carrillo, J.A., Fornasier, M., Toscani, G., Vecil, F.: Particle, kinetic, and hydrodynamic models of swarming. Math. Model. Collect. Behav. Socio-econ. Life Sci. **105**, 297–336 (2010)
14. Chorin, A.J.: Numerical study of slightly viscous flow. J. Fluid Mech. **57**(04), 785–796 (1973)
15. Degond, P.: Global existence of smooth solutions for the Vlasov–Poisson–Fokker–Planck equation in 1 and 2 space dimensions. Annales scientifiques de l'École Normale Supérieure **19**, 519–542 (1986)
16. Dobrushin, R.L.: Vlasov equations. Funct. Anal. Appl. **13**(2), 115–123 (1979)
17. Duerinckx, M.: Mean-field limits for some Riesz interaction gradient flows. SIAM J. Math. Anal. **48**(3), 2269–2300 (2016)
18. Fetecau, R.C., Huang, H., Sun, W.: Propagation of chaos for the Keller–Segel equation over bounded domains. J. Differ. Equ. **266**(4), 2142–2174 (2019)
19. Fournier, N., Hauray, M., Mischler, S.: Propagation of chaos for the 2d viscous vortex model. J. Eur. Math. Soc. **16**(7), 1423–1466 (2014)
20. Freedman, D.: Brownian Motion and Diffusion. Springer Science & Business Media, New York (1983)
21. Girsanov, I.V.: Strongly-feller processes. Theory Probab. Appl. I. General Propert. **5**(1), 5–24 (1959)
22. Goodman, J.: Convergence of the random vortex method. Commun. Pure Appl. Math. **40**(2), 189–220 (1987)
23. Graß, P.: Microscopic derivation of Vlasov equations with singular potentials. PhD thesis, lmu (2019)
24. Hald, O., Del Prete, V.M.: Convergence of vortex methods for Euler's equations. Math. Comput. **32**(143), 791–809 (1978)

25. Hald, O.H.: Convergence of vortex methods for Euler's equations. II. SIAM J. Numer. Anal. **16**(5), 726–755 (1979)
26. Havlak, K.J., Victory Jr., H.D.: The numerical analysis of random particle methods applied to Vlasov–Poisson–Fokker–Planck kinetic equations. SIAM J. Numer. Anal. **33**(1), 291–317 (1996)
27. Henry, D.: Geometric Theory of Semilinear Parabolic Equations, vol. 840. Springer, New York (2006)
28. Huang, H., Liu, J.-G.: Discrete-in-time random particle blob method for the Keller–Segel equation and convergence analysis. Commun. Math. Sci. **15**(7), 1821–1842 (2017)
29. Huang, H., Liu, J.-G.: Error estimate of a random particle blob method for the Keller–Segel equation. Math. Comput. **86**, 2719–2744 (2017)
30. Huang, H., Liu, J.-G., Lu, J.: Learning interacting particle systems: diffusion parameter estimation for aggregation equations. Math. Models Methods Appl. Sci. **29**(01), 1–29 (2019)
31. Jabin, P.-E.: A review of the mean field limits for Vlasov equations. Kinetic Relat Models **7**(4), 661–711 (2014)
32. P.-E. Jabin and M. Hauray. Particles approximations of Vlasov equations with singular forces: Propagation of chaos. In Annales Scientifiques de l'École Normale Supérieure (2015)
33. Jabin, P.-E., Otto, F.: Identification of the dilute regime in particle sedimentation. Commun. Math. Phys. **250**(2), 415–432 (2004)
34. Jabin, P.-E., Wang, Z.: Mean field limit and propagation of chaos for Vlasov systems with bounded forces. J. Funct. Anal. **271**(12), 3588–3627 (2016)
35. Jabin, P.-E., Wang, Z.: Mean field limit for stochastic particle systems. In Active Particles, Volume 1, pp. 379–402. Springer (2017)
36. Jabin, P.-E., Wang, Z.: Quantitative estimates of propagation of chaos for stochastic systems with $w^{-1,\infty}$ kernels. Invent. Math. **214**(1), 523–591 (2018)
37. Jeans, J.: On the theory of star-streaming and the structure of the universe. Mon. Not. R. Astronom. Soc. **76**, 70–84 (1915)
38. Keller, E.F., Segel, L.A.: Initiation of slime mold aggregation viewed as an instability. J. Theor. Biol. **26**(3), 399–415 (1970)
39. Lazarovici, D., Pickl, P.: A mean field limit for the Vlasov–Poisson system. Archive Ration. Mech. Anal. **225**, 1201–1231 (2017)
40. Lévy, P.: Sur certains processus stochastiques homogénes. Compos. Math. **7**, 283–339 (1940)
41. Li, L., Liu, J.-G., Yu, P.: On mean field limit for Brownian particles with Coulomb interaction in 3D. J. Math. Phys. **60**, 11150 (2019)
42. Lions, P.-L., Perthame, B.: Propagation of moments and regularity for the 3-dimensional Vlasov–Poisson system. Invent.Math. **105**(1), 415–430 (1991)
43. Liu, J.-G., Yang, R.: Propagation of chaos for large Brownian particle system with Coulomb interaction. Res. Math. Sci. **3**(1), 40 (2016)
44. Liu, J.G., Zhang, Y.: Convergence of diffusion-drift many particle systems in probability under Sobolev norm. In: Gonçalves, P., Soares, A.J. (eds.) From Particle Systems to Partial Differential Equations-III. Springer, New York (2016)
45. Loeper, G.: Uniqueness of the solution to the Vlasov–Poisson system with bounded density. Journal de Mathématiques Pures et Appliquées **86**(1), 68–79 (2006)
46. Long, D.-G.: Convergence of the random vortex method in two dimensions. J. Am. Math. Soc. **1**(4), 779–804 (1988)
47. Marchioro, C., Pulvirenti, M.: Vortex methods in two-dimensional fluid dynamics. Lect. Notes Phys. **203**, 1–137 (1984)
48. Motsch, S., Tadmor, E.: A new model for self-organized dynamics and its flocking behavior. J. Stat. Phys. **144**(5), 923 (2011)
49. Olla, S., Varadhan, S.: Scaling limit for interacting Ornstein–Uhlenbeck processes. Commun. Math. Phys. **135**(2), 355–378 (1991)
50. Ono, K., Strauss, W.A.: Regular solutions of the Vlasov–Poisson–Fokker–Planck system. Discret. Contin. Dyn. Syst. **6**(4), 751–772 (2000)
51. Osada, H.: Propagation of chaos for the two dimensional Navier–Stokes equation. Series A Mathematical sciences. In: Proceedings of the Japan Academy. **62**(1), 8–11 (1986)
52. Patlak, C.S.: Random walk with persistence and external bias. Bull. Math. Biophys. **15**(3), 311–338 (1953)
53. Pulvirenti, M., Simeoni, C.: l^∞-estimates for the Vlasov–Poisson–Fokker–Planck equation. Math. Methods Appl. Sci. **23**(10), 923–935 (2000)
54. Serfaty, S., Duerinckx, M.: Mean field limit for Coulomb flows. arXiv preprint arXiv:1803.08345 (2018)
55. Soler, J., Carrillo, J.A., Bonilla, L.L.: Asymptotic behavior of an initial-boundary value problem for the Vlasov–Poisson–Fokker–Planck system. SIAM J. Appl. Math. **57**(5), 1343–1372 (1997)

56. Spohn, H.: Dynamics of Charged Particles and Their Radiation Field. Cambridge University Press, Cambridge (2004)
57. Touboul, J., et al.: Propagation of chaos in neural fields. Ann. Appl. Prob. **24**(3), 1298–1328 (2014)
58. Tremoulet, C.: Hydrodynamic limit for interacting Ornstein–Uhlenbeck particles. Stoch. Process. Appl. **102**(1), 139–158 (2002)
59. Victory, H.D., O'Dwyer, B.P.: On classical solutions of Vlasov–Poisson Fokker–Planck systems. Indiana Univ. Math. J. **39**(1), 105–155 (1990)
60. Villani, C.: Optimal Transport: Old and New, vol. 338. Springer Science & Business Media, New York (2008)
61. Vlasov, A.: The vibrational properties of an electron gas
62. Wiener, N.: Differential space. J. Math. Phys. **2**, 131–174 (1923)
63. Yang, R., Liu, J.-G.: Propagation of chaos for the Keller–Segel equation with a logarithmic cut-off. Methods Appl. Anal. **26**(4), 319–348 (2019)