

Propagation of chaos for the Vlasov–Poisson–Fokker–Planck equation with a polynomial cut-off

José A. Carrillo

*Department of Mathematics
Imperial College London, London SW7 2AZ, UK
carrillo@imperial.ac.uk*

Young-Pil Choi

*Department of Mathematics and
Institute of Applied Mathematics, Inha University
Incheon 402-751, Republic of Korea
ypchoi@inha.ac.kr*

Samir Salem

*CEREMADE Paris Dauphine
Place du Maréchal De Lattre De
Tassigny 75775 Paris CEDEX 16, France
salem@ceremade.dauphine.fr*

Received 5 February 2018

Revised 26 April 2018

Accepted 6 May 2018

Published 30 August 2018

We consider a N -particle system interacting through the Newtonian potential with a polynomial cut-off in the presence of noise in velocity. We rigorously prove the propagation of chaos for this interacting stochastic particle system. Taking the cut-off like $N^{-\delta}$ with $\delta < 1/d$ in the force, we provide a quantitative error estimate between the empirical measure associated to that N -particle system and the solutions of the d -dimensional Vlasov–Poisson–Fokker–Planck (VPFP) system. We also study the propagation of chaos for the Vlasov–Fokker–Planck equation with less singular interaction forces than the Newtonian one.

Keywords: Vlasov–Poisson equation; propagation of chaos; concentration inequalities; quantitative estimates; weak–strong stability.

Mathematics Subject Classification 2010: 35Q83, 35Q84, 60K35, 65C35

This is an Open Access article published by World Scientific Publishing Company. It is distributed under the terms of the Creative Commons Attribution 4.0 (CC-BY) License. Further distribution of this work is permitted, provided the original work is properly cited.

1. Introduction

The starting point is the classical Newton dynamics for point particles interacting through the interaction force F in the presence of noise:

$$\begin{cases} dX_t^i = V_t^i dt, \\ dV_t^i = \frac{1}{N} \sum_{j=1}^N F(X_t^i - X_t^j) dt + \sqrt{2\sigma} dB_t^i, \end{cases} \quad i = 1, \dots, N, \quad t > 0, \quad (1.1)$$

where $(\mathcal{X}_t, \mathcal{V}_t) := (X_t^1, \dots, X_t^N, V_t^1, \dots, V_t^N) \in \mathbb{R}^{2dN}$ are positions and velocities of the particles in \mathbb{R}^d , and $\sigma > 0$ is the noise strength. Here, $\{(B_t^i)_{t \geq 0}\}_{i=1}^N$ are N independent d -dimensional Brownian motions constructed over some probability space not specified for simplicity. There are many examples of physical systems of the form (1.1). Most classical examples of interaction kernels are the Coulombian or gravitational forces:

$$F(x) = \xi \frac{x}{|x|^d}, \quad \text{where } \xi = \begin{cases} +1 & \text{for plasma problems,} \\ -1 & \text{for astrophysics problems.} \end{cases} \quad (1.2)$$

We are interested in the rigorous derivation of the Vlasov–Poisson–Fokker–Planck (VPFP) equation as $N \rightarrow \infty$:

$$\begin{cases} \partial_t f_t + v \cdot \nabla_x f_t + (F * \rho_t) \cdot \nabla_v f_t = \sigma \Delta_v f_t, & (x, v) \in \mathbb{R}^d \times \mathbb{R}^d, \quad t > 0, \\ \rho_t(x) = \int_{\mathbb{R}^d} f_t(x, v) dv, \end{cases} \quad (1.3)$$

with the Poisson force $F(x) = \xi x/|x|^d$, $\xi = \pm 1$, where $f = f(x, v, t)$ is the probability density function at phase space (x, v) at time t . Note that the VPFP system is the formal mean-field limit equation associated to (1.1). This is a classical kinetic equation whose well-posedness and qualitative properties were studied in different settings: classical solutions [37, 3], weak solutions [4, 8, 9, 36, 24] and the references therein. We refer to [17, 32] for general theory on these problems.

The rigorous proof of the mean-field limit in the one-dimensional case for (1.3) with the force F given in (1.2) was obtained in [19, 23]. In [26], the propagation of chaos for the system (1.3) with bounded forces, i.e. $F \in L^\infty(\mathbb{R}^d)$ is studied based on relative entropy arguments. The intimately related classical Vlasov–Poisson system has also received lots of attention in the last years. This system corresponds to (1.1) and (1.3) with $\sigma = 0$. The propagation of chaos has been shown in [20, 21] for the less singular case, i.e. $|F(x)| \leq |x|^{-\alpha}$ with $\alpha < 1$. More recently, the physically relevant case of the propagation of chaos for the Newtonian potential with polynomial cut-off has been obtained in [27].

The propagation of chaos and its consequence, the rigorous proof of the mean-field limit, are questions of nowadays importance in many other problems; for

instance, in kinetic models of collective behavior [2, 5, 7, 10–12] with or without noise in velocity, see also [13, 30] for more general types of equations. However, in most of these applications, the singularity of the kernels is much better behaved. The case of Hölder interaction with Hölder exponent greater than $2/3$ was also recently treated in [25]. The main motivation of our present work is to obtain the propagation of chaos in the degenerate diffusion setting, as the noise acts only on velocity, in the classical case of Newtonian interaction with cut-off.

Finally, let us point out that most of the propagation of chaos results for nonlinear nonlocal conservation equations with non-Lipschitz interaction, concern first-order models. Without diffusion, the mean field limit for the aggregation equation with interaction less singular than the Newtonian is studied in [6]. In one dimension, the case of all power-law singularities was treated in [1] (also with noise) while a corresponding result in higher dimensions was recently obtained in [14, 35] for the repulsive case. With diffusion, the case of Hölder interaction is treated in [25] by quantitative means leading to explicit rates of convergence. For the 2D Navier–Stokes equation in vortex formulation, the propagation of chaos is obtained in [16] by compactness, thanks to result of [22]. Similar techniques have been applied to the case of the diffusion dominated 2D Keller–Segel equations in [18, 34] and geometrical constraint interactions with reflecting diffusions in [10].

Let us be more precise about the different approximation levels involved in the propagation of chaos. Consider the regularized particle system

$$\begin{cases} dX_t^{i,N} = V_t^{i,N} dt, \\ dV_t^{i,N} = \frac{1}{N} \sum_{j=1}^N F_\delta^N(X_t^{i,N} - X_t^{j,N}) dt + \sqrt{2\sigma} dB_t^{i,N}, \end{cases} \quad i = 1, \dots, N, \quad t > 0, \tag{1.4}$$

with the initial data $(X^{i,N}(0), V^{i,N}(0)) = (X_0^{i,N}, V_0^{i,N})$ for $i = 1, \dots, N$, where the interaction force $F_\delta^N, \delta > 0$ is given by

$$F_\delta^N(x) := \xi \frac{x}{(|x| \vee N^{-\delta})^d} \quad \text{and} \quad F_\delta^N(0) = 0,$$

where $a \vee b := \max\{a, b\}$. Note that this system of stochastic differential equations (SDEs) (1.4) has a unique strong solution by the standard theorem for SDEs. It is worth mentioning that the sign of force term is not important for the propagation of chaos provided that we analyze it on a finite time interval. Thus, we only focus on the case $\xi = 1$.

Our main purpose is to provide a quantitative error estimates between the solutions to the particle system (1.4) and the VFPF equation (1.3), which extends the work and method in [27] to the case with noise. For this, we need to introduce an intermediate system of independent copies of *nonlinear SDEs with cut-off*

given by

$$\begin{cases} dY_t^{i,N} = W_t^{i,N} dt, & i = 1, \dots, N, \quad t > 0, \\ dW_t^{i,N} = F_\delta^N * \rho_t^N(Y_t^{i,N}) dt + \sqrt{2\sigma} dB_t^{i,N}, & \rho_t^N = \mathcal{L}(Y_t^{i,N}), \\ (Y_0^{i,N}, W_0^{i,N}) = (X_0^{i,N}, V_0^{i,N}) & \text{for } i = 1, \dots, N, \end{cases} \quad (1.5)$$

where $\rho_t^N = \int_{\mathbb{R}^d} f_t^N dv$, and $f_t^N = \mathcal{L}(Y_t^{i,N}, W_t^{i,N})$, for all $i = 1, \dots, N$, is the global-in-time weak solution to the regularized VFPF system

$$\partial_t f_t^N + v \cdot \nabla_x f_t^N + (F_\delta^N * \rho_t^N) \cdot \nabla_v f_t^N = \sigma \Delta_v f_t^N, \quad (x, v) \in \mathbb{R}^d \times \mathbb{R}^d, \quad t > 0, \quad (1.6)$$

with the initial data $f_0^N = \mathcal{L}(Y_0^{i,N}, W_0^{i,N})$, $i = 1, \dots, N$. Due to the regularization, we can easily obtain the global existence of weak solutions to Eq. (1.6) under suitable assumptions on the initial data. Solving the system (1.5) with the given ρ_t^N to get the existence and uniqueness of solutions to the system (1.5) is by now standard.

Before stating our main result in this paper, we need to introduce some notions and notations. We define the empirical measure μ_t^N associated to a solution to the particle system (1.4) and ν_t^N the one associated to the nonlinear independent particle (1.5) as

$$\begin{aligned} \mu_t^N &= \frac{1}{N} \sum_{i=1}^N \delta_{(X_t^{i,N}, V_t^{i,N})} \quad \text{and} \\ \nu_t^N &= \frac{1}{N} \sum_{i=1}^N \delta_{(Y_t^{i,N}, W_t^{i,N})}, \end{aligned}$$

respectively. For a function f and $p \in [1, \infty]$, $\|f\|_p$ represents the usual L^p -norm for functions in $L^p(\mathbb{R}^{2d})$, and $\|f\|_{p,q} := \|f\|_p + \|f\|_q$ for $p, q \in [1, \infty]$. For $T > 0$, $L^p(0, T, E)$ is the set of L^p functions from an interval $(0, T)$ to a Banach space E . $\mathcal{P}(\mathbb{R}^{2d})$ and $\mathcal{P}_p(\mathbb{R}^{2d})$ stand for the sets of all probability measures and probability measures with finite moments of order $p \in [1, \infty)$. We consider the Wasserstein distance of order p , \mathcal{W}_p , $p \in [1, \infty]$, on $\mathcal{P}_p(\mathbb{R}^{2d})$, see [38] for classical definitions and properties of optimal transport distances. We also denote by C a generic positive constant independent of N .

The main result of this work gives an explicit estimate on the decay of the distance between the empirical measure of the interacting particle system μ_t^N and the solution of the VFPF equation f_t in optimal transport distances in the probability sense.

Theorem 1.1. *Let $T > 0, d > 1$ and $\delta < 1/d$. Let $(X_0^{i,N}, V_0^{i,N})_{i=1, \dots, N}$ be N independent random variables with law f_0 . Let f_t and f_t^N be the solutions to the VFPF equation (1.3) and its regularized version (1.6), respectively, up to time $T > 0$.*

Assume further that $f, f^N \in L^\infty(0, T; (L^1 \cap L^\infty)(\mathbb{R}^{2d})) \cap \mathcal{C}([0, T]; \mathcal{P}_q(\mathbb{R}^{2d}))$ with the same initial data $f_0 \in (L^1 \cap L^\infty \cap \mathcal{P}_q)(\mathbb{R}^{2d})$ for some $q \geq 2$, and that their respective spatial densities ρ_t and ρ_t^N satisfy

$$\int_0^T \|\rho_t\|_\infty dt \leq C_0 \quad \text{and} \quad \sup_{N \in \mathbb{N}} \int_0^T \|\rho_t^N\|_\infty dt \leq C_0,$$

with C_0 independent of N . Then, given any $p \in [1, 2q)$, $\gamma < (\frac{1}{2(d \vee p)} \wedge \delta)$, and $\varepsilon \in (0, q - \frac{p}{1-p\gamma})$, the estimate

$$\sup_{0 \leq t \leq T} \mathbb{P}(\mathcal{W}_p(\mu_t^N, f_t) \geq CN^{-\gamma}) \leq CN^{\frac{1}{p}(1 - \frac{(1-p\gamma)(q-\varepsilon)}{p})} \quad \text{for } N \text{ large enough,}$$

holds for some constant $C > 0$ depending only on $d, T, p, q, \varepsilon, f_0$, and C_0 . Here $a \wedge b := \min\{a, b\}$.

We next present the existence and uniqueness of weak solutions for Eq. (1.3), which gives all the required *a priori* hypotheses in Theorem 1.1 for the propagation of chaos, see also Lemma 4.1.

Theorem 1.2. *Let $p \geq 1$ and (Y_0, W_0) be independent of $(B_t)_{t \geq 0}$ with law f_0 . Suppose that f_0 satisfies*

$$f_0 \in (L^1 \cap L^\infty \cap \mathcal{P}_p)(\mathbb{R}^{2d}) \quad \text{and} \quad f_0(x, v) \leq C\langle v \rangle^{-\gamma},$$

for some $\gamma > d$ and $C > 0$, where $\langle v \rangle = \sqrt{1 + |v|^2}$. Then there exists a unique solution to the following nonlinear SDEs:

$$\begin{cases} dY_t = W_t dt, \\ dW_t = (F * \rho_t)(Y_t) dt + \sqrt{2\sigma} dB_t, \quad \rho_t = \mathcal{L}(Y_t), \\ \mathcal{L}(Y_0, W_0) = f_0, \end{cases} \quad (1.7)$$

where $\rho_t = \int_{\mathbb{R}^d} f_t dv$ and $f_t \in L^\infty(0, T_*; (L^1 \cap L^\infty)(\mathbb{R}^{2d})) \cap \mathcal{C}([0, T_*]; \mathcal{P}_p(\mathbb{R}^{2d}))$ is the unique weak solution to the equation (1.3) satisfying $\rho_t \in L^\infty(0, T_*; L^\infty(\mathbb{R}^d))$ for some $T_* > 0$.

Let us briefly explain the strategy of the proof of this theorem.

- First, using the techniques introduced in [27], one can estimate the probability that the error between the empirical measure μ_t^N associated to the particle system (1.1) and the empirical measure ν_t^N associated to the nonlinear independent particle system with cut-off (1.4) exceeds the threshold $N^{-\gamma}$. In this way, it is shown that this probability decreases faster than any negative power of N (see Lemma 3.2).
- Then using a concentration inequality of [15], one can obtain bounds on the probability that the error between ν_t^N empirical measure associated to i.i.d. random variables and their law f_t^N solution at time t to Eq. (1.6) exceeds the threshold

$N^{-\gamma}$. The main results in [15] provide an optimal rate of convergence which is of order of some negative power of N . (See Proposition 2.1).

- Finally, we show that the error between the solution to Eq. (1.6) and the solution to Eq. (1.3) never exceeds the threshold $N^{-\gamma}$, for N large enough. Contrary to [27, Proposition 9.1], this error has to be estimated in \mathcal{W}_p distance with $1 \leq p < \infty$ due to the presence of noise in velocity (See Proposition 3.1).

Our main contributions in comparison to the noiseless case treated in [27] are the following. On the one hand, we provide a convergence result of the solution to Eq. (1.6) to the solution to Eq. (1.3), in \mathcal{W}_p metric for $p \in [1, \infty)$, which is crucial since the support of f_t^N and f_t are not compactly supported in our present case. On the other hand, we provide a well-posedness result for Eq. (1.3). This requires to show that the spatial density ρ_t of solution to this equation lies in $L^1(0, T; L^\infty(\mathbb{R}^d))$. For the case without diffusion, i.e. Vlasov–Poisson system, characteristic methods can be used to get a uniform bound on the spatial density, which leads to a uniqueness of solutions, under suitable assumptions on the initial density [28, 29, 31]. However, the presence of diffusion makes it more complicated. In [33], an uniform-in-time L^∞ -bound of the spatial density is obtained by means of the stochastic characteristic method under the assumptions on compactly supported initial density f_0 in velocity. We also provide a simple proof of the local-in-time L^∞ propagation by employing Feynman–Kac’s formula assuming only that the initial data has a polynomial decay in velocity (see Lemma 4.1 together with Theorem 1.2). Note that obtaining the bound estimate of ρ_t in $L^\infty(\mathbb{R}^d)$ is equivalent to the one for $\|\rho_t^N\|_\infty$ in N .

The rest of this paper is organized as follows. In Sec. 2, we deal with some preliminary materials introduced in [27]. Section 3 contains the new key estimate on the \mathcal{W}_p stability between the solutions the VPFP equation with and without cut-off. Section 4 shows the well-posedness of solution to the VPFP equation with the assumed regularity in Theorem 1.1, i.e. the assumed L^∞ bound on the spatial density. Finally, Sec. 5 contains generalizations of this result for less singular kernels.

2. Preliminaries

In this section, we provide estimates for the force fields in Eqs. (1.3) and (1.6) in $L^1(0, T; L^\infty(\mathbb{R}^d))$ and the q th moment estimate for the solutions of that. We also recall several useful estimates on the force fields whose proofs can be found in [27].

Under the assumptions on the spatial densities ρ_t and ρ_t^N in Theorem 1.1, we can easily find

$$\int_0^T \|F * \rho_t\|_\infty dt < \infty \quad \text{and} \quad \sup_{N \in \mathbb{N}} \int_0^T \|F_\delta^N * \rho_t^N\|_\infty dt < \infty,$$

due to $\|F * \rho_t\|_\infty \leq C\|\rho_t\|_{1,\infty}$ and $\|F_\delta^N * \rho_t^N\|_\infty \leq C\|\rho_t^N\|_{1,\infty}$ with $C > 0$ independent of N . These estimates together with straightforward computations yield that for some $q \geq 2$

$$\begin{aligned} & \frac{d}{dt} \int_{\mathbb{R}^{2d}} (|x|^q + |v|^q) f dx dv \\ &= q \int_{\mathbb{R}^{2d}} |x|^{q-2} x \cdot v f dx dv + q \int_{\mathbb{R}^{2d}} |v|^{q-2} v \cdot (F * \rho) f dx dv \\ & \quad + \sigma q(q-2+d) \int_{\mathbb{R}^{2d}} |v|^{q-2} f dx dv \\ & \leq C + C \int_{\mathbb{R}^{2d}} |x|^q f dx dv + C(1 + \|\rho_t\|_{1,\infty}) \int_{\mathbb{R}^{2d}} |v|^q f dx dv. \end{aligned}$$

Thus, we obtain

$$\sup_{0 \leq t \leq T} \int_{\mathbb{R}^{2d}} (|x|^q + |v|^q) f_t dx dv \leq C \sup_{0 \leq t \leq T} \int_{\mathbb{R}^{2d}} (|x|^q + |v|^q) f_0 dx dv,$$

and similarly, we also have

$$\sup_{N \in \mathbb{N}} \sup_{0 \leq t \leq T} \int_{\mathbb{R}^{2d}} (|x|^q + |v|^q) f_t^N dx dv \leq C \sup_{0 \leq t \leq T} \int_{\mathbb{R}^{2d}} (|x|^q + |v|^q) f_0 dx dv,$$

where $C > 0$ is independent of N .

Let us now recall the interaction force with a cut-off:

$$F_\delta^N(x) := \frac{x}{(|x| \vee N^{-\delta})^d},$$

and define $l_\delta^N(x)$ by

$$l_\delta^N(x) := \begin{cases} \frac{1}{|x|^d} & \text{if } |x| \geq dN^{-\delta}, \\ N^{d\delta} & \text{otherwise.} \end{cases}$$

We drop the subscript δ in F_δ^N and l_δ^N for notational simplicity in the rest of paper, i.e. $F_\delta^N = F^N$ and $l_\delta^N = l^N$.

Lemma 2.1. *Let $d > 1$ be given.*

(1) *There exists a constant C , which depends only on d , such that*

$$|F^N(x) - F^N(x+z)| \leq Cl^N(x)|z|,$$

for any $x, z \in \mathbb{R}^d$ with $|z| \leq (d-1)N^{-\delta}$.

(2) *There exists a constant $C > 0$ independent of N such that*

$$\|l^N * \rho\|_\infty \leq C \ln N \|\rho\|_{1,\infty} \quad \text{and} \quad \|\nabla F^N * \rho\|_\infty \leq C \ln N \|\rho\|_{1,\infty}.$$

Proof. See [27, Lemmas 6.1 and 6.3]. □

Then we recast here some laws of large number like estimates. For $\kappa, \delta > 0$, we set $h : \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that

$$|h(x)| \leq c_0(N^{\kappa\delta} \wedge |x|^{-\kappa}). \tag{2.1}$$

Lemma 2.2. *Let (Y_1, \dots, Y_N) be i.i.d. random variables of law $\rho \in L^\infty(\mathbb{R}^d)$, and define the associated empirical measure $\rho_N = \frac{1}{N} \sum_{i=1}^N \delta_{Y_i}$. Suppose that $0 < \varepsilon := 2\kappa\delta + (1 - d\delta)\mathbf{1}_{1 > d\delta} < 2$. Then, for all integer $m > \frac{1}{2-\varepsilon}$, there exist $\gamma_m, C_m > 0$ such that*

$$\mathbb{E} \left[\sup_{1 \leq i \leq N} |h * \rho_N(Y_i) - h * \rho(Y_i)|^{2m} \right] \leq C_m N^{-\gamma_m},$$

where $\gamma_m = (2 - \varepsilon)m - 1$.

Proof. See the proof of [27, Proposition 7.2]. □

We conclude this section by recalling some concentration inequalities in Proposition 2.1 whose proof can be found in [15, Theorem 2].

Proposition 2.1. *Let $N \geq 1$ and (Y_1, \dots, Y_N) be N i.i.d. random variables of law $\rho \in \mathcal{P}_q(\mathbb{R}^{2d})$ for some $q > 0$. Let $\rho_N := \frac{1}{N} \sum_{i=1}^N \delta_{Y_i}$. Then, for any $p \in (0, q/2), \varepsilon \in (0, q)$, there exist constants $C, c > 0$ depending only on d, p, q, ε and the q -order moment bound of ρ such that*

$$\mathbb{P}(\mathcal{W}_p^p(\rho_N, \rho) \geq x) \leq CN(Nx)^{-\frac{q-\varepsilon}{p}} + a(N, x)\mathbf{1}_{x \leq 1}, \quad \text{for any } x > 0,$$

where $a(N, x)$ is defined by

$$-\log a(N, x) = \begin{cases} cNx^2 & \text{if } p > d, \\ cN \left(\frac{x}{\ln \left(\frac{2+1}{x} \right)} \right)^2 & \text{if } p = d, \\ cNx^{2d/p} & \text{if } 1 \leq p < d. \end{cases}$$

3. Propagation of Chaos: Proof of Theorem 1.1

In this section, we provide the details of the proof of Theorem 1.1. We begin with the following Lipschitz estimates on the force fields F and F^N .

Lemma 3.1. *For $x, y \in \mathbb{R}^d$, we get*

$$|F(x) - F(y)| \leq C_0|x - y| \left(\frac{1}{|x|^d} + \frac{1}{|y|^d} \right),$$

and

$$|F^N(x) - F^N(y)| \leq C_0|x - y| \left(\frac{1}{(|x| \vee N^{-\delta})^d} + \frac{1}{(|y| \vee N^{-\delta})^d} \right),$$

where C_0 is a positive constant depending only on d .

Proof. The first assertion is straightforward. For the proof of the second one, we consider three cases as follows:

(i) $|x|, |y| \geq N^{-\delta}$: In this case, we get $F^N(x) = F(x)$ and thus it is clear to obtain

$$|F^N(x) - F^N(y)| = |F(x) - F(y)| \leq (|\nabla F(x)| \vee |\nabla F(y)|)|x - y|.$$

Note that

$$\nabla F(x) = \frac{1}{|x|^d} I_d - d \frac{x \otimes x}{|x|^{d+2}}, \quad \text{i.e. } |\nabla F(x)| \leq \frac{C}{|x|^d} = \frac{C}{(|x| \vee N^{-\delta})^d},$$

where $C > 0$ depends only on d . This yields

$$|F^N(x) - F^N(y)| \leq \frac{C|x - y|}{(|x| \wedge |y|)^d} \leq C|x - y| \left(\frac{1}{(|x| \vee N^{-\delta})^d} + \frac{1}{(|y| \vee N^{-\delta})^d} \right).$$

(ii) $|x|, |y| \leq N^{-\delta}$: By definition of F^N , we find

$$|F^N(x) - F^N(y)| = \frac{|x - y|}{N^{-\delta d}} \leq |x - y| \left(\frac{1}{(|x| \vee N^{-\delta})^d} + \frac{1}{(|y| \vee N^{-\delta})^d} \right).$$

(iii) $|x| < N^{-\delta} \leq |y|$ or $|y| < N^{-\delta} \leq |x|$: For $|x| < N^{-\delta} \leq |y|$, let us define \tilde{x} as the intersection between the line segment $[x, y]$ and the ball $B(0, N^{-\delta})$. Then we get

$$|x| < |\tilde{x}| = N^{-\delta} \leq |y| \quad \text{and} \quad |x - \tilde{x}| + |\tilde{x} - y| = |x - y|.$$

By employing the similar arguments as in the previous cases (i) and (ii), we have

$$\begin{aligned} |F^N(x) - F^N(\tilde{x})| &\leq \frac{C|x - \tilde{x}|}{N^{-\delta d}} \quad \text{and} \\ |F^N(\tilde{x}) - F^N(y)| &\leq \frac{C|\tilde{x} - y|}{(|\tilde{x}| \wedge |y|)^d} = \frac{C|\tilde{x} - y|}{|\tilde{x}|^d} = \frac{C|\tilde{x} - y|}{N^{-\delta d}}. \end{aligned}$$

This implies that for $|x| < N^{-\delta} \leq |y|$

$$\begin{aligned} |F^N(x) - F^N(y)| &\leq |F^N(x) - F^N(\tilde{x})| + |F^N(\tilde{x}) - F^N(y)| \\ &\leq \frac{C|x - y|}{N^{-\delta d}} = \frac{C|x - y|}{(|x| \vee N^{-\delta})^d}. \end{aligned}$$

Similarly, we have

$$|F^N(x) - F^N(y)| \leq \frac{C|x - y|}{(|y| \vee N^{-\delta})^d} \quad \text{for } |y| < N^{-\delta} \leq |x|.$$

Combining the above all cases, we conclude to the desired result. □

Set

$$G := |X - Y| + |V - W| \quad \text{and} \quad G_N := \sqrt{\ln N}|X - Y| + |V - W|.$$

Proposition 3.1. *Let (X, V) and (Y, W) be two random variables of law f_1 and f_2 respectively, such that their first marginal are ρ_1 and ρ_2 , respectively. Let (\bar{X}, \bar{Y}) be an independent copy of (X, Y) . Suppose $\rho_1, \rho_2 \in L^\infty(\mathbb{R}^d)$ and $N \geq e$. Then we have*

$$\begin{aligned} & \mathbb{E}[|F(X - \bar{X}) - F(Y - \bar{Y})|G^{p-1}] \\ & \leq C(\|\rho_1\|_\infty + \|\rho_2\|_\infty)\mathbb{E}[G^p] \left(1 - \frac{1}{p} \ln^- \mathbb{E}[G^p]\right) \end{aligned}$$

and

$$\mathbb{E}[|F^N(X - \bar{X}) - F^N(Y - \bar{Y})|G_N^{p-1}] \leq C\sqrt{\ln N}(\|\rho_1\|_\infty + \|\rho_2\|_\infty)\mathbb{E}[G_N^p],$$

for $p \geq 1$, where $\ln^-(x) := \ln(x) \wedge 0$ and the constant $C > 0$ depends only on d .

Proof. Let us denote by α_d the surface area of unit ball in \mathbb{R}^d .

Estimate for the non-cut-off force field. First, we note that

$$\begin{aligned} & |F(X - \bar{X}) - F(Y - \bar{Y})| \\ & \leq \begin{cases} C(|X - Y| + |\bar{X} - \bar{Y}|) \left(\frac{1}{|X - \bar{X}|^d} + \frac{1}{|Y - \bar{Y}|^d} \right), \\ \frac{1}{|X - \bar{X}|^{d-1}} + \frac{1}{|Y - \bar{Y}|^{d-1}}. \end{cases} \end{aligned}$$

This yields that for any $r > 0$

$$\begin{aligned} & \mathbb{E}[|F(X - \bar{X}) - F(Y - \bar{Y})|G^{p-1}] \\ & \leq \mathbb{E} \left[\left(\frac{1}{|X - \bar{X}|^{d-1}} + \frac{1}{|Y - \bar{Y}|^{d-1}} \right) \mathbf{1}_{|X - \bar{X}| \wedge |Y - \bar{Y}| \leq r} G^{p-1} \right] \\ & \quad + C\mathbb{E} \left[(|X - Y| + |\bar{X} - \bar{Y}|) \left(\frac{1}{|X - \bar{X}|^d} + \frac{1}{|Y - \bar{Y}|^d} \right) \mathbf{1}_{|X - \bar{X}| \wedge |Y - \bar{Y}| > r} G^{p-1} \right] \\ & =: I_1 + I_2. \end{aligned}$$

◇ Estimate of I_1 : Note that the event $\{|X - \bar{X}| \wedge |Y - \bar{Y}| \leq r\}$ can be partitioned as

$$\begin{aligned} & \{|X - \bar{X}| \wedge |Y - \bar{Y}| \leq r\} \\ &= \{|X - \bar{X}| \vee |Y - \bar{Y}| \leq r\} \cup \{|X - \bar{X}| > r \geq |Y - \bar{Y}|\} \\ & \cup \{|X - \bar{X}| \leq r < |Y - \bar{Y}|\}. \end{aligned}$$

Taking into account this, we split I_1 into three terms:

$$\begin{aligned} I_1 &= \mathbb{E} \left[\left(\frac{1}{|X - \bar{X}|^{d-1}} + \frac{1}{|Y - \bar{Y}|^{d-1}} \right) \mathbf{1}_{|X - \bar{X}| \vee |Y - \bar{Y}| \leq r} G^{p-1} \right] \\ &+ \mathbb{E} \left[\left(\frac{1}{|X - \bar{X}|^{d-1}} + \frac{1}{|Y - \bar{Y}|^{d-1}} \right) \mathbf{1}_{|X - \bar{X}| > r \geq |Y - \bar{Y}|} G^{p-1} \right] \\ &+ \mathbb{E} \left[\left(\frac{1}{|X - \bar{X}|^{d-1}} + \frac{1}{|Y - \bar{Y}|^{d-1}} \right) \mathbf{1}_{|X - \bar{X}| \leq r < |Y - \bar{Y}|} G^{p-1} \right] \\ &=: I_1^1 + I_1^2 + I_1^3. \end{aligned}$$

For the estimate of I_1^1 , by separating in the notations the expectation with respect to the independent copies (\bar{X}, \bar{Y}) , from the expectation with respect to the random variables $(X, V), (Y, W)$, we get

$$\begin{aligned} I_1^1 &= \mathbb{E}_{(X,V),(Y,W)} \left[\mathbb{E}_{\bar{X},\bar{Y}} \left[\left(\frac{1}{|X - \bar{X}|^{d-1}} + \frac{1}{|Y - \bar{Y}|^{d-1}} \right) \mathbf{1}_{|X - \bar{X}| \vee |Y - \bar{Y}| \leq r} G^{p-1} \right] \right] \\ &\leq \mathbb{E}_{(X,V),(Y,W)} \left[\left(\int_{|X-x| \leq r} \frac{1}{|X-x|^{d-1}} \rho_1(dx) \right. \right. \\ & \quad \left. \left. + \int_{|Y-y| \leq r} \frac{1}{|Y-y|^{d-1}} \rho_2(dy) \right) G^{p-1} \right] \\ &\leq \alpha_d (\|\rho_1\|_\infty + \|\rho_2\|_\infty) r \mathbb{E}[G^{p-1}]. \end{aligned}$$

For I_1^2 , we obtain

$$\begin{aligned} I_1^2 &\leq \mathbb{E} \left[\frac{2}{|Y - \bar{Y}|^{d-1}} \mathbf{1}_{|Y - \bar{Y}| \leq r} G^{p-1} \right] \\ &= 2 \mathbb{E}_{(X,V),(Y,W)} \left[\left(\int_{|Y-y| \leq r} \frac{1}{|Y-y|^{d-1}} \rho_2(dy) \right) G^{p-1} \right] \\ &\leq 2 \alpha_d \|\rho_2\|_\infty r \mathbb{E}[G^{p-1}]. \end{aligned}$$

Similarly, we estimate I_1^3 as $I_1^3 \leq 2\alpha_d \|\rho_1\|_\infty r \mathbb{E}[G^{p-1}]$. Combining the above estimates, we have

$$I_1 \leq C(\|\rho_1\|_\infty + \|\rho_2\|_\infty) r \mathbb{E}[G^{p-1}] \leq C(\|\rho_1\|_\infty + \|\rho_2\|_\infty) r \mathbb{E}[G^p]^{(p-1)/p},$$

where $C > 0$ only depends on d .

◇ Estimate of I_2 : We decompose I_2 as

$$\begin{aligned} I_2 &= C \mathbb{E} \left[|X - Y| \left(\frac{1}{|X - \bar{X}|^d} + \frac{1}{|Y - \bar{Y}|^d} \right) \mathbf{1}_{|X - \bar{X}| \wedge |Y - \bar{Y}| > r} G^{p-1} \right] \\ &\quad + C \mathbb{E} \left[|\bar{X} - \bar{Y}| \left(\frac{1}{|X - \bar{X}|^d} + \frac{1}{|Y - \bar{Y}|^d} \right) \mathbf{1}_{|X - \bar{X}| \wedge |Y - \bar{Y}| > r} G^{p-1} \right] \\ &=: I_2^1 + I_2^2. \end{aligned}$$

First, we easily obtain

$$\begin{aligned} I_2^1 &= C \mathbb{E}_{(X,V),(Y,W)} \left[G^p \mathbb{E}_{\bar{X},\bar{Y}} \left[\left(\frac{1}{|X - \bar{X}|^d} + \frac{1}{|Y - \bar{Y}|^d} \right) \mathbf{1}_{|X - \bar{X}| \wedge |Y - \bar{Y}| > r} \right] \right] \\ &\leq C \mathbb{E}_{(X,V),(Y,W)} \left[G^p \left(\int \frac{1}{|X - x|^d} \mathbf{1}_{|X - x| \geq r} \rho_1(dx) \right. \right. \\ &\quad \left. \left. + \int \frac{1}{|Y - y|^d} \mathbf{1}_{|Y - y| \geq r} \rho_2(dy) \right) \right]. \end{aligned}$$

We then consider two cases: $r > 1$ and $0 < r \leq 1$. For $r \leq 1$, we get

$$\begin{aligned} \int \frac{1}{|X - x|^d} \mathbf{1}_{|X - x| \geq r} \rho_1(dx) &= \int \frac{1}{|X - x|^d} \mathbf{1}_{|X - x| > 1} \rho_1(dx) \\ &\quad + \int \frac{1}{|X - x|^d} \mathbf{1}_{|X - x| \in [r, 1]} \rho_1(dx) \\ &\leq \|\rho_1\|_1 + \alpha_d \|\rho_1\|_\infty \int_r^1 u^{-d} du \\ &\leq 1 - \alpha_d \|\rho_1\|_\infty \ln^- r. \end{aligned}$$

For the case $r > 1$, it is clear to obtain

$$\int \frac{1}{|X - x|^d} \mathbf{1}_{|X - x| \geq r} \rho_1(dx) \leq \|\rho_1\|_1 = 1.$$

This yields

$$I_2^1 \leq C(\|\rho_1\|_\infty + \|\rho_2\|_\infty) \mathbb{E}[G^p] (1 - \ln^- r),$$

where $C > 0$ only depends on d . For the term I_2^2 , we use Hölder inequality to find

$$\begin{aligned} I_2^2 &= C\mathbb{E} \left[|\bar{X} - \bar{Y}| \left(\frac{1}{|X - \bar{X}|^d} + \frac{1}{|Y - \bar{Y}|^d} \right)^{1/p} \right. \\ &\quad \left. \times \left(\left(\frac{1}{|X - \bar{X}|^d} + \frac{1}{|Y - \bar{Y}|^d} \right)^{1/p} G \right)^{p-1} \mathbf{1}_{|X - \bar{X}| \wedge |Y - \bar{Y}| > r} \right] \\ &\leq C\mathbb{E} \left[|\bar{X} - \bar{Y}|^p \left(\frac{1}{|X - \bar{X}|^d} + \frac{1}{|Y - \bar{Y}|^d} \right) \mathbf{1}_{|X - \bar{X}| \wedge |Y - \bar{Y}| > r} \right]^{1/p} \\ &\quad \times \mathbb{E} \left[\left(\frac{1}{|X - \bar{X}|^d} + \frac{1}{|Y - \bar{Y}|^d} \right) \mathbf{1}_{|X - \bar{X}| \wedge |Y - \bar{Y}| > r} G^p \right]^{(p-1)/p}. \end{aligned}$$

Similarly as before, we take the expectations on (X, Y) and (\bar{X}, \bar{Y}) for the first and second expectations in the above, respectively, to find

$$\begin{aligned} I_2^2 &\leq C(\|\rho_1\|_\infty + \|\rho_2\|_\infty)(1 - \ln^- r)\mathbb{E}[|X - Y|^p]^{1/p}\mathbb{E}[G^p]^{(p-1)/p} \\ &\leq C(\|\rho_1\|_\infty + \|\rho_2\|_\infty)(1 - \ln^- r)\mathbb{E}[G^p], \end{aligned}$$

due to $|X - Y| \leq G$, where $C > 0$ only depends on d . Thus, by putting all those estimates together and using Hölder inequality, we have for any $r > 0$

$$\begin{aligned} &\mathbb{E}[|F(X - \bar{X}) - F(Y - \bar{Y})|G^{p-1}] \\ &\leq C(\|\rho_1\|_\infty + \|\rho_2\|_\infty)(1 - \ln^- r)\mathbb{E}[G^p] + C(\|\rho_1\|_\infty + \|\rho_2\|_\infty)r\mathbb{E}[G^p]^{(p-1)/p}. \end{aligned}$$

Finally, we choose $r = \mathbb{E}[G^p]^{1/p}$ to obtain the desired result.

Estimate for the cut-off force field. Note that since F^N is continuous, we have

$$\begin{aligned} &\mathbb{E}[|F^N(X - \bar{X}) - F^N(Y - \bar{Y})|G_N^{p-1}] \\ &\leq C\mathbb{E}[(|X - Y| + |\bar{X} - \bar{Y}|)(|\nabla F^N(X - \bar{X})| + |\nabla F^N(Y - \bar{Y})|)G_N^{p-1}] \\ &\leq C\mathbb{E}[|X - Y|(|\nabla F^N(X - \bar{X})| + |\nabla F^N(Y - \bar{Y})|)G_N^{p-1}] \\ &\quad + C\mathbb{E}[|\bar{X} - \bar{Y}|(|\nabla F^N(X - \bar{X})| + |\nabla F^N(Y - \bar{Y})|)G_N^{p-1}] \\ &=: J_1 + J_2. \end{aligned}$$

As in the proof above, we first easily get

$$\begin{aligned} J_1 &\leq C\mathbb{E}_{(X, V), (Y, W)}[|X - Y|G_N^{p-1}\mathbb{E}_{\bar{X}, \bar{Y}}[|\nabla F^N(X - \bar{X})| + |\nabla F^N(Y - \bar{Y})|]] \\ &\leq C(\|\rho_1\|_\infty + \|\rho_2\|_\infty)\ln N\mathbb{E}[|X - Y|G_N^{p-1}] \\ &\leq C(\|\rho_1\|_\infty + \|\rho_2\|_\infty)\sqrt{\ln N}\mathbb{E}[G_N^p], \end{aligned}$$

where we used Lemma 2.1 and $\sqrt{\ln N}|X - Y| \leq G_N$. For the term J_2 , we again use the similar argument as before to find

$$J_2 \leq C\mathbb{E}[|\bar{X} - \bar{Y}|^p (|\nabla F^N(X - \bar{X})| + |\nabla F^N(Y - \bar{Y})|)]^{1/p} \times \mathbb{E}[(|\nabla F^N(X - \bar{X})| + |\nabla F^N(Y - \bar{Y})|)G_N^p]^{(p-1)/p}.$$

Taking the expectations on (X, Y) in the first expectation and on (\bar{X}, \bar{Y}) in the second one leads to the desired result. \square

We next estimate the error between solutions to the nonlinear SDEs and the one with cut-off given by

- *Nonlinear SDEs:*

$$\begin{cases} dY_t = W_t dt, \\ dW_t = (F * \rho_t)(Y_t) dt + \sqrt{2\sigma} dB_t, \quad \rho_t = \mathcal{L}(Y_t), \\ \mathcal{L}(Y_0, W_0) = f_0. \end{cases} \tag{3.1}$$

- *Nonlinear SDEs with cut-off:*

$$\begin{cases} dY_t^N = W_t^N dt, \\ dW_t^N = (F^N * \rho_t^N)(Y_t^N) dt + \sqrt{2\sigma} dB_t, \quad \rho_t^N = \mathcal{L}(Y_t^N), \\ \mathcal{L}(Y_0^N, W_0^N) = f_0. \end{cases} \tag{3.2}$$

Here, $\rho_t^N = \int_{\mathbb{R}^d} f_t^N dv$ and f_t^N is the global-in-time weak solution of Eq. (1.6). As mentioned in Sec. 1, for fixed $N > 0$, the global existence and uniqueness of solutions to (3.2) is ensured due to classical SDEs theory. At the moment, we assume the existence of solutions to the SDEs (3.1) and its associated PDE (1.3) up to a given time $T > 0$. We will give the details of that in Sec. 4.

Proposition 3.2. *For a given $T > 0$, let (Y_t, W_t) and (Y_t^N, W_t^N) be the solutions to the equations (3.1) and (3.2) for the same initial condition on the time interval $[0, T]$, respectively. Suppose that $\rho_t, \rho_t^N \in L^1(0, T; L^\infty(\mathbb{R}^d))$ and $\|\rho_t^N\|_{L^1(0, T; L^\infty)} \leq C$ with $C > 0$ independent of N . Then, for $N \geq e$ and $p \geq 1$, we have*

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} (\sqrt{\ln N} |Y_t^N - Y_t| + |W_t^N - W_t|)^p \right]^{1/p} \leq CN^{-\delta} \exp(C\sqrt{\ln N}),$$

where C is a positive constant independent of p and N .

Proof. For $p \geq 1$, we set

$$D_s := \sqrt{\ln N} |Y_s^N - Y_s| + |W_s^N - W_s| \quad \text{and} \quad \phi_t^N := \sup_{0 \leq s \leq t} D_s^p.$$

Then we estimate ϕ_t^N as

$$\begin{aligned} \phi_t^N &\leq p \int_0^t (\sqrt{\ln N} |W_s^N - W_s| + |F^N * \rho_s^N(Y_s^N) - F * \rho_s(Y_s)|) D_s^{p-1} ds \\ &\leq p \int_0^t \sqrt{\ln N} D_s^p + |F^N * (\rho_s^N(Y_s^N) - \rho_s(Y_s))| D_s^{p-1} \\ &\quad + |(F^N - F) * \rho_s(Y_s)| D_s^{p-1} ds. \end{aligned}$$

Taking the expectation on both sides of the above inequality and using Fubini's Theorem, we obtain

$$\begin{aligned} \mathbb{E}[\phi_t^N] &\leq p\sqrt{\ln N} \int_0^t \mathbb{E}[\phi_s^N] ds + p \int_0^t \mathbb{E}[|F^N * (\rho_s^N(Y_s^N) - \rho_s(Y_s))| D_s^{p-1}] ds \\ &\quad + p \int_0^t \mathbb{E}[|(F^N - F) * \rho_s(Y_s)| D_s^{p-1}] ds \\ &=: I_1 + I_2 + I_3, \end{aligned}$$

where we can directly use the cut-off force field estimate in Lemma 3.1 to estimate I_2 as

$$I_2 \leq Cp\sqrt{\ln N} \int_0^t (\|\rho_s^N\|_\infty + \|\rho_s\|_\infty) \mathbb{E}[D_s^p] ds,$$

where $C > 0$ only depends on d . For the estimate of I_3 , we easily find

$$\begin{aligned} |(F^N - F) * \rho_s(Y_s)| &\leq \int_{|Y_s - y| \leq N^{-\delta}} |Y_s - y| (|Y_s - y|^{-d} - N^{d\delta}) \rho_s(dy) \\ &\leq \int_{|Y_s - y| \leq N^{-\delta}} |Y_s - y|^{-(d-1)} \rho_s(dy) \leq \alpha_d \|\rho_s\|_\infty N^{-\delta}. \end{aligned}$$

This yields

$$\begin{aligned} I_3 &\leq p\alpha_d N^{-\delta} \int_0^t \|\rho_s\|_\infty \mathbb{E}[D_s^{p-1}] ds \leq p\alpha_d N^{-\delta} \int_0^t \|\rho_s\|_\infty \mathbb{E}[\phi_s^N]^{(p-1)/p} ds \\ &\leq (p-1)\alpha_d \int_0^t \|\rho_s\|_\infty \mathbb{E}[\phi_s^N] ds + \alpha_d N^{-p\delta} \int_0^t \|\rho_s\|_\infty ds, \end{aligned}$$

where we used Young's inequality for the last inequality. We then combine the above estimates to have

$$\mathbb{E}[\phi_t^N] \leq pC\sqrt{\ln N} \int_0^t (1 + \|\rho_s\|_\infty + \|\rho_s^N\|_\infty) \mathbb{E}[\phi_s^N] ds + \alpha_d N^{-p\delta} \int_0^t \|\rho_s\|_\infty ds,$$

where $C > 0$ only depends on d . We now apply Lemma A.1.1 with

$$\begin{aligned} f(t) &= \mathbb{E}[\phi_t^N], \quad g(t) = pC\sqrt{\ln N} (1 + \|\rho_s\|_\infty + \|\rho_s^N\|_\infty), \quad \text{and} \\ h(t) &= \alpha_d N^{-p\delta} \int_0^t \|\rho_s\|_\infty ds, \end{aligned}$$

to obtain

$$\begin{aligned} \mathbb{E}[\phi_t^N] &\leq \alpha_d N^{-p\delta} \int_0^t \|\rho_s\|_\infty \exp\left(pC\sqrt{\ln N} \int_s^t (1 + \|\rho_u\|_\infty + \|\rho_u^N\|_\infty) du\right) ds \\ &\leq CtN^{-p\delta} \exp(pC\sqrt{\ln N}t), \end{aligned}$$

due to the uniform bound assumption on both ρ_t and ρ_t^N in $L^1(0, T; L^\infty(\mathbb{R}^d))$. This completes the proof. \square

Remark 3.1. Note that for any $p > 0$

$$\exp(\sqrt{\ln N}) = \exp\left(\frac{\ln N}{\sqrt{\ln N}}\right) = N^{\frac{1}{\sqrt{\ln N}}} = o(N^p).$$

Let us define the functional $J^N : \mathbb{R}^{4dN} \rightarrow \mathbb{R}_+$ by

$$J^N : (\mathcal{X}, \mathcal{V}, \mathcal{Y}, \mathcal{W}) \in \mathbb{R}^{4dN} \mapsto 1 \wedge (\sqrt{\ln N} N^\delta |\mathcal{X} - \mathcal{Y}|_\infty + N^\delta |\mathcal{V} - \mathcal{W}|_\infty).$$

Then by using the same argument as in [27, Theorem 4.2], which is based on Gronwall lemma, Lemma 2.2 and Markov's inequality, we obtain the following estimate on J^N .

Lemma 3.2. *Let $(\mathcal{X}_t^N, \mathcal{V}_t^N)_{t \geq 0}$ be a solution to (1.4) and let $(\mathcal{Y}_t^N, \mathcal{W}_t^N)_{t \geq 0}$ be solutions to (1.5) with the same independent identically distributed initial conditions. Assume that*

$$\sup_{N \in \mathbb{N}} \int_0^T \|\rho_t^N\|_\infty dt < \infty.$$

Then, for any $\delta \in (0, \frac{1}{d})$ and $\beta > 0$, there exists $C_\beta > 0$ such that

$$\mathbb{P}\left(\sup_{0 \leq t \leq T} J_t^N \geq 1\right) \leq \frac{C_\beta}{N^\beta},$$

where $J_t^N := J^N(\mathcal{X}_t^N, \mathcal{V}_t^N, \mathcal{Y}_t^N, \mathcal{W}_t^N)$.

Proof of Theorem 1.1. For any $p \geq 1$, we first estimate $\mathcal{W}_p(\mu_t^N, f_t)$ as

$$\mathcal{W}_p(\mu_t^N, f_t) \leq \mathcal{W}_\infty(\mu_t^N, \nu_t^N) + \mathcal{W}_p(\nu_t^N, f_t^N) + \mathcal{W}_p(f_t^N, f_t),$$

where the empirical measure ν_t^N is associated to N copies solutions to the nonlinear SDEs with cut-off (4.1). It follows from Proposition 3.2 that for any $\gamma < \delta$ and $t \in [0, T]$, we find

$$\mathcal{W}_p(f_t^N, f_t) \leq CN^{-\delta} e^{C\sqrt{\ln N}} \leq N^{-\gamma} \quad \text{for } N \text{ large enough.}$$

This implies

$$\mathbb{P}(\mathcal{W}_p(\mu_t^N, f_t) \geq 3N^{-\gamma}) \leq \mathbb{P}(\mathcal{W}_\infty(\mu_t^N, \nu_t^N) \geq N^{-\gamma}) + \mathbb{P}(\mathcal{W}_p(\nu_t^N, f_t^N) \geq N^{-\gamma}). \tag{3.3}$$

Note that under the event $\{J_t^N < 1\}$, we get

$$\mathcal{W}_\infty(\mu_t^N, \nu_t^N) < N^{-\delta},$$

thus by using Lemma 3.2, we obtain

$$\mathbb{P}(\mathcal{W}_\infty(\mu_t^N, \nu_t^N) \geq N^{-\gamma}) \leq \mathbb{P}(\mathcal{W}_\infty(\mu_t^N, \nu_t^N) \geq N^{-\delta}) \leq \mathbb{P}(J_t^N \geq 1) \leq \frac{C_\beta}{N^\beta},$$

where $C_\beta > 0$ is independent of N . For the estimate of last term in (3.3), we use Proposition 2.1 with $x = N^{-p\gamma}$ to have

$$\mathbb{P}(\mathcal{W}_p^p(\nu_t^N, f_t^N) \geq N^{-p\gamma}) \leq CN^{1-\frac{(1-p\gamma)(q-\varepsilon)}{p}} + a(N, N^{-p\gamma}).$$

Combining the above estimates concludes the desired result. \square

4. Well-Posedness of Nonlinear SDEs and PDE: Proof of Theorem 1.2

In this section, we study the well-posedness of nonlinear SDEs (1.7) which is associated to the VFPF equation (1.3). For this, we use the nonlinear SDEs with cut-off given by

$$\begin{cases} dY_t^N = W_t^N dt, \\ dW_t^N = (F^N * \rho_t^N)(Y_t^N)dt + \sqrt{2\sigma}dB_t, \quad \rho_t^N = \mathcal{L}(Y_t^N), \end{cases} \quad (4.1)$$

where ρ_t^N is the spatial density of solution to (1.6). We first show the uniform-in- N estimate of spatial density ρ_t^N in $L^1(0, T; L^\infty(\mathbb{R}^d))$.

Lemma 4.1. *Let $T > 0$. Assume that the initial data f_0 satisfies $f_0 \in (L^1 \cap L^\infty)(\mathbb{R}^{2d})$. Then there exists a unique weak solution f_t^N to the system (1.6) with the initial data f_0 , such that $f^N \in L^\infty(0, T; (L^1 \cap L^\infty)(\mathbb{R}^{2d}))$. Furthermore, if we assume that for some $C > 0$*

$$f_0(x, v) \leq C\langle v \rangle^{-\gamma}, \quad \text{where } \langle v \rangle = \sqrt{1 + |v|^2},$$

then, for $\gamma > d$, there exists a time $T \geq T_* > 0$ such that

$$\sup_{N \geq 1} \sup_{0 \leq t \leq T_*} \|\rho_t^N\|_\infty < \infty,$$

where ρ_t^N denotes the spatial density of the law of solution at time t to Eq. (4.1).

Proof. Since the existence and uniqueness of solutions f_t^N to Eq. (1.6) is classical due to the regularity of the force fields, we only focus on the uniform-in- N estimate of the spatial density ρ_t^N in the rest of the proof. We divide the proof into two steps.

• **Step A (Feynman–Kac’s representation formula).** Let $(\chi_\varepsilon)_{\varepsilon>0}$ be a family of mollifying kernels. First, we note that $f_t^N = \mathcal{L}(Y_t^N, W_t^N)$, the law of solution

to (4.1), is a solution in the sense of distributions to

$$\partial_t f_t^N + v \cdot \nabla_x f_t^N + (F^N * \rho_t^N) \cdot \nabla_v f_t^N = \sigma \Delta_v f_t^N, \quad (x, v) \in \mathbb{R}^d \times \mathbb{R}^d, \quad t > 0,$$

with the initial data $f_0 = \mathcal{L}(Y_0^N, W_0^N)$. Denote by $f_t^{N,\varepsilon}$ the solution to the same equation with initial condition $f_0 * \chi_\varepsilon$. Since the coefficients of the above equation are Lipschitz and locally bounded, classical existence theory guarantees the global existence and uniqueness of strong solutions. We now fix $t \geq 0$ and consider the following “backward” stochastic integral equations:

$$Y_s^{x,v} = x - \int_0^s W_u^{x,v} du, \quad W_s^{x,v} = v - \int_0^s F^N * \rho_{t-u}^{N,\varepsilon}(Y_u^{x,v}) du + \sqrt{2\sigma} B_s.$$

It is classical that there exists a unique strong solution to the above equations due to the strong regularity of the force fields. We next set

$$\theta_s := f_{t-s}^{N,\varepsilon}(Y_s^{x,v}, W_s^{x,v}),$$

and apply Ito’s rule to θ to find

$$\begin{aligned} \theta_s &= \theta_0 + \int_0^s -(\partial_u f_{t-u}^{N,\varepsilon}(Y_u^{x,v}, W_u^{x,v}) - \langle \nabla_x f_{t-u}^{N,\varepsilon}(Y_u^{x,v}, W_u^{x,v}), W_u^{x,v} \rangle) du \\ &\quad - \int_0^s \langle \nabla_v f_{t-u}^{N,\varepsilon}(Y_u^{x,v}, W_u^{x,v}), F^N * \rho_{t-u}^{N,\varepsilon}(Y_u^{x,v}) \rangle du \\ &\quad + \sqrt{2\sigma} \int_0^s \langle \nabla_v f_{t-u}^{N,\varepsilon}(Y_u^{x,v}, W_u^{x,v}), dB_u \rangle + \sigma \int_0^s \Delta_v f_{t-u}^{N,\varepsilon}(Y_u^{x,v}, W_u^{x,v}) du. \end{aligned}$$

Taking the expectation to the above equation together with

$$\mathbb{E} \left[\int_0^s \langle \nabla_v f_{t-u}^{N,\varepsilon}(Y_u^{x,v}, W_u^{x,v}), dB_u \rangle \right] = 0$$

yields

$$\begin{aligned} \mathbb{E}[\theta_s] &= \mathbb{E}[\theta_0] + \int_0^s \mathbb{E} \left[\left(\underbrace{-\partial_u f_{t-u}^{N,\varepsilon} - v \cdot \nabla_x f_{t-u}^{N,\varepsilon} - (F^N * \rho_{t-u}^{N,\varepsilon})(x) \cdot \nabla_v f_{t-u}^{N,\varepsilon} + \sigma \Delta_v f_{t-u}^{N,\varepsilon}(x, v)}_{=0} \right) \right. \\ &\quad \left. \times (Y_u^{x,v}, W_u^{x,v}) \right] du. \end{aligned}$$

Thus, finally, we choose $s = t$ to have

$$f_t^{N,\varepsilon}(x, v) = \mathbb{E}[f_0 * \chi_\varepsilon(Y_t^{x,v}, W_t^{x,v})].$$

• **Step B (Uniform-in- N estimate).** It follows from the previous bound that

$$f_t^{N,\varepsilon}(x, v) \leq C \mathbb{E} \left[\left\langle v - \int_0^t F^N * \rho_{t-u}^{N,\varepsilon}(Y_u^{x,v}) du + \sqrt{2\sigma} B_t \right\rangle^{-\gamma} \right].$$

Note that for all $v, w \in \mathbb{R}^d$ and $\gamma \geq 1$

$$\langle v - w \rangle^{-1} \leq \sqrt{2} \langle v \rangle^{-1} \langle w \rangle \quad \text{and} \quad \langle v + w \rangle^\gamma \leq C_\gamma (1 + |v|^\gamma + |w|^\gamma).$$

Using those facts, we get

$$f_t^{N,\varepsilon}(x, v) \leq C \langle v \rangle^{-\gamma} \mathbb{E} \left[\left\langle - \int_0^t F^N * \rho_{t-u}^{N,\varepsilon}(Y_u^{x,v}) du + \sqrt{2\sigma} B_t \right\rangle^\gamma \right],$$

and further we find for $\gamma > d$

$$\begin{aligned} \|\rho_t^{N,\varepsilon}\|_\infty &\leq C \left(1 + (\sqrt{2\sigma})^\gamma \mathbb{E}[|B_t|^\gamma] + \mathbb{E} \left[\left| \int_0^t F^N * \rho_{t-u}^{N,\varepsilon}(Y_u^{x,v}) du \right|^\gamma \right] \right) \\ &\leq C \left(1 + t^\gamma + \mathbb{E} \left[\left| \int_0^t F^N * \rho_{t-u}^{N,\varepsilon}(Y_u^{x,v}) du \right|^\gamma \right] \right). \end{aligned}$$

Due to $\|F^N * \rho^{N,\varepsilon}(x)\|_\infty \leq C \|\rho^{N,\varepsilon}\|_{1,\infty}$, we find for any fixed time $T > 0$

$$\|\rho_t^{N,\varepsilon}\|_\infty \leq C \left(1 + \int_0^t \|\rho_s^{N,\varepsilon}\|_\infty^\gamma ds \right) \quad \text{for } 0 \leq t \leq T,$$

where $C > 0$ is independent of N . Finally, we use Lemma A.1.3 to have

$$\sup_{0 \leq t \leq T_*} \|\rho_t^{N,\varepsilon}\|_\infty \leq C \quad \text{with } T_* < \min \left\{ \frac{\gamma - 1}{C^\gamma}, T \right\},$$

where $C > 0$ is independent of N . The result follows from the fact that $\rho_t^{N,\varepsilon}$ converges at least weakly star to ρ_t^N in $L^\infty(\mathbb{R}^d)$ as ε goes to 0. \square

We are now ready to give the details of the proof of Theorem 1.2.

Proof of Theorem 1.2. We split the proof into three steps.

• **Step A (Cauchy estimates).** Let (Y_t^N, W_t^N) be the strong solution to the system (4.1) on the time interval $[0, T]$. Then by using a similar argument as in Proposition 3.2, we can show that for $N, N' \geq e$

$$\begin{aligned} \mathbb{E} \left[\sup_{0 \leq t \leq T} (\sqrt{\ln N} |Y_t^N - Y_t^{N'}| + |W_t^N - W_t^{N'}|)^p \right]^{1/p} \\ \leq C(N^{-\delta} + (N')^{-\delta}) \exp(C\sqrt{\ln N}), \end{aligned}$$

where C is a positive constant independent of p, N' and N .

• **Step B (Existence).** It follows from the previous step that the sequence $(Y_t^N, W_t^N)_{N \in \mathbb{N}}$ of solution to (4.1) is a Cauchy sequence. Thus, there exists a limit process $(Y_t, W_t)_{t \in [0, T]}$ such that (Y_t^N, W_t^N) goes in law to (Y_t, W_t) as N goes to

infinity. Moreover, denoting by $(f_t^N)_{N \in \mathbb{N}}$ the sequence of the law of solution to (4.1), we find

$$\begin{aligned} \sup_{0 \leq t \leq T} \mathcal{W}_p(f_t^N, f_t^{N'}) &\leq \mathbb{E} \left[\sup_{0 \leq t \leq T} (\sqrt{\ln N} |Y_t^N - Y_t^{N'}| + |W_t^N - W_t^{N'}|)^p \right]^{1/p} \\ &\leq C(N^{-\delta} + (N')^{-\delta}) \exp(\sqrt{\ln N}), \end{aligned}$$

where C is a positive constant independent of p and N . This deduces that $(f_t^N)_{N \in \mathbb{N}}$ converges weakly to some $f_t \in C([0, T]; \mathcal{P}_p(\mathbb{R}^d))$ which is the law of $(Y_t, W_t)_{t \in [0, T]}$. It now remains to prove that this process is indeed a solution to (1.7). In order to check this, it is sufficient to prove that $(F^N * \rho_t^N(Y_t^N))_{t \in [0, T]}$ converges \mathbb{P} almost surely (up to a subsequence) to $(F * \rho_t(Y_t))_{t \in [0, T]}$. It follows from Proposition 3.1, Lemma 4.1, and $\|(F^N - F) * \rho_t\|_\infty \leq \|F^N - F\|_1 \|\rho_t\|_\infty \leq CN^{-\delta} \|\rho_t\|_\infty$ that

$$\begin{aligned} &\mathbb{E} \left[\int_0^T |F^N * \rho_t^N(Y_t^N) - F * \rho_t(Y_t)| dt \right] \\ &\leq \mathbb{E} \left[\int_0^T |F^N * \rho_t^N(Y_t^N) - F^N * \rho_t(Y_t)| dt \right] \\ &\quad + \mathbb{E} \left[\int_0^T |F^N * \rho_t(Y_t) - F * \rho_t(Y_t)| dt \right] \\ &\leq C \ln N \int_0^T (\|\rho_t^N\|_\infty + \|\rho_t\|_\infty) \mathbb{E}[|Y_t^N - Y_t|] dt + CN^{-\delta} \int_0^T \|\rho_t\|_\infty dt \\ &\leq C \frac{\ln N}{N^\delta} e^{C\sqrt{\ln N}} \rightarrow 0 \quad \text{as } N \rightarrow \infty, \end{aligned}$$

where $C > 0$ is independent of N .

• **Step C (Uniqueness).** Let $(Y_t^1, W_t^1)_{t \geq 0}$ and $(Y_t^2, W_t^2)_{t \geq 0}$ be two solutions to (1.7) with the same initial data (Y_0, W_0) such that $\rho_t^i \in L^1(0, T; L^\infty(\mathbb{R}^d)), i = 1, 2$. Set

$$\Delta_t := |Y_t^1 - Y_t^2| + |W_t^1 - W_t^2|.$$

Then it follows from Proposition 3.1 that

$$\begin{aligned} \frac{d}{dt} \mathbb{E}[\Delta_t^p] &\leq p \mathbb{E}[|W_t^1 - W_t^2| \Delta_t^{p-1}] + \mathbb{E}[p|F * \rho_t^1(Y_t^1) - F * \rho_t^2(Y_t^2)| \Delta_t^{p-1}] \\ &\leq p \mathbb{E}[\Delta_t^p] + pC(\|\rho_t^1\|_\infty + \|\rho_t^2\|_\infty) \mathbb{E}[\Delta_t^p] \left(1 - \frac{1}{p} \ln^- \mathbb{E}[\Delta_t^p] \right). \end{aligned}$$

Set $Q(t) := \mathbb{E}[\Delta_t^p]$, then we get

$$Q'(t) \leq pQ(t) + Cp(\|\rho_t^1\|_\infty + \|\rho_t^2\|_\infty)Q(t)(1 - \ln^- Q(t)),$$

for $Q(t) \leq e$. On the other hand, since $Q_0 = 0$, applying Lemma A.1.2 yields $Q(t) = 0$ for $t \in [0, T]$. It is very clear that the uniqueness of solutions to (3.1) implies the uniqueness of weak solutions to Eq. (1.3). \square

5. Vlasov–Fokker–Planck Equation with Less Singular Interactions than Newtonian

The previous strategy can directly be applied for the system (1.3) with milder singular interaction forces (see also [27, Sec. 10] for the case without noise). To be more precise, let us consider the following nonlinear Vlasov–Fokker–Planck equation with singular interactions:

$$\partial_t f_t + v \cdot \nabla_x f_t + (F_\alpha * \rho_t) \cdot \nabla_v f_t = \sigma \Delta_v f_t, \quad (x, v) \in \mathbb{R}^d \times \mathbb{R}^d, \quad t > 0, \quad (5.1)$$

where F_α satisfies

$$|F_\alpha(x)| \leq \frac{1}{|x|^\alpha} \quad \text{and} \quad |\nabla F_\alpha(x)| \leq \frac{1}{|x|^{\alpha+1}} \quad \forall x \in \mathbb{R}^d \setminus \{0\},$$

with $F_\alpha(0) = 0$ by definition. Note that $\alpha = d - 1$ corresponds to the Newtonian case (1.2). Concerning the particle approximations, in a similar fashion as before, we consider the following stochastic particle system with cut-off given by

$$\begin{cases} dX_t^{i,N} = V_t^{i,N} dt, \\ dV_t^{i,N} = \frac{1}{N} \sum_{j=1}^N F_{\delta,\alpha}^N(X_t^{i,N} - X_t^{j,N}) dt + \sqrt{2\sigma} dB_t^{i,N}, \end{cases} \quad i = 1, \dots, N, \quad t > 0, \quad (5.2)$$

where the cut-off interaction potential $F_{\delta,\alpha}^N$ is given by $F_{\delta,\alpha}^N(x) = F_\alpha(x)$ for $|x| \geq N^{-\delta}$ and satisfies

$$|F_{\delta,\alpha}^N(x)| \leq N^{\alpha\delta} \quad \text{and} \quad |\nabla F_{\delta,\alpha}^N(x)| \leq N^{(\alpha+1)\delta} \quad \text{for } |x| < N^{-\delta}.$$

Then defining the associated empirical measure

$$\mu_t^N = \frac{1}{N} \sum_{i=1}^N \delta_{X_t^{i,N}, V_t^{i,N}},$$

we can state the following result.

Theorem 5.1. *Let $T > 0$ and $d > 1$. Let $(X_0^{i,N}, V_0^{i,N})_{i=1, \dots, N}$ be N independent variables with law f_0 . Let f_t and f_t^N be the solutions to the nonlinear Vlasov–Fokker–Planck equation (5.1) and its corresponding regularization (1.6) with $F_{\delta,\alpha}^N$ instead of F_δ^N respectively, up to time $T > 0$. Assume that $0 \leq \alpha < d/\ell' - 1$ and that $f, f^N \in L^\infty(0, T; (L^1 \cap L^\ell)(\mathbb{R}^{2d})) \cap \mathcal{C}([0, T]; \mathcal{P}_q(\mathbb{R}^{2d}))$ with the same initial data $f_0 \in (L^1 \cap L^\ell \cap \mathcal{P}_q)(\mathbb{R}^{2d})$ for some $q \geq 2$ and $1/\ell + 1/\ell' = 1$. Furthermore, assume that their respective spatial densities ρ_t and ρ_t^N satisfy*

$$\int_0^T \|\rho_t\|_\ell dt \leq C_0 \quad \text{and} \quad \sup_{N \in \mathbb{N}} \int_0^T \|\rho_t^N\|_\ell dt \leq C_0,$$

where C_0 only depending on the initial data and T . Then, for any $p \in [1, 2q)$, δ satisfying either

$$\frac{\ell'}{d} \leq \delta < \frac{1}{1 + \alpha}$$

or

$$\ell' > \frac{d}{2(1 + \alpha)} \quad \text{and} \quad \delta < \frac{\ell'}{d},$$

and $\varepsilon \in (0, q - \frac{p}{1-p\delta})$, the estimate

$$\sup_{0 \leq t \leq T} \mathbb{P}(\mathcal{W}_p(\mu_t^N, f_t) \geq 3N^{-\delta}) \leq CN^{\frac{1}{p}(1 - \frac{(1-p\delta)(q-\varepsilon)}{p})} \quad \text{for } N \text{ large enough}$$

holds for some constant $C > 0$ depending only on $d, T, p, q, \varepsilon, \alpha, \ell, f_0$, and C_0 .

Remark 5.1. Using almost same argument as in the proof of Theorem 1.2, we can obtain the well-posedness of the nonlinear SDEs (3.1) and the corresponding PDE (5.1) with F_α instead of F under the assumption on the initial data: $f_0 \in (L^1 \cap L^\ell \cap \mathcal{P}_q)(\mathbb{R}^{2d})$ for some $q \geq 2$ and $f_0(x, v) \leq C\langle v \rangle^{-\gamma}$ with $\gamma > d$.

Set

$$l_{\delta, \alpha}^N(x) := \begin{cases} \frac{1}{|x|^{\alpha+1}} & \text{if } |x| \geq (\alpha + 1)N^{-\delta}, \\ N^{(\alpha+1)\delta} & \text{otherwise,} \end{cases}$$

for $\alpha \in [0, d - 1)$. Similarly as before, for notational simplicity, we omit the subscript δ in $F_{\delta, \alpha}^N$ and $l_{\delta, \alpha}^N$ in the rest of this section, i.e. $F_\alpha^N = F_{\delta, \alpha}^N$ and $l_\alpha^N = l_{\delta, \alpha}^N$. In Lemma 5.1, we provide the weak-strong gradient estimate and uniform bound estimate of the gradient of force field in the cut-off parameter N , which can be obtained in the same manner as Lemma 2.1.

Lemma 5.1. *Let $d - 1 > \alpha \geq 0$ be given.*

(1) *There exists a constant C , which depends only on d , such that*

$$|F_\alpha^N(x) - F_\alpha^N(x + z)| \leq Cl_\alpha^N(x)|z|,$$

for any $x, z \in \mathbb{R}^d$ with $|z| \leq \alpha N^{-\delta}$.

(2) *There exists a constant $C > 0$ independent of N such that*

$$\|l_\alpha^N * \rho\|_\infty \leq C\|\rho\|_{1, \ell} \quad \text{and} \quad \|\nabla F_\alpha^N * \rho\|_\infty \leq C\|\rho\|_{1, \ell},$$

where $\ell > 1$ satisfies $d > (\alpha + 1)\ell'$ with $1/\ell' + 1/\ell = 1$.

Then slightly modifying the proof of Lemma 2.2, we have the following lemma.

Lemma 5.2. *Let (Y_1, \dots, Y_N) be i.i.d. random variables of the law $\rho \in L^\ell(\mathbb{R}^d)$, define the associated empirical measure $\rho_N = \frac{1}{N} \sum_{i=1}^N \delta_{Y_i}$. Then define*

$\tilde{\varepsilon} := 2\kappa\delta + (1 - d\delta/\ell')\mathbf{1}_{\ell' > d\delta}$ with $1/\ell + 1/\ell' = 1$. If $\tilde{\varepsilon} < 2$ then for all integer $m > \frac{1}{2-\tilde{\varepsilon}}$ there exist $\tilde{\gamma}_m, C > 0$ such that

$$\mathbb{E} \left[\sup_{1 \leq i \leq N} |h * \rho_N(Y_i) - h * \rho(Y_i)|^2 \right] \leq CN^{-\tilde{\gamma}_m},$$

where $\tilde{\gamma}_m = (2 - \tilde{\varepsilon})m - 1$ and h is defined as in (2.1).

Due to the milder singularity in the interactions than the Newtonian, we can bound the gradient of the force term uniformly in N in Lemma 5.1. This also enables us not to introduce the different weights in position and velocity for the error estimate between solutions of the corresponding nonlinear SDEs and the one with cut-off. More precisely, we have the following proposition which corresponds to Proposition 3.2.

Proposition 5.1. *For a given $T > 0$, let (Y_t, W_t) and (Y_t^N, W_t^N) be the solutions to the equations (3.2) with F_α^N and (3.1) with F_α appeared in (5.1) on the time interval $[0, T]$, respectively. Suppose that $\rho_t, \rho_t^N \in L^1(0, T; L^\ell(\mathbb{R}^d))$ and $\|\rho_t^N\|_{L^1(0, T; L^\ell)} \leq C$ with $C > 0$ independent of N . Then, for $p \geq 1$ and $d > (\alpha + 1)\ell'$, we have*

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} (|Y_t^N - Y_t| + |W_t^N - W_t|)^p \right]^{1/p} \leq CN^{-(d/\ell' - \alpha)\delta},$$

where C is a positive constant independent of p, α, ℓ , and N .

Let us define functional $\tilde{J} : \mathbb{R}^{4dN} \rightarrow \mathbb{R}_+$ by

$$\tilde{J} : (\mathcal{X}, \mathcal{V}, \mathcal{Y}, \mathcal{W}) \in \mathbb{R}^{4dN} \mapsto 1 \wedge (N^\delta |\mathcal{X} - \mathcal{Y}|_\infty + N^\delta |\mathcal{V} - \mathcal{W}|_\infty).$$

In Lemma 5.3, we provide the estimate on \tilde{J}^N whose proof can be obtained by using the similar argument as in Lemma 3.2.

Lemma 5.3. *Let $(\mathcal{X}_t^N, \mathcal{V}_t^N)_{t \geq 0}$ be a solution to (5.2) and let $(\mathcal{Y}_t^N, \mathcal{W}_t^N)_{t \geq 0}$ be N copies solutions to (1.5) with F_α^N instead of F^N and the same independent identically distributed initial conditions. Define \tilde{J}_t^N by*

$$\tilde{J}_t^N := \tilde{J}(\mathcal{X}_t^N, \mathcal{V}_t^N, \mathcal{Y}_t^N, \mathcal{W}_t^N).$$

Suppose $d > (\alpha + 1)\ell'$ and $\|\rho^N\|_{L^1(0, T; L^\ell)} \leq C$. Then for any $\beta > 0$ there is a constant $C_\beta > 0$ such that it holds

$$\mathbb{P} \left(\sup_{0 \leq t \leq T} \tilde{J}_t^N \geq 1 \right) \leq \frac{C_\beta}{N^\beta}.$$

Proof of Theorem 5.1. The strategy of the proof is the same as the one of Theorem 1.1. We first estimate $\mathcal{W}_p(\mu_t^N, f_t)$ as

$$\mathcal{W}_p(\mu_t^N, f_t) \leq \mathcal{W}_\infty(\mu_t^N, \nu_t^N) + \mathcal{W}_p(\nu_t^N, f_t^N) + \mathcal{W}_p(f_t^N, f_t),$$

for any $p \geq 1$, where the empirical measure ν_t^N is associated to N copies solutions to the nonlinear SDEs with cut-off with F_α^N . We now estimate each of the three

terms as follows:

- It follows from Proposition 5.1 that for $t \in [0, T]$, we find that

$$\mathcal{W}_p(f_t^N, f_t) \leq CN^{-(d/\ell' - \alpha)\delta} \leq N^{-\delta} \quad \text{for } N \text{ large enough,}$$

due to $d/\ell' - \alpha > 1$. Then we deduce

$$\mathbb{P}(\mathcal{W}_p(\mu_t^N, f_t) \geq 3N^{-\delta}) \leq \mathbb{P}(\mathcal{W}_\infty(\mu_t^N, \nu_t^N) \geq N^{-\delta}) + \mathbb{P}(\mathcal{W}_p(\nu_t^N, f_t^N) \geq N^{-\delta}).$$

- Note that under the event $\{\tilde{J}_t^N < 1\}$, we get

$$\mathcal{W}_\infty(\mu_t^N, \nu_t^N) < N^{-\delta},$$

thus by Lemma 5.3, we obtain

$$\mathbb{P}(\mathcal{W}_\infty(\mu_t^N, \nu_t^N) \geq N^{-\delta}) \leq \mathbb{P}(\tilde{J}_t^N \geq 1) \leq C_\beta N^{-\beta},$$

for any $\beta > 0$.

- Finally, we use Proposition 2.1 with $x = N^{-p\delta}$ to have

$$\mathbb{P}(\mathcal{W}_p^p(\nu_t^N, f_t^N) \geq N^{-p\delta}) \leq CN^{1 - \frac{(1-p\delta)(q-\varepsilon)}{p}} + a(N, N^{-p\delta}).$$

This completes the proof. □

Appendix A. Gronwall Type Inequalities

In this appendix, we provide several Gronwall type integral and differential inequalities.

Lemma A.1. (1) *Let f, g, h be nonnegative functions satisfying*

$$f(t) \leq h(t) + \int_0^t g(s)f(s)ds, \quad t \geq 0. \tag{A.1}$$

Then we have

$$f(t) \leq h(0)e^{\int_0^t g(s)ds} + \int_0^t h'(s)e^{\int_s^t g(\tau)d\tau} ds, \quad t \geq 0.$$

(2) *Let f be nonnegative function satisfying*

$$f'(t) \leq Cf(t)(1 - \ln^- f(t)) \quad \text{for } t \in [0, T], \tag{A.2}$$

where $C > 0$. If f_0 satisfies

$$f_0 < \min\{\exp(1 - e^{CT}), 1\},$$

then we have

$$f(t) \leq \exp(1 - (1 - \ln f_0)e^{-Ct}) \quad \text{for } t \in [0, T].$$

In particular, if $f_0 = 0$, then $f(t) \equiv 0$ for $t \in [0, T]$.

(3) Let $\gamma > 1$ and f be nonnegative function satisfying

$$f(t) \leq C_0 + C_1 \int_0^t f(s)^\gamma ds,$$

where $C_0, C_1 > 0$. Then we have

$$f(t) \leq \left(C_0^{1-\gamma} - \frac{C_1^2}{\gamma-1} t \right)^{\frac{1}{1-\gamma}} \quad \text{for } t < \frac{C_0^{1-\gamma}(\gamma-1)}{C_1}.$$

Proof. (1) Set

$$F(t) := \int_0^t g(s)f(s)ds,$$

then it follows from (A.1) that

$$F'(t) = g(t)f(t) \leq g(t)(h(t) + F(t)).$$

This yields

$$(F(t)e^{-\int_0^t g(s)ds})' \leq h(t)g(t)e^{-\int_0^t g(s)ds} = -h(t)(e^{-\int_0^t g(s)ds})'.$$

Since $F_0 = 0$, we get

$$F(t)e^{-\int_0^t g(s)ds} \leq -h(t)e^{-\int_0^t g(s)ds} + h(0) + \int_0^t h'(s)e^{-\int_0^s g(\tau)d\tau} ds,$$

i.e.

$$F(t) \leq -h(t) + h(0)e^{\int_0^t g(s)ds} + \int_0^t h'(s)e^{\int_s^t g(\tau)d\tau} ds.$$

Thus, we have

$$f(t) \leq h(t) + F(t) \leq h(0)e^{\int_0^t g(s)ds} + \int_0^t h'(s)e^{\int_s^t g(\tau)d\tau} ds.$$

(2) We first claim that

$$f(t) < 1 \quad \text{for } t \in [0, T].$$

Since f is continuous, there exists a $T_0 > 0$ such that $f(t) < 1$ for $t \in [0, T_0]$. Set

$$T^* = \sup\{t \in [0, T] : f(s) < 1 \text{ for } s \in [0, t]\}.$$

Let us assume that $T^* < T$. Then for $t \in [0, T^*)$, we obtain from (A.2) that

$$(\ln f)' \leq C(1 - \ln f(t)), \quad \text{i.e. } (1 - \ln f)' \geq -C(1 - \ln f(t)).$$

This yields

$$1 - \ln f(t) \geq (1 - \ln f_0)e^{-Ct}, \quad \text{i.e. } \ln f(t) \leq 1 - (1 - \ln f_0)e^{-Ct}.$$

Thus, we get

$$f(t) \leq \exp(1 - (1 - \ln f_0)e^{-ct}) \quad \text{for } t \in [0, T^*].$$

We now let $t \rightarrow T^*$ in the above inequality to find

$$1 = \lim_{t \rightarrow T^*} f(t) \leq \exp(1 - (1 - \ln f_0)e^{-ct}) < 1,$$

due to the assumption on the initial data f_0 . This is a contradiction that $T^* < T$ and implies $T^* = T$. Hence we have $f(t) < 1$ and

$$f(t) \leq \exp(1 - (1 - \ln f_0)e^{-ct}) \quad \text{for } t \in [0, T].$$

(3) Set

$$F := C_0 + C_1 \int_0^t f(s)^\gamma ds,$$

then we get $F(0) = C_0$ and

$$F'(t) = C_1 f(t)^\gamma \leq C_1 F(t)^\gamma.$$

This yields

$$(F^{1-\gamma})' \geq \frac{C_1}{1-\gamma},$$

and for $t < C_0^{1-\gamma}(\gamma - 1)C_1^{-1}$

$$f(t) \leq F(t) \leq \left(F(0)^{1-\gamma} - \frac{C_1}{\gamma-1} t \right)^{\frac{1}{1-\gamma}} = \left(C_0^{1-\gamma} - \frac{C_1}{\gamma-1} t \right)^{\frac{1}{1-\gamma}}. \quad \square$$

Acknowledgments

JAC was partially supported by the EPSRC grant number EP/P031587/1. YPC was supported by NRF grant (Nos. 2017R1C1B2012918 and 2017R1A4A1014735) and POSCO Science Fellowship of POSCO TJ Park Foundation. SS was supported by the Fondation des Sciences Mathématiques de Paris and Université Paris-Sciences-et-Lettres. The authors would like to thank Maxime Hauray for many fruitful discussions.

References

- [1] R. J. Berman and M. Önnheim, Propagation of chaos for a class of first order models with singular mean field interactions, preprint (2016); arXiv:1610.04327.
- [2] F. Bolley, J. A. Cañizo and J. A. Carrillo, Stochastic mean-field limit: Non-Lipschitz forces and swarming, *Math. Models Methods Appl. Sci.* **21**(11) (2011) 2179–2210.
- [3] F. Bouchut, Existence and uniqueness of a global smooth solution for the Vlasov–Poisson–Fokker–Planck system in three dimensions, *J. Funct. Anal.* **111**(1) (1993) 239–258.
- [4] F. Bouchut and J. Dolbeault, On long time asymptotics of the Vlasov–Fokker–Planck equation and of the Vlasov–Poisson–Fokker–Planck system with Coulombic and Newtonian potentials, *Differential Integral Equations* **8**(3) (1995) 487–514.
- [5] J. A. Cañizo, J. A. Carrillo and J. Rosado, A well-posedness theory in measures for some kinetic models of collective motion, *Math. Models Methods Appl. Sci.* **21**(3) (2011) 515–539.

- [6] J. A. Carrillo, Y.-P. Choi and M. Hauray, The derivation of swarming models: Mean-field limit and Wasserstein distances, in *Collective Dynamics from Bacteria to Crowds*, CISM Courses and Lectures, Vol. 553 (Springer, Vienna, 2014), pp. 1–46.
- [7] J. A. Carrillo, Y.-P. Choi, M. Hauray and S. Salem, Mean-field limit for collective behavior models with sharp sensitivity regions, to appear in *J. Eur. Math. Soc.*
- [8] 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, *Math. Methods Appl. Sci.* **18**(10) (1995) 825–839.
- [9] J. A. Carrillo, J. Soler and J. L. Vázquez, Asymptotic behaviour and self-similarity for the three-dimensional Vlasov–Poisson–Fokker–Planck system, *J. Funct. Anal.* **141**(1) (1996) 99–132.
- [10] Y.-P. Choi and S. Salem, Collective behavior models with vision geometrical constraints: Truncated noises and propagation of chaos, preprint (2017); arXiv:1705.01195.
- [11] Y.-P. Choi and S. Salem, Cucker–Smale flocking particles with multiplicative noises: Stochastic mean-field limit and phase transition, preprint (2017); arXiv:1711.10371.
- [12] Y.-P. Choi and S. Salem, Propagation of chaos for aggregation equations with no-flux boundary conditions and sharp sensing zones, *Math. Models Methods Appl. Sci.* **28**(2) (2018) 223–258.
- [13] R. L. Dobrušin, Vlasov equations, *Funktsional. Anal. i Prilozhen.* **13**(2) (1979) 48–58.
- [14] M. Duerinckx, Mean-field limits for some Riesz interaction gradient flows, *SIAM J. Math. Anal.* **48**(3) (2016) 2269–2300.
- [15] N. Fournier and A. Guillin, On the rate of convergence in Wasserstein distance of the empirical measure, *Probab. Theory Related Fields* **162**(3–4) (2015) 707–738.
- [16] N. Fournier, M. Hauray and S. Mischler, Propagation of chaos for the 2D viscous vortex model, *J. Eur. Math. Soc.* **16**(7) (2014) 1423–1466.
- [17] R. T. Glassey, *The Cauchy Problem in Kinetic Theory* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 1996).
- [18] D. Godinho and C. Quíñinao, Propagation of chaos for a subcritical Keller–Segel model, *Ann. Inst. Henri Poincaré Probab. Stat.* **51**(3) (2015) 965–992.
- [19] M. Hauray, Mean field limit for the one dimensional Vlasov–Poisson equation, in *Séminaire Laurent Schwartz — Équations aux dérivées partielles et applications. Année 2012–2013*, Sémin. Équ. Dériv. Partielles, Exp. No. XXI (École Polytech., Palaiseau, 2014), pp. 16.
- [20] M. Hauray and P.-E. Jabin, N -particles approximation of the Vlasov equations with singular potential, *Arch. Ration. Mech. Anal.* **183**(3) (2007) 489–524.
- [21] M. Hauray and P.-E. Jabin, Particle approximation of Vlasov equations with singular forces: Propagation of chaos, *Ann. Sci. Éc. Norm. Supér.* (4) **48**(4) (2015) 891–940.
- [22] M. Hauray and S. Mischler, On Kac’s chaos and related problems, *J. Funct. Anal.* **266**(10) (2014) 6055–6157.
- [23] M. Hauray and S. Salem, Propagation of chaos for the Vlasov–Poisson–Fokker–Planck system in 1D, to appear in *Kinet. Relat. Mod.*
- [24] F. Hérau, Short and long time behavior of the Fokker–Planck equation in a confining potential and applications, *J. Funct. Anal.* **244**(1) (2007) 95–118.
- [25] T. Holding, Propagation of chaos for Hölder continuous interaction kernels via Glivenko–Cantelli, preprint (2016); arXiv:1608.02877.
- [26] P.-E. Jabin and Z. Wang, Mean field limit and propagation of chaos for Vlasov systems with bounded forces, *J. Funct. Anal.* **271**(12) (2016) 3588–3627.
- [27] D. Lazarovici and P. Pickl, A mean field limit for the Vlasov–Poisson system, *Arch. Ration. Mech. Anal.* **225**(3) (2017) 1201–1231.

- [28] P.-L. Lions and B. Perthame, Propagation of moments and regularity for the 3-dimensional Vlasov–Poisson system, *Invent. Math.* **105**(2) (1991) 415–430.
- [29] G. Loeper, Uniqueness of the solution to the Vlasov–Poisson system with bounded density, *J. Math. Pures Appl.* (9) **86**(1) (2006) 68–79.
- [30] H. P. McKean, Jr., Propagation of chaos for a class of non-linear parabolic equations, in *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967)* (Air Force Office Sci. Res., Arlington, Va., 1967), pp. 41–57.
- [31] C. Pallard, Space moments of the Vlasov–Poisson system: Propagation and regularity, *SIAM J. Math. Anal.* **46**(3) (2014) 1754–1770.
- [32] B. Perthame, Mathematical tools for kinetic equations, *Bull. Amer. Math. Soc. (N.S.)* **41**(2) (2004) 205–244.
- [33] M. Pulvirenti and C. Simeoni, L^∞ -estimates for the Vlasov–Poisson–Fokker–Planck equation, *Math. Methods Appl. Sci.* **23**(10) (2000) 923–935.
- [34] S. Salem, Propagation of chaos for some 2 dimensional fractional Keller Segel equations in diffusion dominated and fair competition cases, preprint (2017); arXiv:1712.06677.
- [35] S. Serfaty, Mean field limit for coulomb flows, preprint (2018); arXiv:1803.08345.
- [36] H. D. Victory, Jr., On the existence of global weak solutions for Vlasov–Poisson–Fokker–Planck systems, *J. Math. Anal. Appl.* **160**(2) (1991) 525–555.
- [37] H. D. Victory, Jr., and B. P. O’Dwyer, On classical solutions of Vlasov–Poisson Fokker–Planck systems, *Indiana Univ. Math. J.* **39**(1) (1990) 105–156.
- [38] C. Villani, *Topics in Optimal Transportation*, Graduate Studies in Mathematics, Vol. 58 (American Mathematical Society, Providence, RI, 2003).