Liu and Yang *Res Math Sci(2016)3:40*
DOI 10.1186/s40687-016-0086-5

Research in the Mathematical Sciences

RESEARCH Open Access

Propagation of chaos for large Brownian particle system with Coulomb interaction

Jian-Guo Liu[1*] and Rong Yang[2*]

*Correspondence: jliu@phy.duke.edu; ysihan2010@163.com
[1] Departments of Physics and Mathematics, Duke University, Durham, NC 27708, USA
[2] College of Applied Sciences, Beijing University of Technology, Ping Le Yuan 100, Chaoyang District, Beijing 100124, People's Republic of China

Abstract

We investigate a system of N Brownian particles with the Coulomb interaction in any dimension $d \geq 2$, and we assume that the initial data are independent and identically distributed with a common density ρ_0 satisfying $\int_{\mathbb{R}^d} \rho_0 \ln \rho_0 \, dx < \infty$ and $\rho_0 \in L^{\frac{2d}{d+2}}(\mathbb{R}^d) \cap L^1(\mathbb{R}^d, (1+|x|^2)\, dx)$. We prove that there exists a unique global strong solution for this interacting partsicle system and there is no collision among particles almost surely. For $d = 2$, we rigorously prove the propagation of chaos for this particle system globally in time without any cutoff in the following sense. When $N \to \infty$, the empirical measure of the particle system converges in law to a probability measure and this measure possesses a density which is the unique weak solution to the mean-field Poisson–Nernst–Planck equation of single component.

Keywords: Noncollision among particles, Entropy and Fisher information estimates, Martingale problem, Uniqueness, de Finetti–Hewitt–Savage theorem

1 Background

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, endowed with the standard d-dimensional Brownian motions associated with this space. In this article, we consider the particle system with the following form

$$dX_t^i = \frac{1}{N} \sum_{j \neq i}^{N} F(X_t^i - X_t^j)\, dt + \sqrt{2}\, dB_t^i, \quad i = 1, \ldots, N, \tag{1.1}$$

with the initial data $\{X_0^i\}_{i=1}^N$, where $\{(X_t^i)_{t\geq 0}\}_{i=1}^N$ are the trajectories of N particles ($X_t^i \in \mathbb{R}^d$ for any $t > 0$), and $\{(B_t^i)_{t\geq 0}\}_{i=1}^N$ are a sequence of independent d-dimensional standard Brownian motions. The interparticles force is taken to be the Coulomb interaction, and it is described by the Newtonian potential,

$$F(x) = -\nabla \Phi(x), \qquad \Phi(x) = \begin{cases} \dfrac{C_d}{|x|^{d-2}} & \text{if } d \geq 3, \\ -\dfrac{1}{2\pi} \ln |x| & \text{if } d = 2, \end{cases} \tag{1.2}$$

where $C_d = \dfrac{1}{d(d-2)\alpha_d}$, $\alpha_d = \dfrac{\pi^{d/2}}{\Gamma(d/2+1)}$, i.e., α_d is the volume of d-dimensional unit ball. We recast $F(x) = \frac{C^* x}{|x|^d}$, $\forall x \in \mathbb{R}^d \backslash \{0\}, d \geq 2$, where $C^* = \frac{\Gamma(d/2)}{2\pi^{d/2}}$. The first term

Springer Open

on the right hand in (1.1) represents repulsive force on $(X_t^i)_{t\geq 0}$ by all other particles. The interacting particle system (1.1) is a typical physical model and appears in many applications. For example, in semiconductor, (i) the electrons interact with each other through the Coulomb repulsive force; (ii) the electrons interact with background, and it is modeled by Brownian motions; (iii) the mass of electron is very light and the inertia can be neglected, and the overdamped system of particles is used in (1.1).

Notice that if there exists two particles $(X_t^i)_{t\geq 0}$ and $(X_t^j)_{t\geq 0}$ colliding with each other for some time $t < \infty$, then $X_t^i = X_t^j$, $F(X_t^i - X_t^j) = \infty$, and then the solution to (1.1) breaks up. Fortunately, we will prove that this will not happen. More precisely, when the initial data $X_0^i \neq X_0^j$ almost surely (a.s.) for all $i \neq j$, we will show that there exists a unique global strong solution to system (1.1) and hence there is no collision a.s. among particles in (1.1).

The second object of this paper is to provide a rigorous theory on propagation of chaos for the above system (1.1) for $d = 2$. To do this, we will show the following main result: For any fixed time $T > 0$, there exists a subsequence of the empirical measure $\mu^N := \frac{1}{N}\sum_{i=1}^N \delta_{X_t^i}$ (μ^N are $\mathbf{P}(C([0,T];\mathbb{R}^d))$-valued random variables) converging in law to a deterministic probability measure μ as N goes to infinity, where $\mathbf{P}(C([0,T];\mathbb{R}^d))$ is the set of probability measures over $C([0,T];\mathbb{R}^d)$. Furthermore, the time marginal law μ_t has a density function ρ_t which is the unique weak solution to the Poisson–Nernst–Planck equation (1.4) below.

In this paper, for $k \geq 1$, we denote by $\mathbf{P}_{sym}((\mathbb{R}^d)^k)$ the set of symmetric probability measures on $(\mathbb{R}^d)^k$ (the law of any exchangeable $(\mathbb{R}^d)^k$-valued random variable $X = (X_1, \ldots, X_k)$ belongs to $\mathbf{P}_{sym}((\mathbb{R}^d)^k)$). When $f \in \mathbf{P}_{sym}((\mathbb{R}^d)^k)$ has a density $\rho \in L^1((\mathbb{R}^d)^k)$, we introduce the entropy and the Fisher information of f:

$$H_k(f) := \frac{1}{k}\int_{\mathbb{R}^{kd}} \rho \ln \rho \,\mathrm{d}x \quad \text{and} \quad I_k(f) := \frac{1}{k}\int_{\mathbb{R}^{kd}} \frac{|\nabla \rho|^2}{\rho}\,\mathrm{d}x.$$

Sometimes, we also use $H_k(\rho)$ and $I_k(\rho)$ to present $H_k(f)$ and $I_k(f)$, respectively. If f has no density, we simply put $H_k(f) = +\infty$ and $I_k(f) = +\infty$. Notice that $H_k(f^{\otimes k}) = H_1(f)$ and $I_k(f^{\otimes k}) = I_1(f)$.

We will split the proof of propagation of chaos into three steps. First, we denote f_t^N and ρ_t^N as the joint time marginal distribution and density of $(X_t^1, \ldots, X_t^N)_{0\leq t\leq T}$, respectively, $\mathcal{L}(\mu^N)$ is the law of μ^N and $\Phi^N = \frac{1}{N}\sum_{\substack{i,j=1\\ i\neq j}}^N \Phi(x_i - x_j)$. In Lemma 3.1, when $d = 2$, using the uniform estimate of $\int_0^t I_N(f_s^N)\,\mathrm{d}s$; when $d \geq 3$, using the uniform estimate of $\int_0^t \langle \rho_s^N, |\nabla \Phi^N|^2\rangle\,\mathrm{d}s$, we prove that the sequence $\{\mathcal{L}(\mu^N)\}_{N\geq 2}$ is tight in $\mathbf{P}\big(\mathbf{P}(C([0,T];\mathbb{R}^d))\big)$. (It is well known $C([0,T];\mathbb{R}^d)$ is a polish space and $\mathbf{P}(C([0,T];\mathbb{R}^d))$ is metrizable, and it is also a polish space, see the "Appendix".) Therefore there exists a subsequence of μ^N (without relabeling) and a $\mathbf{P}(C([0,T];\mathbb{R}^d))$-valued random measure μ such that μ^N converges in law to μ as N goes to infinity.

Second, for $d = 2$ and a.s. $\omega \in \Omega$, we prove that $\mu(\omega)$ is exactly a solution to the following self-consistent martingale problem with the initial data f_0 in a new probability space $\big(C([0,T];\mathbb{R}^d), \mathcal{B}, \mu(\omega)\big)$. This definition is the same as the Stroock–Varadhan [17], and it is a variant of the definition of nonlinear martingale problem in [11, p. 40].

Definition 1 In the probability space $(C([0,T];\mathbb{R}^d), \mathcal{B}, \mu, \{\mathcal{B}_t\}_{0\leq t\leq T})$, if a probability measure $\mu \in \mathbf{P}(C([0,T];\mathbb{R}^d))$ with time marginal μ_0 at time $t = 0$ is endowed with a μ-distributed canonical process $(X_t)_{0\le t\le T} \in C([0,T];\mathbb{R}^d)$, then let $\{\mathcal{B}_t\}_{0\le t\le T}$ is the natural filtration generated by $(X_t)_{0\le t\le T}$, i.e.,

$$\mathcal{B}_t = \sigma\{X_s, s \le t\} \tag{1.3}$$

and $\mathcal{B} = \mathcal{B}(C([0,T];\mathbb{R}^d))$ (the σ-algebra of Borel sets over $C([0,T];\mathbb{R}^d)$). μ is called a solution to the $\big(g, C_b^2(\mathbb{R}^d)\big)$-self-consistent martingale problem with the initial distribution μ_0 (meaning that X_0 is distributed according to μ_0), if for any $\varphi \in C_b^2(\mathbb{R}^d)$, $(X_t)_{0\le t\le T}$ induces the following process

$$\mathcal{M}_t := \varphi(X_t) - \varphi(X_0) - \int_0^t g(X_s, \mathcal{L}(X_s))\,\mathrm{d}s \quad \text{for any } \ t \in [0,T],$$

such that $(\mathcal{M}_t)_{0\le t\le T}$ is a martingale with respect to (w.r.t.) the filtration $\{\mathcal{B}_t\}_{0\le t\le T}$, where

$$g(x, \mathcal{L}(X_s)) = \int_{C([0,T];\mathbb{R}^d)} \nabla\varphi(x)\cdot F(x-y_s)\,\mu(\mathrm{d}y) + \triangle\varphi(x) \quad \text{for any } \ s \in [0,T].$$

Indeed, Lemma 4.3 gives a martingale estimate for the N-particle system and Lemma 4.2 states a standard method of checking a process to be a martingale. Then Proposition 4.1 shows that $\mu(\omega)$ is a solution to the above martingale problem for a.s. $\omega \in \Omega$.

Third, denoting $(\mu_t(\omega))_{t\ge 0}$ as the time marginal of $\mu(\omega)$. With the uniform estimates of entropy and the second moments for the particle system (1.1), Lemma 3.2 shows that $(\mu_t(\omega))_{t\ge 0}$ has a density $(\rho_t(\omega))_{t\ge 0}$ a.s.. Using the fact that $\mu(\omega)$ is a.s. a solution to the self-consistent martingale problem in Definition 1, Theorem 5.2 shows that $\rho(\omega)$ is the unique weak solution to the mean-field Poisson–Nernst–Planck (PNP) equations of single component:

$$\begin{cases} \partial_t\rho = \triangle\rho + \nabla\cdot(\rho\nabla c), \quad x \in \mathbb{R}^d, \quad t > 0, \\ -\triangle c = \rho(t,x), \\ \rho(t,x)_{t=0} = \rho_0(x), \end{cases} \tag{1.4}$$

i.e., $\rho(\omega)$ is independent of ω and hence it is deterministic (so does μ), which finishes the proof of propagation of chaos.

The concept of propagation of chaos was originated by Kac [8]. The propagation of chaos for (1.1) with the smooth F has been rigorously proved by McKean in 1970s with a coupling method, and the mean-field equation is a class of nonlinear parabolic equations [16]. For singular interacting kernel, a cutoff parameter is usually introduced to desingularize F by F_ε, and the coupling method sometimes still can be used to prove the propagation of chaos, c.f. [13].

The problem for the Newtonian potential without cutoff parameter is a challenging problem, which is the content of this paper. In this case, the coupling method can no longer be used and we adapt the nonlinear martingale problem method developed by Stroock–Varadhan [17]. Model (1.1) is closely related to the vortex system for the two-dimensional (2D) Navier–Stokes equation. In the vortex system, the interparticles force is given by $F(x) = -\nabla^\perp\Phi(x)$ for $d = 2$, where the operator $\nabla^\perp = \left(-\frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_1}\right)$. In a series papers [18–20], Osada showed that the particles a.s. never encounter, so that the singularity of kernel a.s. never visited. He also studied the propagation of chaos for the Navier–Stokes equation with the random vortex method without regularized parameters. In a recent important work of Fournier et al. [4], the authors significantly improved Osada's result: (i) They proved the propagation of chaos for the 2D viscous vortex model with any positive viscosity coefficient; (ii) the convergence holds in a strong sense, called entropic.

Instead of repulsive force, if the attractive force is used (in this case, the sign of F is changed), then the mean-field equation is the Keller–Segel equation. Much of analysis used in this paper failed due to the change of sign. In fact, recently, there is a deep result proved by Fournier and Jourdain [5, Proposition 4]: For any $N \geq 2$ and $T > 0$, if $\{(X_t^{i,N})_{t\in[0,T]}\}_{i=1}^N$ is the solution to the attractive model, then

$$\mathbb{P}\left(\exists s \in [0,T], \quad \exists 1 \leq i < j \leq N : X_s^{i,N} = X_s^{j,N}\right) > 0,$$

i.e., the singularity is visited and the particle system is not clearly well defined. The sign of F is crucially used in Lemmas 2.2 and 2.3 to achieve the uniform estimates. For a related work, Godinho and Quininao proved propagation of chaos for the subcritical Keller–Segel equations [6]. Some of their frameworks and techniques will be adapted to this paper.

This paper is organized as follows. The well posedness of the N-interacting particle system (1.1) and the uniform estimates for the joint density of those particles are established in Sect. 2. In Sect. 3, we show the tightness of the empirical measures of the trajectories of the N particles. In Sect. 4, we prove that the limiting point of the empirical measures is a.s. solution to the self-consistent martingale problem in Definition 1. In Sect. 5, we provide a simple proof of the uniqueness of weak solution to the PNP equation (1.4), and then, we prove the propagation of chaos results. Finally, in the "Appendix" we provide a metrization of $\mathbf{P}(C([0,T];\mathbb{R}^d))$.

2 Global well posedness of the *N*-interacting particle system in $d \geq 2$

First, we give a definition of the strong solution to (1.1).

Definition 2 For any fixed $T > 0$, initial data $\{X_0^i\}_{i=1}^N$ and given probability space $(\Omega, \mathcal{F}, \mathbb{P})$ endowed with a sequence of independent d-dimensional Brownian motions $\{(B_t^i)_{t\geq 0}\}_{i=1}^N$, if there is a stochastic process $\{(X_t^i)_{t\in[0,T]}\}_{i=1}^N$ adapted to $(\mathcal{F}_t)_{t\in[0,T]}$ such that $\{(X_t^i)_{t\in[0,T]}\}_{i=1}^N$ satisfies (1.1) a.s. in the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ for all $t \in [0,T]$, we say that $\{(X_t^i)_{t\in[0,T]}\}_{i=1}^N$ is a global strong solution to (1.1).

Next, we state some results about the well posedness of the N-interacting particle system (1.1) and the entropy and regularity properties for the density of those particles.

Theorem 2.1 *For any $d \geq 2$, let $N \geq 2$ and $T > 0$. Consider a sequence of independent d-dimensional Brownian motions $\{(B_t^i)_{t\geq 0}\}_{i=1}^N$ and the independent and identically distributed (i.i.d.) initial data $\{X_0^i\}_{i=1}^N$ with a common distribution f_0 satisfying $H_1(f_0) < +\infty$ and a common density $\rho_0 \in L^{\frac{2d}{d+2}}(\mathbb{R}^d) \cap L^1(\mathbb{R}^d, (1+|x|^2)\,\mathrm{d}x)$. Then*

- *(i) There exists a unique global strong solution to (1.1) and thus a.s. $X_t^i \neq X_t^j$ for all $t \in [0,T]$, $i \neq j$.*
- *(ii) Denote by $(f_t^N)_{0\leq t\leq T}$ the joint time marginal distribution function of $(X_t^1, \ldots, X_t^N)_{0\leq t\leq T}$ and assume $\|\rho_0\|_{L^r(\mathbb{R}^d)} < \infty$ for some $r > d \geq 2$. Then $f_t^N(X)$ has a density function $\rho_t^N(X)$, and it is the unique weak solution to the following linear Fokker–Plank equation:*

$$\partial_t \rho^N = \triangle \rho^N + \frac{1}{2}\nabla \cdot (\rho^N \nabla \Phi^N), \tag{2.1}$$

where $\Phi^N = \frac{1}{N}\sum_{\substack{i,j=1\\ i\neq j}}^N \Phi(x_i - x_j)$.

(iii) *Denote* $m_2(\rho) := \int_{\mathbb{R}^d} |x|^2 \rho \, \mathrm{d}x$. *For all* $t > 0$,

$$H_N(f_t^N) + \int_0^t I_N(f_s^N)\, \mathrm{d}s \le H_1(f_0), \quad \text{for } d \ge 2; \tag{2.2}$$

$$\left\langle \rho_t^N, \Phi^N \right\rangle + \frac{1}{2}\int_0^t \left\langle \rho_s^N, |\nabla \Phi^N|^2 \right\rangle \mathrm{d}s \le (N-1)C(d)\|\rho_0\|^2_{L^{\frac{2d}{d+2}}} \quad \text{for } d \ge 3; \tag{2.3}$$

$$\sup_{1\le i\le N} \mathbb{E}[|X_t^i|^2] \le \begin{cases} m_2(\rho_0) + \left(4 + \dfrac{1}{2\pi}\right) t & \text{if } d = 2, \\ 3m_2(\rho_0) + \dfrac{3t}{2} C(d)\|\rho_0\|^2_{L^{\frac{2d}{d+2}}} + 6td & \text{if } d \ge 3, \end{cases} \tag{2.4}$$

where $C(d) = \frac{1}{d(d-2)\pi}\left\{\frac{\Gamma(d)}{\Gamma(\frac{d}{2})}\right\}^{\frac{2}{d}}$.

(iv) *For any* $d \ge 2$ *and* $1 < p < \infty$,

$$\sup_{t\in[0,T]} \|\rho_t^N\|^p_{L^p(\mathbb{R}^{Nd})} + \frac{4(p-1)}{p}\|\nabla((\rho^N)^{\frac{p}{2}})\|^2_{L^2([0,T]\times\mathbb{R}^{Nd})} \le \|\rho_0\|^{Np}_{L^p(\mathbb{R}^d)}, \tag{2.5}$$

and there exists a constant C (depending only on T and the radius of the support of ρ^N*) such that*

$$\|\nabla \rho^N\|^2_{L^2([0,T]\times\mathbb{R}^{Nd})} \le \frac{1}{2}\|\rho_0\|^{2N}_{L^2(\mathbb{R}^d)}; \tag{2.6}$$

$$\|\partial_t \rho^N\|^2_{L^2(0,T;W^{-2,\infty}_{loc}(\mathbb{R}^{Nd}))} \le C\|\rho_0\|^{2N}_{L^2} \quad \text{for } d = 2, 3; \tag{2.7}$$

$$\|\partial_t \rho^N\|^2_{L^2(0,T;W^{-1,\infty}_{loc}(\mathbb{R}^{Nd}))} \le C(\|\rho_0\|^{2N}_{L^2} + \|\rho_0\|^{2N}_{L^q}) \quad \text{for all } d \ge 4 \text{ and some } q > d. \tag{2.8}$$

Additionally, the definition of weak solution to Eq. (2.1) is given as follows.

Definition 3 (*Weak solution*) Let the initial data $\rho_0^N \in L^1_+ \cap L^{\frac{2d}{d+2}}(\mathbb{R}^{Nd})$ and $T > 0$, we shall say that ρ^N is a weak solution to (2.1) with the initial data ρ_0^N if it satisfies:

1. integrability and time regularity:

$$\rho^N \in L^\infty(0,T; L^1 \cap L^{\frac{2d}{d+2}}(\mathbb{R}^{Nd})), \qquad (\rho^N)^{\frac{d}{d+2}} \in L^2(0,T; H^1(\mathbb{R}^{Nd}))$$
$$\partial_t \rho^N \in L^{k_2}(0,T; W^{-1,k_1}_{loc}(\mathbb{R}^{Nd})) \quad \text{for some } \; k_1, k_2 \ge 1;$$

2. for all $\varphi \in C_c^\infty(\mathbb{R}^{Nd})$, $0 < t \le T$, the following holds:

$$\int_{\mathbb{R}^{Nd}} \rho^N(\cdot,t)\varphi \,\mathrm{d}X = \int_{\mathbb{R}^{Nd}} \rho_0^N \varphi \,\mathrm{d}X - \frac{1}{2}\int_0^t \int_{\mathbb{R}^{Nd}} \nabla\varphi \cdot \nabla\Phi^N \rho^N \,\mathrm{d}X\mathrm{d}s + \int_0^t \int_{\mathbb{R}^{Nd}} \rho^N \Delta\varphi \,\mathrm{d}X\mathrm{d}s. \tag{2.9}$$

Next, we will split into two subsections to prove Theorem 2.1.

2.1 Noncollision among particles for the system (1.1)

Since the interacting force F of (1.1) is singular, we regularize F firstly. We directly recall below a lemma stated in [13, Lemma 2.1.], which collects some useful properties of the regularization. In addition, we add (iv) for a estimate on Φ_ε.

Lemma 2.1 *Suppose $J(x) \in C^2(\mathbb{R}^d)$, supp $J(x) \subset B(0,1)$, $J(x) = J(|x|)$ and $J(x) \geq 0$. Let $J_\varepsilon(x) = \frac{1}{\varepsilon^d} J(\frac{x}{\varepsilon})$ and $\Phi_\varepsilon(x) = J_\varepsilon * \Phi(x)$ for $x \in \mathbb{R}^d$, $F_\varepsilon(x) = -\nabla \Phi_\varepsilon(x)$, then $F_\varepsilon(x) \in C^1(\mathbb{R}^d)$, $\nabla \cdot F_\varepsilon(x) = J_\varepsilon(x)$ and*

(i) $F_\varepsilon(0) = 0$ and $F_\varepsilon(x) = F(x) g\left(\frac{|x|}{\varepsilon}\right)$ for any $x \neq 0$, where $g(r) = \frac{1}{C^*} \int_0^r J(s) s^{d-1}\,\mathrm{d}s$, $C^* = \dfrac{\Gamma(d/2)}{2\pi^{d/2}}$, $d \geq 2$ and $g(r) = 1$ for $r \geq 1$;

(ii) $|F_\varepsilon(x)| \leq \min\left\{\frac{C|x|}{\varepsilon^d}, |F(x)|\right\}$ and $|\nabla F_\varepsilon(x)| \leq \frac{C}{\varepsilon^d}$;

(iii) For any bounded domain B and some $1 < q < \frac{d}{d-1}$, $\|F_\varepsilon\|_{L^q(B)}$ is uniformly bounded in ε;

(iv) when $d \geq 3$, $\Phi_\varepsilon(x) = \Phi(x)$ for any $|x| \geq \varepsilon > 0$; when $d = 2$ and $0 < \varepsilon \leq 1$, $\Phi_\varepsilon(x) = \Phi(x) + \Phi_\varepsilon(1)$ for any $|x| \geq \varepsilon$. And

$$\Phi_\varepsilon(\varepsilon) \to +\infty, \quad \text{as} \;\; \varepsilon \to 0^+ \quad \text{for} \;\; d \geq 2. \tag{2.10}$$

Proof of (iv): Let $r = |x|$. By the proof of (i), one knows that

$$r^{d-1} \partial_r \Phi_\varepsilon(r) = -\int_0^r J_\varepsilon(s) s^{d-1}\,\mathrm{d}s = -\int_0^{\frac{r}{\varepsilon}} J(s) s^{d-1}\,\mathrm{d}s = -C^* g\left(\frac{r}{\varepsilon}\right). \tag{2.11}$$

Then for any $r \geq \varepsilon$, we integrate the above equality and use the fact that $g(r) = 1$ for $r \geq 1$,

$$\begin{aligned} -\Phi_\varepsilon(r) &= \int_r^\infty \partial_s \Phi_\varepsilon(s)\,\mathrm{d}s = -C^* \int_r^\infty \frac{g\left(\frac{s}{\varepsilon}\right)}{s^{d-1}}\,\mathrm{d}s \\ &= -C^* \int_r^\infty \frac{1}{s^{d-1}}\,\mathrm{d}s = -\frac{C^*}{(d-2) r^{d-2}} \quad \text{for} \;\; d \geq 3; \end{aligned} \tag{2.12}$$

$$\begin{aligned} \Phi_\varepsilon(r) - \Phi_\varepsilon(1) &= \int_1^r \partial_s \Phi_\varepsilon(s)\,\mathrm{d}s = -C^* \int_1^r \frac{g\left(\frac{s}{\varepsilon}\right)}{s}\,\mathrm{d}s \\ &= -\frac{1}{2\pi} \int_1^r \frac{1}{s}\,\mathrm{d}s = -\frac{1}{2\pi} \ln r \quad \text{for} \;\; d = 2, \varepsilon \leq 1. \end{aligned} \tag{2.13}$$

□

In this article, we take a cutoff function $J(x) \geq 0$, $J(x) \in C_0^3(\mathbb{R}^d)$,

$$J(x) = \begin{cases} C(1 + \cos \pi |x|)^2 & \text{if } |x| \leq 1, \\ 0 & \text{if } |x| > 1, \end{cases}$$

where C is a constant such that $C|\mathbb{S}^{d-1}| \int_0^1 (1 + \cos \pi r)^2 r^{d-1}\,\mathrm{d}r = 1$ and $|\mathbb{S}^{d-1}| = \frac{2\pi^{d/2}}{\Gamma(d/2)}$.

Proof (i) *of Theorem* 2.1: First, we consider the following N-interacting particle system via the regularized force:

$$\begin{cases} \mathrm{d}X_t^{i,\varepsilon} = \frac{1}{N} \sum_{j \neq i}^N F_\varepsilon(X_t^{i,\varepsilon} - X_t^{j,\varepsilon})\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t^i, \\ X_t^{i,\varepsilon}|_{t=0} = X_0^i, \end{cases} \tag{2.14}$$

which has a unique global strong solution $\{(X_t^{i,\varepsilon})_{t \geq 0}\}_{i=1}^N$.

Define a random variable

$$A_\varepsilon(t) := \inf_{0 \leq s \leq t} \min_{i \neq j} |X_s^{i,\varepsilon} - X_s^{j,\varepsilon}|.$$

Fix $T > 0$, define the stopping time

$$\tau_\varepsilon := \begin{cases} 0 & \text{if } \varepsilon \geq A_\varepsilon(0); \\ \sup\{t \wedge 2T : A_\varepsilon(t) \geq \varepsilon\} & \text{if } \varepsilon < A_\varepsilon(0). \end{cases} \tag{2.15}$$

The key step is to prove that

$$\lim_{\varepsilon\to 0}\mathbb{P}(\tau_\varepsilon \le T) = 0, \tag{2.16}$$

We adapt the techniques of [6,24] to prove (2.16). Define a random process $(\Phi_t^{\varepsilon,N})_{0\le t\le 2T}$ as

$$\Phi_t^{\varepsilon,N} := \frac{1}{N}\sum_{\substack{i,j=1\\ i\neq j}}^{N}\Phi(x_i-x_j)\Phi_\varepsilon\left(X^{i,\varepsilon}_{t\wedge\tau_\varepsilon}-X^{j,\varepsilon}_{t\wedge\tau_\varepsilon}\right). \tag{2.17}$$

Then one has the following basic fact

$$\{\tau_\varepsilon \le T\}\subset\left\{\sup_{t\in[0,T]}\Phi_t^{\varepsilon,N}\ge\Phi_{\tau_\varepsilon}^{\varepsilon,N}\right\}, \tag{2.18}$$

and the proof of (2.16) is divided into three steps as follows.

Step 1 We show that

$$\Phi_t^{\varepsilon,N}=\Phi_0^N+M_{t\wedge\tau_\varepsilon}-\frac{2}{N^2}\int_0^{t\wedge\tau_\varepsilon}\sum_{i=1}^N\left(\sum_{\substack{j=1\\ j\neq i}}^N F_\varepsilon(X_s^{i,\varepsilon}-X_s^{j,\varepsilon})\right)^2\,\mathrm{d}s, \tag{2.19}$$

where

$$M_{t\wedge\tau_\varepsilon}:=-\frac{\sqrt{2}}{N}\sum_{\substack{i,j=1\\ i\neq j}}^N\int_0^{t\wedge\tau_\varepsilon}F_\varepsilon(X_s^{i,\varepsilon}-X_s^{j,\varepsilon})\cdot(\mathrm{d}B_s^i-\mathrm{d}B_s^j);$$

$$\Phi_0^N:=\frac{1}{N}\sum_{\substack{i,j=1\\ i\neq j}}^N\Phi(X_0^i-X_0^j) \tag{2.20}$$

and we prove $(M_{t\wedge\tau_\varepsilon})_{0\le t\le T}$ is a martingale w.r.t. the filtration generated by the Brownian motions $\{(B_t^i)_{0\le t\le T}\}_{i=1}^N$.

Using the Itô's formula and the fact $\Delta\Phi_\varepsilon(x)=-J_\varepsilon(x)=0$ on $|x|\ge\varepsilon$, one has

$$\begin{aligned}\Phi_\varepsilon\left(X^{i,\varepsilon}_{t\wedge\tau_\varepsilon}-X^{j,\varepsilon}_{t\wedge\tau_\varepsilon}\right)&=\Phi\left(X_0^i-X_0^j\right)-\sqrt{2}\int_0^{t\wedge\tau_\varepsilon}F_\varepsilon(X_s^{i,\varepsilon}-X_s^{j,\varepsilon})\cdot(\mathrm{d}B_s^i-\mathrm{d}B_s^j)\\ &-\int_0^{t\wedge\tau_\varepsilon}F_\varepsilon(X_s^{i,\varepsilon}-X_s^{j,\varepsilon})\cdot\left[\frac{1}{N}\sum_{k\neq i}^N F_\varepsilon(X_s^{i,\varepsilon}-X_s^{k,\varepsilon})-\frac{1}{N}\sum_{k\neq j}^N F_\varepsilon(X_s^{j,\varepsilon}-X_s^{k,\varepsilon})\right]\mathrm{d}s.\end{aligned} \tag{2.21}$$

Summing (2.21) together, we obtain (2.19) and thus only need to show that $(M_{t\wedge\tau_\varepsilon})_{0\le t\le T}$ is a martingale. Using the fact $|F_\varepsilon(x)|\le\frac{C|x|}{\varepsilon^d}$ in Lemma 2.1 and $\left\{(X_t^{i,\varepsilon})_{0\le t\le T}\right\}_{i=1}^N$ are exchangeable, one has

$$\begin{aligned}\int_0^T\mathbb{E}[|F_\varepsilon(X_t^{i,\varepsilon}-X_t^{j,\varepsilon})|^2]\,\mathrm{d}t&\le\frac{C\int_0^T\mathbb{E}[|X_t^{i,\varepsilon}-X_t^{j,\varepsilon}|^2]\,\mathrm{d}t}{\varepsilon^{2d}}\\ &\le\frac{C\int_0^T(\mathbb{E}[|X_t^{i,\varepsilon}|^2]+\mathbb{E}[|X_t^{j,\varepsilon}|^2])\,\mathrm{d}t}{\varepsilon^{2d}}\\ &=\frac{2C\int_0^T(\mathbb{E}[|X_t^{i,\varepsilon}|^2])\,\mathrm{d}t}{\varepsilon^{2d}}.\end{aligned} \tag{2.22}$$

If one can prove that

$$\int_0^T \mathbb{E}|X_t^{i,\varepsilon}|^2 \,\mathrm{d}t < C(\varepsilon, T), \tag{2.23}$$

then by [17, Corollary 3.2.6], $\{\int_0^{t\wedge\tau_\varepsilon} F_\varepsilon(X_s^{i,\varepsilon} - X_s^{j,\varepsilon}) \cdot (\mathrm{d}B_s^i - \mathrm{d}B_s^j)\}_{0\le t\le T}$ is a martingale w.r.t. the filtration generated by the Brownian motions $(B_t^i)_{0\le t\le T}$ and $(B_t^j)_{0\le t\le T}$, and then $M_{t\wedge\tau_\varepsilon}$ is a martingale w.r.t. the filtration generated by the Brownian motions $\{(B_t^i)_{0\le t\le T}\}_{i=1}^N$. Below we prove (2.23).

According to Eq. (2.14) and the fact $(\sum_{i=1}^N a_i)^2 \le N\sum_{i=1}^N a_i^2$, one has

$$\begin{aligned}
\mathbb{E}[|X_t^{i,\varepsilon}|^2] &= \mathbb{E}[|X_0^i + \frac{1}{N}\int_0^t \sum_{j\ne i}^N F_\varepsilon(X_s^{i,\varepsilon} - X_s^{j,\varepsilon})\,\mathrm{d}s + \sqrt{2}B_t^i|^2] \\
&\le 3\mathbb{E}[|X_0^i|^2] + \frac{3t}{N^2}\mathbb{E}\left[\int_0^t \left(\sum_{j\ne i}^N F_\varepsilon(X_s^{i,\varepsilon} - X_s^{j,\varepsilon})\right)^2 \mathrm{d}s\right] + 6\mathbb{E}[|B_t^i|^2] \\
&\le 3\mathbb{E}[|X_0^i|^2] + \frac{3t}{N}\mathbb{E}\left[\int_0^t \sum_{j\ne i}^N F_\varepsilon^2(X_s^{i,\varepsilon} - X_s^{j,\varepsilon})\,\mathrm{d}s\right] + 6td.
\end{aligned}$$

Since $\{(X_t^{i,\varepsilon})_{t\ge0}\}_{i=1}^N$ are exchangeable, one has

$$\begin{aligned}
\mathbb{E}[|X_t^{i,\varepsilon}|^2] &\le 3\mathbb{E}[|X_0^i|^2] + \frac{3Ct}{N\varepsilon^{2d}}\int_0^t \sum_{j\ne i}^N \left(\mathbb{E}[|X_s^{i,\varepsilon}|^2] + \mathbb{E}[|X_s^{j,\varepsilon}|^2]\right) \mathrm{d}s + 6td \\
&\le 3\mathbb{E}[|X_0^i|^2] + \frac{6Ct}{\varepsilon^{2d}}\int_0^t \mathbb{E}[|X_s^{i,\varepsilon}|^2]\,\mathrm{d}s + 6td.
\end{aligned}$$

Hence by Gronwall's lemma, one obtains (2.23).

Step 2 We prove that there exists a constant C (depending only on $H_1(\rho_0)$, $m_2(\rho_0)$, $\|\rho_0\|_{L^{\frac{2d}{d+2}}}$, d, T and N) such that for any $R > 0$ and small enough ε,

$$\mathbb{P}(\tau_\varepsilon \le T) \le \frac{C}{R} + \mathbb{P}\left(\inf_{t\in[0,T]} M_{t\wedge\tau_\varepsilon} > -R,\ \sup_{t\in[0,T]} M_{t\wedge\tau_\varepsilon} \ge \frac{1}{N}\Phi_\varepsilon(\varepsilon) - R\right) \tag{2.24}$$

and split the proof into two cases.

Case1 ($d \ge 3$): Using the fact $\Phi_\varepsilon(x) > 0$, if $\tau_\varepsilon \le T$, then $\Phi_{\tau_\varepsilon}^{\varepsilon,N} \ge \frac{1}{N}\Phi_\varepsilon(\varepsilon)$. Combining (2.18), one has

$$\mathbb{P}(\tau_\varepsilon \le T) \le \mathbb{P}\left(\sup_{t\in[0,T]} \Phi_t^{\varepsilon,N} \ge \frac{1}{N}\Phi_\varepsilon(\varepsilon)\right) =: I_1. \tag{2.25}$$

From (2.19), one also has

$$0 < \Phi_t^{\varepsilon,N} \le \Phi_0^N + M_{t\wedge\tau_\varepsilon}. \tag{2.26}$$

Directly from (2.25) and (2.26), one has

$$\begin{aligned}
I_1 &\le \mathbb{P}\left(\sup_{t\in[0,T]} M_{t\wedge\tau_\varepsilon} \ge \frac{1}{N}\Phi_\varepsilon(\varepsilon) - \Phi_0^N\right) \\
&= \mathbb{P}\left(\inf_{t\in[0,T]} M_{t\wedge\tau_\varepsilon} \ge -\Phi_0^N,\ \sup_{t\in[0,T]} M_{t\wedge\tau_\varepsilon} \ge \frac{1}{N}\Phi_\varepsilon(\varepsilon) - \Phi_0^N\right)
\end{aligned} \tag{2.27}$$

Then for any $R > 0$,

$$I_1 \le \mathbb{P}(-\Phi_0^N \le -R) + \mathbb{P}\left(\inf_{t\in[0,T]} M_{t\wedge\tau_\varepsilon} > -R,\ \sup_{t\in[0,T]} M_{t\wedge\tau_\varepsilon} \ge \frac{1}{N}\Phi_\varepsilon(\varepsilon) - R\right). \tag{2.28}$$

Using the Markov's inequality to the first term of (2.28) and combining (2.25), one has

$$\mathbb{P}(\tau_\varepsilon \le T) \le \frac{\mathbb{E}[|\Phi_0^N|]}{R} + \mathbb{P}\left(\inf_{t\in[0,T]} M_{t\wedge\tau_\varepsilon} > -R, \sup_{t\in[0,T]} M_{t\wedge\tau_\varepsilon} \ge \frac{1}{N}\Phi_\varepsilon(\varepsilon) - R \right). \tag{2.29}$$

Moreover, $\mathbb{E}[|\Phi_0^N|]$ can be controlled by

$$\begin{aligned} \mathbb{E}[|\Phi_0^N|] &= \langle \rho_0^N, \Phi^N \rangle = (N-1)\int_{\mathbb{R}^{2d}} \rho_0(x)\rho_0(y)\Phi(x-y)\,\mathrm{d}x\mathrm{d}y \\ &\le (N-1)C(d)\|\rho_0\|^2_{L^{\frac{2d}{d+2}}}, \end{aligned} \tag{2.30}$$

where $C(d) = \frac{1}{d(d-2)\pi}\left\{\frac{\Gamma(d)}{\Gamma(\frac{d}{2})}\right\}^{\frac{2}{d}}$, and the last inequality comes from the Hardy–Littlewood–Sobolev inequality. Plugging (2.30) into (2.29) gives (2.24) for $d \ge 3$.

Case2 ($d = 2$): For small enough ε, using the fact that $\Phi_\varepsilon(x) = -\frac{1}{2\pi}\ln|x| + \Phi_\varepsilon(1) > -\frac{1}{2\pi}|x|$ for any $|x| \ge \varepsilon$ by (iv) in Lemma 2.1, one has

$$\Phi_{\tau_\varepsilon}^{\varepsilon,N} \ge \frac{1}{N}\Phi_\varepsilon(\varepsilon) - \frac{1}{\pi}\sum_{i=1}^N |X_{\tau_\varepsilon}^{i,\varepsilon}| \ge \frac{1}{N}\Phi_\varepsilon(\varepsilon) - \frac{1}{\pi}\sum_{i=1}^N \sup_{t\in[0,T]} |X_t^{i,\varepsilon}| \quad \text{if } \tau_\varepsilon \le T. \tag{2.31}$$

Combining (2.18), one has

$$\mathbb{P}(\tau_\varepsilon \le T) \le \mathbb{P}\left(\sup_{t\in[0,T]} \Phi_t^{\varepsilon,N} \ge \frac{1}{N}\Phi_\varepsilon(\varepsilon) - \frac{1}{\pi}\sum_{i=1}^N \sup_{t\in[0,T]} |X_t^{i,\varepsilon}| \right) =: I_2. \tag{2.32}$$

From (2.19) and the fact that $\Phi_\varepsilon(x) > -\frac{1}{2\pi}|x|$ for any $|x| \ge \varepsilon$ and small enough ε, one also has

$$-\frac{1}{\pi}\sum_{i=1}^N \left|X_{t\wedge\tau_\varepsilon}^{i,\varepsilon}\right| < \Phi_t^{\varepsilon,N} \le \Phi_0^N + M_{t\wedge\tau_\varepsilon}. \tag{2.33}$$

Denote $Y := \Phi_0^N + \frac{1}{\pi}\sum_{i=1}^N \sup_{t\in[0,T]} |X_t^{i,\varepsilon}|$. From (2.33), one has

$$\begin{aligned} I_2 &\le \mathbb{P}\left(\sup_{t\in[0,T]} M_{t\wedge\tau_\varepsilon} \ge \frac{1}{N}\Phi_\varepsilon(\varepsilon) - \Phi_0^N - \frac{1}{\pi}\sum_{i=1}^N \sup_{t\in[0,T]} |X_t^{i,\varepsilon}| \right) \\ &= \mathbb{P}\left(\inf_{t\in[0,T]} M_{t\wedge\tau_\varepsilon} \ge -\Phi_0^N - \inf_{t\in[0,T]}\left\{ \frac{1}{\pi}\sum_{i=1}^N \left|X_{t\wedge\tau_\varepsilon}^{i,\varepsilon}\right| \right\}, \sup_{t\in[0,T]} M_{t\wedge\tau_\varepsilon} \ge \frac{1}{N}\Phi_\varepsilon(\varepsilon) - Y \right). \end{aligned} \tag{2.34}$$

Combining (2.34) and (2.32), for any $R > 0$, one has

$$\begin{aligned} &\mathbb{P}(\tau_\varepsilon \le T) \\ &\quad \le \mathbb{P}\left(\inf_{t\in[0,T]} M_{t\wedge\tau_\varepsilon} \ge -\Phi_0^N - \inf_{t\in[0,T]}\left\{ \frac{1}{\pi}\sum_{i=1}^N \left|X_{t\wedge\tau_\varepsilon}^{i,\varepsilon}\right| \right\}, \sup_{t\in[0,T]} M_{t\wedge\tau_\varepsilon} \ge \frac{1}{N}\Phi_\varepsilon(\varepsilon) - Y \right) \\ &\quad \le \mathbb{P}\left(\inf_{t\in[0,T]} M_{t\wedge\tau_\varepsilon} \ge -Y, \sup_{t\in[0,T]} M_{t\wedge\tau_\varepsilon} \ge \frac{1}{N}\Phi_\varepsilon(\varepsilon) - Y \right) \\ &\quad \le \mathbb{P}(-Y \le -R) + \mathbb{P}\left(\inf_{t\in[0,T]} M_{t\wedge\tau_\varepsilon} > -R, \sup_{t\in[0,T]} M_{t\wedge\tau_\varepsilon} \ge \frac{1}{N}\Phi_\varepsilon(\varepsilon) - R \right). \end{aligned} \tag{2.35}$$

The first term of (2.35) is given by the Markov's inequality

$$\mathbb{P}(-Y \le -R) \le \frac{\mathbb{E}\left[\left|\Phi_0^N + \frac{1}{\pi}\sum_{i=1}^N \sup_{t\in[0,T]} |X_t^{i,\varepsilon}|\right|\right]}{R} \le \frac{\mathbb{E}[|\Phi_0^N|] + \frac{1}{\pi}\sum_{i=1}^N \mathbb{E}\left[\sup_{t\in[0,T]} |X_t^{i,\varepsilon}|\right]}{R}. \tag{2.36}$$

Therefore we need to evaluate $\mathbb{E}[\sup_{t\in[0,T]} |X_t^{i,\varepsilon}|]$.

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X_t^{i,\varepsilon}|\right] \le \left\{\mathbb{E}\left[\sup_{t\in[0,T]} |X_t^{i,\varepsilon}|^2\right]\right\}^{\frac{1}{2}}. \tag{2.37}$$

By the Itô formula, one has

$$d|X_t^{i,\varepsilon}|^2 = 2X_t^{i,\varepsilon} \cdot \left(\frac{1}{N}\sum_{j\neq i}^N F_\varepsilon(X_t^{i,\varepsilon} - X_t^{j,\varepsilon})\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t^i\right) + 4td. \tag{2.38}$$

Since $x \cdot F_\varepsilon(x) = x \cdot F(x)g(\frac{|x|}{\varepsilon}) = \frac{1}{2\pi}g(\frac{|x|}{\varepsilon}) \le \frac{1}{2\pi}$ by Lemma 2.1, then $\sum_{i=1}^N 2X_t^{i,\varepsilon} \cdot \left(\frac{1}{N}\sum_{j\neq i}^N F_\varepsilon(X_t^{i,\varepsilon} - X_t^{j,\varepsilon})\right) = \frac{1}{N}\sum_{\substack{i,j=1\\ i\neq j}}^N \Phi(x_i - x_j)(X_t^{i,\varepsilon} - X_t^{j,\varepsilon}) \cdot F_\varepsilon(X_t^{i,\varepsilon} - X_t^{j,\varepsilon}) \le \frac{N-1}{2\pi}$.

Summing (2.38) and integrating in time, one has

$$\sum_{i=1}^N |X_t^{i,\varepsilon}|^2 \le \sum_{i=1}^N |X_0^i|^2 + \left(4 + \frac{1}{2\pi}\right)Nt + 2\sqrt{2}\sum_{i=1}^N \int_0^t X_s^{i,\varepsilon} \cdot \mathrm{d}B_s^i. \tag{2.39}$$

Taking expectation, one has

$$\begin{aligned}\mathbb{E}\left[\sup_{r\in[0,t]} |X_r^{i,\varepsilon}|^2\right] &\le \mathbb{E}\left[\sup_{r\in[0,t]}\left(\sum_{i=1}^N |X_r^{i,\varepsilon}|^2\right)\right] \\ &\le \sum_{i=1}^N \mathbb{E}[|X_0^i|^2] + \left(4 + \frac{1}{2\pi}\right)Nt \\ &\quad + 2\sqrt{2}\mathbb{E}\left[\sup_{r\in[0,t]}\left|\sum_{i=1}^N \int_0^r X_s^{i,\varepsilon} \cdot \mathrm{d}B_s^i\right|\right] \\ &\le \sum_{i=1}^N \mathbb{E}[|X_0^i|^2] + \left(4 + \frac{1}{2\pi}\right)Nt \\ &\quad + 2\sqrt{2}\left(\mathbb{E}\left[\sup_{r\in[0,t]}\left|\sum_{i=1}^N \int_0^r X_s^{i,\varepsilon} \cdot \mathrm{d}B_s^i\right|^2\right]\right)^{\frac{1}{2}}.\end{aligned} \tag{2.40}$$

From (2.23), $\left(\sum_{i=1}^N \int_0^t X_s^{i,\varepsilon} \cdot \mathrm{d}B_s^i\right)_{0\le t\le T}$ is a martingale w.r.t. the filtration generated by the Brownian motions $\{(B_t^i)_{t\ge 0}\}_{i=1}^N$. Then using the Doob's inequality for martingale [10, see p. 203, Theorem 7.31], one has

$$\mathbb{E}\left[\sup_{r\in[0,t]}\left|\sum_{i=1}^N \int_0^r X_s^{i,\varepsilon} \cdot \mathrm{d}B_s^i\right|^2\right] \le 4\mathbb{E}\left[\left|\sum_{i=1}^N \int_0^t X_s^{i,\varepsilon} \cdot \mathrm{d}B_s^i\right|^2\right]. \tag{2.41}$$

Combining the exchangeability of $\{(X_t^{i,\varepsilon})_{t\ge 0}\}_{i=1}^N$, we obtain that

$$\begin{aligned}\mathbb{E}\left[\left|\sum_{i=1}^N \int_0^t X_s^{i,\varepsilon} \cdot \mathrm{d}B_s^i\right|^2\right] &= \mathbb{E}\left[\sum_{i=1}^N \int_0^t |X_s^{i,\varepsilon}|^2\,\mathrm{d}s\right] = N\mathbb{E}\left[\int_0^t |X_s^{i,\varepsilon}|^2\,\mathrm{d}s\right] \\ &\le N\mathbb{E}\left[\int_0^t \sup_{r\in[0,s]} |X_r^{i,\varepsilon}|^2\,\mathrm{d}s\right].\end{aligned} \tag{2.42}$$

Combining (2.40), (2.41) and (2.42) together, one has

$$\begin{aligned}\mathbb{E}\left[\sup_{r\in[0,t]}|X_r^{i,\varepsilon}|^2\right] &\leq \sum_{i=1}^N \mathbb{E}[|X_0^i|^2] + \left(4+\frac{1}{2\pi}\right)Nt \\ &\quad + 4\sqrt{2N}\left(\int_0^t \mathbb{E}\left[\sup_{r\in[0,s]}|X_r^{i,\varepsilon}|^2\right]\mathrm{d}s\right)^{\frac{1}{2}} \\ &\leq \sum_{i=1}^N \mathbb{E}[|X_0^i|^2] + \left(4+\frac{1}{2\pi}\right)Nt \\ &\quad + 8N + \int_0^t \mathbb{E}\left[\sup_{r\in[0,s]}|X_r^{i,\varepsilon}|^2\right]\mathrm{d}s. \end{aligned} \tag{2.43}$$

By Gronwall's lemma, one has

$$\mathbb{E}\left[\sup_{t\in[0,T]}|X_t^{i,\varepsilon}|^2\right] \leq C(T,N). \tag{2.44}$$

Plugging (2.44) into (2.36), one has

$$\mathbb{P}(-Y\leq -R) \leq \frac{\mathbb{E}[|\Phi_0^N|] + \frac{NC(T,N)}{\pi}}{R}. \tag{2.45}$$

Plugging (2.45) into (2.35), one has

$$\begin{aligned}\mathbb{P}(\tau_\varepsilon \leq T) &\leq \frac{\mathbb{E}[|\Phi_0^N|] + \frac{NC(T,N)}{\pi}}{R} \\ &\quad + \mathbb{P}\left(\inf_{t\in[0,T]} M_{t\wedge\tau_\varepsilon} > -R, \sup_{t\in[0,T]} M_{t\wedge\tau_\varepsilon} \geq \frac{1}{N}\Phi_\varepsilon(\varepsilon) - R\right). \end{aligned} \tag{2.46}$$

For $\mathbb{E}[|\Phi_0^N|] = \frac{N-1}{2\pi}\int_{\mathbb{R}^{2d}} \rho_0(x)\rho_0(y)|\ln|x-y||\,\mathrm{d}x\,\mathrm{d}y$, using the logarithmic Hardy–Littlewood–Sobolev inequality (see [21, p. 173, Lemma 6.8]), one has

$$H_1(\rho_0) + 2\int_{\mathbb{R}^{2d}} \rho_0(x)\rho_0(y)\ln|x-y|\,\mathrm{d}x\,\mathrm{d}y \geq -1-\ln\pi. \tag{2.47}$$

On the other hand,

$$\int_{\mathbb{R}^{2d}} \rho_0(x)\rho_0(y)\ln|x-y|\,\mathrm{d}x\,\mathrm{d}y \leq \int_{\mathbb{R}^{2d}} \rho_0(x)\rho_0(y)(x^2+y^2)\,\mathrm{d}x\,\mathrm{d}y = 2m_2(\rho_0). \tag{2.48}$$

Combining (2.47) and (2.48), one knows that $\mathbb{E}[|\Phi_0^N|]$ can be controlled by $H_1(\rho_0)$ and $m_2(\rho_0)$. Thus (2.24) holds for $d=2$.

Step 3 Setting $T_a := \inf\{t\geq 0, M_{t\wedge\tau_\varepsilon} = a\}$. Then from this definition, for small enough ε such that $\Phi_\varepsilon(\varepsilon) > R$, one directly has

$$\begin{aligned}\mathbb{P}\left(\inf_{t\in[0,T]} M_{t\wedge\tau_\varepsilon} > -R, \sup_{t\in[0,T]} M_{t\wedge\tau_\varepsilon} \geq \frac{1}{N}\Phi_\varepsilon(\varepsilon) - R\right) &\leq \mathbb{P}\left(T_{\frac{1}{N}\Phi_\varepsilon(\varepsilon)-R} \leq T < T_{-R}\right) \\ &\leq \mathbb{P}\left(T_{\frac{1}{N}\Phi_\varepsilon(\varepsilon)-R} \leq T_{-R}\right). \end{aligned} \tag{2.49}$$

Using the classical results on martingale [2, see p. 395, Theorem 5.3], one has

$$\mathbb{P}(T_{\frac{1}{N}\Phi_\varepsilon(\varepsilon)-R} \leq T_{-R}) = 1 - \mathbb{P}(T_{-R} < T_{\frac{1}{N}\Phi_\varepsilon(\varepsilon)-R}) = \frac{NR}{\Phi_\varepsilon(\varepsilon)}. \tag{2.50}$$

Combining (2.24), (2.49) and (2.50) together, we have

$$\mathbb{P}(\tau_\varepsilon \leq T) \leq \frac{C}{R} + \frac{NR}{\Phi_\varepsilon(\varepsilon)}. \tag{2.51}$$

Taking $R^2 = \Phi_\varepsilon(\varepsilon)$ and letting $\varepsilon \to 0$ in the above inequality, combining the fact $\Phi_\varepsilon(\varepsilon) \xrightarrow{\varepsilon\to 0^+} +\infty$ by (iv) in Lemma 2.1, one obtains (2.16) immediately.

Below, we define $\{(X_t^i)_{t\geq 0}\}_{i=1}^N$ as a limit of $\{(X_t^{i,\varepsilon})_{t\geq 0}\}_{i=1}^N$ and show that it is the unique strong solution to (1.1).

Since τ_ε is decreasing with respect to ε, (2.16) implies that

$$\mathbb{P}(\lim_{\varepsilon\to 0}\tau_\varepsilon > T) = \lim_{\varepsilon\to 0}\mathbb{P}(\tau_\varepsilon > T) = 1. \tag{2.52}$$

In other words, for a.s. $\omega \in \Omega$, there exists a $\varepsilon_0(\omega)$ such that if $\varepsilon \leq \varepsilon_0(\omega)$,

$$\tau_\varepsilon(\omega) \geq T.$$

Since $F_\varepsilon(x) = F(x)$ for any $|x| \geq \varepsilon$ and $|X_t^{i,\varepsilon} - X_t^{j,\varepsilon}| > \varepsilon$ for $t \in [0, T]$, we know that $\{(X_t^{i,\varepsilon})_{t\geq 0}\}_{i=1}^N$ satisfies the following equation on $t \in [0, T]$,

$$X_t^{i,\varepsilon}(\omega) = X_0^i(\omega) + \frac{1}{N}\sum_{j\neq i}^N \int_0^t F(X_s^{i,\varepsilon}(\omega) - X_s^{j,\varepsilon}(\omega))\,\mathrm{d}s + \sqrt{2}B_t^i(\omega), \quad i = 1, \ldots, N. \tag{2.53}$$

Since $F(x)$ is Lipschitz continuous in $\{x \in \mathbb{R}^d,\ |x| > \varepsilon\}$, using the uniqueness of the above ODE, then the solution on $t \in [0, T]$ is unique, i.e.,

$$X_t^{i,\varepsilon}(\omega) \equiv X_t^{i,\varepsilon_0}(\omega) \quad \text{for any } \ \varepsilon \leq \varepsilon_0,\ t \leq T,\ i = 1, \ldots, N. \tag{2.54}$$

If we define $X_t^i := \lim_{\varepsilon\to 0} X_t^{i,\varepsilon}$, then X_t^i is exactly the unique strong solution to (1.1) on $t \in [0, T]$. Since T is arbitrary, the global existence and uniqueness of strong solution to the system (1.1) can be achieved immediately. □

2.2 A uniform priori estimates for the density of N-interacting particle system

First, we start from the regularized system of (1.1) to achieve the uniform estimates of entropy and the second moments. Notice that the sign of F is crucially used in this section. For example, we used the positivity of J_ε to prove (2.55), (2.56) and (2.70).

Lemma 2.2 *Let $\{(X_t^{i,\varepsilon})_{t\geq 0}\}_{i=1}^N$ be the unique strong solution to (2.14) and $(f_t^{N,\varepsilon})_{t\geq 0}$ be its joint time marginal distribution with density $(\rho_t^{N,\varepsilon})_{t\geq 0}$. We have the uniform estimates for entropy:*

$$H_N(f_t^{N,\varepsilon}) + \int_0^t I_N(f_s^{N,\varepsilon})\,\mathrm{d}s \leq H_1(f_0) \qquad \text{for } \ d \geq 2, \tag{2.55}$$

$$\left\langle \rho_t^{N,\varepsilon},\ \Phi^{N,\varepsilon}\right\rangle + \frac{1}{2}\int_0^t \left\langle \rho_s^{N,\varepsilon},\ |\nabla\Phi^{N,\varepsilon}|^2\right\rangle \mathrm{d}s \leq (N-1)C(d)\|\rho_0\|^2_{L^{\frac{2d}{d+2}}} \quad \text{for } \ d \geq 3, \tag{2.56}$$

where $\Phi^{N,\varepsilon}(x) = \frac{1}{N}\sum_{\substack{i,j=1\\ i\neq j}}^N \Phi(x_i - x_j)\Phi_\varepsilon(x_i - x_j)$, $C(d) = \frac{1}{d(d-2)\pi}\left\{\frac{\Gamma(d)}{\Gamma(\frac{d}{2})}\right\}^{\frac{2}{d}}$. We also have the second moment estimates:

$$\sup_{1\leq i\leq N} \mathbb{E}[|X_t^{i,\varepsilon}|^2] \leq \begin{cases} m_2(\rho_0) + \left(4 + \dfrac{1}{2\pi}\right)t & \text{if } d = 2, \\ 3m_2(\rho_0) + \dfrac{3t}{2}C(d)\|\rho_0\|^2_{L^{\frac{2d}{d+2}}} + 6td & \text{if } d \geq 3. \end{cases} \tag{2.57}$$

Proof Denote by $\mathbf{X}_t^{N,\varepsilon} = (X_t^{1,\varepsilon}, \ldots, X_t^{N,\varepsilon})$. For any $\varphi \in C_b^2(\mathbb{R}^{Nd})$, applying the Itô formula one deduces that

$$\varphi(\mathbf{X}_t^{N,\varepsilon}) = \varphi(\mathbf{X}_0^{N,\varepsilon}) + \int_0^t \frac{1}{N} \sum_{\substack{i,j=1\\ i\neq j}}^N F_\varepsilon(X_s^{i,\varepsilon} - X_s^{j,\varepsilon}) \cdot \nabla_{x_i}\varphi(\mathbf{X}_s^{N,\varepsilon})\, \mathrm{d}s$$
$$+ \sqrt{2}\sum_{i=1}^N \int_0^t \nabla_{x_i}\varphi(\mathbf{X}_s^{N,\varepsilon}) \cdot \mathrm{d}B_s^i + \int_0^t \triangle\varphi(\mathbf{X}_s^{N,\varepsilon})\, \mathrm{d}s.$$

Taking expectation, one has

$$\int_{\mathbb{R}^{Nd}} \varphi\rho_t^{N,\varepsilon}\, \mathrm{d}X = \int_{\mathbb{R}^{Nd}} \varphi\rho_0^N\, \mathrm{d}X + \int_0^t \int_{\mathbb{R}^{Nd}} \left(\frac{1}{N} \sum_{\substack{i,j=1\\ i\neq j}}^N F_\varepsilon(x_i - x_j) \cdot \nabla_{x_i}\varphi \right) \rho_s^{N,\varepsilon}\, \mathrm{d}X\mathrm{d}s$$
$$+ \int_0^t \int_{\mathbb{R}^{Nd}} \triangle\varphi\rho_s^{N,\varepsilon}\, \mathrm{d}X\mathrm{d}s.$$

Thus $\rho_t^{N,\varepsilon}$ is a classical positive solution to

$$\partial_t\rho_t^{N,\varepsilon} = \frac{1}{2}\nabla \cdot (\rho_t^{N,\varepsilon}\nabla\Phi^{N,\varepsilon}) + \triangle\rho_t^{N,\varepsilon}, \quad t > 0, \tag{2.58}$$

where $\Phi^{N,\varepsilon}(x) = \frac{1}{N}\sum_{\substack{i,j=1\\ i\neq j}}^N \Phi(x_i - x_j)\Phi_\varepsilon(x_i - x_j)$.

We compute the entropy:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}H_N(f_t^{N,\varepsilon}) &= \frac{1}{N}\int_{\mathbb{R}^{Nd}} (1 + \ln\rho_t^{N,\varepsilon})\partial_t\rho_t^{N,\varepsilon}\, \mathrm{d}X \\
&= \frac{1}{2N}\int_{\mathbb{R}^{Nd}} (1 + \ln\rho_t^{N,\varepsilon})\nabla \cdot (\rho_t^{N,\varepsilon}\nabla\Phi^{N,\varepsilon})\, \mathrm{d}X + \frac{1}{N}\int_{\mathbb{R}^{Nd}} (1 + \ln\rho_t^{N,\varepsilon})\triangle\rho_t^{N,\varepsilon}\, \mathrm{d}X \\
&= -\frac{1}{2N}\int_{\mathbb{R}^{Nd}} \nabla\Phi^{N,\varepsilon} \cdot \nabla\rho_t^{N,\varepsilon}\, \mathrm{d}X - I_N(f_t^{N,\varepsilon}) \\
&= -\frac{1}{N^2}\int_{\mathbb{R}^{Nd}} \sum_{\substack{i,j=1\\ i\neq j}}^N J_\varepsilon(x_i - x_j)\rho_t^{N,\varepsilon}\, \mathrm{d}X - I_N(f_t^{N,\varepsilon}).
\end{aligned} \tag{2.59}$$

By the fact $J_\varepsilon(x_i - x_j) \geq 0$ in Lemma 2.1 and the symmetry of $\rho_t^{N,\varepsilon}$, one has

$$\begin{aligned}
&H_N(f_t^{N,\varepsilon}) + \int_0^t I_N(f_s^{N,\varepsilon})\, \mathrm{d}s + \frac{N-1}{N}\int_0^t \int_{\mathbb{R}^{2d}} J_\varepsilon(x_1 - x_2)\rho_s^{(2),N,\varepsilon}(x_1, x_2)\, \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}s \\
&= H_N(f_0^N)
\end{aligned} \tag{2.60}$$

where $\rho_s^{(2),N,\varepsilon}$ is the second marginal density. Since $\{X_0^i\}_{i=1}^N$ are i.i.d. with common distribution f_0, one has $H_N(f_0^N) = H_1(f_0)$. Then combining the positivity of J_ε, (2.55) is obtained.

Next, multiplying (2.58) with $\Phi^{N,\varepsilon}(x)$ and integrating, one has

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left\langle \rho_t^{N,\varepsilon}, \Phi^{N,\varepsilon} \right\rangle &= \left\langle \frac{1}{2}\nabla \cdot (\rho_t^{N,\varepsilon}\nabla\Phi^{N,\varepsilon}) + \triangle\rho_t^{N,\varepsilon}, \Phi^{N,\varepsilon} \right\rangle \\
&= -\left\langle \rho_t^{N,\varepsilon}, \frac{1}{2}|\nabla\Phi^{N,\varepsilon}|^2 \right\rangle - \left\langle \rho_t^{N,\varepsilon}, \frac{2}{N}\sum_{\substack{i,j=1\\ i\neq j}}^N J_\varepsilon(x_i - x_j) \right\rangle.
\end{aligned} \tag{2.61}$$

When $d \geq 3$, combining the positivity of J_ε and (2.30), one has

$$
\begin{aligned}
&\left\langle \rho_t^{N,\varepsilon},\, \Phi^{N,\varepsilon} \right\rangle + \frac{1}{2}\int_0^t \left\langle \rho_s^{N,\varepsilon},\, |\nabla \Phi^{N,\varepsilon}|^2 \right\rangle \mathrm{d}s \\
&\quad \leq -\int_0^t \left\langle \rho_s^{N,\varepsilon},\, \frac{2}{N}\sum_{\substack{i,j=1\\ i\neq j}}^N J_\varepsilon(x_i - x_j) \right\rangle \mathrm{d}s + \left\langle \rho_0^N,\, \Phi^N \right\rangle \\
&\quad \leq \left\langle \rho_0^N,\, \Phi^N \right\rangle \leq (N-1)C(d)\|\rho_0\|^2_{L^{\frac{2d}{d+2}}},
\end{aligned} \tag{2.62}
$$

which means (2.56) is true.

Finally, we prove the second moment estimates. Combining the fact that $\left(\sum_{i=1}^N \int_0^t X_s^{i,\varepsilon} \cdot \mathrm{d}B_s^i\right)_{t\geq 0}$ is a martingale and taking expectation of (2.39), one has

$$
\mathbb{E}\left[\sum_{i=1}^N \left|X_t^{i,\varepsilon}\right|^2\right] \leq \mathbb{E}\left[\sum_{i=1}^N |X_0^i|^2\right] + \left(4 + \frac{1}{2\pi}\right) Nt \quad \text{for} \quad d = 2. \tag{2.63}
$$

Since $\{X_0^i\}_{i=1}^N$ are i.i.d. with common density ρ_0, one has $\frac{1}{N}\mathbb{E}[\sum_{i=1}^N |X_0^i|^2] = m_2(\rho_0)$. Combining the fact that $\{(X_t^{i,\varepsilon})_{t\geq 0}\}_{i=1}^N$ are exchangeable, one obtains the second moment estimates for two dimension.

For $d \geq 3$, since

$$
X_t^{i,\varepsilon} = X_0^i + \frac{1}{N}\int_0^t \sum_{j\neq i}^N F_\varepsilon(X_s^{i,\varepsilon} - X_s^{j,\varepsilon})\,\mathrm{d}s + \sqrt{2}B_t^i, \quad \text{for} \ i = 1, \ldots, N, \tag{2.64}
$$

then

$$
|X_t^{i,\varepsilon}|^2 \leq 3|X_0^i|^2 + \frac{3t}{N^2}\int_0^T \left(\sum_{j\neq i}^N F_\varepsilon(X_s^{i,\varepsilon} - X_s^{j,\varepsilon})\right)^2 \mathrm{d}s + 6|B_t^i|^2. \tag{2.65}
$$

By the exchangeability of $\{(X_t^{i,\varepsilon})_{t\geq 0}\}_{i=1}^N$, one has

$$
\begin{aligned}
&\mathbb{E}\left[\frac{1}{N^2}\int_0^T \left(\sum_{j\neq i}^N F_\varepsilon(X_s^{i,\varepsilon} - X_s^{j,\varepsilon})\right)^2 \mathrm{d}s\right] \\
&\quad = \int_0^T \mathbb{E}\left[\frac{1}{N^3}\sum_{i=1}^N \left(\sum_{\substack{j=1\\ j\neq i}}^N F_\varepsilon(X_t^{i,\varepsilon} - X_t^{j,\varepsilon})\right)^2\right] \mathrm{d}t \\
&\quad = \int_0^T \left\langle \rho_t^{N,\varepsilon},\, \frac{1}{N^3}\sum_{i=1}^N \left(\sum_{\substack{j=1\\ j\neq i}}^N F_\varepsilon(x_i - x_j)\right)^2 \right\rangle \mathrm{d}t
\end{aligned} \tag{2.66}
$$

Using the identity:

$$|\nabla \Phi^{N,\varepsilon}|^2 = \sum_{i=1}^N |\partial_i \Phi^{N,\varepsilon}|^2 = \frac{4}{N^2}\sum_{i=1}^N \left(\sum_{\substack{j=1\\ j\neq i}}^N F_\varepsilon(x_i - x_j) \right)^2, \tag{2.67}$$

and combining (2.56), one has

$$\int_0^T \left\langle \rho_t^{N,\varepsilon}, \frac{1}{N^3}\sum_{i=1}^N \left(\sum_{\substack{j=1\\ j\neq i}}^N F_\varepsilon(x_i - x_j) \right)^2 \right\rangle \mathrm{d}t = \frac{1}{N}\int_0^T \left\langle \rho_t^{N,\varepsilon}, \frac{1}{4}|\nabla \Phi^{N,\varepsilon}|^2 \right\rangle \mathrm{d}t$$

$$\leq \frac{1}{2} C(d) \|\rho_0\|^2_{L^{\frac{2d}{d+2}}}. \tag{2.68}$$

Taking expectation of (2.65), using the fact $\mathbb{E}[|B_t^i|^2] = td$ and combining (2.66), (2.68), one has

$$\mathbb{E}\left[|X_t^{i,\varepsilon}|^2 \right] \leq 3m_2(\rho_0) + \frac{3t}{2} C(d) \|\rho_0\|^2_{L^{\frac{2d}{d+2}}} + 6td. \tag{2.69}$$

Combining (2.63) and (2.69) together, one obtains (2.57). □

Starting from the regularized system of (1.1), we also have a uniform priori regularity estimates.

Lemma 2.3 *Let* $\{(X_t^{i,\varepsilon})_{t\geq 0}\}_{i=1}^N$ *be the unique strong solution to* (2.14) *and* $(\rho_t^{N,\varepsilon})_{t\geq 0}$ *be its joint time marginal density. We have the uniform regularity estimates: For any* $d \geq 2$ *and* $1 < p < \infty$,

$$\sup_{t\in[0,T]} \left\| \rho_t^{N,\varepsilon} \right\|^p_{L^p(\mathbb{R}^{Nd})} + \frac{4(p-1)}{p} \|\nabla((\rho^{N,\varepsilon})^{\frac{p}{2}})\|^2_{L^2([0,T]\times\mathbb{R}^{Nd})} \leq \|\rho_0\|^{Np}_{L^p(\mathbb{R}^d)}, \tag{2.70}$$

and there exists a constant C (depending only on T and the radius of the support of $\rho^{N,\varepsilon}$*) such that*

$$\|\nabla \rho^{N,\varepsilon}\|^2_{L^2([0,T]\times\mathbb{R}^{Nd})} \leq \frac{1}{2}\|\rho_0\|^{2N}_{L^2(\mathbb{R}^d)}, \tag{2.71}$$

$$\|\partial_t \rho^{N,\varepsilon}\|^2_{L^2(0,T;W^{-2,\infty}_{loc}(\mathbb{R}^{Nd}))} \leq C\|\rho_0\|^{2N}_{L^2} \quad \textit{for } d = 2,3; \tag{2.72}$$

$$\|\partial_t \rho^{N,\varepsilon}\|^2_{L^2(0,T;W^{-1,\infty}_{loc}(\mathbb{R}^{Nd}))} \leq C(\|\rho_0\|^{2N}_{L^2} + \|\rho_0\|^{2N}_{L^q}) \quad \textit{for all } d \geq 4 \textit{ and some } q > d. \tag{2.73}$$

Proof For any $p > 1$, multiplying (2.58) with $p(\rho^{N,\varepsilon})^{p-1}$ and integrating, one has

$$\frac{d}{\mathrm{d}t}\int_{\mathbb{R}^{Nd}} (\rho_t^{N,\varepsilon})^p \,\mathrm{d}X + \frac{4(p-1)}{p}\int_{\mathbb{R}^{Nd}} \left|\nabla((\rho^{N,\varepsilon})^{\frac{p}{2}})\right|^2 \mathrm{d}X$$

$$+\frac{p-1}{N}\int_{\mathbb{R}^{Nd}} \sum_{\substack{i,j=1\\ i\neq j}}^N (\rho^{N,\varepsilon})^p J_\varepsilon(x_i - x_j)\,\mathrm{d}X = 0. \tag{2.74}$$

By the positivity of J_ε, we have

$$\int_{\mathbb{R}^{Nd}} (\rho_t^{N,\varepsilon})^p \,\mathrm{d}X + \frac{4(p-1)}{p}\int_0^t \int_{\mathbb{R}^{Nd}} \left|\nabla\left((\rho^{N,\varepsilon})^{\frac{p}{2}}\right)\right|^2 \mathrm{d}X\mathrm{d}s$$

$$\leq \int_{\mathbb{R}^{Nd}} \left(\rho_0^N\right)^p \mathrm{d}X = \left(\int_{\mathbb{R}^d} \rho_0^p \,\mathrm{d}x\right)^N, \tag{2.75}$$

which implies (2.70). Taking $p = 2$ in (2.75), then

$$2\int_0^t \int_{\mathbb{R}^{Nd}} |\nabla(\rho^{N,\varepsilon})|^2 \,\mathrm{d}X\mathrm{d}s \le \int_{\mathbb{R}^{Nd}} (\rho_0^N)^2 \,\mathrm{d}X = \|\rho_0\|_{L^2(\mathbb{R}^d)}^{2N}. \tag{2.76}$$

i.e., (2.71) holds. From (2.75), one also has

$$\sup_{t\in[0,T]} \|\rho_t^{N,\varepsilon}\|_{L^p(\mathbb{R}^{Nd})} \le \|\rho_0^N\|_{L^p(\mathbb{R}^{Nd})} = \|\rho_0\|_{L^p(\mathbb{R}^d)}^N. \tag{2.77}$$

For any $B_R \subset \mathbb{R}^{Nd}$, multiplying (2.58) with test function $\varphi(x) \in C_0^\infty(B_R)$ and integrating in space, one has

$$\int_{\mathbb{R}^{Nd}} \partial_t \rho^{N,\varepsilon} \varphi \,\mathrm{d}X = \int_{\mathbb{R}^{Nd}} -\left(\frac{1}{2}\nabla\Phi_\varepsilon^N \rho^{N,\varepsilon} + \nabla\rho^{N,\varepsilon}\right)\cdot\nabla\varphi \,\mathrm{d}X \tag{2.78}$$

$$= \int_{\mathbb{R}^{Nd}} \left(\frac{1}{2}\Phi_\varepsilon^N \nabla\rho^{N,\varepsilon}\cdot\nabla\varphi + \frac{1}{2}\Phi_\varepsilon^N \rho^{N,\varepsilon}\Delta\varphi - \nabla\rho^{N,\varepsilon}\cdot\nabla\varphi\right)\mathrm{d}X \tag{2.79}$$

For $d = 2, 3$, since $\|\Phi_\varepsilon\|_{L^{q'}(B)}$ is uniformly bounded in ε for any bounded domain B and some $1 < q' < \frac{d}{d-2}$, then $\|\Phi^{N,\varepsilon}\|_{L^{q'}(B)}$ is uniformly bounded too. Then from (2.79), one has

$$\begin{aligned}\left|\int_{\mathbb{R}^{Nd}} \partial_t\rho^{N,\varepsilon}\varphi \,\mathrm{d}X\right| &\le \frac{1}{2}\|\Phi^{N,\varepsilon}\|_{L^2}\big(\|\nabla\rho^{N,\varepsilon}\|_{L^2}\|\nabla\varphi\|_{L^\infty} \\ &\quad + \|\rho^{N,\varepsilon}\|_{L^2}\|\Delta\varphi\|_{L^\infty}\big) + \|\nabla\rho^{N,\varepsilon}\|_{L^2}\|\nabla\varphi\|_{L^2} \\ &\le C_R(\|\nabla\varphi\|_{L^\infty} + \|\Delta\varphi\|_{L^\infty})(\|\rho^{N,\varepsilon}\|_{L^2} + \|\nabla\rho^{N,\varepsilon}\|_{L^2}).\end{aligned} \tag{2.80}$$

For $d \ge 4$, by the fact that $\|F_\varepsilon\|_{L^{q'}(B)}$ is uniformly bounded in ε for any bounded domain B and some $1 < q' < \frac{d}{d-1}$ by Lemma 2.1 (*iii*), then $\|\nabla\Phi^{N,\varepsilon}\|_{L^{q'}(B)}$ is uniformly bounded too. Using (2.78), it holds that

$$\begin{aligned}\left|\int_{\mathbb{R}^{Nd}} \partial_t\rho^{N,\varepsilon}\varphi \,\mathrm{d}X\right| &\le \frac{1}{2}\|\rho^{N,\varepsilon}\|_{L^q}\|\nabla\Phi^{N,\varepsilon}\|_{L^{q'}}\|\nabla\varphi\|_{L^\infty} + \|\nabla\rho^{N,\varepsilon}\|_{L^2}\|\nabla\varphi\|_{L^2} \\ &\le C_R\|\nabla\varphi\|_{L^\infty}(\|\rho^{N,\varepsilon}\|_{L^q} + \|\nabla\rho^{N,\varepsilon}\|_{L^2}),\end{aligned} \tag{2.81}$$

where $\frac{1}{q} + \frac{1}{q'} = 1$ and $q > d \ge 4$. Combining (2.80) and (2.81) derives that for any $t \in [0, T]$,

$$\|\partial_t\rho^{N,\varepsilon}\|_{W^{-2,\infty}(B_R)} \le \sup_{\varphi\in C_0^\infty(B_R)} \frac{\left|\int_{\mathbb{R}^{Nd}} \partial_t\rho^{N,\varepsilon}\varphi \,\mathrm{d}x\right|}{\|\varphi\|_{W^{2,\infty}(B_R)}} \le C_R(\|\rho^{N,\varepsilon}\|_{L^2} + \|\nabla\rho^{N,\varepsilon}\|_{L^2}) \quad \text{for } d = 2, 3; \tag{2.82}$$

$$\|\partial_t\rho^{N,\varepsilon}\|_{W^{-1,\infty}(B_R)} \le \sup_{\varphi\in C_0^\infty(B_R)} \frac{\left|\int_{\mathbb{R}^{Nd}} \partial_t\rho^{N,\varepsilon}\varphi \,\mathrm{d}x\right|}{\|\varphi\|_{W^{1,\infty}(B_R)}} \le C_R(\|\rho^{N,\varepsilon}\|_{L^q} + \|\nabla\rho^{N,\varepsilon}\|_{L^2}) \quad \text{for } d \ge 4. \tag{2.83}$$

Combining (2.71), (2.77), (2.82) and (2.83) together, there exists a constant C (depending only on T and R) such that

$$\int_0^T \|\partial_t\rho^{N,\varepsilon}\|_{W^{-2,\infty}(B_R)} \,\mathrm{d}t \le C\|\rho_0\|_{L^2}^{2N} \quad \text{for } d = 2, 3; \tag{2.84}$$

$$\int_0^T \|\partial_t\rho^{N,\varepsilon}\|_{W^{-1,\infty}(B_R)} \,\mathrm{d}t \le C(\|\rho_0\|_{L^2}^{2N} + \|\rho_0\|_{L^q}^{2N}) \quad \text{for all } d \ge 4 \quad \text{and some } q > d, \tag{2.85}$$

which finishes the proof of (2.72) and (2.73). □

Next, we finish the rest proof of Theorem 2.1.

Proof (*ii*) *of Theorem* 2.1: Using the uniform estimates for the joint distribution of strong solution to (2.14), we split into three steps to study the joint distribution of strong solution to (1.1).

Step 1 We show that $\rho_t^{N,\varepsilon}$ is relatively compact.

Combining (2.55) and (2.57), then there exists a constant C independent of ε such that

$$\int_{\mathbb{R}^{Nd}} \rho_t^{N,\varepsilon} |\ln \rho_t^{N,\varepsilon}| \, \mathrm{d}X < C \quad \text{for any } \; t \in [0,T]. \tag{2.86}$$

Hence, it holds

$$\lim_{K\to\infty} \int_{\rho_t^{N,\varepsilon} \geq K} \rho_t^{N,\varepsilon} \, \mathrm{d}X \leq \lim_{K\to\infty} \frac{1}{\ln K} \int_{\mathbb{R}^{Nd}} \rho_t^{N,\varepsilon} |\ln \rho_t^{N,\varepsilon}| \, \mathrm{d}X = 0, \tag{2.87}$$

which means that $\rho^{N,\varepsilon}$ is uniformly integrable in $L^1(\mathbb{R}^{Nd})$. Combining the tightness of $\rho^{N,\varepsilon}$ according to (2.57) and the Dunford-Pettis theorem [22] together, we have the following classical compactness: There exists a subsequence of $\{\rho_t^{N,\varepsilon}\}_{\varepsilon>0}$ (without relabeling) such that

$$\rho_t^{N,\varepsilon} \rightharpoonup \rho_t^N \quad \text{in } \; L^1(\mathbb{R}^{Nd}) \quad \text{as } \; \varepsilon \to 0. \tag{2.88}$$

Step 2 We show that ρ^N obtained above is the unique weak solution to (2.1).

For any $\varphi \in C_c^\infty(\mathbb{R}^{Nd})$, $\rho_t^{N,\varepsilon}(X)$ satisfies the following equation:

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathbb{R}^{Nd}} \rho^{N,\varepsilon} \varphi \, \mathrm{d}X = \frac{1}{N} \sum_{\substack{i,j=1 \\ i\neq j}}^{N} \int_{\mathbb{R}^{Nd}} \nabla_{x_i}\varphi \cdot F_\varepsilon(x_i - x_j) \rho^{N,\varepsilon} \, \mathrm{d}X + \int_{\mathbb{R}^{Nd}} \triangle\varphi \rho^{N,\varepsilon} \, \mathrm{d}X. \tag{2.89}$$

Based on the uniform estimates (2.71), (2.72), (2.73) and the Lions–Aubin lemma, there exists a subsequence of $\{\rho^{N,\varepsilon}\}_{\varepsilon>0}$ (without relabeling) such that for any ball $B_R \subset \mathbb{R}^{Nd}$,

$$\rho^{N,\varepsilon} \to \rho^N \quad \text{in } \; L^2\left(0,T;L^2(B_R)\right) \quad \text{as } \; \varepsilon \to 0. \tag{2.90}$$

Direct computation shows that

$$\begin{aligned} &\left| \int_0^t \int_{\mathbb{R}^{Nd}} \nabla_{x_i}\varphi \cdot F_\varepsilon(x_i - x_j)\rho^{N,\varepsilon} \, \mathrm{d}X\mathrm{d}s - \int_0^t \int_{\mathbb{R}^{Nd}} \nabla_{x_i}\varphi \cdot F(x_i - x_j)\rho^N \, \mathrm{d}X\mathrm{d}s \right| \\ &\quad \leq \int_0^t \int_{\mathbb{R}^{Nd}} |\nabla_{x_i}\varphi| \left(|F_\varepsilon(x_i - x_j)||\rho^{N,\varepsilon} - \rho^N| + |F_\varepsilon(x_i - x_j) - F(x_i - x_j)|\rho^N \right) \mathrm{d}X\mathrm{d}s, \end{aligned} \tag{2.91}$$

and then

$$\begin{aligned} &\int_0^t \int_{\mathbb{R}^{Nd}} |\nabla_{x_i}\varphi||F_\varepsilon(x_i - x_j)||\rho^{N,\varepsilon} - \rho^N| \, \mathrm{d}X\mathrm{d}s \\ &\quad \leq C_R \|\nabla\varphi\|_{L^\infty} \|F_\varepsilon\|_{L^{q'}(B_{2R})} \int_0^t \|\rho^{N,\varepsilon} - \rho^N\|_{L^q(\mathbb{R}^{Nd})} \, \mathrm{d}s, \end{aligned} \tag{2.92}$$

where $\frac{1}{q'} + \frac{1}{q} = 1$, $1 < q' < \frac{d}{d-1}$ and $d < q < r$. Here r is a constant given in (*ii*) of Theorem 2.1. Below, we estimate $\int_0^t \|\rho^{N,\varepsilon} - \rho^N\|_{L^q} \, \mathrm{d}s$. Since $\sup_{t\in[0,T]} \|\rho^{N,\varepsilon}\|_{L^r} \leq \|\rho_0\|_{L^r}^N$ for any $r > d$ by (2.70), then $\sup_{t\in[0,T]} \|\rho^N\|_{L^r} \leq \|\rho_0\|_{L^r}^N$. By the interpolation inequality, for any $2 \le d < q < r$,

$$\int_0^t \|\rho^{N,\varepsilon} - \rho^N\|_{L^q}\,\mathrm{d}s \le \int_0^t \|\rho^{N,\varepsilon} - \rho^N\|_{L^r}^{\theta}\|\rho^{N,\varepsilon} - \rho^N\|_{L^2}^{1-\theta}\,\mathrm{d}s$$
$$\le 2^\theta \|\rho_0\|_{L^r(\mathbb{R}^d)}^{\theta N}\int_0^t \|\rho^{N,\varepsilon} - \rho^N\|_{L^2}^{1-\theta}\,\mathrm{d}s$$
$$\le 2^\theta \|\rho_0\|_{L^r(\mathbb{R}^d)}^{\theta N} C(T)\left(\int_0^t \|\rho^{N,\varepsilon} - \rho^N\|_{L^2}^{2}\,\mathrm{d}s\right)^{\frac{1-\theta}{2}} \tag{2.93}$$

where $\frac{\theta}{r} + \frac{1-\theta}{2} = \frac{1}{q}$. Since

$$\int_0^t \int_{\mathbb{R}^{Nd}} |\nabla_{x_i}\varphi||F_\varepsilon(x_i - x_j) - F(x_i - x_j)|\rho^N\,\mathrm{d}X\mathrm{d}s$$
$$\le C_R\|\nabla\varphi\|_{L^\infty}\|F_\varepsilon - F\|_{L^{q'}(B_{2R})}\int_0^t \|\rho^N\|_{L^q}\,\mathrm{d}s$$
$$\le C_R\|\nabla\varphi\|_{L^\infty}\|F_\varepsilon - F\|_{L^{q'}(B_{2R})} T\|\rho_0\|_{L^q}^N \tag{2.94}$$

Combining (2.90), (2.91), (2.92), (2.93) and (2.94), letting $\varepsilon \to 0$ in (2.89), one has ρ^N satisfies the following equation

$$\int_{\mathbb{R}^{Nd}} \rho^N(\cdot,t)\varphi\,\mathrm{d}X = \int_{\mathbb{R}^{Nd}} \rho_0^N\varphi\,\mathrm{d}X + \frac{1}{N}\sum_{\substack{i,j=1\\ i\neq j}}^{N}\int_0^t\int_{\mathbb{R}^{Nd}} \nabla_{x_i}\varphi\cdot F(x_i - x_j)\rho^N\,\mathrm{d}X\mathrm{d}s$$
$$+ \int_0^t\int_{\mathbb{R}^{Nd}} \triangle\varphi\rho^N\,\mathrm{d}X\mathrm{d}s. \tag{2.95}$$

By $F(x) = -\nabla\Phi(x)$, we have

$$\int_{\mathbb{R}^{Nd}} \rho^N(\cdot,t)\varphi\,\mathrm{d}X = \int_{\mathbb{R}^{Nd}} \rho_0^N\varphi\,\mathrm{d}X - \frac{1}{2}\int_0^t\int_{\mathbb{R}^{Nd}} \nabla\varphi\cdot\nabla\Phi^N\rho^N\,\mathrm{d}X\mathrm{d}s$$
$$+ \int_0^t\int_{\mathbb{R}^{Nd}} \triangle\varphi\rho^N\,\mathrm{d}X\mathrm{d}s. \tag{2.96}$$

Combining the regularity of ρ^N from Lemma 2.3, we obtain that ρ^N is exactly a weak solution to (2.1).

Suppose $\bar{\rho}^N$ is another weak solution to (2.1) with the same initial data. One has

$$\|\rho_t^N - \bar{\rho}_t^N\|_2^2 = \frac{1}{2}\int_0^t\int_{\mathbb{R}^{Nd}} (\rho^N - \bar{\rho}^N)^2\triangle\Phi^N\,\mathrm{d}X\mathrm{d}s$$
$$- \frac{1}{2}\int_0^t \|\nabla(\rho^N - \bar{\rho}^N)\|_2^2\,\mathrm{d}s \le 0, \tag{2.97}$$

which means $\rho^N \equiv \bar{\rho}^N$.

Step 3 Finally, we prove $\rho_t^N(X)$ is the density of $f_t^N(X)$.

By (2.16), one has

$$1 = \mathbb{P}\left\{\omega : \lim_{\varepsilon\to 0}\sup_{t\in[0,T]} |\mathbf{X}_t^{N,\varepsilon} - \mathbf{X}_t^N| = 0\right\}. \tag{2.98}$$

It can be deduced that

$$\mathbf{X}_t^{N,\varepsilon} \to \mathbf{X}_t^N \quad \text{a.s. for any } 0 \le t \le T \quad \text{as } \varepsilon \to 0. \tag{2.99}$$

Therefore,

$$f_t^{N,\varepsilon} \rightharpoonup f_t^N \quad \text{narrowly.} \tag{2.100}$$

From Step 1 and Step 2, we know that all the limited subsequence of $\{\rho_t^{N,\varepsilon}\}_{\varepsilon>0}$ weakly converges to ρ^N. Combining the fact $\mathrm{d}f_t^{N,\varepsilon}(X) = \rho_t^{N,\varepsilon}\,\mathrm{d}X$ and (2.100), then $\mathrm{d}f_t^N(X) = \rho_t^N\,\mathrm{d}X$. □

Proof (iii) and (iv) of Theorem 2.1: (*iv*) comes from Lemma 2.3 and (2.90) by the standard method. Now we prove (*iii*). Combining (2.55) and the fact that the functionals H and I both are lower semicontinuous [4, Lemma 4.2.], one has

$$\begin{aligned} H_N\left(f_t^N\right) + \int_0^t I_N(f_s^N)\,\mathrm{d}s &\le \liminf_{\varepsilon\to 0}\left\{H_N(f_t^{N,\varepsilon}) + \int_0^t I_N(f_s^{N,\varepsilon})\,\mathrm{d}s\right\} \\ &\le H_1(f_0) \quad \text{for any } d \ge 2, \end{aligned} \tag{2.101}$$

which gives (2.2).

Recalling (2.99) and the fact $\inf_{0\le s\le T}\min_{i\neq j}|X_s^i - X_s^j| > 0$ a.s. from the proof of (*i*) of Theorem 2.1, then combining (2.56) and using the Fatou's Lemma, for $d \ge 3$, one has

$$\begin{aligned} &\left\langle \rho_t^N,\, \Phi^N\right\rangle + \frac{1}{2}\int_0^t \left\langle \rho_s^N,\, |\nabla\Phi^N|^2\right\rangle \mathrm{d}s \\ &= \mathbb{E}\left[\Phi^N(\mathbf{X}_t^N) + \frac{1}{2}\int_0^t |\nabla\Phi^N(\mathbf{X}_t^N)|^2\,\mathrm{d}s\right] \\ &\le \liminf_{\varepsilon\to 0}\mathbb{E}\left[\Phi^{N,\varepsilon}(\mathbf{X}_t^{N,\varepsilon}) + \frac{1}{2}\int_0^t |\nabla\Phi^{N,\varepsilon}(\mathbf{X}_t^{N,\varepsilon})|^2\,\mathrm{d}s\right] \\ &= \liminf_{\varepsilon\to 0}\left\{\langle\rho_t^{N,\varepsilon},\, \Phi^{N,\varepsilon}\rangle + \frac{1}{2}\int_0^t \langle\rho_s^{N,\varepsilon},\, |\nabla\Phi^{N,\varepsilon}|^2\rangle\,\mathrm{d}s\right\} \\ &\le (N-1)C(d)\|\rho_0\|^2_{L^{\frac{2d}{d+2}}}, \end{aligned} \tag{2.102}$$

which gives (2.3).

Since

$$\mathbb{E}[|X_t^i|^2] \le \liminf_{\varepsilon\to 0}\mathbb{E}[|X_t^{i,\varepsilon}|^2], \tag{2.103}$$

then combining (2.57), one has (2.4). We have concluded the proof of Theorem 2.1 so far. □

3 Tightness of the empirical measures

Lemma 3.1 *For any $N \ge 2$ and $d \ge 2$, let $\{(X_t^{i,N})_{0\le t\le T}\}_{i=1}^N$ be the unique solution to (1.1) with the i.i.d initial data $\{X_0^{i,N}\}_{i=1}^N$. Suppose the common density $\rho_0(x) \in L^{\frac{2d}{d+2}}(\mathbb{R}^d) \cap L^1(\mathbb{R}^d, (1+|x|^2)dx)$ and $H_1(\rho_0) < +\infty$. Set $\mu^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N}}$, then*

- *(i) The sequence $\{\mathcal{L}(X^{1,N})\}$ is tight in $\mathbf{P}(C([0,T];\mathbb{R}^d))$.*
- *(ii) The sequence $\{\mathcal{L}(\mu^N)\}$ is tight in $\mathbf{P}\left(\mathbf{P}(C([0,T];\mathbb{R}^d))\right)$.*

Proof For $d = 2$, we directly cite the proof of Lemma 5.2 in [4].

For $d \ge 3$, in order to prove (i), it means that for fixed $\eta > 0$, $T > 0$, one should find a compact subset $K_{\eta,T}$ of $C([0,T];\mathbb{R}^d)$ such that $\sup_{N\ge 2}\mathbb{P}\left\{(X_t^{1,N})_{t\in[0,T]} \notin K_{\eta,T}\right\} \le \eta$. Considering the particle system (1.1), for any $0 \le s < t \le T$, one has

$$X_t^{1,N} - X_s^{1,N} = \frac{1}{N}\int_s^t \sum_{j\neq 1}^N F(X_r^{1,N} - X_r^{j,N})\,\mathrm{d}r + \sqrt{2}(B_t^1 - B_s^1). \tag{3.1}$$

A direct computation shows the time regularity of the Brownian motion term:

$$\mathbb{E}[\sqrt{2}|B_t^1 - B_s^1|] \le \sqrt{2}(\mathbb{E}[|B_t^1 - B_s^1|^2])^{\frac{1}{2}} = \sqrt{2d}\,|t-s|^{\frac{1}{2}}. \tag{3.2}$$

The estimate for the drift term is given by

$$\begin{aligned}\left|\frac{1}{N}\int_s^t \sum_{j\neq 1}^N F(X_r^1 - X_r^j)\,\mathrm{d}r\right| &\le \frac{(t-s)^{\frac{1}{2}}}{N}\left\{\int_s^t \left(\sum_{j\neq 1}^N F(X_r^1 - X_r^j)\right)^2 \mathrm{d}r\right\}^{\frac{1}{2}} \\ &\le \frac{(t-s)^{\frac{1}{2}}}{N}\left\{\int_0^T \left(\sum_{j\neq 1}^N F(X_t^1 - X_t^j)\right)^2 \mathrm{d}t\right\}^{\frac{1}{2}}. \end{aligned}\tag{3.3}$$

Denoting $U_T^N := \frac{1}{N}\left\{\int_0^T (\sum_{j\neq 1}^N F(X_t^1 - X_t^j))^2\,\mathrm{d}t\right\}^{\frac{1}{2}}$, $B(s,t) := \frac{\sqrt{2}|B_t^1 - B_s^1|}{|t-s|^{\frac{1}{2}}}$. Then combining (3.1) and (3.3) together, one has for any $0 \le s, t \le T$,

$$|X_t^{1,N} - X_s^{1,N}| \le (t-s)^{\frac{1}{2}}(U_T^N + B(s,t)). \tag{3.4}$$

By the exchangeability of $\{(X_t^{i,N})_{0\le t\le T}\}_{i=1}^N$, one has

$$\begin{aligned}\mathbb{E}[(U_T^N)^2] &= \int_0^T \mathbb{E}\left[\frac{1}{N^3}\sum_{i=1}^N \left(\sum_{\substack{j=1\\ j\neq i}}^N F(X_t^i - X_t^j)\right)^2\right]\mathrm{d}t \\ &= \int_0^T \left\langle \rho_t^N, \frac{1}{N^3}\sum_{i=1}^N \left(\sum_{\substack{j=1\\ j\neq i}}^N F(x_i - x_j)\right)^2\right\rangle \mathrm{d}t. \end{aligned}\tag{3.5}$$

Using the identity:

$$|\nabla \Phi^N|^2 = \sum_{i=1}^N |\partial_i \Phi^N|^2 = \frac{4}{N^2}\sum_{i=1}^N \left(\sum_{\substack{j=1\\ j\neq i}}^N F(x_i - x_j)\right)^2, \tag{3.6}$$

and combining (2.3), one has

$$\mathbb{E}[(U_T^N)^2] = \frac{1}{N}\int_0^T \left\langle \rho_t^N, \frac{1}{4}|\nabla\Phi^N|^2\right\rangle \mathrm{d}t \le \frac{1}{2}\bar{C}(d)\|\rho_0\|^2_{L^{\frac{2d}{d+2}}} < \infty \quad \text{for } d \ge 3. \tag{3.7}$$

Hence by the Markov's inequality, combining (3.2) and (3.7), for any $\eta > 0$, one can find a constant $R_\eta > 0$ (depending only on d and $\|\rho_0\|_{L^{\frac{2d}{d+2}}}$) such that

$$\sup_{N\ge 2}\mathbb{P}\{U_T^N + B(s,t) \ge R_\eta\} \le \frac{\mathbb{E}\left[U_T^N + B(s,t)\right]}{R_\eta} \le \frac{\eta}{2}. \tag{3.8}$$

Since $E[|X_0^{1,N}|^2] < \infty$, then one can find a constant $a_\eta > 0$ (depending only on $m_2(\rho_0)$) such that

$$\sup_{N\ge 2}\mathbb{P}\{|X_0^{1,N}| > a_\eta\} \le \frac{\eta}{2}. \tag{3.9}$$

Now we construct the following set

$$K_{\eta,T} := \big\{X \in C([0,T];\mathbb{R}^d), |X_0| \le a_\eta, |X_t - X_s| \le R_\eta (t-s)^{\frac{1}{2}},$$
$$\forall\, 0 \le s < t \le T, d \ge 3\big\},$$

which is a compact subset of $C([0,T];\mathbb{R}^d)$ by Ascoli's theorem. Combining (3.4), (3.8) and (3.9), one has

$$\begin{aligned}
&\sup_{N\ge 2} \mathbb{P}\big\{(X_t^{1,N})_{t\in[0,T]} \notin K_{\eta,T}\big\} \\
&\le \sup_{N\ge 2} \mathbb{P}\big\{|X_0^{1,N}| > a_\eta\big\} + \sup_{N\ge 2} \mathbb{P}\big\{|X_t^{1,N} - X_s^{1,N}| > R_\eta (t-s)^{\frac{1}{2}}\big\} \\
&\le \sup_{N\ge 2} \mathbb{P}\big\{|X_0^{1,N}| > a_\eta\big\} + \sup_{N\ge 2} \mathbb{P}\big\{B(s,t) + U_T^N > R_\eta\big\} \le \eta,
\end{aligned} \tag{3.10}$$

which finishes the proof of (i). (ii) follows from the exchangeability of $\{(X_t^{i,N})_{0\le t\le T}\}_{i=1}^N$, see [23, Proposition 2.2] or [15, Lemma 4.5].

From the tightness of $\{\mathcal{L}(\mu^N)\}$ in $\mathbf{P}\left(\mathbf{P}(C([0,T];\mathbb{R}^d))\right)$ by Lemma 3.1, one has that there exists a subsequence of $\mu^N \in \mathbf{P}(C([0,T];\mathbb{R}^d))$ (without relabeling) and a random measure $\mu \in \mathbf{P}(C([0,T];\mathbb{R}^d))$ such that

$$\mu^N \to \mu \quad \text{in law as } N \to \infty. \tag{3.11}$$

Next, we prove that the limited measure-valued process μ has a density a.s..

Lemma 3.2 *For any $N \ge 2$ and $d \ge 2$, let $\{(X_t^{i,N})_{t\ge 0}\}_{i=1}^N$ be the unique strong solution to (1.1) with the i.i.d. initial data $\{X_0^{i,N}\}_{i=1}^N$ such that $\mathcal{L}(X_0^{i,N}) = f_0$, $\mathrm{d}f_0 = \rho_0(x)\,\mathrm{d}x$. Denote by $(f_t^N)_{t\ge 0}$ the joint time marginal distribution of $\{(X_t^{i,N})_{t\ge 0}\}_{i=1}^N$ and $f_t^{(j),N}$ be the j-th marginal of f_t^N for any $j \ge 1$. If $\rho_0(x) \in L^{\frac{2d}{d+2}}(\mathbb{R}^d) \cap L^1(\mathbb{R}^d, (1+|x|^2)\,\mathrm{d}x)$ and $H_1(\rho_0) < +\infty$, then*

(i) $f_t^{(j),N}$ has a density $\rho_t^{(j),N}$ and there exists a subsequence $\rho_t^{(j),N}$ (without relabeling) weakly converging to ρ^j in $L^1(\mathbb{R}^{dj})$ as $N \to \infty$ with the following regularity:

$$H_j(f_t^j) + \int_0^t I_j(f_s^j)\,\mathrm{d}s \le H_1(f_0), \quad \int_{\mathbb{R}^{dj}} |x|^2 \rho_t^j(x)\,\mathrm{d}x < \infty, \tag{3.12}$$

where $\mathrm{d}f_t^j = \rho_t^j\,\mathrm{d}x$.

(ii) The limited measure-valued process $(\mu_t)_{t\ge 0}$ of the subsequence processes $\mu^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N}}$ (without relabeling) has a density $(\rho_t)_{t\ge 0}$ a.s.. At time $t = 0$, ρ_t takes the initial density ρ_0.

Proof Step 1 By Theorem 2.1, we know that f_t^N has a density ρ_t^N satisfying the entropy inequality (2.2). Then $f_t^{(j),N}$ also has a density $\rho_t^{(j),N}$. Combining (2.2) and Lemma 3.3 in [7], one has

$$H_j(f_t^{(j),N}) + \int_0^t I_j(f_s^{(j),N})\,\mathrm{d}s \le H_N(f_t^N) + \int_0^t I_N(f_s^N)\,\mathrm{d}s \le H_1(f_0) < \infty. \tag{3.13}$$

Combining (2.4) and the exchangeability of $\{(X_t^{i,N})_{t\ge 0}\}_{i=1}^N$, one has

$$\int_{\mathbb{R}^{dj}} |x|^2 \rho_t^{(j),N}(x)\,\mathrm{d}x = j\mathbb{E}[|X_t^1|^2] < \infty. \tag{3.14}$$

Similarly with (2.87), one has the following uniformly integrable property of $\rho^{(j),N}$ in $L^1(\mathbb{R}^{dj})$,

$$\lim_{K\to\infty} \int_{\rho^{(j),N} \ge K} \rho^{(j),N}\,\mathrm{d}x \le \lim_{K\to\infty} \frac{1}{\ln K} \int_{\mathbb{R}^{dj}} \rho^{(j),N} |\ln \rho^{(j),N}|\,\mathrm{d}x = 0. \tag{3.15}$$

And then using the Dunford-Pettis theorem, there exists a subsequence of $\rho_t^{(j),N}$ (without relabeling) and $\rho_t^j \in L^1(\mathbb{R}^{dj})$ such that

$$\rho_t^{(j),N} \rightharpoonup \rho_t^j \quad \text{in} \quad L^1(\mathbb{R}^{dj}) \quad \text{weakly as} \quad N \to \infty, \tag{3.16}$$

and ρ_t^j satisfies (3.12).

Combining $\int_{\mathbb{R}^{dj}} \rho_t^{(j),N} \, \mathrm{d}x \equiv 1$, we also obtain that

$$\int_{\mathbb{R}^{dj}} \rho_t^j(x) \, \mathrm{d}x \equiv 1. \tag{3.17}$$

Then combining (3.16) and (3.17), one has

$$\int_{\mathbb{R}^{dj}} \varphi(x)\rho_t^{(j),N}(x) \, \mathrm{d}x \xrightarrow{N\to\infty} \int_{\mathbb{R}^{dj}} \varphi(x)\rho_t^j(x) \, \mathrm{d}x \quad \text{for any} \quad \varphi \in C_b(\mathbb{R}^{dj}). \tag{3.18}$$

Step 2 For any $\varphi \in C_b(\mathbb{R}^{dj})$, we show that

$$\mathbb{E}\left[\langle \varphi, \mu_t \otimes \cdots \otimes \mu_t \rangle\right] = \int_{\mathbb{R}^{dj}} \varphi \rho_t^j \, \mathrm{d}x. \tag{3.19}$$

Define $I_t^{j,N} := \mathbb{E}\left[\langle \varphi, \mu_t^N \otimes \cdots \otimes \mu_t^N \rangle\right] = \mathbb{E}\left[\frac{1}{N^j} \sum_{i_1,\ldots,i_j=1}^N \varphi(X_t^{i_1,N}, \ldots, X_t^{i_j,N})\right]$, then by the exchangeability of $\{(X_t^{i,N})_{t\ge 0}\}_{i=1}^N$, we have

$$\begin{aligned} I_t^{j,N} &= \frac{1}{N^j} \sum_{\substack{i_1,\ldots,i_j=1 \\ i_k \neq i_\ell}}^N \mathbb{E}\left[\varphi\left(X_t^{i_1,N}, \ldots, X_t^{i_j,N}\right)\right] \\ &\quad + \frac{1}{N^j} \sum_{\substack{i_1,\ldots,i_j=1 \\ i_k = i_\ell \ \text{for some} \ k \neq \ell}}^N \mathbb{E}\left[\varphi\left(X_t^{i_1,N}, \ldots, X_t^{i_j,N}\right)\right] \\ &=: I_1 + I_2. \end{aligned} \tag{3.20}$$

By (3.18), one has

$$I_1 = \frac{N!}{(N-j)!N^j} \mathbb{E}[\varphi(X_t^{1,N}, \ldots, X_t^{j,N})] \to \int_{\mathbb{R}^{dj}} \varphi \rho_t^j \, \mathrm{d}x \quad \text{as} \quad N \to \infty. \tag{3.21}$$

Since $|\varphi| \le C$, one also has

$$|I_2| \le \frac{C C_j^2 N^{j-1}}{N^j} \to 0 \quad \text{as} \quad N \to \infty. \tag{3.22}$$

Let $N \to \infty$ in (3.20) and combining (3.21), (3.22), there exists a subsequence of $I_t^{j,N}$ (without relabeling) such that

$$I_t^{j,N} \to \int_{\mathbb{R}^{dj}} \varphi \rho_t^j \, \mathrm{d}x \quad \text{as} \quad N \to \infty. \tag{3.23}$$

On the other hand, for any $\varphi \in C_b(\mathbb{R}^{dj})$ and $m \in \mathbf{P}(\mathbb{R}^d)$, define $\Psi(m) := \langle \varphi, m \otimes \cdots \otimes m \rangle$. By induction from (4.20) below in Lemma 4.4, one can deduce that $\Psi \in C_b(\mathbf{P}(\mathbb{R}^d))$. Since $\mu_t^N \to \mu_t$ in law as $N \to \infty$, then

$$I_t^{j,N} \to \mathbb{E}\left[\langle \varphi, \mu_t \otimes \cdots \otimes \mu_t \rangle\right] \quad \text{as} \quad N \to \infty. \tag{3.24}$$

Combining (3.23) and (3.24), we obtain (3.19).

Step 3 Now we prove that μ_t has a density ρ_t a.s. for any time $t \ge 0$.

By strong law of large numbers, for any $\varphi \in C_b(\mathbb{R}^d)$, one has

$$\text{a.s. w.r.t } (\Omega, \mathcal{F}, \mathbb{P}) \quad \langle \mu_0^N, \varphi \rangle \to \langle f_0, \varphi \rangle \quad \text{as} \quad N \to \infty. \tag{3.25}$$

Since $\langle \mu_0^N, \varphi \rangle$ is uniformly bounded a.s., then

$$\mathbb{E}[|\langle \mu_0^N - f_0, \varphi \rangle|] \to 0 \quad \text{as} \quad N \to \infty, \tag{3.26}$$

which implies that μ_0^N converges in law to the constant random variable f_0 by Proposition 3.3. in [14]. Since ρ_0 is the density of f_0, then $\mu_0 = f_0$ has a density ρ_0.

For $t > 0$, let $\pi_t = \mathcal{L}(\mu_t) \in \mathbf{P}\left(\mathbf{P}(\mathbb{R}^d)\right)$ and define the projection $\pi_t^j = \int_{\mathbf{P}(\mathbb{R}^d)} g^{\otimes j}\, \pi_t(\mathrm{d}g) \in \mathbf{P}(\mathbb{R}^{dj})$ for any $j \geq 1$ in the following sense

$$\forall \varphi \in C_b(\mathbb{R}^{dj}), \quad \langle \pi_t^j, \varphi \rangle := \int_{g \in \mathbf{P}(\mathbb{R}^d)} \int_{\mathbb{R}^{dj}} \varphi(X)\, g^{\otimes j}(\mathrm{d}X) \pi_t(\mathrm{d}g).$$

Then $\mathbb{E}\left[\langle \varphi, \mu_t \otimes \cdots \otimes \mu_t \rangle\right] = \langle \pi_t^j, \varphi \rangle$. From Step 2, we know that $f_t^{(j),N}$ narrowly converges to π_t^j as $N \to \infty$ for all $j \geq 1$. Then combining the uniform estimates (2.4) and applying Theorem 4.1 in [4] (a refined version of the de Finetti–Hewitt–Savage theorem), μ_t has a density denoted by ρ_t a.s. such that

$$\mathbb{E}[H_1(\mu_t)] = \int_{\mathbf{P}(\mathbb{R}^d)} H_1(g)\, \pi_t(\mathrm{d}g) = \sup_{j \geq 1} H_j(\pi_t^j) \leq \liminf_{N \to \infty} H_N(f_t^N) < \infty, \tag{3.27}$$

$$\mathbb{E}[I_1(\mu_t)] = \int_{\mathbf{P}(\mathbb{R}^d)} I_1(g)\, \pi_t(\mathrm{d}g) = \sup_{j \geq 1} I_j(\pi_t^j) \leq \liminf_{N \to \infty} I_N(f_t^N) \tag{3.28}$$

where the last inequality of (3.27) comes from (2.2) and (2.4).

4 The self-consistent martingale problem

As a preparatory work, recalling directly from the definition of time marginal law and the probability measure on the path space for a stochastic process, we have the following lemma.

Lemma 4.1 *Let $(X_t)_{0 \leq t \leq T} \in C([0,T];\mathbb{R}^d)$ be a stochastic process, $\mu \in \mathbf{P}(C([0,T];\mathbb{R}^d))$ be the law of (X_t), and $\mu_t(x)$ be the time marginal law of (X_t) on the space $\mathbb{R}^d$. Then for any $\psi \in C_b(\mathbb{R}^d)$ and $t \in [0,T]$,*

$$\int_{C([0,T];\mathbb{R}^d)} \psi(X_t)\, d\mu(X) = \int_{\mathbb{R}^d} \psi(x)\, d\mu_t(x).$$

The following lemma gives a standard method of checking a stochastic process to be a solution to the martingale problem in Definition 1, and it is stated in [3, p. 174] without a proof. For completeness, we give a detail proof below.

Lemma 4.2 *A probability measure $\mu \in \mathbf{P}(C([0,T];\mathbb{R}^d))$ with time marginal μ_0 at $t = 0$, endowed with a μ-distributed canonical process $(X_t)_{0 \leq t \leq T} \in C([0,T];\mathbb{R}^d)$, is a solution to the $\left(g, C_b^2(\mathbb{R}^d)\right)$-self-consistent martingale problem with the initial distribution μ_0 in Definition 1 if and only if*

$$\mathbb{E}\left[\prod_{k=1}^{n} h_k(X_{t_k}) \left(\varphi(X_t) - \varphi(X_{t_n}) - \int_{t_n}^{t} g(X_r, \mathcal{L}(X_r))\, \mathrm{d}r\right)\right] = 0 \tag{4.1}$$

for all $\varphi \in C_b^2(\mathbb{R}^d)$, whenever $0 \leq t_1 < \cdots < t_n < t \leq T$, $h_1, \ldots, h_n \in B(\mathbb{R}^d)$ (or equivalently $h_1, \ldots, h_n \in C_b(\mathbb{R}^d)$), where $B(\mathbb{R}^d)$ is the space of bounded Borel measurable functions.

Proof (*i*) If $(X_t)_{0\le t\le T}$ is a solution to the $(g, C_b^2(\mathbb{R}^d))$-self-consistent martingale problem with the initial distribution μ_0 in Definition 1, i.e., let $\mathcal{M}_t = \varphi(X_t) - \varphi(X_0) - \int_0^t g(X_r, \mathcal{L}(X_r))\,\mathrm{d}r$, then $(\mathcal{M}_t)_{0\le t\le T}$ is a martingale w.r.t. the filtration $\{\mathcal{B}_t\}_{0\le t\le T}$, and then

$$\begin{aligned}
&\mathbb{E}\left[\prod_{k=1}^n h_k(X_{t_k})\left(\varphi(X_t)-\varphi(X_{t_n})-\int_{t_n}^t g(X_r,\mathcal{L}(X_r))\,\mathrm{d}r\right)\right]\\
&=\mathbb{E}\left[\mathbb{E}\left[\prod_{k=1}^n h_k(X_{t_k})\left(\varphi(X_t)-\varphi(X_{t_n})-\int_{t_n}^t g(X_r,\mathcal{L}(X_r))\,\mathrm{d}r\right)\,\Big|\,\mathcal{B}_{t_n}\right]\right]\\
&=\mathbb{E}\left[\prod_{k=1}^n h_k(X_{t_k})\mathbb{E}\left[\left(\mathcal{M}_t-\mathcal{M}_{t_n}\right)|\mathcal{B}_{t_n}\right]\right]=0,
\end{aligned}\tag{4.2}$$

where the first and second equalities come from Theorem B.2. *b*) and *e*) in [17], respectively.

(*ii*) By the definition of martingale, in order to prove $(\mathcal{M}_t)_{0\le t\le T}$ is a martingale w.r.t. the filtration $\{\mathcal{B}_t\}_{0\le t\le T}$, one need to show that for any $0 < s < t \le T$,

$$\mathbb{E}[(\mathcal{M}_t - \mathcal{M}_s)|\mathcal{B}_s] = 0. \tag{4.3}$$

For any $0 \le t_1 < \cdots < t_n = s < t \le T$ and $A_k \in \mathcal{B}(\mathbb{R}^d)$ $(k = 1, \ldots, n)$, taking $\{h_k(x)\}_{k=1}^n$ as n indicator functions $\{1_{(x\in A_k)}\}_{k=1}^n$, then $h_1, \ldots, h_n \in B(\mathbb{R}^d)$. If (4.1) holds, then

$$\mathbb{E}\left[\prod_{k=1}^n 1_{(X_{t_k}\in A_k)}\left(\varphi(X_t)-\varphi(X_{t_n})-\int_{t_n}^t g(X_r,\mathcal{L}(X_r))\,\mathrm{d}r\right)\right]=0. \tag{4.4}$$

Now we show that (4.4) implies (4.3) by the contradiction method. If (4.3) does not hold, then there exists $0 < s < t \le T$ such that $\mathbb{E}\left[(\mathcal{M}_t - \mathcal{M}_s)|\mathcal{B}_s\right] \ne 0$, i.e.,

$$\mu\left((X_t) : \mathbb{E}\left[(\mathcal{M}_t - \mathcal{M}_s)|\mathcal{B}_s\right] > 0\right) > 0 \quad \text{or} \quad \mu\left((X_t) : \mathbb{E}\left[(\mathcal{M}_t - \mathcal{M}_s)|\mathcal{B}_s\right] < 0\right) > 0 \tag{4.5}$$

Without loss of generality, we assume that

$$\mu\left((X_t) : \mathbb{E}\left[(\mathcal{M}_t - \mathcal{M}_s)|\mathcal{B}_s\right] > 0\right) > 0. \tag{4.6}$$

Since $\mu\left((X_t) : \mathbb{E}[(\mathcal{M}_t - \mathcal{M}_s)|\mathcal{B}_s) \ge 1/k]\right)$ is an increase sequence and has the following inequality

$$\lim_{k\to\infty} \mu\left((X_t) : \mathbb{E}[(\mathcal{M}_t - \mathcal{M}_s)|\mathcal{B}_s)] \ge 1/k\right) = \mu\left((X_t) : \mathbb{E}\left[(\mathcal{M}_t - \mathcal{M}_s)|\mathcal{B}_s\right] > 0\right) > 0,$$

then there is a k_0 such that

$$\mu\left((X_t) : \mathbb{E}[(\mathcal{M}_t - \mathcal{M}_s)|\mathcal{B}_s)] \ge 1/k_0\right) > 0.$$

In other words,

$$B := \left\{(X_t) : \mathbb{E}[(\mathcal{M}_t - \mathcal{M}_s)|\mathcal{B}_s)] \ge 1/k_0\right\} \in \mathcal{B}_s \quad \text{and} \quad \mu(B) > 0. \tag{4.7}$$

From the definition of σ-complete algebra $\mathcal{B}_s$, there exists a sequence of $0 \le \bar{t}_1 < \cdots < \bar{t}_n = s < t \le T$ and $\bar{A}_k \in \mathcal{B}(\mathbb{R}^d)$ $(k = 1, \ldots, n)$ such that

$$\widetilde{B} := \{(X_t) : X_{\bar{t}_k} \in \bar{A}_k, k = 1, \ldots, n\} \subset B \quad \text{and} \quad \mu(\widetilde{B}) > 0. \tag{4.8}$$

we have

$$\begin{aligned}
&\mathbb{E}\left[\prod_{k=1}^{n} 1_{(X_{\tilde{t}_k}\in \tilde{A}_k)}\left(\varphi(X_t)-\varphi(X_{\tilde{t}_n})-\int_{\tilde{t}_n}^{t} g(X_r,\mathcal{L}(X_r))\,\mathrm{d}r\right)\right]\\
&=\mathbb{E}\left[1_{((X_t)\in\widetilde{B})}\left(\mathcal{M}_t-\mathcal{M}_s\right)\right]=\mathbb{E}\left[\mathbb{E}[1_{((X_t)\in\widetilde{B})}\left(\mathcal{M}_t-\mathcal{M}_s\right)|\mathcal{B}_s]\right]\\
&=\mathbb{E}\left[1_{((X_t)\in\widetilde{B})}\mathbb{E}\left[\left(\mathcal{M}_t-\mathcal{M}_s\right)|\mathcal{B}_s\right]\right]\geq\frac{1}{k_0}\mathbb{E}[1_{((X_t)\in\widetilde{B})}]=\frac{1}{k_0}\mu(\widetilde{B})>0,
\end{aligned}\tag{4.9}$$

which is a contradiction to (4.4).

By the fact that any bounded Borel measurable function can be approximated by a sequence of bounded continuous functions and using the dominated convergence theorem, one knows that (4.1) holds for any $h_1,\ldots,h_n \in B(\mathbb{R}^d)$ is equivalent for any $h_1,\ldots,h_n \in C_b(\mathbb{R}^d)$. □

From Lemma 4.2, for solving the martingale problem in Definition 1, we just need to prove (4.1). Therefore we construct a functional ψ on $C([0,T];\mathbb{R}^d)\times C([0,T];\mathbb{R}^d)$ in the following way: For any $0\leq t_1<\cdots<t_n<t\leq T$, $\varphi\in C_b^2(\mathbb{R}^d)$, $h_1,\ldots,h_n\in C_b(\mathbb{R}^d)$, $X,Y\in C([0,T];\mathbb{R}^d)$, define

$$\begin{aligned}
\psi(X,Y) &:= \prod_{k=1}^{n} h_k(X_{t_k})\left[\varphi(X_t)-\varphi(X_{t_n})-\int_{t_n}^{t}\nabla\varphi(X_s)\cdot F(X_s-Y_s)\,\mathrm{d}s\right.\\
&\qquad\left.-\int_{t_n}^{t}\triangle\varphi(X_s)\,\mathrm{d}s\right]
\end{aligned}\tag{4.10}$$

$$\begin{aligned}
\psi_\varepsilon(X,Y) &:= \prod_{k=1}^{n} h_k(X_{t_k})\left[\varphi(X_t)-\varphi(X_{t_n})-\int_{t_n}^{t}\nabla\varphi(X_s)\cdot F_\varepsilon(X_s-Y_s)\,\mathrm{d}s\right.\\
&\qquad\left.-\int_{t_n}^{t}\triangle\varphi(X_s)\,\mathrm{d}s\right]
\end{aligned}\tag{4.11}$$

We also define a functional on $\mathbf{P}(C([0,T];\mathbb{R}^d))$ below, for any $Q\in\mathbf{P}(C([0,T];\mathbb{R}^d))$,

$$\mathcal{K}_\psi(Q)=\int_{C\times C}\psi(X,Y)\,Q(\mathrm{d}X)Q(\mathrm{d}Y),\tag{4.12}$$

then we have the following martingale estimate lemma.

Lemma 4.3 *For $N\geq 2$ and $d\geq 2$, let $\left\{(X_t^{i,N})_{t\geq 0}\right\}_{i=1}^N$ be the unique solution to (1.1) with the i.i.d. initial random variables $\{X_0^{i,N}\}_{i=1}^N$. Set $\mu^N=\frac{1}{N}\sum_{i=1}^N\delta_{X_t^{i,N}}$, then there exists a constant C (depending only on $\|\varphi\|_{C_b^1(\mathbb{R}^d)}$,s $\|h_1\|_{C_b(\mathbb{R}^d)},\ldots,\|h_n\|_{C_b(\mathbb{R}^d)}$ and T) such that*

$$\mathbb{E}[(\mathcal{K}_\psi(\mu^N))^2]\leq\frac{C}{N}\quad\text{for all } 0<t\leq T.\tag{4.13}$$

Proof By the definition of $\mathcal{K}_\psi(Q)$, simple computation shows that

$$\begin{aligned}
\mathcal{K}_\psi(\mu^N) &= \int_{C^2}\psi(X,Y)\,\mu^N(\mathrm{d}X)\mu^N(\mathrm{d}Y)=\frac{1}{N^2}\sum_{i,j=1}^{N}\psi(X_t^{i,N},X_t^{j,N})\\
&=\frac{1}{N^2}\sum_{i,j=1}^{N}\left\{\prod_{k=1}^{n}h_k\left(X_{t_k}^{i,N}\right)\left[\varphi(X_t^{i,N})-\varphi(X_{t_n}^{i,N})\right.\right.\\
&\qquad\left.\left.-\int_{t_n}^{t}\nabla\varphi(X_s^{i,N})\cdot F(X_s^{i,N}-X_s^{j,N})\,\mathrm{d}s-\int_{t_n}^{t}\triangle\varphi(X_s^{i,N})\,\mathrm{d}s\right]\right\}.
\end{aligned}\tag{4.14}$$

Using the Itô formula, for any $1 \le i \le N$, $\varphi \in C_b^2(\mathbb{R}^d)$, one has

$$\varphi(X_t^{i,N}) = \varphi(X_0^{i,N}) + \frac{C^*}{N}\sum_{j=1}^{N}\int_0^t \frac{\nabla\varphi(X_s^{i,N})\cdot(X_s^{i,N}-X_s^{j,N})}{|X_s^{i,N}-X_s^{j,N}|^d}\,\mathrm{d}s + \int_0^t \triangle\varphi(X_s^{i,N})\,\mathrm{d}s \\ + \sqrt{2}\int_0^t \nabla\varphi(X_s^{i,N})\cdot \mathrm{d}B_s^i. \tag{4.15}$$

Plugging (4.15) into (4.14), one has

$$\mathcal{K}_\psi(\mu^N) = \frac{\sqrt{2}}{N}\sum_{i=1}^{N}\prod_{k=1}^{n} h_k(X_{t_k}^{i,N})\int_{t_n}^t \nabla\varphi(X_s^{i,N})\cdot \mathrm{d}B_s^i. \tag{4.16}$$

Then one has

$$\mathbb{E}[|\mathcal{K}_\psi(\mu^N)|^2] = \frac{2}{N^2}\mathbb{E}\left[\left|\sum_{i=1}^{N} h_1(X_{t_1}^{i,N})\ldots h_n(X_{t_n}^{i,N})\int_{t_n}^t \nabla\varphi(X_s^{i,N})\cdot \mathrm{d}B_s^i\right|^2\right]. \tag{4.17}$$

Denoting $M^i = h_1(X_{t_1}^{i,N})\ldots h_n(X_{t_n}^{i,N})\int_{t_n}^t \nabla\varphi(X_s^{i,N})\cdot \mathrm{d}B_s^i$. Since the Brownian motions $\{(B_t^i)_{t\ge 0}\}_{i=1}^N$ are independent, then when $i \neq j$,

$$\mathbb{E}\left[M^iM^j\right] = 0,$$

and then

$$\mathbb{E}\left[\left|\sum_{i=1}^{N} M^i\right|^2\right] = \mathbb{E}\left[\sum_{i=1}^{N}|M^i|^2\right] \le NC, \tag{4.18}$$

where C depends only on T, $\|\varphi\|_{C_b^1(\mathbb{R}^d)}$ and $\|h_1\|_{C_b(\mathbb{R}^d)},\ldots,\|h_n\|_{C_b(\mathbb{R}^d)}$. Plugging (4.18) into (4.17), one can achieve (4.13) immediately. □

Lemma 4.4 *Let E be a polish space. Assume a sequence of $\mathbf{P}(E)$-valued random variables μ^N converge in law to a random measure μ. For any $\psi(x,y) \in C_b(E\times E)$ and $Q \in \mathbf{P}(E)$, define a functional $\mathcal{K}_\psi : \mathbf{P}(E) \to \mathbb{R}$, $Q \mapsto \mathcal{K}_\psi(Q) = \int_{E^2}\psi(x,y)\,Q(dx)Q(dy)$. Then*

$$\mathcal{K}_\psi(\mu^N) \to \mathcal{K}_\psi(\mu) \quad \textit{in law as } \; N\to\infty. \tag{4.19}$$

Proof For any $Q \in \mathbf{P}(E)$, $\psi(x,y) \in C_b(E\times E)$ and $\varphi \in C_b(\mathbb{R})$, define a functional Γ : $\mathbf{P}(E) \to \mathbb{R}$, $Q \mapsto \Gamma(Q) = \varphi\left(\mathcal{K}_\psi(Q)\right)$. We prove that

$$\Gamma \in C_b(\mathbf{P}(E)). \tag{4.20}$$

Here, the space $\mathbf{P}(E)$ is endowed with a metric induced by the narrowly convergence, and it is a Polish space too. Note that $\varphi \in C_b(\mathbf{P}(E))$ if and only if a sequence $\mu^N(\in \mathbf{P}(E))$ narrowly converge to μ as $N\to\infty \Rightarrow \varphi(\mu^N)$ converges to $\varphi(\mu)$ as $N\to\infty$.

For any sequence $Q^N(\in \mathbf{P}(E))$ narrowly converge to Q, by [1, p. 23, Theorem 2.8], the following convergence result holds,

$$\mathcal{K}_\psi(Q^N) \to \mathcal{K}_\psi(Q),$$

hence $\varphi\left(\mathcal{K}_\psi(Q^N)\right) \to \varphi\left(\mathcal{K}_\psi(Q)\right)$, i.e., (4.20) holds.

Since the sequence μ^N converges in law to μ, then

$$\mathbb{E}\left[\varphi\left(\mathcal{K}_\psi(\mu^N)\right)\right] \to \mathbb{E}\left[\varphi\left(\mathcal{K}_\psi(\mu)\right)\right] \quad \text{as } \; N\to\infty,$$

which gives (4.19). □

Proposition 4.1 *For $d = 2$, let $\big\{(X_t^{i,N})_{0\le t\le T}\big\}_{i=1}^N$ be the unique solution to (1.1) with the i.i.d initial data $\{X_0^{i,N}\}_{i=1}^N$ and the common initial distribution f_0 and density ρ_0 satisfies $H(\rho_0) < \infty, m_2(\rho_0) < \infty$. Suppose μ is the limited $\mathbf{P}(C([0,T];\mathbb{R}^d))$-valued random variable of a subsequence of empirical measures $\mu^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N}}$. Then there exists a μ-distributed canonical process $(X_t)_{0\le t\le T} \in C([0,T];\mathbb{R}^d)$, and μ is a.s. solution to the $\big(g, C_b^2(\mathbb{R}^d)\big)$-self-consistent martingale problem with the initial distribution f_0 in Definition 1.*

Proof Since $C([0,T];\mathbb{R}^d)$ is a metric space, then for any $\mu \in \mathbf{P}(C([0,T];\mathbb{R}^d))$, there exists a probability space $\big(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}}\big)$ and μ-distributed random variable $(X_t)_{0\le t\le T} \in C([0,T];\mathbb{R}^d)$. One can take the probability space as $(C([0,T];\mathbb{R}^d), \mathcal{B}, \mu)$ as defined in Definition 1 and the random variable as the identity map.

For the $\mathbf{P}(C([0,T];\mathbb{R}^d))$-valued random variable μ in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, recalling $\mathcal{M}_t$ in Definition 1, let

$$\begin{aligned}\mathcal{M}_t(\mu) = \varphi(X_t) - \varphi(X_0) &- \int_0^t \int_{C([0,T];\mathbb{R}^d)} \nabla\varphi(X_s)\cdot F(X_s - Y_s)\mu(\mathrm{d}Y)\,\mathrm{d}s \\ &- \int_0^t \triangle\varphi(X_s)\,\mathrm{d}s,\end{aligned}$$

where $\mathcal{L}(X_0) = f_0$.

To verify $(\mathcal{M}_t(\mu))_{0\le t\le T}$ is a martingale w.r.t. the filtration $\{\mathcal{B}_t\}_{0\le t\le T}$ a.s. for μ in $(\Omega, \mathcal{F}, \mathbb{P})$, by Lemma 4.2, one only needs to show that (4.1) holds a.s. w.r.t $(\Omega, \mathcal{F}, \mathbb{P})$. Then by the definition of the function ψ in (4.10) and the functional $\mathcal{K}_\psi$ in (4.12), (4.1) equals the following equality

$$\mathcal{K}_\psi(\mu) = 0 \quad \text{a.s. w.r.t} \quad (\Omega, \mathcal{F}, \mathbb{P}) \tag{4.21}$$

holds.

Following the spirit of [4], one has

$$\mathbb{E}[|\mathcal{K}_\psi(\mu)|] \le \mathbb{E}[|\mathcal{K}_\psi(\mu) - \mathcal{K}_{\psi_\varepsilon}(\mu)|] + \mathbb{E}[|\mathcal{K}_{\psi_\varepsilon}(\mu)|]. \tag{4.22}$$

here ψ_ε and $\mathcal{K}_{\psi_\varepsilon}$ are defined by (4.11) and (4.12).

It is obvious that $\psi_\varepsilon \in C_b(C([0,T];\mathbb{R}^d) \times C([0,T];\mathbb{R}^d))$ for any fixed $\varepsilon > 0$. Then combining (3.11) and using Lemma 4.4, we obtain that

$$\lim_{N\to\infty} \mathbb{E}[|\mathcal{K}_{\psi_\varepsilon}(\mu^N)|] = \mathbb{E}\left[\mathcal{K}_{\psi_\varepsilon}(\mu)|\right], \tag{4.23}$$

Define $A_1(\varepsilon) = \mathbb{E}[|\mathcal{K}_\psi(\mu) - \mathcal{K}_{\psi_\varepsilon}(\mu)|]$ and $A_2(\varepsilon, N) = \mathbb{E}[|\mathcal{K}_{\psi_\varepsilon}(\mu^N)|]$, (4.22) equals to

$$\mathbb{E}[|\mathcal{K}_\psi(\mu)|] \le A_1(\varepsilon) + \lim_{N\to\infty} A_2(\varepsilon, N). \tag{4.24}$$

Combining the fact $|F_\varepsilon(x)| \le |F(x)|$, $F_\varepsilon(x) = F(x)$ for $|x| \ge \varepsilon$ by Lemma 2.1 and Lemma 4.1, there exists a constant C (depending only on d, $\|\varphi\|_{C_b^1(\mathbb{R}^d)}, \|h_1\|_{C_b(\mathbb{R}^d)}, \ldots, \|h_n\|_{C_b(\mathbb{R}^d)}$ and T) such that

$$\begin{aligned}A_1(\varepsilon) &= \mathbb{E}\left[\left|\int_{t_n}^t \int_{C\times C} \prod_{k=1}^n h_k(X_{t_k})\nabla\varphi(X_s)\cdot [F_\varepsilon(X_s - Y_s) - F(X_s - Y_s)]\ \mu(\mathrm{d}X)\mu(\mathrm{d}Y)\mathrm{d}s\right|\right] \\ &\le C\mathbb{E}\left[\int_{t_n}^t \int_{C\times C} \big|F_\varepsilon(X_s - Y_s) - F(X_s - Y_s)\big|\ \mu(\mathrm{d}X)\mu(\mathrm{d}Y)\mathrm{d}s\right] \\ &\le 2C\mathbb{E}\left[\int_{t_n}^t \int_{|X-Y|\le\varepsilon} \big|F(X - Y)\big|\ \mu_s(\mathrm{d}X)\mu_s(\mathrm{d}Y)\mathrm{d}s\right].\end{aligned}$$

Since μ_s has a density ρ_s a.s. by Lemma 3.2, then

$$A_1(\varepsilon) \le C\mathbb{E}\left[\int_{t_n}^t \int_{|X-Y|\le\varepsilon} \frac{\rho_s(X)\rho_s(Y)}{|X-Y|^{d-1}}\,\mathrm{d}X\,\mathrm{d}Y\,\mathrm{d}s\right]. \tag{4.25}$$

By Lemma 4.3, there exists a constant C (depending only on $\|\varphi\|_{C_b^1(\mathbb{R}^d)}, \|h_1\|_{C_b(\mathbb{R}^d)}, \ldots, \|h_n\|_{C_b(\mathbb{R}^d)}$ and T) such that

$$\begin{aligned} A_2(\varepsilon, N) = \mathbb{E}[|\mathcal{K}_{\psi_\varepsilon}(\mu^N)|] &\le \mathbb{E}[|\mathcal{K}_{\psi_\varepsilon}(\mu^N) - \mathcal{K}_\psi(\mu^N)|] + \mathbb{E}[|\mathcal{K}_\psi(\mu^N)|] \\ &\le \mathbb{E}[|\mathcal{K}_{\psi_\varepsilon}(\mu^N) - \mathcal{K}_\psi(\mu^N)|] + \frac{C}{\sqrt{N}}. \end{aligned} \tag{4.26}$$

From (4.14), one has

$$\begin{aligned} |\mathcal{K}_{\psi_\varepsilon}(\mu^N) - \mathcal{K}_\psi(\mu^N)| &= \Bigg| \frac{1}{N^2}\sum_{i,j=1}^N \prod_{k=1}^n h_k\left(X_{t_k}^{i,N}\right) \int_{t_n}^t \nabla\varphi(X_s^{i,N}) \\ &\qquad \cdot \left[F_\varepsilon(X_s^{i,N} - X_s^{j,N}) - F(X_s^{i,N} - X_s^{j,N})\right]\mathrm{d}s \Bigg| \\ &\le \frac{C}{N^2}\sum_{i,j=1}^N \int_{t_n}^t \left|F_\varepsilon(X_s^{i,N} - X_s^{j,N}) - F(X_s^{i,N} - X_s^{j,N})\right|\mathrm{d}s, \end{aligned} \tag{4.27}$$

where C is a constant depending only on $\|\varphi\|_{C_b^1(\mathbb{R}^d)}, \|h_1\|_{C_b(\mathbb{R}^d)}, \ldots, \|h_n\|_{C_b(\mathbb{R}^d)}$.

Then by the exchangeability of $\{(X_t^i)_{t\ge0}\}_{i=1}^N$ and the fact $|F_\varepsilon(x)| \le |F(x)|$, $F_\varepsilon(x) = F(x)$ for $|x| \ge \varepsilon$, we have

$$\mathbb{E}[|\mathcal{K}_{\psi_\varepsilon}(\mu^N) - \mathcal{K}_\psi(\mu^N)|] \le C\int_{t_n}^t \int_{|x-y|<\varepsilon} \frac{\rho_s^{(2),N}(x,y)}{|x-y|^{d-1}}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s, \tag{4.28}$$

When $d = 2$, similarly with the proof of Lemma 3.3. in [4], one obtains that

$$\int_{|x-y|<\varepsilon} \frac{\rho_s^{(2),N}(x,y)}{|x-y|}\,\mathrm{d}x\mathrm{d}y \le C\varepsilon^{\frac{2-q}{q}}\{I_2(\rho_s^{(2),N})\}^{\frac{1}{q}}, \tag{4.29}$$

$$\int_{|X-Y|\le\varepsilon} \frac{\rho_s(X)\rho_s(Y)}{|X-Y|}\,\mathrm{d}X\mathrm{d}Y \le C\varepsilon^{\frac{2-q}{q}}\{I_2(\rho_s^{\otimes 2})\}^{\frac{1}{q}} = C\varepsilon^{\frac{2-q}{q}}\{I_1(\rho_s)\}^{\frac{1}{q}} \tag{4.30}$$

where $0 < q < 2$ and C is a constant depending only on q.

Plugging (4.29) and (4.30) into (4.28) and (4.25), respectively, one has

$$\begin{aligned} \mathbb{E}[|\mathcal{K}_{\psi_\varepsilon}(\mu^N) - \mathcal{K}_\psi(\mu^N)|] &\le C\varepsilon^{\frac{2-q}{q}} \int_{t_n}^t \{I_2(\rho_s^{(2),N})\}^{\frac{1}{q}}\,\mathrm{d}s \\ &\le C\varepsilon^{\frac{2-q}{q}}\left\{\int_{t_n}^t I_2(\rho_s^{(2),N})\,\mathrm{d}s\right\}^{\frac{1}{q}}, \end{aligned} \tag{4.31}$$

and

$$A_1(\varepsilon) \le C\varepsilon^{\frac{2-q}{q}}\mathbb{E}\left[\int_{t_n}^t \{I_1(\rho_s)\}^{\frac{1}{q}}\,\mathrm{d}s\right] \le C\varepsilon^{\frac{2-q}{q}}\left\{\int_{t_n}^t \mathbb{E}[I_1(\rho_s)]\,\mathrm{d}s\right\}^{\frac{1}{q}}. \tag{4.32}$$

where C is a constant depending only on q and T.

Using (3.13) and (3.14), there exists a constant C (depending only on T, $H_1(f_0)$ and $m_2(\rho_0)$) such that

$$\int_0^t I_2(\rho_s^{(2),N})\,\mathrm{d}s \le C \quad \text{for all } t \in [0, T]; \tag{4.33}$$

from (3.28), one also has

$$\int_0^t \mathbb{E}[I_1(\rho_s)]\,\mathrm{d}s \le \int_0^t \liminf_{N\to\infty} I_N(f_s^N)\,\mathrm{d}s \le C \quad \text{for all } \ t\in[0,T]. \tag{4.34}$$

Combining (4.26), (4.31) and (4.33), there exists a constant C (depending only on q, T, $H_1(f_0)$ and $m_2(\rho_0)$) such that

$$A_2(\varepsilon,N) \le C\varepsilon^{\frac{2-q}{q}} + \frac{C}{\sqrt{N}} \quad \text{for all } \ t\in[0,T], \quad 0<q<2. \tag{4.35}$$

Plugging (4.34) into (4.32), there exists a constant C (depending only on q, T, $H_1(f_0)$ and $m_2(\rho_0)$) such that

$$A_1(\varepsilon) \le C\varepsilon^{\frac{2-q}{q}} \quad \text{for all } \ t\in[0,T], \quad 0<q<2. \tag{4.36}$$

Plugging (4.35) and (4.36) into (4.24)

$$\mathbb{E}[|\mathcal{K}_\psi(\mu)|] \le C\varepsilon^{\frac{2-q}{q}} \quad \text{for all } \ t\in[0,T], \quad 0<q<2. \tag{4.37}$$

Let ε goes to 0, one obtains that

$$\mathbb{E}[|\mathcal{K}_\psi(\mu)|] = 0, \tag{4.38}$$

which means (4.21) holds. □

5 Propagation of chaos for 2D

5.1 The refined hyper-contractivity and uniqueness for the mean-field Poisson–Nernst–Planck equations 1.4

In this subsection, we prove the uniqueness of weak solution to 1.4 by the standard semigroup method, see [12]. We use the following definition of weak solution to (1.4).

Definition 4 (*Weak solution*) Let the initial data $\rho_0(x)\in L_+^1\cap L^{\frac{2d}{d+2}}(\mathbb{R}^d)$ and $T>0$. c is the potential associated with ρ and is given by $c(t,x)=\Phi*\rho(t,x)$. We shall say that $\rho(t,x)$ is a weak solution to (1.4) with the initial data $\rho_0(x)$ if it satisfies:

1. Regularity:

$$\rho\in L^\infty\left(0,T;L^1\cap L^{\frac{2d}{d+2}}(\mathbb{R}^d)\right), \qquad \rho^{\frac{d}{d+2}}\in L^2(0,T;H^1(\mathbb{R}^d)) \tag{5.1}$$

$$\text{and } \ \partial_t\rho\in L^p(0,T;W_{loc}^{-1,q}(\mathbb{R}^d)) \quad \text{for some } \ p,q\ge 1. \tag{5.2}$$

2. For all $\varphi\in C_0^\infty(\mathbb{R}^d)$ and $0<t\le T$, the following holds,

$$\begin{aligned}&\int_{\mathbb{R}^d}\rho(t,x)\varphi(x)\,\mathrm{d}x - \int_{\mathbb{R}^d}\rho_0(x)\varphi(x)\,\mathrm{d}x + \int_0^t\int_{\mathbb{R}^d}\nabla\varphi(x)\cdot\nabla\rho(s,x)\,\mathrm{d}x\,\mathrm{d}s\\ &\quad = \int_0^t\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\rho(s,x)\rho(s,y)F(x-y)\cdot\nabla\varphi(x)\,\mathrm{d}y\mathrm{d}x\,\mathrm{d}s.\end{aligned} \tag{5.3}$$

Remark 5.1 Notice that the regularity of $\rho(t,x)$ is enough to make sense of each term in (5.3). By the Hardy–Littlewood–Sobolev inequality, one has

$$\begin{aligned}&\left|\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\rho(s,x)\rho(s,y)F(x-y)\cdot\nabla\varphi(x)\,\mathrm{d}y\,\mathrm{d}x\right|\\ &\quad = \frac{C^*}{2}\left|-\int_{\mathbb{R}^{2d}}\frac{(\varphi(x)-\varphi(y))\cdot(x-y)}{|x-y|^2}\frac{\rho(s,x)\rho(s,y)}{|x-y|^{d-2}}\,\mathrm{d}x\,\mathrm{d}y\right|\\ &\quad \le \frac{C^*}{2}\int_{\mathbb{R}^{2d}}\frac{\rho(s,x)\rho(s,y)}{|x-y|^{d-2}}\,\mathrm{d}x\,\mathrm{d}y \le C(d)\|\rho(\cdot,s)\|^2_{L^{\frac{2d}{d+2}}(\mathbb{R}^d)}.\end{aligned} \tag{5.4}$$

Theorem 5.1 *The global weak solution to* (1.4) *is unique if the initial data* ρ_0 *satisfy the following conditions*

(i) $m_2(\rho_0) < \infty$ *and* $H_1(\rho_0) < \infty$ *for* $d = 2$*;*
(ii) $\|\rho_0\|_{L^{\frac{d}{2}+\gamma}} < \infty$ *for any* $0 < \gamma < 1$, $d \geq 3$.

Proof Follows the spirit of [12], we outline the proof briefly.

Step 1 From Eq. (1.4), for any $T > 0$, there is a uniform in time bound estimates:

$$\sup_{t\in[0,T]} H(\rho_t) \leq H(\rho_0), \quad \sup_{t\in[0,T]} m_2(\rho_t) \leq C(T, d, m_2(\rho_0)), \quad \sup_{t\in[0,T]} \|\rho_t\|_{L^p} \leq \|\rho_0\|_{L^p}.$$

There, one has the following estimates:

$$\sup_{t\in[0,T]} \int_{\mathbb{R}^2} \rho(t,x)|\ln\rho(t,x)|\,\mathrm{d}x \leq C(T, m_2(\rho_0), H_1(\rho_0)) \tag{5.5}$$

and
$$\sup_{t\in[0,T]} \|\rho(t,\cdot)\|_{L^{\frac{d}{2}+\gamma}} \leq \|\rho_0\|_{L^{\frac{d}{2}+\gamma}} \quad \text{for any } d \geq 3. \tag{5.6}$$

Step 2 For any fixed $T > 0$, $0 < \gamma < 1$, by (5.5) and (5.6), the refined hyper-contractivity holds,

$$t^{q-1}|\ln t|^{1-\gamma}\|\rho(\cdot,t)\|_{L^q}^q \leq C(d, q, T, m_2(\rho_0), H_1(\rho_0)) \\ \text{for any } t \in (0,1], q \geq 1, d = 2; \tag{5.7}$$

$$t^{q-\frac{d}{2}-\frac{2\gamma q}{d+2\gamma}}\|\rho(\cdot,t)\|_{L^q}^q \leq C(d, q, T, \|\rho_0\|_{L^{\frac{d}{2}+\gamma}}) \\ \text{for any } t \in (0,T], q \geq \frac{d}{2}+\gamma, d \geq 3. \tag{5.8}$$

Combining the above properties of refined hyper-contractivity with the standard semi-group theory, one can prove that there exists a time $0 < t_1 < T$ (depending only on $C(d, q, T, m_2(\rho_0), H_1(\rho_0))$ or $C(d, q, T, \|\rho_0\|_{L^{\frac{d}{2}+\gamma}})$) such that the weak solution to (1.4) is unique in $t \in [0, t_1]$.

Step 3 Finally, since t_1 is a constant only depending on $C(d, q, T, m_2(\rho_0), H_1(\rho_0))$ or $C(d, q, T, \|\rho_0\|_{L^{\frac{d}{2}+\gamma}})$, taking t_1 as a new initial time, repeating the above process, we have that model (1.4) has a unique weak solution in $t \in [t_1, 2t_1]$. One can continue this process and obtain a unique global solution in $[0, T)$. □

5.2 Propagation of chaos result

First, for $d = 2$, we show that the limited measure-valued random variable μ satisfies that: For any $\varphi \in C_b^2(\mathbb{R}^d)$ and $t \in [0, T]$, the time marginal measure $\mu_t \in \mathbf{P}(\mathbb{R}^d)$ a.s. solves the following equation

$$\langle \mu_t, \varphi\rangle = \langle f_0, \varphi\rangle + \int_0^t \int_{\mathbb{R}^{2d}} \nabla\varphi(x)\cdot F(x-y)\,\mu_s(\mathrm{d}x)\mu_s(\mathrm{d}y)\mathrm{d}s + \int_0^t \langle \mu_s, \triangle\varphi\rangle\,\mathrm{d}s. \tag{5.9}$$

Proposition 5.1 *For* $N \geq 2$ *and* $d = 2$*, let* $\{(X_t^{i,N})_{0\leq t\leq T}\}_{i=1}^N$ *be the unique solution to* (1.1) *with the i.i.d initial data* $\{X_0^{i,N}\}_{i=1}^N$ *and the common density* ρ_0 *satisfies* $H(\rho_0) < \infty$, $m_2(\rho_0) < \infty$*. Then the limited measure-valued process* μ_t *of the subsequence processes* $\mu_t^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N}}$ *(without relabeling) a.s. satisfies* (5.9)*.*

Proof From Proposition 4.1, we know that

$$\mathcal{M}_t(\mu) = \varphi(X_t) - \varphi(X_0) - \int_0^t \int_{C([0,T];\mathbb{R}^d)} \nabla\varphi(X_s) \cdot F(X_s - Y_s)\mu(\mathrm{d}Y)\,\mathrm{d}s$$
$$- \int_0^t \triangle\varphi(X_s)\,\mathrm{d}s.$$

is a martingale w.r.t. the filtration $\{\mathcal{B}_t\}_{0\le t\le T}$ a.s. in $(\Omega, \mathcal{F}, \mathbb{P})$, i.e.,

$$\mathbb{E}_\mu[\mathcal{M}_t(\mu)] = 0 \quad \text{a.s. w.r.t} \quad (\Omega, \mathcal{F}, \mathbb{P}), \tag{5.10}$$

where $\mathbb{E}_\mu$ means taking the expectation in the probability space $(C([0,T];\mathbb{R}^d), \mathcal{B}, \mathcal{B}_t, \mu)$. Applying Lemma 4.1, we have

$$\langle \mu_t, \varphi\rangle - \langle f_0, \varphi\rangle - \int_0^t \int_{\mathbb{R}^{2d}} \nabla\varphi(x) \cdot F(x-y)\,\mu_s(\mathrm{d}x)\mu_s(\mathrm{d}y)\,\mathrm{d}s - \int_0^t \langle \mu_s, \triangle\varphi\rangle\,\mathrm{d}s = 0 \quad \text{a.s.,} \tag{5.11}$$

one obtains (5.9) immediately. □

Next, we recall the following standard equivalent notions of propagation of chaos from the lecture of Sznitman [23, Proposition 2.2].

Definition 5 Let E be a polish space. A sequence of symmetric probability measures f^N on E^N are said to be f-chaotic; f is a probability measure on E, if one of three following equivalent conditions is satisfied:

(i) The sequence of second marginals $f^{2,N} \rightharpoonup f \otimes f$ as $N \to \infty$;
(ii) For all $j \ge 1$, the sequence of j-th marginals $f^{j,N} \rightharpoonup f^{\otimes j}$ as $N \to \infty$;
(iii) The empirical measure $\frac{1}{N}\sum_{i=1}^N \delta_{X^{i,N}}$ ($X^{i,N}$, $i = 1, \ldots, N$ are canonical coordinates on E^N) converges in law to the constant random variable f as $N \to \infty$.

Finally, putting together some results above, we have the following propagation of chaos result.

Theorem 5.2 *For $d = 2$, let $\{(X_t^{i,N})_{0\le t\le T}\}_{i=1}^N$ be the unique solution to (1.1) with the i.i.d initial data $\{X_0^{i,N}\}_{i=1}^N$ and the common initial density ρ_0 satisfies $H(\rho_0) < \infty$, $m_2(\rho_0) < \infty$. Then the empirical measure $\frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N}}$ goes in probability to a deterministic measure μ in $\mathbf{P}(C([0,T];\mathbb{R}^d))$ as $N \to \infty$. Furthermore, $(\mu_t)_{t\ge 0}$ has a density $(\rho_t)_{t\ge0}$, ρ_t takes the initial density ρ_0 at time $t = 0$, and ρ_t is the unique weak solution to (1.4).*

Proof Let $\mu^N := \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N}}$. First, by Lemma 3.1, one knows the sequence $\mathcal{L}(\mu^N)$ is tight in $\mathbf{P}\left(\mathbf{P}(C([0,T];\mathbb{R}^d))\right)$. Denote μ as a limiting point of a subsequence of μ^N. Then by Proposition 5.1, one knows that μ_t satisfies (5.9) a.s.. And Lemma 3.2 shows that $(\mu_t)_{t\ge0}$ has a density $(\rho_t)_{t\ge0}$ a.s. and ρ_t takes the initial density ρ_0 at time $t = 0$. Recalling equation (5.9), we deduce that

$$\int_{\mathbb{R}^d} \varphi\rho_t\,\mathrm{d}x = \int_{\mathbb{R}^d} \varphi\rho_0\,\mathrm{d}x + \frac{1}{2\pi}\int_0^t \int_{\mathbb{R}^{2d}} \nabla\varphi(x) \cdot \frac{x-y}{|x-y|^2}\rho_s(x)\rho_s(y)\,\mathrm{d}x\mathrm{d}y\mathrm{d}s$$
$$+ \int_0^t \int_{\mathbb{R}^d} \triangle\varphi\rho_s\,\mathrm{d}x\mathrm{d}s, \tag{5.12}$$

i.e., ρ_t a.s. is a weak solution to (1.4) with the initial data ρ_0. Finally, by the uniqueness of weak solution to (1.4) from Theorem 5.1, ρ_t is deterministic, which completes the proof of Theorem 5.2 immediately. □

Finally, we make a remark on the possible using stochastic PDE method.

Remark 5.2 For any test function $\varphi \in C_b(\mathbb{R}^d)$, setting $F(0) = 0$ and using the fact from (*i*) of Theorem 2.1 $X_t^i \neq X_t^j$ a.s. for all $t \in [0, T]$, $i \neq j$, by Itô's formula, one has the following stochastic equation

$$\langle \mu_t^N - \mu_0^N, \varphi \rangle = \frac{1}{2} \int_0^t \int_{\mathbb{R}^{2d}} (\nabla\varphi(X) - \nabla\varphi(Y)) \cdot F(X - Y)\, \mu_s^N(\mathrm{d}X)\mu_s^N(\mathrm{d}Y)\mathrm{d}s$$
$$+ \int_0^t \int_{\mathbb{R}^d} \triangle\varphi(X)\mu_s^N(\mathrm{d}X)\mathrm{d}s + \frac{\sqrt{2}}{N} \sum_{i=1}^N \int_0^t \nabla\varphi(X_s^i) \cdot \mathrm{d}B_s^i. \tag{5.13}$$

Then passing to the limit $N \to \infty$, the nonlinear term $\int_0^t \int_{\mathbb{R}^{2d}} (\nabla\varphi(X) - \nabla\varphi(Y)) \cdot F(X - Y)\, \mu_s^N(\mathrm{d}X)\mu_s^N(\mathrm{d}Y)\mathrm{d}s$ is the difficult one. Set $\psi(X, Y) = (\nabla\varphi(X) - \nabla\varphi(Y)) \cdot F(X - Y)$ in (4.19) of Lemma 4.4. Notice that ψ has a singularity $\frac{1}{|X-Y|^{d-2}}$, it requires more delicate estimates to pass $\mathcal{K}_\psi(\mu^N) \to \mathcal{K}_\psi(\mu)$ in law as $N \to \infty$. We leave this question for future.

Acknowledgements
We would like to thank Yu Gao and Hui Huang for their careful proofreading and some helpful discussions. The research of J.-G. L. was partially supported by KI-Net NSF RNMS Grant No. 1107291 and NSF DMS Grant No. 1514826.
Dedicated to the seventieth birthday of Professor Björn Engquist.

Appendix: Metrization of $\mathbb{P}(C([0,T];\mathbb{R}^d))$

First, by the Stone-Weierstrass theorem, it is well known that $C([0, T]; \mathbb{R}^d)$ is a separable space; hence, it is a polish space. Then there is a dense sequence $(\varphi_n)_{n\in\mathbb{N}}$ in $C_0\left(C([0, T]; \mathbb{R}^d)\right)$. One can define the weak-$*$ distance [25, page 98],

$$\text{for any } g_1, g_2 \in \mathbf{P}(C([0, T]; \mathbb{R}^d)), \qquad d_1(g_1, g_2) := \sum_{n\in\mathbb{N}} \frac{1}{2^n}(1 \wedge |\langle g_1 - g_2, \varphi_n\rangle|),$$

then $(\mathbf{P}(C([0, T]; \mathbb{R}^d)), d_1)$ is a Polish space [9, Section 15.7].

Received: 5 May 2016 Accepted: 27 September 2016
Published online: 14 November 2016

References

1. Billingsley, P.: Convergence of Probability Measures, 2nd edn. Wiley-Interscience, New York (1999)
2. Durrett, R.: Probability: Theory and Examples, 3rd edn. Engage Learning (2007) (ISBN-10: 0534424414, ISBN-13: 9780534424411)
3. Ethier, S.N., Thomas, G.K.: Markov Processes: Characterization and Convergence. Wiley, New York (1986)
4. Fournier, N., Hauray, M., Mischler, S.: Propagation of chaos for the 2D viscous vortex model. J. Eur. Math. Soc. **16**(7), 1425–1466 (2014)
5. Fournier, N., Jourdain, B.: Stochastic particle approximation of the Keller–Segel equation and two-dimensional generalization of Bessel processes (2015). arXiv:1507.01087
6. Godinho, D., Quininao, C.: Propagation of chaos for a subcritical Keller–Segel model. Ann. Inst. Henri Poincaré Probab. Stat. Inst. Henri Poincaré **51**(3), 965–992 (2015)
7. Hauray, M., Mischler, S.: On Kac's chaos and related problems. J. Funct. Anal. **266**, 6055–6157 (2014)
8. Kac, M.: Foundation of kinetic theory. In: Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, vol. 3, pp. 171–197. University of California, California (1956)
9. Kallenberg, O.: Random Measures, 3rd edn. Akademie-Verlag, Berlin (1983)
10. Klebaner, F.C.: Interaction to Stochastic Calculus with Applications, 3rd edn. Imperial College Press, London (2012)
11. Kolokoltsov, V.N.: Nonlinear Markov Processes and Kinetic Equations. Cambridge University Press, New York (2010)
12. Liu, J.-G., Wang, J.: Refined hyper-contractivity and uniqueness for the Keller–Segel equations. Appl. Math. Lett. **52**, 212–219 (2016)
13. Liu, J.-G., Yang, R.: Propagation of chaos for the Keller–Segel equation with a logarithmic cut-off (preprint)
14. Liu, J.-G., Yang, R.: A random particle blob method for the Keller–Segel equation and convergence analysis. Math. Comp. (to appear)

15. Méléard, S.: Asymptotic behaviour of some interacting particle systems; McKean–Vlasov and Boltzmann models. In: Probabilistic Models for Nonlinear Partial Differential Equations (Montecatini Terme, 1995). Lecture Notes in Math., vol. 1627, pp. 42–95. Springer, Berlin (1996)
16. McKean, H. P.: Propagation of chaos for a class of non-linear parabolic equation. In: Lecture Series in Differential Equations, Session 7, pp. 177–194. Catholic University, Washington (1967)
17. Øksendal, B.: Stochastic Differential Equations: An Introduction with Applications, 6th edn. Springer, Berlin (2003)
18. Osada, H.: A stochastic differential equation arising from the vortex problem. Proc. Japan Acad. Ser. A Math. Sci. **61**(10), 333–336 (1985)
19. Osada, H.: Propagation of chaos for the two-dimensional Navier–Stokes equation. Proc. Japan Acad. Ser. A Math. Sci. **62**(1), 8–11 (1986)
20. Osada, H., Propagation of chaos for the two dimensional Navier–Stokes equation. In: Probabilistic Methods in Mathematical Physics (Katata/Kyoto,1985), pp. 303–334. Academic Press, Boston (1987)
21. Perthame, B.: Transport Equations in Biology. Birkhauser Verlag, Basel (2007)
22. Royden, H.L., Fitzpatrick, P.M.: Real Analysis, 4th edn. Pearson, New York (2010)
23. Sznitman, A.-S.: Topics in propagation of chaos. In: Ecole d'Eté de Probabilités de Saint-Flour XIX-1989. Lecture Notes in Math., vol. 1464, pp. 165–251. Springer, Berlin (1991)
24. Takanobu, S.: On the existence and uniqueness of SDE describing an n-particle system interacting via a singular potential. Proc. Japan Acad. Ser. A Math. Sci. **6**, 287–290 (1985)
25. Villani, C.: Optimal Transport, Old and New. Grundlehren Math. Wiss. 338. Springer, Berlin (2009)