On the mean-field limit for the Vlasov-Poisson-Fokker-Planck system

Hui Huang*, Jian-Guo Liu†, Peter Pickl‡§

April 20, 2018

Abstract

We devise and study a random particle blob method for approximating the Vlasov-Poisson-Fokkker-Planck (VPFP) equations by a N-particle system subject to the Brownian motion in $\mathbb{R}^3$ space. More precisely, we show that maximal distance between the exact microscopic and the mean-field trajectories is bounded by $N^{-\frac{1}{3}+\varepsilon}$ ($\frac{1}{63} \leq \varepsilon < \frac{1}{36}$) for a system with blob size $N^{-\delta}$ ($\frac{1}{3} \leq \delta < \frac{19}{54} - \frac{2\varepsilon}{3}$) up to a probability $1 - N^{-\alpha}$ for any $\alpha > 0$, which improves the cut-off in [10]. Our result thus leads to a derivation of VPFP equations from the microscopic N-particle system. In particular we prove the convergence rate between the empirical measure associated to the particle system and the solution of the VPFP equations. The technical novelty of this paper is that our estimates crucially rely on the randomness coming from the initial data *and* from the Brownian motion.

Keywords: Coupling method, propagation of chaos, concentration inequality, Wasserstein metric.

1 Introduction

The research to be carried out in this manuscript is a microscopic derivation of the Vlasov-Poisson-Fokker-Planck (VPFP) system. The VPFP system is the kinetic description of the Brownian motion of a large system of particles in a surrounding bath. For example, in the mathematical model for an electrostatic plasma, when the collisions between the electron distribution and a surrounding medium are taken into account, the time evolution of the electron distribution function $f(x, v, t): (x, v, t) \in \mathbb{R}^3 \times \mathbb{R}^3 \times [0, \infty) \to \mathbb{R}^+$ satisfies the following VPFP equations

$$\begin{cases} \partial_t f(x,v,t) + v \cdot \nabla_x f(x,v,t) + k * \rho(x,t) \cdot \nabla_v f(x,v,t) = \sigma \Delta_v f(x,v,t), \\ f(x,v,0) = f_0(x,v), \end{cases} \tag{1}$$

where k is the Coulomb kernel

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

for some real number a and

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

is the charge density introduced by the distribution $f(x, v, t)$. The case $a > 0$ corresponds to the electrostatic (repulsive) interaction of charged particles in a plasma, while $a < 0$ describes the attraction between massive particles subject to gravitation. We denote by $E(x, t) := k * \rho(x, t)$ the Coulombic or gravitational force field.

The VPFP equations are based on particle-like description of a set of large individuals, or we can call it individual-based model. Denote by $x_i \in \mathbb{R}^3$ and $v_i \in \mathbb{R}^3$ the position and verlocity of

*Department of Mathematics, Simon Fraser University, Burnaby, BC, Canada. Email: hha101@sfu.ca
†Departments of Physics and Mathematics, Duke University, Durham, NC, USA. Email: jliu@phy.duke.edu
‡Mathematisches Institut, Universität München, München, Germany. Email: pickl@math.lmu.de
§Duke Kunshan University, Kunshan, Jiangsu, China. Email: peter.pickl@duke.edu

particle i. The evolution of the system is given by the following stochastic differential equations (SDEs),

$$dx_i = v_i dt, \quad dv_i = \frac{1}{N-1}\sum_{j\neq i}^{N} k(x_i - x_j)dt + \sqrt{2\sigma}dB_i, \quad i = 1, \cdots, N, \tag{4}$$

where $k(x)$ models the pairwise interaction between the individuals and B_i are independent realizations of Brownian motions which count for extrinsic random perturbations.

The analysis of the scaling limits of interacting particle systems is usually called the *mean-field limit*, which passes the limit from the microscopic discrete particle system to the macroscopic continuum model. When $\sigma = 0$, namely there is no randomness coming from the noise, classical results for globally Lipschitz forces was obtained by Braun and Hepp [7] and Dobrushin [16]. Then Bolley, Cañizo and Carrilo [4] presented an extension of the classical theory to the particle system with only locally Lipschitz interacting force. 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 [29] 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 perform the mean-field limit for typical initial data and the cut-off that can be chosen to be $N^{-\frac{1}{2d}}$. For $\alpha < 1$, they prove molecular chaos without cut-off. Unfortunately, their method fails precisely at the Coulomb threshold when $\alpha = d-1$. More recently, Boers and Pickl [3] 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 the cut-off as small as $N^{-\frac{1}{d}}$. Furthermore, Lazarovici and Pickl [33] extended the method in [3] to include the Coulomb singularity and they obtained a microscopic derivation of the Vlasov-Poisson equations with a cut-off $N^{-\delta}$ $(0 < \delta < \frac{1}{d})$. When $\sigma > 0$, Jabin and Wang [30] rigorously justified the mean-filed limit and propagation of chaos for VPFP system with bounded forces by using a relative entropy method. Most recently, Carrilo *et.al.* [10] investigated the VPFP system with the singular force and obtained the propagation of chaos through a cut-off $N^{-\delta}$ $(0 < \delta < \frac{1}{d})$. For a general overview of this topic we refer readers to [11, 28, 31, 45].

When the interacting kernel k is singular, it poses problems for both the 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 one hand, the larger 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 to the actual system. On the other hand, the smaller 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 significant. The analysis we review above tries to justify that. In this manuscript, we set $\sigma > 0$. The main technical innovation of this paper is that we fully use the randomness coming from the initial conditions *and* the Brownian motion to significantly improve the cut-off: we can assume it to be smaller than $N^{-\frac{1}{d}}$ (see Remark 1.2), which is a sort of average minimal distance between N particles in dimension d. This manuscript significantly improves the ideas presented in [8]. 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 we in addition use 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 based on the randomness coming from the initial condition. This is carried out in Lemma 3.2, the proof of which can be found in section 5. [8] 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 (4), 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

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

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

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

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

This leads to the 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*. [21, 22], Beale and Majda [1, 2]. 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 equation has been given by [20, 38, 39, 41] in 1980s. For a more recent result we refer to [19]. Another well-known application of the system (5) is to choose the interaction to be the Poisson kernel

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

Now the system (5) coincides with the particle models to approximate the classical Keller-Segel (KS) equation for chemotaxis [32, 42]. We refer mainly to [8, 17, 25, 26, 27, 36] for the mean-field limit of the KS system. Many techniques used in this manuscript are adapted from these papers. For the Poisson-Nernst-Planck equation (k is set to be repulsion), [35] proved the propagation of chaos without cut-off.

Introduction in the following will be split into three parts: We start with introducing the microscopic random particle system underlining VPFP equations in Section 1.1. Then we present the result of the existence to macroscopic mean-field VPFP equations in Section 1.2. Last, our main theorem will be stated in Section 1.3, where we obtain the degree of the approximation of solutions to VPFP equations by random evolving 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 manuscript is described by a trajectory on phase space, i.e. a time dependent $\Phi_t : \mathbb{R} \to \mathbb{R}^{6N}$. We use the notation

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

where x_j^t stands for the position of the j^{th} particle at time t, v_j^t for the momentum of the j^{th} particle at time t. The system is a Newtonian system with a noise term coupled to the momentum, whose evolution is governed by a system of SDEs of the type

$$\begin{cases} dx_i^t = v_i^t dt, \\ dv_i^t = \dfrac{1}{N-1} \displaystyle\sum_{j \neq i}^{N} k(x_i^t - x_j^t) dt + \sqrt{2\sigma} dB_i^t , \end{cases} \tag{9}$$

where k is the Coulomb kernel (2) modeling interaction between particles and B_i^t are independent realization of Brownian motions. All masses will be set to one, thus we do not distinguish between momentum and velocity.

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

$$k^N(x) = k * \psi_\delta^N. \tag{10}$$

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

$$\begin{cases} dx_i^t = v_i^t dt, \\ dv_i^t = \dfrac{1}{N-1} \displaystyle\sum_{i \neq j}^{N} k^N(x_i^t - x_j^t) dt + \sqrt{2\sigma} dB_i^t , \end{cases} \tag{11}$$

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

$$\begin{cases} \partial_t f^N(x,v,t) + v\cdot\nabla_x f^N(x,v,t) + k^N * \rho^N(x,t)\cdot\nabla_v f^N(x,v,t) = \sigma\Delta_v f^N(x,v,t), \\ \rho^N(x,t) = \int_{\mathbb{R}^3} f^N(x,v,t)dv, \\ f^N(x,v,0) = f_0(x,v). \end{cases} \tag{12}$$

1.2 Existence of the Vlasov-Poisson-Fokker-Planck system

The existence of weak and classical solutions to VPFP equations (1) and related systems have 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 electrostatic case. Later on, Bouchut [5, 6] extended the result to three dimensions when the electric field was coupled through a Poisson equation, and the results were given in both electrostatic and gravitational cases. Also, Victory and O'Dwyer [46] derived the classical solutions for VPFP equations when spacial dimension is less or equal to two, but the local existence for all other dimensions. Then, Carrillo and Soler [12] proved the global existence of weak solutions for the VPFP equations in three dimensions with an $L^1\cap L^p$ initial data. Furthermore, they [13] considered the distribution of particles to be the measures with special decay contained in Morrey spaces and the existence of weak solutions, local and global in time solutions for small initial distribution of particles were obtained. The long time behavior of the VPFP system was studied by Ono and Strauss [40], Carpio [9] and Carrillo *et al.* [44].

For our purposes, we summarize the following existence result from Bouchut [5], where his proof relied on the techniques introduced by Lions and Perthame [34] concerning regularity for the Vlasov-Poisson system (the VPFP system without diffusion term, namely $\sigma = 0$).

Theorem 1.1. *(Bouchut) Given $f_0 \geq 0$, $f_0 \in L^1\cap L^\infty(\mathbb{R}^6)$ satisfying*

a) there exists a $m_0 > 6$, such that

$$\iint_{\mathbb{R}^6} (|x|^{m_0} + |v|^{m_0}) f_0(x,v)dxdv < +\infty; \tag{13}$$

b) f_0 has a compact support in the v-variable, namely $f_0(x,v) = 0$ when $|v| > Q_v$.

Then VPFP equations (1) has a unique continuous, bounded solution satisfying

$$f(\cdot,\cdot,t) \in C(\mathbb{R}^+; L^p(\mathbb{R}^6)) \cap L^\infty(\mathbb{R}^+; L^1\cap L^\infty(\mathbb{R}^6)), \quad \text{for } 1 \leq p < +\infty, \tag{14}$$

and for all $T < +\infty$, $m < m_0$

$$\sup_{t\in[0,T]} \iint_{\mathbb{R}^6} (|x|^m + |v|^m) f(x,v,t)dxdv < +\infty, \tag{15}$$

$$\sup_{t\in[0,T]} \|\rho(\cdot,t)\|_\infty < +\infty, \quad \sup_{t\in[0,T]} \|E(\cdot,t)\|_\infty < +\infty, \tag{16}$$

where $\rho(x,t)$ is the charge density and $E(x,t) = k * \rho(x,t)$ is the force field.

Remark 1.1. The uniform-in-time L^∞- bound of the charge density ρ was obtained in [43] by means of the stochastic characteristic method under the assumption the f_0 is compactly supported in velocity. And [10] provided a proof of the local-in-time L^∞ bound for ρ by employing Feynman-Kac's formula assuming the initial data is polynomial decay.

Notice that Theorem 1.1 shows in particular the force field $E \in L^\infty([0,T]\times\mathbb{R}^3)$ for $T > 0$. By a classical result [24], this implies that initial smooth data remain smooth for all times. In this manuscript, we assume the initial data f_0 satisfying the following assumption:

Assumption 1.1. *The initial density $0 \leq f_0(x) \in W^{1,1}\cap W^{1,\infty}(\mathbb{R}^6)$ satisfies*

1. *there exists a* $m_0 > 6$*, such that* $(|x|^{m_0} + |v|^{m_0})f_0 \in W^{1,1}(\mathbb{R}^6)$.

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

The above assumption makes sure we have the regularity needed for this article:

$$f(\cdot,\cdot,t) \in C(\mathbb{R}^+; W^{1,p}(\mathbb{R}^6)) \cap L^\infty(\mathbb{R}^+; W^{1,1} \cap W^{1,\infty}(\mathbb{R}^6)), \quad \text{for } 1 \le p < +\infty, \tag{17}$$

and for all $T < +\infty$, $m < m_0$

$$\begin{aligned} \sup_{t\in[0,T]} \|\rho(\cdot,t)\|_{W^{1,\infty}(\mathbb{R}^3)} &+ \sup_{t\in[0,T]} \|f(\cdot,\cdot,t)\|_{W^{1,1}\cap W^{1,\infty}(\mathbb{R}^6)} \\ &+ \sup_{t\in[0,T]} \|(|\cdot|^m + |\cdot|^m)f(\cdot,\cdot,t)\|_{W^{1,1}(\mathbb{R}^6)} < C_{f_0}, \end{aligned} \tag{18}$$

where C_{f_0} depends only on $\|f_0\|_{W^{1,1}\cap W^{1,\infty}(\mathbb{R}^6)}$, $\|(|\cdot|^{m_0} + |\cdot|^{m_0})f_0\|_{W^{1,1}(\mathbb{R}^6)}$ and Q_v. Notice that one can obtain the bound estimate for f^N or ρ^N uniformly in N equivalently.

1.3 The statement of main results

Our objective is to derive the macroscopic mean-field PDE (1) from the microscopic particle system (11). This we will do by using probabilistic methods as in [8, 26, 25, 33]. More precisely, we shall prove the convergence rate between the solution of VPFP equations (1) and the empirical measure associated to the particle system Φ_t satisfying (11). We assume that the initial condition Φ_0 of the system is independently, identically distributed (i.i.d.) with common probability density given by f_0.

Given the solution f^N to the mean-field equation (12), we first construct an auxiliary trajectory Ψ_t from (12). Then we prove the closeness between Φ_t and Ψ_t. For the auxiliary trajectory $\Psi_t := (\overline{X}_t, \overline{V}_t) = (\overline{x}_1^t, \dots, \overline{x}_N^t, \overline{v}_1^t, \dots \overline{v}_N^t)$, 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 $k^N * \rho^N(x,t)$

$$\begin{cases} d\overline{x}_i^t = \overline{v}_i^t dt, \\ d\overline{v}_i^t = \displaystyle\int_{\mathbb{R}^3} k^N(\overline{x}_i^t - x)\rho^N(x,t)dxdt + \sqrt{2\sigma}dB_i^t, \end{cases} \tag{19}$$

here we let Ψ_t has the same initial condition as Φ_t (i.i.d. with common density f_0). Since the particles are subject to an external field, 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 (12) is Kolmogorov's forward equation for any solution of (19), and in particular their probability distribution f^N solves (12). 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 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 $\sup_{t\in[0,T]} \|\Phi_t - \Psi_t\|_\infty$ on $[0,T]$ exceeds $N^{-\lambda_2}$ decreases exponentially with the number of particles N, as N grows to infinity. Here the distance $\|\Phi_t - \Psi_t\|_\infty$ is measured by

$$\|\Phi_t - \Psi_t\|_\infty := \sqrt{\log(N)}\|X_t - \overline{X}_t\|_\infty + \|V_t - \overline{V}_t\|_\infty. \tag{20}$$

Theorem 1.2. *Assume that trajectories* $\Phi_t = (X_t, V_t)$*,* $\Psi_t = (\overline{X}_t, \overline{V}_t)$ *satisfy* (11) *and* (19) *respectively with the initial data* $\Phi_0 = \Psi_0$*, which is i.i.d. sharing the common density* f_0 *satisfying 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* $N_0 > 0$ *depending only on* α*,* T *and* C_{f_0}*, such that for* $N \ge N_0$*, the following estimate holds with the cut-off index* $\delta \in \left[\frac{1}{3}, \min\left\{\frac{\lambda_1+3\lambda_2+1}{6}, \frac{1-\lambda_2}{2}\right\}\right)$

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

where $\|\Phi_t - \Psi_t\|_\infty$ *is defined in* (20).

Remark 1.2. In particular, for any $\frac{1}{63} \le \varepsilon < \frac{1}{36}$, 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 the cut-off size $N^{-\delta}$ $(\frac{1}{3} \le \delta < \frac{19}{54} - \frac{2\varepsilon}{3})$. In other words, the cut-off parameter δ can be chosen very close to $\frac{19}{54}$ in particular larger than $\frac{1}{3}$, which is a significant improvement to previous results in the literature.

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 space

$$\mathcal{P}_p(\mathbb{R}^d) = \{\mu | \ \mu \text{ is a probability measure on } \mathbb{R}^d \text{ and } \int_{\mathbb{R}^d} |x|^p d\mu(x) < +\infty\}. \tag{21}$$

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

$$\mathcal{W}_p^p(\mu,\nu) = \inf_{\pi \in \Lambda(\mu,\ \nu)} \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^p d\pi(x,y) \right\} = \inf_{X\sim\mu, Y\sim\nu} \Big\{ \mathbb{E}[|X-Y|^p] \Big\}, \tag{22}$$

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 notational 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 $\mathcal{W}_p^p(\rho_1, \rho_2)$. For further details, we refer reader to the book of Villani [47].

Following the same argument as [33, Corollary 4.3], Theorem 1.2 implies molecular chaos in the following sense:

Corollary 1.1. *Let $F_0^N := \otimes^N f_0$ and F_t^N be the N-particle distribution evolving with the microscopic flow (11) 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 (1) with $f_t|_{t=0} = f_0$. More precisely, under the assumptions of Theorem 1.2, Then 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 $N \ge N_0$, the following estimate holds

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

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 (1) from the microscopic random particle system (11). We define the empirical measure associated to the microscopic N-particle systems (11) and (19) respectively as

$$\mu_\Phi(t) := \frac{1}{N}\sum_{i=1}^N \delta(x - x_i^t)\delta(v - v_i^t), \quad \mu_\Psi(t) := \frac{1}{N}\sum_{i=1}^N \delta(x - \overline{x}_i^t)\delta(v - \overline{v}_i^t). \tag{23}$$

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 (1) 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 VPFP equations (1) with the initial data satisfying Assumption 1.1 and $\mu_\Phi(t)$ be the empirical measure defined in (23) with Φ_t being the particle flow solving (11). Let $p \in [1,\infty)$ and assume that there exists $m > 2p$ such that $\iint_{\mathbb{R}^6} (|x|^m + |v|^m) f_0(x,v)dxdv < +\infty$. 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 $N \ge e^{\left(\frac{C_1}{1-3\lambda_2}\right)^2}$ it holds that*

$$\begin{aligned} \mathbb{P}\bigg(\sup_{t\in[0,T]} W_p(f_t, \mu_\Phi(t)) \le N^{-\kappa+1-3\lambda_2} + N^{-\lambda_2} \bigg) \\ \ge 1 - C_2\left(e^{-C_3 N^{1-\max\{6,2p\}\kappa}} + N^{1-\frac{m}{2p}} \right). \end{aligned} \tag{24}$$

where δ and λ_2 are used in Theorem 1.2.

This theorem provides a derivation of VPFP equations from a interacting N-particle system, bridging the gap between a microscopic description in terms of agent based model 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 main theorems. Throughout this manuscript, the generic constants will be denoted generically by C, even if it is different from line to line. We use $\|\cdot\|_p$ for the L^p $(1 \le p \le \infty)$ norm of a function. Moreover if $v = (v_1, \cdots, v_N)$ is a vector, then $\|v\|_\infty := \sup_{i=1,\cdots,N} |v_i|$.

2.1 Local Lipschitz bound

First let us recall some estimates of the regularized kernel k^N defined in (10):

Lemma 2.1. *([25, Lemma 2.1])*

(i) $k^N(0) = 0$, $k^N(x) = k(x)$ *for any* $|x| \ge N^{-\delta}$ *and* $|k^N(x)| \le |k(x)|$*;*

(ii) $|\partial^\beta k^N(x)| \le CN^{(2+|\beta|)\delta}$, *for any* $x \in \mathbb{R}^3$*;*

(iii) $\|k^N\|_2 \le CN^{\frac{\delta}{2}}$.

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

Definition 2.1. Let

$$\ell^N(x) = \begin{cases} \dfrac{6^3}{|x|^3}, & \text{if } |x| \ge 6N^{-\delta}, \\ N^{3\delta}, & \text{else }, \end{cases} \tag{25}$$

and $L^N : \mathbb{R}^{6N} \to \mathbb{R}^N$ be defined by $(L^N(X_t))_i := \frac{1}{N-1} \sum_{i \ne j} \ell^N(x_i^t - x_j^t)$. 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* $|x - y| \le N^{-\lambda_2} \gg N^{-\delta}$ $(\lambda_2 < \delta)$ *the following holds:*

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

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

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

$$\frac{|\nabla k^N(x)|}{\ell^N(y)} \le \frac{N^{3\delta}}{\ell^N(2N^{-\lambda_2})} = CN^{3(\delta-\lambda_2)}, \tag{26}$$

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

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

$$\frac{|\nabla k^N(x)|}{\ell^N(y)} \le \frac{C|x|^{-3}}{|y|^{-3}} \le C\frac{(|y| - N^{-\lambda_2})^{-3}}{|y|^{-3}} \le C, \tag{27}$$

where we used in the last step that $|x| \ge (|y| - N^{-\lambda_2}) \ge \frac{|y|}{2}$ for $|y| \ge 2N^{-\lambda_2}$. Collecting (26) and (27) finishes the proof. $\square$

Denote

$$(K^N(\overline{X}_t))_i := \frac{1}{N-1}\sum_{j\neq i} k^N(\overline{x}_i^t - \overline{x}_j^t), \tag{28}$$

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

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

$$\left\|K^N(X_t) - K^N(\overline{X}_t)\right\|_\infty \le C\|L^N(\overline{X}_t)\|_\infty \|X_t - \overline{X}_t\|_\infty, \tag{29}$$

for some $C > 0$ independent of N.

Proof. For any $\xi \in \mathbb{R}^3$ with $|\xi| < 4N^{-\delta}$, it follows from [33, Lemma 6.3] that

$$|k^N(x+\xi) - k^N(x)| \le C\ell^N(x)|\xi|, \tag{30}$$

where $\ell^N(q)$ is defined in (25). Therefore

$$\begin{aligned}
|(K^N(X_t))_i - (K^N(\overline{X}_t))_i| &\le \frac{1}{N-1}\sum_{j\neq i}^N |k^N(x_i^t - x_j^t) - k^N(\overline{x}_i^t - \overline{x}_j^t)| \\
&\le \frac{1}{N-1}\sum_{j\neq i}^N C\ell^N(\overline{x}_i^t - \overline{x}_j^t)|x_i^t - x_j^t - \overline{x}_i^t + \overline{x}_j^t| \\
&\le C(L^N(\overline{X}_t))_i \|X_t - \overline{X}_t\|_\infty \le C\|L^N(\overline{X}_t)\|_\infty \|X_t - \overline{X}_t\|_\infty,
\end{aligned} \tag{31}$$

which leads to (29). ∎

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*

$$\|\ell^N * \rho\|_\infty \le C\log(N)(\|\rho\|_1 + \|\rho\|_\infty), \quad \|(\ell^N)^2 * \rho\|_\infty \le CN^{(3\delta)}(\|\rho\|_1 + \|\rho\|_\infty); \tag{32}$$

and

$$\|k^N * \rho\|_\infty \le C(\|\rho\|_1 + \|\rho\|_\infty), \quad \|\nabla k^N * \rho\|_\infty \le C(\|\nabla\rho\|_1 + \|\nabla\rho\|_\infty). \tag{33}$$

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

$$\begin{aligned}
&\|\ell^N * \rho\|_\infty = \left\|\int_{\mathbb{R}^3} \ell^N(x-y)\rho(y)dy\right\|_\infty \\
&\le \left\|\int_{|x-y|<6N^{-\delta}} \ell^N(x-y)\rho(y)dy\right\|_\infty + \left\|\int_{6N^{-\delta}\le|x-y|\le 1} \ell^N(x-y)\rho(y)dy\right\|_\infty \\
&\quad + \left\|\int_{1\le|x-y|} \ell^N(x-y)\rho(y)dy\right\|_\infty \\
&\le \|\rho\|_\infty N^{3\delta}|B(6N^{-\delta})| + \|\rho\|_{L^\infty}\int_{6N^{-\delta}\le|x-y|\le 1}\frac{C}{|y|^3}dy + C\|\rho\|_1 \\
&\le C\log(N)(\|\rho\|_\infty + \|\rho\|_1),
\end{aligned} \tag{34}$$

where $B(r)$ denotes the ball with radius r in $\mathbb{R}^3$, which concludes the proof. ∎

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

$$\mathbb{P}\left(|\bar{Z}| \ge \frac{C_\alpha \sqrt{g(N)} \log(N)}{\sqrt{N}}\right) \le N^{-\alpha}, \tag{35}$$

where C_α depends only on C and α.

The proof can be seen in [20, Lemma 1], which is a direct result of the Taylor expansion and the Markov's inequality.

Denote

$$(\overline{K}^N(\overline{X}_t))_i := \int_{\mathbb{R}^3} k^N(\overline{x}_i^t - x)\rho^N(x,t)dx, \tag{36}$$

then we can introduce the following lemma of 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 (19), K^N and $\overline{K}^N$ are defined in (28) and (36) respectively, L^N and $\overline{L}^N$ are showed in Definition 2.1. For any $\alpha > 0$ and $\frac{1}{3} \le \delta < 1$, there exist a constant $C_{1,\alpha} > 0$ depending only on α, T and C_{f_0} such that*

$$\mathbb{P}\left(\left\|K^N(\overline{X}_t) - \overline{K}^N(\overline{X}_t)\right\|_\infty \ge C_{1,\alpha} N^{2\delta-1}\log(N)\right) \le N^{-\alpha}, \tag{37}$$

and

$$\mathbb{P}\left(\left\|L^N(\overline{X}_t) - \overline{L}^N(\overline{X}_t)\right\|_\infty \ge C_{1,\alpha} N^{3\delta-1}\log(N)\right) \le N^{-\alpha}. \tag{38}$$

Proof. We can prove this lemma by using Lemma 2.5. Due to the exchangeability of the particles, we are ready to bound

$$(K^N(\overline{X}_t))_1 - (\overline{K}^N(\overline{X}_t))_1 = \frac{1}{N-1}\sum_{j=2}^N k^N(\overline{x}_1^t - \overline{x}_j^t) - \int_{\mathbb{R}^3} k^N(\overline{x}_1^t - x)\rho^N(x,t)dx = \frac{1}{N-1}\sum_{j=2}^N Z_j, \tag{39}$$

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 $j \ne 1$ 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

$$\begin{aligned}\mathbb{E}'\left[k^N(\overline{x}_1^t - \overline{x}_j^t)\right] &= \iint_{\mathbb{R}^6} k^N(\overline{x}_1^t - x) f^N(x,v,t)dxdv \\ &= \int_{\mathbb{R}^3} k^N(\overline{x}_1^t - x)\rho^N(x,t)dx.\end{aligned} \tag{40}$$

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

$$\mathbb{E}'\left[|Z_j|^2\right] = \mathbb{E}'\left[\left|k^N(\overline{x}_1^t - \overline{x}_j^t) - \int_{\mathbb{R}^3} k^N(\overline{x}_1^t - x)\rho^N(x,t)dx\right|^2\right]. \tag{41}$$

Since it follows from Lemma 2.4 that

$$\int_{\mathbb{R}^3} k^N(\overline{x}_1^t - x)\rho^N(x,t)dx \le C(\|\rho^N\|_1 + \|\rho^N\|_\infty), \tag{42}$$

it suffices to bound

$$\mathbb{E}'\left[k^N(\overline{x}_1^t - \overline{x}_j^t)\right] = \int_{\mathbb{R}^3} k^N(\overline{x}_1^t - x)\rho^N(x,t)dx \le C(\|\rho^N\|_1 + \|\rho^N\|_\infty) \le C(T, C_{f_0}), \tag{43}$$

and

$$\mathbb{E}'\left[k^N(\overline{x}_1^t-\overline{x}_j^t)^2\right]=\int_{\mathbb{R}^3}k^N(\overline{x}_1^t-x)^2\rho^N(x,t)dx\le\|\rho^N\|_\infty\|k^N\|_2^2\le C(T,C_{f_0})N^{\delta},\tag{44}$$

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

$$\mathbb{E}'\left[|Z_j|^2\right]\le CN^{\delta}.\tag{45}$$

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 $|Z_j|\le CN^{2\delta}\le C\sqrt{Ng(N)}$. Hence, using Lemma 2.5, we have the probability bound

$$\mathbb{P}\left(\left|(K^N(\overline{X}_t))_1-(\overline{K}^N(\overline{X}_t))_1\right|\ge C(\alpha,T,C_{f_0})N^{2\delta-1}\log(N)\right)\le N^{-\alpha}.\tag{46}$$

Similarly, the same bound must also apply hold to other term with $i=2,\cdots,N$, which leads to

$$\mathbb{P}\left(\left\|K^N(\overline{X}_t)-\overline{K}^N(\overline{X}_t)\right\|_\infty\ge C(\alpha,T,C_{f_0})N^{2\delta-1}\log(N)\right)\le N^{1-\alpha}.\tag{47}$$

Let $C_{1,\alpha}$ be the constant in (47), we conclude (37).

To prove (38), we follow the same procedure above

$$(L^N(\overline{X}_t))_1-(\overline{L}^N(\overline{X}_t))_1=\frac{1}{N-1}\sum_{j=2}^N\ell^N(\overline{x}_1^t-\overline{x}_j^t)-\int_{\mathbb{R}^3}\ell^N(\overline{x}_1^t-x)\rho^N(x,t)dx=\frac{1}{N-1}\sum_{j=2}^N Z_j,\tag{48}$$

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

$$\mathbb{E}'\left[\ell^N(\overline{x}_1^t-\overline{x}_j^t)\right]=\int_{\mathbb{R}^3}\ell^N(\overline{x}_1^t-x)\rho^N(x,t)dx\le C\log(N)(\|\rho\|_1+\|\rho\|_\infty)\le C(T,C_{f_0})\log(N),\tag{49}$$

and

$$\mathbb{E}'\left[\ell^N(\overline{x}_1^t-\overline{x}_j^t)^2\right]=\int_{\mathbb{R}^3}\ell^N(\overline{x}_1^t-x)^2\rho^N(x,t)dx\le CN^{3\delta}(\|\rho\|_1+\|\rho\|_\infty)\le C(T,C_{f_0})N^{3\delta},\tag{50}$$

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

$$\mathbb{E}'\left[|Z_j|^2\right]\le CN^{3\delta}.\tag{51}$$

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 $|Z_j|\le CN^{3\delta}\le C\sqrt{Ng(N)}$. Hence, we have the probability bound

$$\mathbb{P}\left(\left|(L^N(\overline{X}_t))_1-(\overline{L}^N(\overline{X}_t))_1\right|\ge C(\alpha,T,C_{f_0})N^{3\delta-1}\log(N)\right)\le N^{-\alpha},\tag{52}$$

by Lemma 2.5, which leads to

$$\mathbb{P}\left(\left\|L^N(\overline{X}_t)-\overline{L}^N(\overline{X}_t)\right\|_\infty\ge C(\alpha,T,C_{f_0})N^{3\delta-1}\log(N)\right)\le N^{1-\alpha}.\tag{53}$$

Thus, (38) follows from (53).

□

3 Proof of Theorem 1.2

We do the proof by following the idea in [25, 26], which is that consistency and stability imply convergence.

3.1 Consistency

In order to obtain the consistency error in all time, we divide $[0,T]$ into $M+1$ sunintervals with length $\Delta t = N^{-\frac{\gamma}{3}}$ for some $\gamma > 3$ and $t_n = n\Delta t$, $n = 0, \cdots, M$. The choice of γ will be clear from the discussion below. Here the choice of Δt is only for the purpose of proving consistency and it is different from the one in the proof of stability in the next subsection.

First, we summarize the following lemma by using only the randomness coming from the Brownian motion.

Lemma 3.1. *Assume that $(\overline{X}_t, \overline{V}_t)$ satisfies the mean-field dynamics (19). There exists some $\gamma > 3$, such that it holds*

$$\mathbb{P}\left(\sup_n \sup_{t\in[t_n,t_{n+1}]} \|\overline{X}_t - \overline{X}_{t_n}\|_\infty \geq C_B N^{-\frac{\gamma-1}{3}}\right) \leq C_B N^{\frac{2+\gamma}{3}} \exp(-C_B N^{\frac{1}{3}}), \tag{54}$$

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

Proof. Notice that

$$\begin{aligned}\overline{X}_t - \overline{X}_{t_n} &= \int_{t_n}^t \overline{V}_s ds = \int_{t_n}^t \int_{t_n}^s \overline{K}^N(\overline{X}_\tau) d\tau ds + \sqrt{2\sigma} \int_{t_n}^t (B(s) - B(t_n)) ds + \int_{t_n}^t \overline{V}_{t_n} ds, \\ &=: I_1^n(t) + I_2^n(t) + I_3^n(t),\end{aligned} \tag{55}$$

where

$$\overline{V}_{t_n} = V_0 + \int_0^{t_n} \overline{K}^N(\overline{X}_s) ds + \sqrt{2\sigma} B(t_n). \tag{56}$$

The estimation of $I_1^n(t)$ follows from Lemma 2.4

$$\int_{t_n}^t \int_{t_n}^s \overline{K}^N(\overline{X}_\tau) d\tau ds \leq (\Delta t)^2 \|\overline{K}^N\|_\infty \leq C N^{-\frac{2\gamma}{3}}. \tag{57}$$

So we have

$$\sup_n \sup_{t\in[t_n,t_{n+1}]} \|I_1^n(t)\|_\infty \leq C N^{-\frac{2\gamma}{3}}. \tag{58}$$

To estimate $I_2^n(t)$, recall a basic property of the Brownian motion [26, Lemma 2.7]:

$$\mathbb{P}\left(\sup_{t\leq s\leq t+\Delta t} \|B(s) - B(t)\|_\infty \geq b\right) \leq C_1(\sqrt{\Delta t}/b) \exp(-C_2 b^2/\Delta t), \tag{59}$$

which leads to

$$\mathbb{P}\left(\sup_{t\in[t_n,t_{n+1}]} \|B(t) - B(t_n)\|_\infty \geq N^{-\frac{1}{3}}\right) \leq C_1 N^{-\frac{\gamma-2}{6}} \exp(-C_2 N^{\frac{\gamma-2}{3}}). \tag{60}$$

Since $\sup_{t\in[t_n,t_{n+1}]} \|I_2^n(t)\|_\infty \leq \Delta t\sqrt{2\sigma} \sup_{t\in[t_n,t_{n+1}]} \|B(t) - B(t_n)\|_\infty$, it follows from (60) that

$$\mathbb{P}\left(\sup_{t\in[t_n,t_{n+1}]} \|I_2^n(t)\|_\infty \geq C N^{-\frac{\gamma+1}{3}}\right) \leq C_1 N^{-\frac{\gamma-2}{6}} \exp(-C_2 N^{\frac{\gamma-2}{3}}), \tag{61}$$

which leads to

$$\mathbb{P}\left(\sup_n \sup_{t\in[t_n,t_{n+1}]} \|I_2^n(t)\|_\infty \geq C N^{-\frac{\gamma+1}{3}}\right) \leq C_1 N^{\frac{\gamma+2}{6}} \exp(-C_2 N^{\frac{\gamma-2}{3}}). \tag{62}$$

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

$$\int_0^{t_n} \overline{K}^N(\overline{X}_s) ds \leq \Delta t \|\overline{K}^N\|_\infty \leq C N^{-\frac{\gamma}{3}}, \tag{63}$$

and it follows from (59) that

$$\mathbb{P}(\|B(t_n)\|_\infty \geq N^{\frac{1}{3}}) \leq C_1 N^{-\frac{1}{3}}\sqrt{T}\exp(-C_2 N^{\frac{2}{3}}/T). \tag{64}$$

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

$$f_0^v(v) = \int_{\mathbb{R}^3} f_0(x,v)dx = 0, \text{ when } |v| > Q_v \tag{65}$$

one has

$$\mathbb{P}(\|V_0\|_\infty \geq N^{\frac{1}{3}}) = \int_{|v|\geq N^{\frac{1}{3}}} f_0^v(v)dv = 0, \text{ when } N > Q_v^3. \tag{66}$$

Recall that

$$\sup_{t\in[t_n,t_{n+1}]} \|I_3^n(t)\|_\infty = \int_{t_n}^t \|\overline{V}_{t_n}\|_\infty ds \leq N^{-\frac{\gamma}{3}}(\|V_0\|_\infty + \sqrt{2\sigma}\|B(t_n)\|_\infty) + CN^{-\frac{2\gamma}{3}}, \tag{67}$$

then it yields

$$\begin{aligned}
&\mathbb{P}\left(\sup_{t\in[t_n,t_{n+1}]} \|I_3^n(t)\|_\infty \geq 3N^{-\frac{\gamma-1}{3}}\right) \\
\leq & \mathbb{P}(\|V_0\|_\infty \geq N^{\frac{1}{3}}) + \mathbb{P}(\|B(t_n)\|_\infty \geq N^{\frac{1}{3}}) \\
\leq & 0 + CN^{\frac{2}{3}}\exp(-CN^{\frac{1}{3}}) \leq CN^{\frac{2}{3}}\exp(-CN^{\frac{1}{3}}),
\end{aligned} \tag{68}$$

which leads to

$$\mathbb{P}\left(\sup_n \sup_{t\in[t_n,t_{n+1}]} \|I_3^n(t)\|_\infty \geq 3N^{-\frac{\gamma-1}{3}}\right) \leq CN^{\frac{2+\gamma}{3}}\exp(-CN^{\frac{1}{3}}), \tag{69}$$

Then it follows from (58), (62) and (69) that

$$\begin{aligned}
&\mathbb{P}\left(\sup_n \sup_{t\in[t_n,t_{n+1}]} \left\|\overline{X}_t - \overline{X}_{t_n}\right\|_\infty \geq CN^{-\frac{\gamma-1}{3}}\right) \\
\leq & C_1 N^{\frac{\gamma+2}{6}}\exp(-C_2 N^{\frac{\gamma-2}{3}}) + CN^{\frac{2+\gamma}{3}}\exp(-CN^{\frac{1}{3}}) \leq CN^{\frac{2+\gamma}{3}}\exp(-CN^{\frac{1}{3}}), (\gamma > 3)
\end{aligned}$$

which completes the proof. □

Now we can prove the consistency error in all time.

Proposition 3.1. *(Consistency) Let $(\overline{X}_t, \overline{V}_t)$ satisfying the mean-field dynamics (19) with initial density $f_0(x,v)$, K^N and $\overline{K}^N$ be defined in (28) and (36) respectively. For any $\alpha > 0$ and $\frac{1}{3} \leq \delta < 1$, there exist a constant $C_{2,\alpha} > 0$ depending only on depends on α, T and C_{f_0} such that*

$$\mathbb{P}\left(\sup_{t\in[0,T]} \left\|K^N(\overline{X}_t) - \overline{K}^N(\overline{X}_t)\right\|_\infty \geq C_{2,\alpha} N^{2\delta-1}\log(N)\right) \leq N^{-\alpha}, \tag{70}$$

and

$$\mathbb{P}\left(\sup_{t\in[0,T]} \left\|L^N(\overline{X}_t) - \overline{L}^N(\overline{X}_t)\right\|_\infty \geq C_{2,\alpha} N^{3\delta-1}\log(N)\right) \leq N^{-\alpha}. \tag{71}$$

Proof. Denote events:

$$\mathcal{H} := \left\{\sup_n \sup_{t\in[t_n,t_{n+1}]} \left\|\overline{X}_t - \overline{X}_{t_n}\right\|_\infty \leq C_B N^{-\frac{\gamma-1}{3}}\right\}, \tag{72}$$

and

$$\mathcal{C}_{t_n} := \left\{ \left\| K^N(\overline{X}_{t_n}) - \overline{K}^N(\overline{X}_{t_n}) \right\|_\infty \le C_{1,\alpha} N^{2\delta-1} \log(N) \right\}, \tag{73}$$

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

$$\mathbb{P}(\mathcal{C}_{t_n}^c) \le N^{-\alpha}, \quad \mathbb{P}(\mathcal{H}^c) \le C_B N^{\frac{2+\gamma}{3}} \exp(-C_B N^{\frac{1}{3}}), \tag{74}$$

for any $\alpha > 0$.

Furthermore, we denote

$$\mathcal{B}_{t_n} := \left\{ \left\| L^N(\overline{X}_{t_n}) - \overline{L}^N(\overline{X}_{t_n}) \right\|_\infty \le C_{1,\alpha} N^{3\delta-1} \log(N) \right\}, \tag{75}$$

then under the event $\mathcal{B}_{t_n}$, it holds that

$$\|L^N(\overline{X}_{t_n})\|_\infty \le \|\overline{L}^N(\overline{X}_{t_n})\|_\infty + C_{1,\alpha} N^{3\delta-1} \log(N) \le C(\alpha, T, C_{f_0}) N^{3\delta-1} \log(N). \tag{76}$$

and $\mathbb{P}(\mathcal{B}_{t_n}^c) \le N^{-\alpha}$ by Lemma 2.6.

For all $t \in [t_n, t_{n+1}]$, under the event $\mathcal{B}_{t_n} \cap \mathcal{C}_{t_n} \cap \mathcal{H}$, we obtain

$$\begin{aligned}
&\left\| K^N(\overline{X}_t) - \overline{K}^N(\overline{X}_t) \right\|_\infty \\
\le& \left\| K^N(\overline{X}_t) - K^N(\overline{X}_{t_n}) \right\|_\infty + \left\| K^N(\overline{X}_{t_n}) - \overline{K}^N(\overline{X}_{t_n}) \right\|_\infty + \left\| \overline{K}^N(\overline{X}_{t_n}) - \overline{K}^N(\overline{X}_t) \right\|_\infty \\
\le& C \|L^N(\overline{X}_{t_n})\|_\infty \|\overline{X}_t - \overline{X}_{t_n}\|_\infty + C_{1,\alpha} N^{2\delta-1} \log(N) + C \log(N) \left\| \overline{X}_t - \overline{X}_{t_n} \right\|_\infty \\
\le& C(\alpha, T, C_{f_0}) N^{3\delta-1} \log(N) N^{-\frac{\gamma-1}{3}} + C_{1,\alpha} N^{2\delta-1} \log(N) \\
\le& C(\alpha, T, C_{f_0}) N^{2\delta-1} \log(N), \quad (3\delta + 1 < \gamma)
\end{aligned} \tag{77}$$

where in the second inequality we have used the local Lipschitz bound of K^N

$$\left\| K^N(X_t) - K^N(\overline{X}_{t_n}) \right\|_\infty \le C \|L^N(\overline{X}_{t_n})\|_\infty \|X_t - \overline{X}_{t_n}\|_\infty, \tag{78}$$

under the event $\mathcal{H}$ (see in Lemma 2.3). It yields that

$$\sup_{t \in [0,T]} \left\| K^N(\overline{X}_t) - \overline{K}^N(\overline{X}_t) \right\|_\infty \le C(\alpha, T, C_{f_0}) N^{2\delta-1} \log(N), \tag{79}$$

holds under the event $\bigcap_{n=0}^{M} \mathcal{B}_{t_n} \cap \mathcal{C}_{t_n} \cap \mathcal{H}$. Therefore

$$\begin{aligned}
&\mathbb{P}\left(\sup_{t \in [0,T]} \left\| K^N(\overline{X}_t) - \overline{K}^N(\overline{X}_t) \right\|_\infty \ge C(\alpha, T, C_{f_0}) N^{2\delta-1} \log(N) \right) \\
\le& \sum_{n=0}^{M} P(\mathcal{B}_{t_n}^c) + \sum_{n=0}^{M} P(\mathcal{C}_{t_n}^c) + P(\mathcal{H}^c) \\
\le& T N^{-\frac{3\alpha-\gamma}{3}} + T N^{-\frac{3\alpha-\gamma}{3}} + C_B N^{\frac{2+\gamma}{3}} \exp(-C_B N^{\frac{1}{3}}) \le N^{-\alpha'}.
\end{aligned} \tag{80}$$

Denote $C_{2,\alpha'}$ to be the constant $C(\alpha, T, C_{f_0})$ in (80). Since $\alpha > 0$ is arbitrary and so is α', hence (70) holds true. The proof of (71) can be done similarly. $\square$

3.2 Stability

In this subsection we obtain the stability result under certain conditions.

Proposition 3.2. *(Stability) Assume that trajectories $\Phi_t = (X_t, V_t)$, $\Psi_t = (\overline{X}_t, \overline{V}_t)$ satisfy (11) and (19) respectively with the initial data $\Phi_0 = \Psi_0$ which is i.i.d. sharing the common density*

f_0 satisfying Assumption 1.1, K^N is showed in (28). For any $0<\lambda_2<\frac{1}{3}$, $0<\lambda_1<\frac{\lambda_2}{3}$ and $\frac{1}{3}\le\delta<1$, we denote events:

$$\mathcal{A} := \left\{ \sup_{t\in[0,T]} \sqrt{\log(N)}\left\|X_t-\overline{X}_t\right\|_\infty + \left\|V_t-\overline{V}_t\right\|_\infty \le N^{-\lambda_2} \right\}. \tag{81}$$

and

$$\begin{aligned}\mathcal{S}(\Lambda) := \Big\{ &\|K^N(X_t)-K^N(\overline{X}_t)\|_\infty \le \Lambda\log(N)\left\|X_t-\overline{X}_t\right\|_\infty \\ &+\Lambda\log(N)(N^{6\delta-1-\lambda_1-4\lambda_2}+\log^2(N)N^{3\lambda_1-2\lambda_2}+N^{2\delta-1}),\ \forall\, t\in[0,T]\Big\}.\end{aligned} \tag{82}$$

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

$$\mathcal{A}\cap\mathcal{B}_2\cap\mathcal{B}_3\cap(\bigcap_{n=0}^{M'}\mathcal{G}_{t_n})\subset S(C_{3,\alpha}), \tag{83}$$

where events $\mathcal{B}_2$, $\mathcal{B}_3$ and $\mathcal{G}_n$ are defined in (93), (101) and (105) respectively. Here the event $\mathcal{S}(C_{3,\alpha})$ can be seen as the stability result and the events $\mathcal{A}$, $\mathcal{B}_2$, $\mathcal{B}_3$ and $\mathcal{G}_n$ can be treated as the stability conditions.

Proof. First, we split $\mathcal{S}(\Lambda)$ into the intersection of non-overlapping sets $\{\mathcal{S}_n(\Lambda)\}_{n=1}^{M'}$, where

$$\begin{aligned}\mathcal{S}_n(\Lambda) := \Big\{ &\|K^N(X_t)-K^N(\overline{X}_t)\|_\infty \le \Lambda\log(N)\left\|X_t-\overline{X}_t\right\|_\infty \\ &+\Lambda\log(N)(N^{6\delta-1-\lambda_1-4\lambda_2}+\log^2(N)N^{3\lambda_1-2\lambda_2}+N^{2\delta-1}),\ \forall\, t\in[t_n,t_{n+1}]\Big\}.\end{aligned} \tag{84}$$

with $\Delta t := t_{n+1}-t_n = N^{-\lambda_1}$, then $\mathcal{S}=\bigcap_{n=0}^{M'}\mathcal{S}_n$. Notice that here the choice of Δt is for the purpose of proving stability and it is different from the one 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

$$K_1^N := K^N-K_2^N, \quad k_2^N = k*\psi_{\lambda_2}^N, \tag{85}$$

which means that we choose $\delta=\lambda_2$ in (10) to be k_2^N and respectively in (25) to be ℓ_2^N.

Following the approach in [8], we introduce the following auxiliary trajectory in time $t\in[t_n,t_{n+1}]$

$$\begin{cases} d\widetilde{X}_t = \widetilde{V}_t dt, \\ d\widetilde{V}_t = \overline{K}^N(\widetilde{X}_t)dt+\sqrt{2\sigma}dB^t\ , \end{cases} \tag{86}$$

starting from the initial phase

$$(\widetilde{X}_{t_n},\widetilde{V}_{t_n}) = (X_{t_n},V_{t_n}), \tag{87}$$

where (X_{t_n},V_{t_n}) satisfies (11) at time t_n.

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 motion for these trajectories we have

$$\frac{d}{dt}\|\overline{X}_t-\widetilde{X}_t\|_\infty = \|\overline{V}_t-\widetilde{V}_t\|_\infty, \tag{88}$$

and

$$
\begin{aligned}
\frac{d}{dt}\|\overline{V}_t - \widetilde{V}_t\|_\infty =&\|\overline{K}^N(\overline{X}_t) - \overline{K}^N(\widetilde{X}_t)\|_\infty \\
\leq& \max_{1\leq j\leq N} |k^N * \rho(\cdot,t)(\overline{x}_j) - k^N * \rho(\cdot,t)(\widetilde{x}_j)| \\
\leq& \max_{1\leq j\leq N} |\overline{x}_j - \widetilde{x}_j| \|\nabla k^N * \rho(\cdot,t)\|_\infty \\
\leq& C(\|\nabla\rho\|_1 + \|\nabla\rho\|_\infty)\|\overline{X}_t - \widetilde{X}_t\|_\infty \\
\leq& C\|\overline{X}_t - \widetilde{X}_t\|_\infty \,,
\end{aligned}
\tag{89}
$$

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

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

Using Gronwall's inequality it follows that

$$
\begin{aligned}
\sup_{t_n\leq t\leq t_{n+1}}\left(\|\overline{X}_t - \widetilde{X}_t\|_\infty + \|\overline{V}_t - \widetilde{V}_t\|_\infty\right) &\leq \exp(C\Delta t)(\|\overline{X}_{t_n} - X_{t_n}\|_\infty + \|\overline{V}_{t_n} - V_{t_n}\|_\infty) \\
&\leq \exp(CN^{-\lambda_1})N^{-\lambda_2} \leq CN^{-\lambda_2},
\end{aligned}
\tag{90}
$$

under the event $\mathcal{A}$.

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

$$
\begin{aligned}
&\|K^N(X_t) - K^N(\overline{X}_t)\|_\infty \\
\leq&\|K_2^N(X_t) - K_2^N(\overline{X}_t)\|_\infty + \|K_1^N(X_t) - K_1^N(\widetilde{X}_t)\|_\infty + \|K_1^N(\widetilde{X}_t) - K_1^N(\overline{X}_t)\|_\infty \\
=:&\mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3,
\end{aligned}
\tag{91}
$$

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

$$
\|K_2^N(X_t) - K_2^N(\overline{X}_t)\|_\infty \leq C\|L_2^N(\overline{X}_t)\|_\infty \left\|X_t - \overline{X}_t\right\|_\infty, \tag{92}
$$

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

$$
\mathcal{B}_2 := \left\{\sup_{t\in[0,T]}\left\|L_2^N(\overline{X}_t) - \overline{L}_2^N(\overline{X}_t)\right\|_\infty \leq C_{2,\alpha}N^{3\lambda_2-1}\log(N)\right\}, \tag{93}
$$

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

$$
\|L_2^N(\overline{X}_t)\|_\infty \leq \|\overline{L}_2^N(\overline{X}_t)\|_\infty + C_{2,\alpha}N^{3\lambda_2-1}\log(N) \leq C\log(N) \quad (\lambda_2 < \frac{1}{3}), \tag{94}
$$

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

$$
\|K_2^N(X_t) - K_2^N(\overline{X}_t)\|_\infty \leq C\log(N)\left\|X_t - \overline{X}_t\right\|_\infty, \quad \forall\, t \in [t_n, t_{n+1}], \tag{95}
$$

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

To estimate $\mathcal{I}_2$, notice that by triangle inequality and (90) one has

$$
\begin{aligned}
\|X_t - \widetilde{X}_t\|_\infty \leq \int_{t_n}^t \|V_s - \widetilde{V}_s\|_\infty ds \leq& \int_{t_n}^t \|V_s - \overline{V}_s\|_\infty + \|\overline{V}_s - \widetilde{V}_s\|_\infty ds && (96)\\
\leq&\Delta t \sup_{s\in[t_n,t]}\left(\|V_s - \overline{V}_s\|_\infty + \|\overline{V}_s - \widetilde{V}_s\|_\infty\right) && (97)\\
\leq&CN^{-\lambda_1-\lambda_2}, && (98)
\end{aligned}
$$

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

$$\begin{aligned}\|K_1^N(X_t) - K_1^N(\widetilde{X}_t)\|_\infty \le & (\|\nabla K_1^N(X_t)\|_\infty + \|\nabla K_1^N(\widetilde{X}_t)\|_\infty)\|X_t - \widetilde{X}_t\|_\infty \\ \le & CN^{3(\delta-\lambda_2)}\|L^N(\overline{X}_t)\|_\infty\|X_t - \widetilde{X}_t\|_\infty \le CN^{3\delta-\lambda_1-4\lambda_2}\|L^N(\overline{X}_t)\|_\infty. \end{aligned} \tag{99}$$

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

$$\left\|X_t - \overline{X}_t\right\|_\infty \le N^{-\lambda_2} \gg N^{-\delta}. \tag{100}$$

And the similar estimate goes to $\frac{\|\nabla K_1^N(\widetilde{X}_t)\|_\infty}{\|L^N(\overline{X}_t)\|_\infty} \le CN^{3(\delta-\lambda_2)}$.

We denote the event

$$\mathcal{B}_3 := \left\{ \sup_{t\in[0,T]} \left\|L^N(\overline{X}_t) - \overline{L}^N(\overline{X}_t)\right\|_\infty \le C_{2,\alpha}N^{3\delta-1}\log(N) \right\}, \tag{101}$$

it has been proved in Proposition 3.1 that

$$\mathbb{P}(\mathcal{B}_3^c) \le N^{-\alpha}. \tag{102}$$

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

$$\|L^N(\overline{X}_t)\|_\infty \le \|\overline{L}^N(\overline{X}_t)\|_\infty + C_{2,\alpha}N^{3\delta-1}\log(N) \le CN^{3\delta-1}\log(N), \tag{103}$$

thus we have

$$\mathcal{I}_2 \le CN^{6\delta-1-\lambda_1-4\lambda_2}\log(N). \tag{104}$$

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

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

$$\begin{aligned}\mathcal{G}_n := \Big\{ & \sup_{t\in[t_n,t_{n+1}]} \left\|K_1^N(\widetilde{X}_t) - K_1^N(\overline{X}_t)\right\|_\infty \le C_{4,\alpha}N^{2\delta-1}\log(N) \\ & + C_{4,\alpha}\log^2(N)N^{3\lambda_1}(\|X_{t_n} - \overline{X}_{t_n}\|_\infty + \|V_{t_n} - \overline{V}_{t_n}\|_\infty)\|k_1^N\|_1\Big\}, \end{aligned} \tag{105}$$

then one has

$$\begin{aligned}&\|K_1^N(\widetilde{X}_t) - K_1^N(\overline{X}_t)\|_\infty \\ \le & C\log^2(N)N^{3\lambda_1-2\lambda_2} + C_{4,\alpha}N^{2\delta-1}\log(N) \\ \le & C(T,C_{f_0})(\log^2(N)N^{3\lambda_1-2\lambda_2} + N^{2\delta-1}\log(N)), \end{aligned} \tag{106}$$

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

$$\|k_1^N\|_1 = \|k^N - k_2^N\|_1 \le C\int_{0\le|x|\le N^{-\lambda_2}} \frac{1}{|x|^2}dx \le CN^{-\lambda_2}. \tag{107}$$

Collecting (95), (104) and (106), it yields that

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

under the event $\mathcal{B}_2\cap\mathcal{B}_3\cap\mathcal{A}\cap\mathcal{G}_n$. Denote $C(\alpha,T,C_{f_0})$ in the above inequality as $C_{3,\alpha}$. This implies $\mathcal{B}_2\cap\mathcal{B}_3\cap\mathcal{A}\cap\mathcal{G}_n \subset \mathcal{S}_n(C_{3,\alpha})$, which yields that

$$\mathcal{B}_2\cap\mathcal{B}_3\cap\mathcal{A}\cap(\bigcap_{n=0}^{M'}\mathcal{G}_n) \subset \left(\bigcap\nolimits_{n=0}^{M'} S_n(C_{3,\alpha})\right) = S(C_{3,\alpha}). \tag{108}$$

Thus, the proposition has been proved. □

Lemma 3.2. *Let K_1^N be defined in (85) and consider two trajectories $(\widetilde{X}_t, \widetilde{V}_t)$, $(\overline{X}_t, \overline{V}_t)$ satisfying (86) and (19) respectively with two different initial phases (X_{t_n}, V_{t_n}) and $(\overline{X}_{t_n}, \overline{X}_{t_n})$. Then for any $\alpha > 0$, there exists $C_{4,\alpha} > 0$ depending only on depends on α, T and C_{f_0} such that for N sufficiently large it holds that*

$$\mathbb{P}\Bigg(\sup_{t\in[t_n,t_{n+1}]} \left\| K_1^N(\widetilde{X}_t) - K_1^N(\overline{X}_t) \right\|_\infty \geq C_{4,\alpha} N^{2\delta-1}\log(N) \\ + C_{4,\alpha}\log^2(N) N^{3\lambda_1}(\|X_{t_n} - \overline{X}_{t_n}\|_\infty + \|V_{t_n} - \overline{V}_{t_n}\|_\infty)\|k_1^N\|_1 \Bigg) \leq N^{-\alpha}, \tag{109}$$

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

This lemma plays a crucial role in improving the cut-off. The proof will be carried out in Section 5.

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. Denote the event

$$\mathcal{C} := \left\{ \sup_{t\in[0,T]} \left\| K^N(\overline{X}_t) - \overline{K}^N(\overline{X}_t) \right\|_\infty \leq C_{2,\alpha} N^{2\delta-1}\log(N) \right\}. \tag{110}$$

Consider the quantity $e(t)$ defined as

$$e(t) := \sqrt{\log(N)} \left\| X_t - \overline{X}_t \right\|_\infty + \left\| V_t - \overline{V}_t \right\|_\infty. \tag{111}$$

Computing under the event $\mathcal{C} \cap \mathcal{S}(C_{3,\alpha})$ and using the fact $\frac{d\|x\|_\infty}{dt} \leq \|\frac{dx}{dt}\|_\infty$, one has

$$\begin{aligned}
\frac{de(t)}{dt} &\leq \sqrt{\log(N)} \left\| V_t - \overline{V}_t \right\|_\infty + \left\| K^N(X_t) - \overline{K}^N(\overline{X}_t) \right\|_\infty \\
&\leq \sqrt{\log(N)} \left\| V_t - \overline{V}_t \right\|_\infty + \left\| K^N(X_t) - K^N(\overline{X}_t) \right\|_\infty + \left\| K^N(\overline{X}_t) - \overline{K}^N(\overline{X}_t) \right\|_\infty \\
&\leq \sqrt{\log(N)} \left\| V_t - \overline{V}_t \right\|_\infty \\
&\quad + C_{3,\alpha}\log(N) \left\| X_t - \overline{X}_t \right\|_\infty + C_{3,\alpha}\log(N)(N^{6\delta-1-\lambda_1-4\lambda_2} + \log(N)N^{3\lambda_1-2\lambda_2} + N^{2\delta-1}) \\
&\quad + C_{2,\alpha}N^{2\delta-1}\log(N) \\
&\leq C(\alpha, T, C_{f_0})\sqrt{\log(N)}e(t) \\
&\quad + C(\alpha, T, C_{f_0})\log(N)(N^{6\delta-1-\lambda_1-4\lambda_2} + \log(N)N^{3\lambda_1-2\lambda_2} + N^{2\delta-1}).
\end{aligned} \tag{112}$$

Using Gronwall's inequality with $e(0) = 0$, it follows from (112) that

$$\sup_{t\in[0,T]} e(t) \leq CTe^{C\sqrt{\log(N)}T}\log(N)(N^{6\delta-1-\lambda_1-4\lambda_2} + \log(N)N^{3\lambda_1-2\lambda_2} + N^{2\delta-1}). \tag{113}$$

If we denote the event

$$\mathcal{M} := \left\{ \sup_{t\in[0,T]} e(t) \leq CTe^{C\sqrt{\log(N)}T}\log(N)(N^{6\delta-1-\lambda_1-4\lambda_2} + \log(N)N^{3\lambda_1-2\lambda_2} + N^{2\delta-1}) \right\}. \tag{114}$$

then it follows from Proposition (3.2) that

$$\mathcal{C} \cap \mathcal{B}_2 \cap \mathcal{B}_3 \cap \mathcal{A} \cap (\bigcap_{n=0}^{M'} \mathcal{G}_{t_n}) \subset \mathcal{C} \cap \mathcal{S}(C_{3,\alpha}) \subset \mathcal{M}. \tag{115}$$

Notice that for

$$0 < \lambda_2 < 1/3;\ 0 < \lambda_1 < \frac{\lambda_2}{3};\ \frac{1}{3} \leq \delta < \min\left\{ \frac{\lambda_1 + 3\lambda_2 + 1}{6}, \frac{1-\lambda_2}{2} \right\}, \tag{116}$$

there exists some N_0 depending only on α, T and C_{f_0}, such that for $N \geq N_0$

$$\begin{aligned}&\sup_{t\in[0,T]} \sqrt{\log(N)}\left\|X_t - \overline{X}_t\right\|_\infty + \left\|V_t - \overline{V}_t\right\|_\infty \leq \sup_{t\in[0,T]} e(t)\\ \leq &CTe^{C\sqrt{\log(N)}T}\log(N)(N^{6\delta-1-\lambda_1-4\lambda_2} + \log(N)N^{3\lambda_1-2\lambda_2} + N^{2\delta-1}) \leq \frac{1}{2}N^{-\lambda_2} < N^{-\lambda_2}. \quad (117)\end{aligned}$$

Since $e(t)$ is a continuous function and it vanishes at $t = 0$, it can never reach $N^{-\lambda_2}$. So the condition $\mathcal{A}$ defined in (81) has never been used. The above argument is a standard *a-priori* estimate in PDE analysis, which has been used in [20, 26, 27, 38]. Thus it follows from (115) that

$$\mathcal{C} \cap \mathcal{B}_2 \cap \mathcal{B}_3 \cap (\bigcap_{n=0}^{M'} \mathcal{G}_{t_n}) \subset \mathcal{C} \cap \mathcal{S}(C_{3,\alpha}) \subset \mathcal{M}, \tag{118}$$

which concludes that

$$\begin{aligned}\mathbb{P}\left(\sup_{0\leq t\leq T} e(t) \geq N^{-\lambda_2}\right) &\leq \mathbb{P}((\mathcal{M})^c) \leq \sum_{n=0}^{M'} \mathbb{P}(\mathcal{G}_{t_n}^c) + \mathbb{P}(\mathcal{B}_2^c) + \mathbb{P}(\mathcal{B}_3^c) + \mathbb{P}(\mathcal{C}^c)\\ &\leq CN^{-\alpha+\lambda_1} + 3N^{-\alpha} \leq N^{-\alpha'},\end{aligned} \tag{119}$$

by using Proposition 3.1, Proposition 3.2 and Lemma 3.2. Since $\alpha > 0$ is arbitrary and so is α', we have proved Theorem 1.2.

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

$$W_p(f_t, \mu_\Phi(t)) \leq W_p(f_t, f_t^N) + W_p(f_t^N, \mu_\Psi(t)) + W_p(\mu_\Psi(t), \mu_\Phi(t)). \tag{120}$$

Then the idea of the proof of Theorem 1.3 is to obtain the error estimates of those three parts respectively.

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 (12) approximate solutions of the original VPFP equations (1) as the width of the cut-off goes to zero. It follows from [10, Lemma 3.2] that

$$\sup_{t\in[0,T]} W_p(f_t, f_t^N) \leq N^{-\delta} e^{C_1\sqrt{\log(N)}}, \tag{121}$$

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 [37]. Note that here we can't follow the method in [33] 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 [33, Corollary 9.4] by using the concentration estimate of Fournier and Guillin [18]. We summarize the result as following: 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

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

Then there exists a constants C_2 and C_3 such that it holds

$$\begin{aligned}\mathbb{P}\Bigg(\sup_{t\in[0,T]} W_p(f_t^N, \mu_\Psi(t)) \leq \sqrt{\log(N)}N^{-\kappa}e^{C_2\sqrt{\log(N)}}\Bigg)&\\ \geq 1 - C_3\left(e^{-C_4N^{1-\max\{6,2p\}\kappa}} + N^{1-\frac{m}{2p}}\right).&\end{aligned} \tag{122}$$

• *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 [33, Lemma 5.2] that for all $p\in[0,\infty]$

$$\sup_{t\in[0,T]} W_p(\mu_\Psi(t),\mu_\Phi(t)) \le \sup_{t\in[0,T]} \|\Psi(t)-\Phi(t)\|_\infty. \tag{123}$$

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

$$\mathbb{P}\left(\sup_{t\in[0,T]} W_p(\mu_\Psi(t),\mu_\Phi(t)) \le N^{-\lambda_2}\right) \ge 1-N^{1-\frac{m}{2p}}. \tag{124}$$

• *Convergence of* $W_p(f_t,\mu_\Phi(t))$. Collecting estimates (121), (122) and (124) and choosing $\kappa<\min\{\delta,\frac{1}{6},\frac{1}{2p}\}$, it yields that

$$\begin{aligned}\mathbb{P}\bigg(\sup_{t\in[0,T]} W_p(f_t,\mu_\Phi(t)) &\le (1+\sqrt{\log(N)})N^{-\kappa}e^{C_5\sqrt{\log(N)}}+N^{-\lambda_2}\bigg)\\ &\ge 1-C_6\left(e^{-C_7N^{1-\max\{6,2p\}\kappa}}+N^{1-\frac{m}{2p}}\right),\end{aligned} \tag{125}$$

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 $N\ge e^{\left(\frac{2C_5}{1-3\lambda_2}\right)^2}$, which yields $N^{1-3\lambda_2}\ge(1+\sqrt{\log(N)})e^{C_5\sqrt{\log(N)}}$. Hence we concludes that

$$\begin{aligned}\mathbb{P}\bigg(\sup_{t\in[0,T]} W_p(f_t,\mu_\Phi(t)) &\le N^{-\kappa+1-3\lambda_2}+N^{-\lambda_2}\bigg)\\ &\ge 1-C_6\left(e^{-C_7N^{1-\max\{6,2p\}\kappa}}+N^{1-\frac{m}{2p}}\right).\end{aligned} \tag{126}$$

□

5 The proof of Lemma 3.2

In this section, we present the proof of Lemma 3.2, where it provides the difference between the actions of the force k_1^N on solutions of (19) starting at different points X_{t_n} and $\overline{X}_{t_n}$. First, let us consider the fundamental solution $G(x,v,t)$ of the equation

$$\partial_t G+v\cdot\nabla_x G=\Delta_v G, \tag{127}$$

which can be calculated explicitly as

$$G(x,v,t)=C\frac{1}{t^6}\exp\left(-\frac{|v|^2}{4t}-\frac{3|x-tv/2|^2}{t^3}\right), \tag{128}$$

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

Lemma 5.1. *Let*

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

and $p\in[1,\infty]$, *there exists a* C_p *such that for any* $j\in\mathbb{N}_0$ *the following holds*

$$\left\||x|^j\nabla_v G\right\|_{p,1}\le C_p t^{\frac{-10p+3jp+9}{2p}}\ ,\quad \left\||x|^j\nabla_x G\right\|_{p,1}\le C_p t^{\frac{-12p+3jp+9}{2p}}\ , \tag{129}$$

and

$$\left\|G\left(\cdot-\frac{1}{2}(a-b)\right)-G\left(\cdot-\frac{1}{2}(b-a)\right)\right\|_{p,1}\le C_p|a-b|\left(t^{\frac{-12p+9}{2p}}+t^{\frac{-10p+9}{2p}}\right)\ , \tag{130}$$

as well as

$$\left\| |\cdot| \left(G\left(\cdot - \frac{1}{2}(a-b)\right) - G\left(\cdot - \frac{1}{2}(b-a)\right)\right)\right\|_{p,1} \leq C_p|a-b| \left(t^{\frac{-7p+9}{2p}} + t^{\frac{-9p+9}{2p}}\right) . \tag{131}$$

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

$$\|f\|_{p,q} := \left(\int_{\mathbb{R}^3} \left(\int_{\mathbb{R}^3} |f(x,v)|^q dv\right)^{p/q} dx\right)^{1/p} . \tag{132}$$

Proof. It is easy to compute that

$$G = C\frac{1}{t^6} \exp\left(-\frac{3|x|^2}{4t^3}\right) \exp\left(-\frac{|v-\frac{3x}{2t}|^2}{t}\right) \tag{133}$$

and

$$\nabla_v G = C\frac{1}{t^6} \exp\left(-\frac{3|x|^2}{4t^3}\right) \exp\left(-\frac{|v-\frac{3x}{2t}|^2}{t}\right)\left(-\frac{2v}{t}+\frac{3x}{t^2}\right). \tag{134}$$

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

$$\begin{aligned}\int_{\mathbb{R}^3}|G|dv &= C\frac{1}{t^6} \exp\left(-\frac{3|x|^2}{4t^3}\right)\int_{\mathbb{R}^3} \exp\left(-\frac{|v-\frac{3x}{2t}|^2}{t}\right) dv \\ &\leq C\frac{1}{t^{9/2}} \exp\left(-\frac{3|x|^2}{4t^3}\right).\end{aligned} \tag{135}$$

respectively

$$\begin{aligned}\int_{\mathbb{R}^3}|\nabla_v G|dv &= C\frac{1}{t^6} \exp\left(-\frac{3|x|^2}{4t^3}\right)\int_{\mathbb{R}^3} \exp\left(-\frac{|v-\frac{3x}{2t}|^2}{t}\right)\left(-\frac{2v}{t}+\frac{3x}{t^2}\right) dv \\ &\leq C\frac{1}{t^5} \exp\left(-\frac{3|x|^2}{4t^3}\right)\int_{\mathbb{R}^3} u\exp\left(-u^2\right) du \leq C\frac{1}{t^5}\exp\left(-\frac{3|x|^2}{4t^3}\right).\end{aligned} \tag{136}$$

As a direct result from (135) and (136), one has

$$\left\| |\cdot|^j G\right\|_{\infty,1} \leq C\frac{1}{t^{(9-3j)/2}} \qquad \left\| |\cdot|^j \nabla_v G\right\|_{\infty,1} \leq C\frac{1}{t^{5-3j/2}}. \tag{137}$$

For $1 \leq p < \infty$

$$\begin{aligned}\left\| |\cdot|^j G\right\|_{p,1} &\leq Ct^{-9/2}\left(\int_{\mathbb{R}^3} |x|^{pj} \exp\left(-\frac{3px^2}{4t^3}\right) dx\right)^{\frac{1}{p}} \\ &\leq C_p t^{-\frac{9}{2}+\frac{9+3pj}{2p}} \left(\int_{\mathbb{R}^3} |y|^j \exp\left(-y^2\right) dy\right)^{\frac{1}{p}} \leq C_p t^{\frac{-9p+3jp+9}{2p}},\end{aligned} \tag{138}$$

and

$$\begin{aligned}\left\| |\cdot|^j \nabla_v G\right\|_{p,1} &\leq Ct^{-5}\left(\int_{\mathbb{R}^3} |x|^{j} \exp\left(-\frac{3px^2}{4t^3}\right) dx\right)^{\frac{1}{p}} \\ &\leq C_p t^{-5+\frac{9+3jp}{2p}} \left(\int_{\mathbb{R}^3} |y|^j \exp\left(-y^2\right) dy\right)^{\frac{1}{p}} \leq C_p t^{\frac{-10p+3jp+9}{2p}},\end{aligned} \tag{139}$$

We also have

$$\nabla_x G = \frac{1}{t^6} \exp\left(-\frac{3|x|^2}{4t^3}\right) \exp\left(-\frac{|v-\frac{3x}{2t}|^2}{t}\right)\left(-\frac{6x}{t^2}+\frac{3v}{t}\right), \tag{140}$$

which leads to

$$\begin{aligned}
\int_{\mathbb{R}^3} |\nabla_x G| dv &= C\frac{1}{t^6}\exp\left(-\frac{3|x|^2}{4t^3}\right)\int_{\mathbb{R}^3}\exp\left(-\frac{|v-\frac{3x}{2t}|^2}{t}\right)\left(-\frac{6x}{t^3}+\frac{3v}{t^2}\right)dv\\
&\le C\frac{1}{t^8}\exp\left(-\frac{3|x|^2}{4t^3}\right)\int_{\mathbb{R}^3}\exp\left(-\frac{|v-\frac{3x}{2t}|^2}{t}\right)\left(-\frac{2x}{t}+v\right)dv\\
&\le Ct^{-8+\frac{3}{2}}\exp\left(-\frac{3|x|^2}{4t^3}\right)\int_{\mathbb{R}^3}(\sqrt{t}u-\frac{x}{2t})\exp\left(-u^2\right)du\\
&\le C\frac{1}{t^6}\exp\left(-\frac{3|x|^2}{4t^3}\right)+C\frac{1}{t^6}\frac{x}{t^{\frac{3}{2}}}\exp\left(-\frac{3|x|^2}{4t^3}\right).
\end{aligned}\tag{141}$$

It follows from above that

$$\big\||\cdot|^j\nabla_x G\big\|_{\infty,1}\le C\frac{1}{t^{6-3j/2}}.\tag{142}$$

For $1\le p<\infty$

$$\begin{aligned}
&\||\cdot|^j\nabla_x G\|_{p,1}\\
\le& C\frac{1}{t^6}\left(\left(\int_{\mathbb{R}^3}|x|^j\exp\left(-\frac{3px^2}{4t^3}\right)dx\right)^{\frac{1}{p}}+\left(\int_{\mathbb{R}^3}|x|^j\exp\left(-\frac{3p|x|^2}{4t^3}\right)\left(\frac{x}{t^{\frac{3}{2}}}\right)^p dx\right)^{\frac{1}{p}}\right)\\
\le& C_p t^{\frac{-12p+3jp+9}{2p}}\left(\left(\int_{\mathbb{R}^3}|y|^j\exp\left(-y^2\right)dy\right)^{\frac{1}{p}}+\left(\int_{\mathbb{R}^3}|y|^j\exp\left(-py^2\right)|y|^p dy\right)^{\frac{1}{p}}\right)\\
\le& C_p t^{\frac{-12p+3jp+9}{2p}},
\end{aligned}\tag{143}$$

which concludes the proof of (129).

As a direct result from (129), we can prove (130). Indeed,

$$\begin{aligned}
&\left|G\left(\cdot-\frac{1}{2}(a-b)\right)-G\left(\cdot-\frac{1}{2}(b-a)\right)\right|\\
\le&|a-b|\int_0^1\left|\nabla G\left(\cdot-\frac{1}{2}(b-a)+s(b-a)\right)\right|ds\\
\le&|a-b|\int_0^1\left|\nabla_v G\left(\cdot-\frac{1}{2}(b-a)+s(b-a)\right)\right|ds\\
&+|a-b|\int_0^1\left|\nabla_x G\left(\cdot-\frac{1}{2}(b-a)+s(b-a)\right)\right|ds.
\end{aligned}\tag{144}$$

which leads to

$$\begin{aligned}
&\left\|\left(G\left(\cdot-\frac{1}{2}(a-b)\right)-G\left(\cdot-\frac{1}{2}(b-a)\right)\right)\right\|_{p,1}\\
\le& C|a-b|\left(\|\nabla_v G\|_{p,1}+\|\nabla_x G\|_{p,1}\right)\le C_p|a-b|\left(t^{\frac{-12p+9}{2p}}+t^{\frac{-10p+9}{2p}}\right).
\end{aligned}\tag{145}$$

Next we prove (131):

$$\begin{aligned}
&|\cdot|\left(G\left(\cdot-\frac{1}{2}(a-b)\right)-G\left(\cdot-\frac{1}{2}(b-a)\right)\right)\\
\le&\left(|\cdot-\frac{1}{2}(a-b)|G\left(\cdot-\frac{1}{2}(a-b)\right)-|\cdot-\frac{1}{2}(b-a)|G\left(\cdot-\frac{1}{2}(b-a)\right)\right)\\
&+\frac{1}{2}|a-b|\left(G\left(\cdot-\frac{1}{2}(a-b)\right)+G\left(\cdot-\frac{1}{2}(b-a)\right)\right)
\end{aligned}\tag{146}$$

In view of (138), the $(p,1)$-norms of the terms in the last line have the right bound. With the other term we proceed as above, using the function $H = |\cdot|G$:

$$
\begin{aligned}
&\left(|\cdot - \frac{1}{2}(a-b)|G\left(\cdot - \frac{1}{2}(a-b)\right) - |\cdot - \frac{1}{2}(b-a)|G\left(\cdot - \frac{1}{2}(b-a)\right)\right)\\
\leq &|a-b|\int_0^1 \left|\nabla H\left(\cdot - \frac{1}{2}(b-a)+s(b-a)\right)\right| ds\\
\leq &|a-b|\int_0^1 \left|\nabla_v H\left(\cdot - \frac{1}{2}(b-a)+s(b-a)\right)\right| ds\\
&+|a-b|\int_0^1 \left|\nabla_x H\left(\cdot - \frac{1}{2}(b-a)+s(b-a)\right)\right| ds.
\end{aligned}
\tag{147}
$$

It follows from our estimates in (129) that

$$
\left\| |\cdot - \frac{1}{2}(a-b)|G\left(\cdot - \frac{1}{2}(a-b)\right) - |\cdot - \frac{1}{2}(b-a)|G\left(\cdot - \frac{1}{2}(b-a)\right)\right\|_{p,1} \tag{148}
$$

$$
\begin{aligned}
\leq &C|a-b|\left(\||\cdot|\nabla_v G\|_{p,1} + \||\cdot|\nabla_x G\|_{p,1} + \|G\|_{p,1}\right)\\
\leq &C_p|a-b|\left(t^{\frac{-7p+9}{2p}} + t^{\frac{-9p+9}{2p}}\right),
\end{aligned}
\tag{149}
$$

which leads to (131). □

Next we need to introduce a process: For time $0 \leq s \leq t$ and $a := (a_x, a_v) \in \mathbb{R}^{6N}$, let $Z^{a,N}_{t,s} := (Z^{a,N}_{x,t,s}, Z^{a,N}_{v,t,s})$ 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 $\overline{K}^N$:

$$
\begin{cases}
dZ^{a,i,N}_{x,t,s} &= Z^{a,i,N}_{v,t,s}dt,\\
dZ^{a,i,N}_{v,t,s} &= \int_{\mathbb{R}^3} k^N(Z^{a,i,N}_{x,t,s} - x)\rho^N(x,t)dx + \sqrt{2\sigma}dB^t_i, \quad i = 1,\cdots,N,
\end{cases}
\tag{150}
$$

and

$$
(Z^{a,i,N}_{x,s,s}, Z^{a,i,N}_{v,s,s}) = (a^i_x, a^i_v). \tag{151}
$$

Furthermore $(Z^{a,N}_{x,t,s}, Z^{a,N}_{v,t,s})$ has the strong Feller property, implying in particular that it has a transition probability density $u^{a,N}_{t,s}$ is given by the product $u^{a,N}_{t,s} := \prod_{i=1}^N u^{a,i,N}_{t,s}$. Hence each term $u^{a,i,N}_{t,s}$ is the transition probability density of $(Z^{a,i,N}_{x,t,s}, Z^{a,i,N}_{v,t,s})$ and also the solution to the linearized equation:

$$
\partial_t u^{a,i,N}_{t,s} + v\cdot\nabla_x u^{a,i,N}_{t,s} + k^N * \rho^N \cdot \nabla_v u^{a,i,N}_{t,s} = \Delta_v u^{a,i,N}_{t,s}, \quad u^{a,i,N}_{s,s} = \delta_{a_i}, \tag{152}
$$

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

Consider now the process $Z^{a,N}_{t,s}$ and $Z^{b,N}_{t,s}$ for two different starting points $a, b \in \mathbb{R}^{6N}$. It is intuitively clear that the probability density $u^{a,i,N}_{t,s}$ and $u^{b,i,N}_{t,s}$ 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 $t-s$.

Lemma 5.2. *There exists a positive constant C depending 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 $0 < t \leqslant T$ the following estimates for the transition probability densities $u^{a,N}_{t,s}$ resp. $u^{b,N}_{t,s}$ of the processes $Z^{a,N}_{t,s}$ resp. $Z^{b,N}_{t,s}$ given by (150) hold for $t-s < 1$:*

(i) $\|u^{a,N}_{t,s}\|_{\infty,1} \leqslant C(t-s)^{-\frac{9}{2}}$,

(ii) $\|u^{a,N}_{t,s} - u^{b,N}_{t,s}\|_{\infty,1} \leqslant C(t-s)^{-6}|a-b|$.

Proof. The proof of the estimates follows the ideas of [8, Lemma 2]. However, the evolution equation for the present system is more difficult to handle, in particular the spacial overlap is suppressed for short times 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 $t < 1$.

Note that the force $\overline{k}_t^N(x) := k^N * \rho_t^N$ we consider is globally Lipschitz and L^∞ because of (33), thus there exists a $C > 0$ independent of N such that

$$\sup_{0\leq t\leq T; x,y\in\mathbb{R}^3} \frac{|\overline{k}_t(x) - \overline{k}_t(y)|}{|x-y|} \leq C\,. \tag{153}$$

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

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

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

$$w_{t,0}^{a,N}((x,v)) := u_{t,0}^{a,N}((x,v) + c_t)\,. \tag{154}$$

From the evolution equation of $u_{t,s}^{a/b,N}$ and c_t one gets directly

$$\frac{\partial}{\partial t} w_{t,0}^{a/b,N}(x,v) := \Delta_v w_{t,0}^{a/b,N}(x,v) - \nabla_x w_{t,0}^{a/b,N}\cdot v - \nabla_v w_{t,0}^{a/b,N}\cdot\left(\overline{k}_t^N(x + x_t^c) - \overline{k}_t^N(x_t^c)\right). \tag{155}$$

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

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

$$\|u_{t,0}^{a,N} - u_{t,0}^{b,N}\|_{p,1} = \|w_{t,0}^{a,N} - w_{t,0}^{b,N}\|_{p,1}\,. \tag{156}$$

Before proceeding we would like to explain the advantage of looking at w instead of v first. The difficulties arise when dealing with short times. There the $u^{a/b}$ are peaked around the center roughly given by $\frac{1}{2}(a+b)$, respectively the $w^{a/b}$ are peaked roughly at 0. Here the force term of w – which is zero at $x = 0$ – suppresses the last term of (155). Thus w will be very close to the heat-kernel G_t of our time evolution.

Using (155) and the properties of the heat kernel we get

$$\begin{aligned} w_{t,0}^{a,N} =& G_t * \delta\left(\cdot - \left(\frac{1}{2}(a-b)\right)\right) - \int_0^t G_{t-s} * \left(\nabla_v w_{s,0}^{a,N}\cdot\left(\overline{k}_s(\cdot + x_s^c) - \overline{k}_s(x_s^c)\right)\right) ds \\ =& G_t\left(\cdot - \frac{1}{2}(a-b)\right) - \int_0^t \nabla_v G_{t-s} * \left(w_{s,0}^{a,N}\left(\overline{k}_s(\cdot + x_s^c) - \overline{k}_s(x_s^c)\right)\right) ds \end{aligned} \tag{157}$$

and

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

thus

$$\begin{aligned} w_{t,0}^{a,N} - w_{t,0}^{b,N} =& \left(G_t\left(\cdot - \frac{1}{2}(a-b)\right) - G_t\left(\cdot - \frac{1}{2}(b-a)\right)\right) \\ &- \int_0^t \nabla_v G_{t-s} * \left(\left(w_{s,0}^{a,N} - w_{s,0}^{b,N}\right)\left(\overline{k}_s(\cdot + x_s^c) - \overline{k}_s(x_s^c)\right)\right) ds \end{aligned} \tag{158}$$

Defining $\eta^N_{t,0} : \mathbb{R}^6 \to \mathbb{R}^+_0$ by $\eta^N_{t,0}(x,v) := |(x,v)| \left| w^{a,N}_{t,0} - w^{b,N}_{t,0} \right|$ and using (153), we can find a constant C such that

$$\begin{aligned} \eta^N_{t,0} \leq & |\cdot| \left| G_t\left(\cdot - \frac{1}{2}(a-b)\right) - G_t\left(\cdot - \frac{1}{2}(b-a)\right) \right| \\ & + C \left| \int_0^t \nabla_v G_{t-s} * \eta^N_{s,0} ds \right| \end{aligned} \tag{159}$$

Using the properties of the heat kernel (129), (131) and Young's inequality in (159), we get

$$\left\| \eta^N_{t,0} \right\|_{1,1} \leq C|a-b| + C \int_0^t (t-s)^{-1/2} \left\| \eta^N_{s,0} \right\|_{1,1} ds. \tag{160}$$

Applying a generalized Gronwall's inequality with weak singularities [23, Lemma 7.1.1] leads to

$$\|\eta^N_{t,0}\|_{1,1} \leq C|a-b| \qquad \text{uniform in } t \in [0,T] \,. \tag{161}$$

Further (159) gives for any $1 \leq p \leq \infty$

$$\begin{aligned} \left\| \eta^N_{t,0} \right\|_{p,1} \leq & \left\| |\cdot| G_t\left(\cdot - \frac{1}{2}(a-b)\right) - G_t\left(\cdot - \frac{1}{2}(b-a)\right) \right\|_{p,1} \\ & + C \int_0^{t/2} \left\| \nabla_v G_{t-s} * \eta^N_{s,0} \right\|_{p,1} ds + C \int_{t/2}^t \left\| \nabla_v G_{t-s} * \eta^N_{s,0} \right\|_{p,1} ds \end{aligned} \tag{162}$$

Using Young's inequality we get for $1 + p^{-1} = \frac{9}{10} + q^{-1}$.

$$\begin{aligned} \left\| \eta^N_{t,0} \right\|_{p,1} \leq & C|a-b| t^{\frac{-9p+9}{2p}} + C \int_0^{t/2} \left\| \nabla_v G_{t-s} \right\|_{p,1} \left\| \eta^N_{s,0} \right\|_{1,1} ds \\ & + C \int_{t/2}^t \left\| \nabla_v G_{t-s} \right\|_{10/9,1} \left\| \eta^N_{s,0} \right\|_{q,1} ds \end{aligned}$$

Due to (129), one has $\|\nabla_v G_{t-s}\|_{10/9,1} \leq C(t-s)^{-19/20}$. This and (161) give

$$\left\| \eta^N_{t,0} \right\|_{p,1} \leq C|a-b| t^{\frac{-9p+9}{2p}} + C|a-b| \int_0^{t/2} \left\| \nabla_v G_{t-s} \right\|_{p,1} ds + C \sup_{t/2 \leq s \leq t} \left\| \eta^N_{s,0} \right\|_{q,1} . \tag{163}$$

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 $\left\| \eta^{a,N}_{t,0} \right\|_{1,1}$ (see (161)) we can then iteratively estimate the L^p norms of $\eta^N_{t,0}$ for higher exponents, i.e.

$$\left\| \eta^N_{t,0} \right\|_{p_{k+1},1} \leq C|a-b| t^{\frac{-9p_{k+1}+9}{2p_{k+1}}} + C|a-b| \int_0^{t/2} \left\| \nabla_v G_{t-s} \right\|_{p_{k+1},1} ds + C \sup_{t/2 \leq s \leq t} \left\| \eta^N_{s,0} \right\|_{p_k,1} . \tag{164}$$

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

$$\left\| \eta^N_{t,0} \right\|_{\infty,1} \leq C|a-b| t^{\frac{-9}{2}} \,. \tag{165}$$

Having good control of $\|\eta_{t,0}^N\|_{\infty,1}$ we can now estimate $w_{t,0}^{a,N} - w_{t,0}^{b,N}$ using (158):

$$\begin{aligned}
\left\|w_{t,0}^{a,N} - w_{t,0}^{b,N}\right\|_{\infty,1} \le & \left\|G_t\left(\cdot - \frac{1}{2}(a-b)\right) - G_t\left(\cdot - \frac{1}{2}(b-a)\right)\right\|_{\infty,1} \qquad (166)\\
&+ \int_0^t \left\|\nabla_v G_{t-s} * \left(\left(w_{s,0}^{a,N} - w_{s,0}^{b,N}\right)\left(\overline{k}_s\left(\cdot + x_s^c\right) - \overline{k}_s\left(x_s^c\right)\right)\right)\right\|_{\infty,1} ds\\
\le & C|a-b|t^{-6} + C\int_0^{t/2} \|\nabla_v G_{t-s}\|_{\infty,1} \left\|\eta_{s,0}^N\right\|_{1,1} ds\\
&+ \int_{t/2}^t \|\nabla_v G_{t-s}\|_{1,1} \left\|\eta_{s,0}^N\right\|_{\infty,1} ds\\
\le & C|a-b|t^{-6} + C|a-b|\int_0^{t/2}(t-s)^{-5}ds + C|a-b|\int_{t/2}^t (t-s)^{-\frac{1}{2}}s^{-\frac{9}{2}}ds\\
\le & C|a-b|\left(t^{-6}+t^{-4}\right) \le C|a-b|t^{-6}. \qquad (167)
\end{aligned}$$

With (156) statement (ii) of the lemma follows.

$\square$

Next we define the random sets

$$M_{t_n} := \left\{2 \le j \le N : \left|x_1^{t_n} - x_j^{t_n} + (t-t_n)(v_1^{t_n} - v_j^{t_n})\right| \le N^{-\lambda_2} + \log N(t-t_n)^{3/2}\right\}, \qquad (168)$$

and

$$\overline{M}_{t_n} := \left\{2 \le j \le N : \left|\overline{x}_1^{t_n} - \overline{x}_j^{t_n} + (t-t_n)(\overline{v}_1^{t_n} - \overline{v}_j^{t_n})\right| \le 3N^{-\lambda_2} + \log N(t-t_n)^{3/2}\right\}, \qquad (169)$$

where M_{t_n} is the indies of particles which are in the ball of radius $N^{-\lambda_2} + \log N(t-t_n)^{3/2}$ around $x_1^{t_n}$ and $\overline{M}_{t_n}$ is an intermediate set introduced to help control M_{t_n}.

We also random sets

$$\mathcal{S}_{t_n} = \left\{\omega : |M_{t_n}(\omega)| < 2C_* N\left(3N^{-\lambda_2} + \log N(t-t_n)^{3/2}\right)^2\right\}, \qquad (170)$$

and

$$\overline{\mathcal{S}}_{t_n} = \left\{\omega : \left|\overline{M}_{t_n}(\omega)\right| < 2C_* N\left(3N^{-\lambda_2} + \log N(t-t_n)^{3/2}\right)^2\right\}, \qquad (171)$$

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

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

Lemma 5.3. *Let the indices set M_{t_n} satisfy (168), then for any $\alpha > 0$, there exists some constant $N_0 > 0$ depending only on α, T and C_{f_0} such that it holds*

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

which indicates that for particles outside M_{t_n}, $\widetilde{x}_j^t$ and $\overline{x}_j^t$ contribute to the interaction of k_1 with low probability (almost zero).

Proof. Let $a^t, b^t \in \mathbb{R}^6$ be given by $a^{t_n} = 0$ and $b^{t_n} = 0$ and we use the notation $a^t = (a^t_x, a^t_v)$, $b^t = (b^t_x, b^t_v)$, which satisfy stochastic differential equations $da^t_x = a^t_v dt$, $da^t_v = \sqrt{2\sigma} dB^t_1$, $db^t_x = b^t_v dt$ and $db^t_v = \sqrt{2\sigma} dB^t_j$. It follows with the evolution equation (86) that

$$d\left(\widetilde{x}^t_1 - a^t_x\right) = \left(\widetilde{v}^t_1 - a^t_v\right) dt \text{ and } d\left(\widetilde{v}^t_1 - a^t_v\right) = \overline{k}^N(\widetilde{x}^t_1),$$

and

$$d\left(\widetilde{x}^t_j - b^t_x\right) = \left(\widetilde{v}^t_j - b^t_v\right) dt \text{ and } d\left(\widetilde{v}^t_j - b^t_v\right) = \overline{k}^N\left(\widetilde{x}^t_j\right),$$

where $\overline{k}^N = k^N * \rho^N$. Integrating twice we get for any $s \geq t_n$

$$\left(\widetilde{v}^s_j - b^s_v\right) = \widetilde{v}^{t_n}_j + \int_{t_n}^s \overline{k}^N(\widetilde{x}^\tau_j) d\tau$$

and

$$\left(\widetilde{x}^t_j - b^t_x\right) = \widetilde{x}^{t_n}_j + \int_{t_n}^t \left(\widetilde{v}^{t_n}_j + \int_{t_n}^s \overline{k}^N(\widetilde{x}^\tau_j) d\tau\right) ds .$$

And by the same argument

$$\widetilde{x}^t_1 - \widetilde{x}^t_j - (a^t_x - b^t_x) = \widetilde{x}^{t_n}_1 - \widetilde{x}^{t_n}_j + \int_{t_n}^t \left(\widetilde{v}^{t_n}_1 - \widetilde{v}^{t_n}_j + \int_{t_n}^s \overline{k}^N(\widetilde{x}^\tau_1) d\tau - \int_{t_n}^s \overline{k}^N(\widetilde{x}^\tau_j) d\tau\right) ds .$$

Since $\overline{k}(\widetilde{x}^\tau_j)$ is bounded, it follows that there is a constant $0 < C < \infty$ such that

$$\left|\widetilde{x}^t_1 - \widetilde{x}^t_j\right| \geq \left|\widetilde{x}^{t_n}_1 - \widetilde{x}^{t_n}_j + (t - t_n)(\widetilde{v}^{t_n}_1 - \widetilde{v}^{t_n}_j)\right| - \left|a^t_x - b^t_x\right| - C(t - t_n)^2.$$

Thus for $(t - t_n) < 1$ it follows that $j \in M^c_{t_n}$, i.e. $\left|\widetilde{x}^{t_n}_1 - \widetilde{x}^{t_n}_j - (t - t_n)(\widetilde{v}^{t_n}_1 - \widetilde{v}^{t_n}_j)\right| \geq N^{-\lambda_2} + \log N (t - t_n)^{3/2}$, together with $\inf_{t \in [t_n, t_{n+1}]} \left|\widetilde{x}^t_1 - \widetilde{x}^t_j\right| < N^{-\lambda_2}$ implies that $\sup_{t \in [t_n, t_{n+1}]} |a^t_x - b^t_x| > (\ln N - C)(t - t_n)^{3/2}$. Hence

$$\begin{aligned}
&\mathbb{P}_{j \in M^c_{t_n}}\left(\inf_{t \in [t_n, t_{n+1}]} \left|\widetilde{x}^t_1 - \widetilde{x}^t_j\right| < N^{-\lambda_2}\right) \\
\leq &\mathbb{P}\left(\sup_{t \in [t_n, t_{n+1}]} \left|a^t_x - b^t_x\right| > (\ln N - C)(t - t_n)^{3/2}\right) \\
\leq &\mathbb{P}\left(\sup_{t \in [t_n, t_{n+1}]} \left|a^t_v - b^t_v\right| > (\ln N - C)(t - t_n)^{1/2}\right),
\end{aligned} \tag{172}$$

where we used that $a^t_x = \int_0^t a^s_v ds$ and $b^t_x = \int_0^t b^s_v ds$ in the second step. In the same way we can argue that

$$\begin{aligned}
&\mathbb{P}_{j \in M^c_{t_n}}\left(\inf_{t \in [t_n, t_{n+1}]} \left|\overline{x}^t_1 - \overline{x}^t_j\right| < N^{-\lambda_2}\right) \\
\leq &\mathbb{P}\left(\sup_{t \in [t_n, t_{n+1}]} \left|a^t_v - b^t_v\right| > (\ln N - C)(t - t_n)^{1/2}\right) .
\end{aligned} \tag{173}$$

Due to independence the difference $c^t_v = (c^t_1, c^t_2, c^t_3) = a^t_v - b^t_v$ is itself a Wiener process. Splitting up this Wiener process into its three spacial components and using the reflection principle we get

$$\begin{aligned}
&\mathbb{P}\left(\sup_{t \in [t_n, t_{n+1}]} \left|a^t_v - b^t_v\right| > (\ln N - C)(t - t_n)^{1/2}\right) \\
\leq &3\mathbb{P}\left(\sup_{t \in [t_n, t_{n+1}]} \left|c^t_1\right| > (\ln N - C)(t - t_n)^{1/2}\right) \\
\leq &6\mathbb{P}\left(\sup_{t \in [t_n, t_{n+1}]} c^t_1 > (\ln N - C)(t - t_n)^{1/2}\right) \\
= &12\mathbb{P}\left(c^{t_{n+1}}_1 > (\ln N - C)(t - t_n)^{1/2}\right) .
\end{aligned}$$

Recall that the time evolution of a_v^t and b_v^t are standard Brownian motions, i.e. the density is a Gaussian with standard deviation $\sigma_t = \sigma(t-t_n)^{1/2}$. Due to independence of a_v^t and b_v^t also c_1^t is normal distributed with standard deviation of order $(t-t_n)^{1/2}$. Hence there is for N sufficiently large, it holds that

$$\mathbb{P}\left(\sup_{t\in[t_n,t_{n+1}]} \left|a_v^t - b_v^t\right| > (\ln N - C)(t-t_n)^{1/2}\right) \le N^{-\alpha} \, .$$

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

$\square$

Now we have all the estimates needed for the proof of Lemma 3.2.

Proof of Lemma 3.2. We show that for any $\alpha > 0$ there exists a C_α depending only on α, T and C_{f_0} such that

$$\mathbb{P}\Bigg(\sup_{t\in[t_n,t_{n+1}]} \left|\frac{1}{N-1}\sum_{j\neq 1}^{N} k_1(\widetilde{x}_1^t - \widetilde{x}_j^t) - k_1(\overline{x}_1^t - \overline{x}_j^t)\right| \ge C_\alpha N^{2\delta-1}\log(N)$$
$$+ C\log^2(N) N^{3\lambda_1}(\|X_{t_n} - \overline{X}_{t_n}\|_\infty + \|V_{t_n} - \overline{V}_{t_n}\|_\infty)\|k_1^N\|_1\Bigg) \le N^{-\alpha}. \tag{174}$$

This is done in three steps:

(1) We prove that the number of particles inside M_{t_n} is smaller than

$$2C_* N\left(3N^{-\lambda_2} + \log N (t-t_n)^{3/2}\right)^2$$

with probability almost one, namely

$$\mathbb{P}(\mathcal{A}\cap \mathcal{S}_{t_n}^c) \le \mathbb{P}(\mathcal{A}\cap \overline{\mathcal{S}}_{t_n}^c) \le \mathbb{P}(\overline{\mathcal{S}}_{t_n}^c) \le N^{-\alpha}. \tag{175}$$

(2) We prove that particles outside M_{t_n} do not contribute interaction of k_1 with probability almost one, namely

$$\mathbb{P}\Bigg(\mathcal{A}\cap \sup_{t\in[t_n,t_{n+1}]} \left|\frac{1}{N-1}\sum_{j\in M_{t_n}^c} k_1(\widetilde{x}_1^t - \widetilde{x}_j^t) - k_1(\overline{x}_1^t - \overline{x}_j^t)\right| > 0\Bigg) \le N^{-\alpha}. \tag{176}$$

(3) Since particles outside M_{t_n} do not contribute interaction of k_1 with high probability, we only consider particle inside M_{t_n}. And we know already from (a) that the number of particles inside M_{t_n} is small with high probability. Hence we can prove

$$\mathbb{P}\Bigg(\mathcal{A}\cap \mathcal{S}_{t_n}\cap \sup_{t\in[t_n,t_{n+1}]} \left|\frac{1}{N-1}\sum_{j\in M_{t_n}} k_1(\widetilde{x}_1^t - \widetilde{x}_j^t) - k_1(\overline{x}_1^t - \overline{x}_j^t)\right| \ge C_\alpha N^{2\delta-1}\log(N)$$
$$+ C\log^2(N) N^{3\lambda_1}(\|X_{t_n} - \overline{X}_{t_n}\|_\infty + \|V_{t_n} - \overline{V}_{t_n}\|_\infty)\|k_1^N\|_1\Bigg) \le N^{-\alpha}. \tag{177}$$

•*Step 1:* To prove the first part of (175), note that on the set $\mathcal{A}$ and assuming that $|t_n - t| < 1$

$$\left|x_1^{t_n} - x_j^{t_n} + (t-t_n)(v_1^{t_n} - v_j^{t_n})\right| \le N^{-\lambda_2} + \log N (t-t_n) \, ,$$

implies

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

Hence $M_{t_n} \subset \overline{M}_{t_n}$ and thus for any $R > 0$, $\left|\overline{M}_{t_n}\right| < R$ implies that $|M_{t_n}| \le \left|\overline{M}_{t_n}\right| < R$, consequently $\mathcal{S}_{t_n} \supset \overline{\mathcal{S}}_{t_n}$, i.e. $\mathcal{S}_{t_n}^c \subset \overline{\mathcal{S}}_{t_n}^c$.

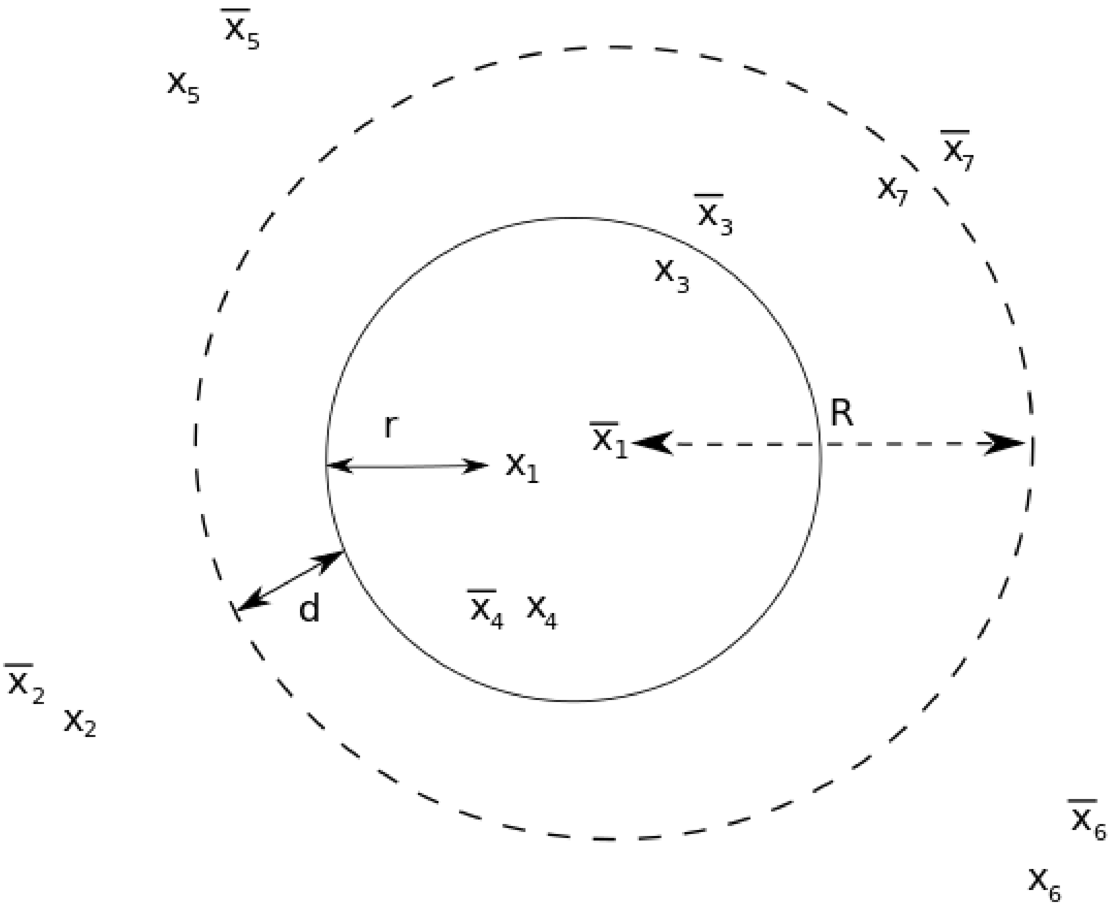

Figure 1: Illustration of the sets M_{t_n} and $\overline{M}_{t_n}$: The set M_{t_n} contains all indices of particles which are in the ball of radius $r = N^{-\lambda_2} + \log N(t - t_n)^{3/2}$ around x_1. In the figure this is the ball with solid boundary. Since on the set $\mathcal{A}$ the distance d of the particles x_1 and $\overline{x}_1$ can not be larger than $N^{-\lambda_2}$, it follows that a particle $\overline{x}_j$ is in the solid ball only if the particle x_j is in the ball with discontinued boundary, i.e. with radius $R = 3N^{-\lambda_2} + \log N(t - t_n)^{3/2}$ around x_1 (see for example particles x_3 and $\overline{x}_3$). Controlling M_{t_n} by $\overline{M}_{t_n}$ will be helpful to estimate the number of particles inside these sets. The $\overline{x}_j$ are i.i.d., the probability to find 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 practically not contribute to the interaction k_1. This comes from the fact that in order to get a sufficiently small distance to x_1 to interact, they have to travel a long distance during the short time interval $(t - t_n)$: the distance $\log N(t - t_n)^{3/2}$ (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 (2). The main contribution thus comes from (3). Knowing that the number of particles in M_{t_n} is quite small helps to estimate this term.

The second part of (175) is trivial, for the third part we use independence of the $\overline{x}$-particles. For any $j \in \{2, \dots, N\}$ the probability to find $j \in \overline{M}_{t_n}$ is given by

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

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

Define $g^N(x, v, s) := f^N(x - vs, v, t_n)$, then it satisfies the following transport equation

$$\begin{cases} \partial_s g^N + v \cdot \nabla_x g^N = 0, \\ g^N(x, v, 0) = f^N(x, v, t_n). \end{cases} \tag{178}$$

It follows that the probability to find $j \in \overline{M}_{t_n}$ is given by

$$\mathbb{P}\left(j \in \overline{M}_{t_n}\right) = \int_{\mathbb{R}^3} \int_{B_R(\mu_0)} g^N(x, v, t - t_n) dx dv, \tag{179}$$

where the center μ_0 of the ball is given by $\mu_0 = \overline{x}_1^{t_n} + (t - t_n)\overline{v}_1^{t_n}$, in particular the integration area is independent of v.

Next, We computes

$$\begin{aligned} \bar{\rho}^N(x, s) &:= \int_{\mathbb{R}^3} g^N(x, v, s) dv = \int_{|v| \le r} g^N(x, v, s) dv + \int_{|v| > r} g^N(x, v, s) dv \\ &\le C_1 \|g^N(\cdot, \cdot, s)\|_\infty r^3 + \frac{1}{r^6} \int_{|v| > r} |v|^6 g^N(x, v, s) dv \\ &= 2C_1^{\frac{2}{3}} \|g^N(\cdot, \cdot, s)\|_\infty^{\frac{2}{3}} \left(\int_{|v| > r} |v|^6 g^N(x, v, s) dv \right)^{\frac{1}{3}}, \end{aligned} \tag{180}$$

where we have chosen

$$r = \left(\frac{\int_{|v| > r} |v|^6 g^N(x, v, s) dv}{C_1 \|g^N(\cdot, \cdot, s)\|_\infty} \right)^{\frac{1}{9}}. \tag{181}$$

It follows that

$$\begin{aligned} \int_{\mathbb{R}^3} |\bar{\rho}^N(x, s)|^3 dx &\le 8C_1^2 \|g^N(\cdot, \cdot, s)\|_\infty^2 \iint_{\mathbb{R}^6} |v|^6 g^N(x, v, s) dx dv \\ &\le C\left(\|f^N(\cdot, \cdot, s)\|_{W^{1,1} \cap W^{1,\infty}(\mathbb{R}^6)}, \ \|(|\cdot|^6 + |\cdot|^6) f^N(\cdot, \cdot, s)\|_{L^1(\mathbb{R}^6)} \right), \end{aligned} \tag{182}$$

where the bound of $\|g^N(\cdot, \cdot, s)\|_\infty$ and $\iint_{\mathbb{R}^6} |v|^6 g^N(x, v, s) dx dv$ can be found in [34], which concludes that

$$\|\bar{\rho}^N\|_3 \le C_2, \tag{183}$$

where C_2 depends only on T, and C_{f_0}. It follows from (179) that

$$\begin{aligned} \mathbb{P}\left(j \in \overline{M}_{t_n}\right) &= \int_{B_R(\mu_0)} \bar{\rho}^N(x, t - t_n) dx \le \|\bar{\rho}^N\|_3 |B_R(\mu_0)|^{\frac{2}{3}} \\ &\le C_2 (\frac{4}{3}\pi)^{\frac{2}{3}} \left(3N^{-\lambda_2} + \log N(t - t_n)^{3/2} \right)^2 \\ &= C_* \left(3N^{-\lambda_2} + \log N(t - t_n)^{3/2} \right)^2 =: p, \end{aligned} \tag{184}$$

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

The probability to find k particles inside the set $\overline{M}_{t_n}$ is thus bounded from above by the binomial probability mass function with parameter p at position k, i.e. for any natural number $0 \le A \le N$

$$\mathbb{P}\left(|\overline{M}_{t_n}| \ge A\right) \le \sum_{j=A}^{N} \binom{N}{j} p^j (1 - p)^{N-j}.$$

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 $\overline{M}_{t_n}$ is exponentially small in a, i.e. there is a sufficiently large N for any $\alpha > 0$ such that

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

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 to find 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 $\overline{M}_{t_n}$ is smaller than any polynomial in N, i.e. there is a C_α for any $\alpha > 0$ such that

$$\mathbb{P}(\mathcal{S}^c_{t_n}) = \mathbb{P}\left(|\overline{M}_{t_n}| \geq 2Np\right) \leq N^{-\alpha}.$$

•*Step 2:* For (176) it is sufficient to show that for any $\alpha > 0$ there is a sufficiently large N such that for some $j \notin M_{t_n}$

$$\mathbb{P}\Big(\mathcal{A} \cap \sup_{t\in[t_n,t_{n+1}]} \left|k_1(\widetilde{x}^t_1 - \widetilde{x}^t_j) - k_1(\overline{x}^t_1 - \overline{x}^t_j)\right| > 0\Big) \leq N^{-\alpha}$$

The total probability we have to control in (176) is at maximum the N-fold value of this. The key to prove that is Lemma 5.3. To have an interaction $k_1(\widetilde{x}^t_1 - \widetilde{x}^t_j) \neq 0$ 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 $k_1(\widetilde{x}^t_1 - \widetilde{x}^t_j) \neq 0$, the probability is smaller than any polynomial in N (see Lemma 5.3).The same holds true for $k_1(\overline{x}^t_1 - \overline{x}^t_j)$.

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

$$\begin{aligned}
&\mathbb{P}\Big(\mathcal{A} \cap \sup_{t\in[t_n,t_{n+1}]} \left|\frac{1}{N-1}\sum_{j\in M^c_{t_n}} k_1(\widetilde{x}^t_1 - \widetilde{x}^t_j) - k_1(\overline{x}^t_1 - \overline{x}^t_j)\right| > 0\Big)\\
&\leq \mathbb{P}\Big(\mathcal{A} \cap \sup_{t\in[t_n,t_{n+1}]} \left|\frac{1}{N-1}\sum_{j\in M^c_{t_n}} k_1(\widetilde{x}^t_1 - \widetilde{x}^t_j)\right| > 0\Big)\\
&\quad + \mathbb{P}\Big(\mathcal{A} \cap \sup_{t\in[t_n,t_{n+1}]} \left|\frac{1}{N-1}\sum_{j\in M^c_{t_n}} k_1(\overline{x}^t_1 - \overline{x}^t_j)\right| > 0\Big)\\
&\leq N\left(\mathbb{P}_{j\in M^c_{t_n}}\left(\inf_{t\in[t_n,t_{n+1}]} |\widetilde{x}^t_1 - \widetilde{x}^t_j| < N^{-\lambda_2}\right) + \mathbb{P}_{j\in M^c_{t_n}}\left(\inf_{t\in[t_n,t_{n+1}]} |\overline{x}^t_1 - \overline{x}^t_j| < N^{-\lambda_2}\right)\right).
\end{aligned}$$

With Lemma 5.3 we get the bound for (176).

•*Step 3:* For (177), let $\Phi_{t_n} = \left(\widetilde{X}^{t_n}, \widetilde{V}^{t_n}\right)$ and $\Psi_{t_n} = \left(\overline{X}^{t_n}, \overline{V}^{t_n}\right)$ be given. We assume that

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

Note that the Brownian motion on the time interval $[t_n, t_{n+1}]$ is independent of Φ_{t_n} and Ψ_{t_n}. To get (177) we prove that for any natural number

$$\begin{aligned}
0 \leq M &\leq 2C_* N\left(3N^{-\lambda_2} + \log(N)(t-t_n)^{3/2}\right)^2\\
&\leq 2C_* N\left(3N^{-\lambda_2} + \log(N)N^{-\frac{3}{2}\lambda_1}\right)^2\\
&\leq 4C_* N\log^2(N)N^{-3\lambda_1}, \quad (0 < \lambda_1 < \frac{2}{3}\lambda_2),
\end{aligned}$$

one has

$$\mathbb{P}_{|M_{t_n}|=M}\Bigg(\sup_{t\in[t_n,t_{n+1}]}\Bigg|\sum_{j\in M_{t_n}} k_1(\widetilde{x}_1^t-\widetilde{x}_j^t)-k_1(\overline{x}_1^t-\overline{x}_j^t)\Bigg|\geq C_\alpha N^{2\delta-1}\log(N)$$
$$+C\log^2(N)N^{3\lambda_1}(\|X_{t_n}-\overline{X}_{t_n}\|_\infty+\|V_{t_n}-\overline{V}_{t_n}\|_\infty)\|k_1^N\|_1\Bigg)\leq N^{-\alpha}. \tag{185}$$

It suffices to prove (185) for $M=CN\log^2(N)N^{-3\lambda_1}$.

We use Lemma 2.5 which we repeat below for easier reference:

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

$$\mathbb{P}\left(|\bar{Z}|\geq\frac{C_\alpha\sqrt{g(M)}\log(M)}{\sqrt{M}}\right)\leq M^{-\alpha}, \tag{186}$$

where C_α depends only on C and α.

We choose $Z_j:=\frac{M}{N-1}k_1(\widetilde{x}_1^t-\widetilde{x}_j^t)-\frac{M}{N-1}\mathbb{E}[k_1(\widetilde{x}_1^t-\widetilde{x}_j^t)]$ and $g(M):=CMN^{4\delta-2}$, then following the same argument as in (45), the conditions

$$\mathbb{E}[Z_j^2]\leq C\frac{M^2}{(N-1)^2}N^\delta\|u_{t,t_n}^{a,N}\|_{\infty,1}\leq CM^2N^{\delta-2}N^{\frac{9}{2}\lambda_1}\leq g(M),$$

is satisfied, where we use the fact $\|u_{t,t_n}^{a,N}\|_{\infty,1}\leq CN^{\frac{9}{2}\lambda_1}$ from (i) in Lemma 5.2. We can also deduce that

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

Thus one has

$$\mathbb{P}\left(\sup_{t\in[t_n,t_{n+1}]}\left|\frac{1}{N-1}\sum_{j\in M_{t_n}}\big(k_1(\widetilde{x}_1^t-\widetilde{x}_j^t)-\mathbb{E}[k_1(\widetilde{x}_1^t-\widetilde{x}_j^t)]\big)\right|\geq C_\alpha N^{2\delta-1}\log N\right)\leq N^{-\alpha}, \tag{187}$$

and, in a similar manner,

$$\mathbb{P}\left(\sup_{t\in[t_n,t_{n+1}]}\left|\frac{1}{N-1}\sum_{j\in M_{t_n}}\big(k_1(\overline{x}_1^t-\overline{x}_j^t)-\mathbb{E}[k_1(\overline{x}_1^t-\overline{x}_j^t)]\big)\right|\geq C_\alpha N^{2\delta-1}\log N\right)\leq N^{-\alpha}. \tag{188}$$

It is left to control the difference

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

which can be done by using Lemma 5.2. Writing $a = \widetilde{X}^{t_n}$ and $b = \overline{X}^{t_n}$ it follows that

$$\begin{aligned}
&\left|\frac{1}{N-1}\sum_{j\in M_{t_n}}\left(\mathbb{E}[k_1(\widetilde{x}_1^t-\widetilde{x}_j^t)]-\mathbb{E}[k_1(\overline{x}_1^t-\overline{x}_j^t)]\right)\right| \\
=&\frac{1}{N-1}\Big|\sum_{j\in M_{t_n}}\int k_1^N(x_1-x_j)\big(u_{t,s}^{a,1,N}(x_1,v_1)u_{t,s}^{a,j,N}(x_j,v_j) \\
&\qquad -u_{t,s}^{b,1,N}(x_1,v_1)u_{t,s}^{b,j,N}(x_j,v_j)\big)dx_1dv_1dx_jdv_j\Big| \\
\leq&\frac{1}{N-1}\sum_{j\in M_{t_n}}\left|\int k_1^N(x_1-x_j)u_{t,s}^{a,1,N}(x_1,v_1)\left(u_{t,s}^{a,j,N}(x_j,v_j)-u_{t,s}^{b,j,N}(x_j,v_j)\right)dx_1dv_1dx_jdv_j\right| \\
&+\frac{1}{N-1}\sum_{j\in M_{t_n}}\left|\int k_1^N(x_1-x_j)u_{t,s}^{b,j,N}(x_1,v_1)\left(u_{t,s}^{a,1,N}(x_j,v_j)-u_{t,s}^{b,1,N}(x_j,v_j)\right)dx_1dv_1dx_jdv_j\right| \\
\leq&\frac{1}{N-1}\sum_{j\in M_{t_n}}\left(\|u_{t,s}^{a,j,N}-u_{t,s}^{b,j,N}\|_{\infty,1}\|k_1^N*\rho_{t,s}^{a,1,N}\|_1+\|u_{t,s}^{a,1,N}-u_{t,s}^{b,1,N}\|_{\infty,1}\|k_1^N*\rho_{t,s}^{b,j,N}\|_1\right) \\
\leq&\frac{1}{N-1}\sum_{j\in M_{t_n}}C(t-s)^{-6}|a-b|(\|k_1^N\|_1\|\rho_{t,s}^{a,1,N}\|_1+\|k_1^N\|_1\|\rho_{t,s}^{b,j,N}\|_1) \\
\leq&C\log^2(N)N^{3\lambda_1}|a-b|\|k_1^N\|_1, \qquad (189)
\end{aligned}$$

by Lemma 5.2, where $\rho_{t,s}^{a,1,N}(x_1)=\int_{\mathbb{R}^3}u_{t,s}^{a,1,N}(x_1,v_1)dv_1$. Collecting (187), (188) and (189) we get (185). which finishes the proof of (177).

•*Step 4:* Using the results (175), (176) and (177) from *Steps* $1-3$ we have for any $\alpha>0$ there exists constants $C_\alpha>0$ and $N_0>0$ depending only on α, T and C_{f_0} such that for $N\geq N_0$ it holds

$$\begin{aligned}
\mathbb{P}\Bigg(&\sup_{t\in[t_n,t_{n+1}]}\left|\frac{1}{N-1}\sum_{j\neq 1}k_1(\widetilde{x}_1^t-\widetilde{x}_j^t)-k_1(\overline{x}_1^t-\overline{x}_j^t)\right|\geq C_\alpha N^{2\delta-1}\log(N) \\
&+C\log^2(N)N^{3\lambda_1}(\|X_{t_n}-\overline{X}_{t_n}\|_\infty+\|V_{t_n}-\overline{V}_{t_n}\|_\infty)\|k_1^N\|_1\Bigg)\leq N^{-\alpha}. \qquad (190)
\end{aligned}$$

Since the particles are exchangeable, the same result holds for $\widetilde{x}_i^t$ and $\overline{x}_i^t$ when $i=2,\cdots,N$, which concludes the proof of Lemma 3.2. □

Acknowledgments: H.H. is partially supported by NSFC (Grant No. 11771237). The research of J.-G. L. is partially supported by KI-Net NSF RNMS (Grant No. 1107444) and NSF DMS (Grant No. 1514826).

References

[1] J. T. Beale and A. Majda. Vortex methods. I: Convergence in three dimensions. *Mathematics of Computation*, 39(159):1–27, 1982.

[2] J. T. Beale and A. Majda. Vortex methods. II: Higher order accuracy in two and and three dimensions. *Mathematics of Computation*, 39(159):29–52, 1982.

[3] N. Boers and P. Pickl. On mean field limits for dynamical systems. *Journal of Statistical Physics*, 164(1):1–16, 2016.

[4] F. Bolley, J. A. Canizo, and J. A. Carrillo. Stochastic mean-field limit: non-Lipschitz forces and swarming. *Mathematical Models and Methods in Applied Sciences*, 21(11):2179–2210, 2011.

[5] F. Bouchut. Existence and uniqueness of a global smooth solution for the Vlasov-Poisson-Fokker-Planck system in three dimensions. *Journal of functional analysis*, 111(1):239–258, 1993.

[6] F. Bouchut. Smoothing effect for the non-linear Vlasov-Poisson-Fokker-Planck system. *Journal of differential equations*, 122(2):225–238, 1995.

[7] W. Braun and K. Hepp. The Vlasov dynamics and its fluctuations in the $\frac{1}{N}$ limit of interacting classical particles. *Communications in mathematical physics*, 56(2):101–113, 1977.

[8] A. Cañizares-García and P. Pickl. Microscopic derivation of the Keller-Segel equation in the sub-critical regime. *arXiv preprint arXiv:1703.04376*, 2017.

[9] A. Carpio. Long-time behaviour for solutions of the Vlasov-Poisson-Fokker-Planck equation. *Mathematical methods in the applied sciences*, 21(11):985–1014, 1998.

[10] J. A. Carrillo, Y.-P. Choi, and S. Salem. Propagation of chaos for the VPFP equation with a polynomial cut-off. *arXiv preprint arXiv:1802.01929*, 2018.

[11] J. A. Carrillo, M. Fornasier, G. Toscani, and F. Vecil. Particle, kinetic, and hydrodynamic models of swarming. *Mathematical modeling of collective behavior in socio-economic and life sciences*, pages 297–336, 2010.

[12] J. A. Carrillo and J. Soler. On the initial value problem for the Vlasov-Poisson-Fokker-Planck system with initial data in l^p spaces. *Mathematical methods in the applied sciences*, 18(10):825–839, 1995.

[13] J. A. Carrillo and J. Soler. On the Vlasov-Poisson-Fokker-Planck equations with measures in Morrey spaces as initial data. *Journal of Mathematical Analysis and Applications*, 207(2):475–495, 1997.

[14] A. J. Chorin. Numerical study of slightly viscous flow. *Journal of fluid mechanics*, 57(04):785–796, 1973.

[15] P. Degond. Global existence of smooth solutions for the Vlasov-Fokker-Planck equation in 1 and 2 space dimensions. In *Annales scientifiques de l'École Normale Supérieure*, volume 19, pages 519–542, 1986.

[16] R. L. Dobrushin. Vlasov equations. *Functional Analysis and Its Applications*, 13(2):115–123, 1979.

[17] R. C. Fetecau, H. Huang, and W. Sun. Propagation of chaos for the Keller-Segel equation over bounded domains. *arXiv preprint arXiv:1802.09668*, 2018.

[18] N. Fournier and A. Guillin. On the rate of convergence in wasserstein distance of the empirical measure. *Probability Theory and Related Fields*, 162(3-4):707–738, 2015.

[19] N. Fournier, M. Hauray, and S. Mischler. Propagation of chaos for the 2d viscous vortex model. *Journal of the European Mathematical Society*, 16(7):1423–1466, 2014.

[20] J. Goodman. Convergence of the random vortex method. *Communications on Pure and Applied Mathematics*, 40(2):189–220, 1987.

[21] O. Hald and V. M. Del Prete. Convergence of vortex methods for Euler's equations. *Mathematics of Computation*, 32(143):791–809, 1978.

[22] O. H. Hald. Convergence of vortex methods for Euler's equations. II. *SIAM Journal on Numerical Analysis*, 16(5):726–755, 1979.

[23] D. Henry. *Geometric theory of semilinear parabolic equations*, volume 840. Springer, 2006.

[24] E. Horst. Global strong solutions of Vlasov's equation-necessary and sufficient conditions for their existence. *Banach Center Publications*, 1(19):143–153, 1987.

[25] H. Huang and J.-G. Liu. Discrete-in-time random particle blob method for the Keller-Segel equation and convergence analysis. *Communication in Mathematical Sciences*, 15(7):1821–1842, 2017.

[26] H. Huang and J.-G. Liu. Error estimate of a random particle blob method for the Keller-Segel equation. *Mathematics of Computation*, 86:2719–2744, 2017.

[27] H. Huang, J.-G. Liu, and J. Lu. Learning interacting particle systems: diffusion parameter estimation for aggregation equations. *arXiv preprint arXiv:1802.02267*, 2018.

[28] P.-E. Jabin. A review of the mean field limits for Vlasov equations. *Kinet. Relat. Models*, 7(4):661–711, 2014.

[29] 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.

[30] P.-E. Jabin and Z. Wang. Mean field limit and propagation of chaos for Vlasov systems with bounded forces. *Journal of Functional Analysis*, 271(12):3588–3627, 2016.

[31] P.-E. Jabin and Z. Wang. Mean field limit for stochastic particle systems. In *Active Particles, Volume 1*, pages 379–402. Springer, 2017.

[32] E. F. Keller and L. A. Segel. Initiation of slime mold aggregation viewed as an instability. *Journal of Theoretical Biology*, 26(3):399–415, 1970.

[33] D. Lazarovici and P. Pickl. A mean field limit for the Vlasov-Poisson system. *Archive for Rational Mechanics and Analysis*, 225:1201–1231, 2017.

[34] P.-L. Lions and B. Perthame. Propagation of moments and regularity for the 3-dimensional Vlasov-Poisson system. *Inventiones mathematicae*, 105(1):415–430, 1991.

[35] J.-G. Liu and R. Yang. Propagation of chaos for large Brownian particle system with Coulomb interaction. *Research in the Mathematical Sciences*, 3(1):40, 2016.

[36] J.-G. Liu and Y. Zhang. Convergence of diffusion-drift many particle systems in probability under Sobolevnorm. In P. Gonçalves and A. J. Soares, editors, *From Particle Systems to Partial Differential Equations-III.* Springer, 2016.

[37] G. Loeper. Uniqueness of the solution to the Vlasov-Poisson system with bounded density. *Journal de mathématiques pures et appliquées*, 86(1):68–79, 2006.

[38] D.-G. Long. Convergence of the random vortex method in two dimensions. *Journal of the American Mathematical Society*, 1(4):779–804, 1988.

[39] C. Marchioro and M. Pulvirenti. Vortex methods in two-dimensional fluid dynamics. *Lecture notes in physics*, 203:1–137, 1984.

[40] K. Ono and W. A. Strauss. Regular solutions of the Vlasov-Poisson-Fokker-Planck system. *Discrete and Continuous Dynamical Systems*, 6(4):751–772, 2000.

[41] H. Osada. Propagation of chaos for the two dimensional Navier-Stokes equation. *Proceedings of the Japan Academy. Series A Mathematical sciences*, 62(1):8–11, 1986.

[42] C. S. Patlak. Random walk with persistence and external bias. *The bulletin of mathematical biophysics*, 15(3):311–338, 1953.

[43] M. Pulvirenti and C. Simeoni. l^∞-estimates for the Vlasov-Poisson-Fokker-Planck equation. *Mathematical methods in the applied sciences*, 23(10):923–935, 2000.

[44] J. Soler, J. A. Carrillo, and L. L. Bonilla. Asymptotic behavior of an initial-boundary value problem for the Vlasov-Poisson-Fokker-Planck system. *SIAM Journal on Applied Mathematics*, 57(5):1343–1372, 1997.

[45] H. Spohn. *Dynamics of charged particles and their radiation field*. Cambridge university press, 2004.

[46] H. D. Victory and B. P. O'Dwyer. On classical solutions of Vlasov-Poisson Fokker-Planck systems. *Indiana University mathematics journal*, 39(1):105–155, 1990.

[47] C. Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008.