A random particle blob method for the Keller-Segel equation and convergence analysis

Jian-Guo Liu * Rong Yang †

October 29, 2015

Abstract

In this paper, we introduce a random particle blob method for the Keller-Segel equation (with dimension $d \geq 2$) and establish a rigorous convergence analysis.

Keywords: Interacting Brownian particle system, Newtonian aggregation, propagation of chaos, mean-field nonlinear stochastic differential equation, chemotaxis, Dobrushins type stability in Wasserstein distance.

AMS Subject Classifications: 60H10, 65M75, 35Q92, 35K55.

1 Introduction

The vortex method, pioneered by Chorin in 1973 [3], is one of the most successful computational methods for fluid dynamics and other related fields. When the effect of viscosity is important, the random vortex method is used to replace the vortex method. The success of the random vortex method was exemplified when it was shown to accurately compute flow past a cylinder at the Reynolds numbers up to 9500 in the 1990s [13]. The convergence analysis for the random vortex method for the Navier-Stokes equation was given by [8, 16, 18] in the 1980s. We refer to the book [4] for theoretical and practical use of the vortex methods, refer to [5] for recent progress on a blob method for the aggregation

*Departments of Physics and Mathematics, Duke University, Durham, NC 27708, USA. Email:jliu@phy.duke.edu

†College of Applied Sciences, Beijing University of Technology, Ping Le Yuan100, Chaoyang District, Beijing, 100124, People's Republic of China. Email: ysihan2010@163.com

equation, and also refer to [7, 9, 19] for many exciting recent developments on the theory of propagation of chaos.

In this paper, analog to the random vortex blob method, we introduce a random particle blob method (we will abbreviate it to a random blob method) for the Keller-Segel (KS) equation, which reads,

$$\begin{cases} \partial_t \rho = \nu \triangle \rho - \nabla \cdot (\rho \nabla c), \quad x \in \mathbb{R}^d, \quad t > 0, \\ -\triangle c = \rho(t,x), \\ \rho(0,x) = \rho_0(x), \end{cases} \tag{1.1}$$

where ν is a positive constant, and $\operatorname{supp} \rho_0(x) \subset D$, where D is a bounded domain. $\int_D \rho_0(x)dx = 1$ and $\rho_0 \in L^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d, (1+|x|)dx)$. In the context of biological aggregation, $\rho(t,x)$ represents the bacteria density and $c(t,x)$ represents the chemical substance concentration. In this paper, we also provide a rigorous convergence proof of this random blob method for the KS equation. We point out that the convergence rate given in this paper is far from sharp. Sharp convergence rates have been found for the random vortex method when applied to the Navier-Stokes equation by Goodman [8] and Long [16]. We will investigate the question on the sharp rate of the random blob method in the future.

In Section 2, we introduce a random blob method for the KS equation (1.1) and it is given by the following stochastic particle system of N particle pathes $\{(M_i, X_t^{i,\varepsilon})\}_{i=1}^N$

$$X_t^{i,\varepsilon} = X_0^i + \frac{1}{N-1} \sum_{j \neq i}^{N} M_j \int_0^t \nabla \big(J_\varepsilon * \Phi(X_s^{i,\varepsilon} - X_s^{j,\varepsilon}) \big) ds + \sqrt{2\nu} B_t^i, \quad i = 1, \cdots, N, \tag{1.2}$$

where $\{B_t^i\}_{i=1}^N$ are N independent Brownian motions and Φ is the Newtonian potential which we consider to be attractive. J_ε is a blob function with size ε (see Lemma 2.1). For the initial data ρ_0, we can chose a probability measure g_0 on $\mathbb{R}_+ \times \mathbb{R}^d$ such that $\rho_0(x)dx = \int_{\mathbb{R}_+} m g_0(dm, dx)$. Let $\{(M_i, X_0^i)\}_{i=1}^N$ be N independent and identically g_0-distributed random variables. M_i is the mass of the particle $X_t^{i,\varepsilon}$, which is sampled from the initial distribution and remains constant throughout the time evolution. The purpose of introducing this variable mass M_i is to coarse grain many small particles. In other words, a particle with large mass, M_i, can be viewed as a cluster of many particles with small mass, and thus we are effectively taking a coarse grain approximation of the small particles, which leads to a gain in the efficiency and stability of the computational method. Because the KS equation favors aggregation, this coarse-grained approximation does not lead to a significant loss in accuracy.

In Section 3, we prove the convergence of the random blob method. To do this, we show that there exists regularized parameters ε going to zero as N goes to infinity, and the stochastic particle system (1.2) converges to the following mean-field nonlinear stochastic differential equation

$$X_t = X_0 + \int_0^t \int_{\mathbb{R}^d} F(X_s - y)\rho(s,y)dyds + \sqrt{2\nu}B_t, \tag{1.3}$$

where $\rho(t,x)dx = \int_{\mathbb{R}_+} mg_t(dm,dx)$, $g_t(m,x) = \mathcal{L}(M,X_t)$, (M,X_0) is a g_0-distributed random variable that is independent of B_t, $\mathcal{L}(M,X_t)$ denotes the law of (M,X_t) and $F(x) = \nabla\Phi(x)$. The global existence and uniqueness of strong solution (see Definition 2) to (1.3) is almost same as a previous result of [15, Theorem 1.1] for the case of $\rho(t,x)dx = g_t(dx)$ and $M \equiv 1$. Thanks to the Itô formula [20, Theorem 4.1.2], we derive that ρ is a weak solution to (1.1) with the initial density ρ_0.

In Subsection 3.2, a coupling method is used to estimate the difference between the stochastic path $(M_i, X_t^{i,\varepsilon})$ for the particle method (1.2) and the stochastic path (M_i, X_t^i) of (1.3), where both paths begin with the same initial data $\{(M_i, X_0^i)\}_{i=1}^N$ and $\{B_t^i\}_{i=1}^N$ as (1.2). We show that there exists regularized parameters $\varepsilon(N) \sim (\ln N)^{-\frac{1}{d}} \to 0$ as $N \to \infty$ such that for any $1 \leq i \leq N$,

$$\lim_{N\to\infty} \mathbb{E}\big[\sup_{t\in[0,T]} M_i|X_t^{i,\varepsilon(N)} - X_t^i|\big] = 0. \tag{1.4}$$

In Subsection 3.3, we show that the empirical measures

$$\mu^N := \frac{1}{N}\sum_{i=1}^N M_i \delta_{X^{i,\varepsilon}} \to f \ \text{ in law}, \tag{1.5}$$

where $f_t(dx) = \rho_t(x)dx$ and ρ is the unique weak solution to (1.1) with the initial density ρ_0. μ^N are positive Radon measures with total variation $\frac{1}{N}\sum_{i=1}^N M_i$. By the strong law of large numbers [6, see pp.55 (7.1)], $\frac{1}{N}\sum_{i=1}^N M_i \to 1$ almost surely (a.s.) as $N \to \infty$, which implies f is a probability measure on $\mathbb{R}^d$. To obtain (1.5), we first prove in Proposition 3.3 that (1.5) is equivalent to showing that $\mathbb{E}[|\langle \mu^N - f, \varphi\rangle|] \to 0$ for any $\varphi \in C_b(\mathbb{R}^d)$. Then, following Sznitman [21] and by the exchangeability of $\{(M_i, X_t^{i,\varepsilon})\}_{i=1}^N$, we have

$$\mathbb{E}[\langle \mu^N - f, \varphi\rangle^2] = \frac{1}{N}\mathbb{E}\big[M_1^2\varphi(X_t^{1,\varepsilon})^2\big] + \frac{N-1}{N}\mathbb{E}\big[M_1\varphi(X_t^{1,\varepsilon})M_2\varphi(X_t^{2,\varepsilon})\big] - 2\langle f,\varphi\rangle\mathbb{E}\big[M_1\varphi(X_t^{1,\varepsilon})\big] + \langle f,\varphi\rangle^2.$$

Therefore, demonstrating the convergence of μ^N is reduced to show that (i) $\mathbb{E}\big[M_1\varphi(X_t^{1,\varepsilon})\big]$ converges to $\langle f,\varphi\rangle$, and (ii) two-particle's pairwise correlation converges to 0, i.e $\mathbb{E}\big[M_1\varphi(X_t^{1,\varepsilon})M_2\varphi(X_t^{2,\varepsilon})\big]$ converges to $\langle f,\varphi\rangle^2$. Both (i) and (ii) can be proved by (1.4).

Besides the above mentioned main techniques, we also mention that some standard techniques are also used to achieve our goal. By using the uniform estimates and the Lions-Aubin lemma [17, Lemma 10.4.], we give a sufficient condition (Assumption 1) of the initial data ρ_0 for the existence of the global weak solution to (1.1) in Section 2.

Finally, in Section 4, we provide some practical algorithms and their convergence results.

2 The random blob method and the main results

We begin by introducing the topology of the 1-Wasserstein space which will be used for the well-posedness of the KS equation. Let (E, d) be a Polish space. Consider the space of probability measures,

$$\mathcal{P}_1(E) = \big\{ f \,|\; f \text{ is a probability measure on } E \text{ and } \int_E d(0,x)df(x) < +\infty \big\}.$$

We define the Kantorovich-Rubinstein distance in $\mathcal{P}_1(E)$ as follows

$$\mathcal{W}_1(f,\ g) = \inf_{\pi \in \Lambda(f,\ g)} \Big\{ \int_{E\times E} d(x,y)d\pi(x,y) \Big\},$$

where $\Lambda(f,\ g)$ is the set of joint probability measures on $E \times E$ with marginals f and g. When f, g have densities ρ^1, ρ^2 respectively, we also denote the distance as $\mathcal{W}_1(\rho^1, \rho^2)$. In [22, Theorem 6.18], it has been proven that $\mathcal{P}_1(E)$ endowed with this distance is a complete metric space. And the following proposition holds by [22, Theorem 6.9].

Proposition 2.1. *(Wasserstein distances metrize weak convergence) If (E, d) is a Polish space, then for a given sequence $\{f_k\}_{k=1}^\infty$ and f in $\mathcal{P}_1(E)$, the convergence of $\{f_k\}_{k=1}^\infty$ to f in the 1-Wasserstein distance can deduce the narrow convergence of $\{f_k\}_{k=1}^\infty$, i.e.*

$$\mathcal{W}_1(f_k, f) \xrightarrow{k\to\infty} 0 \qquad \Rightarrow \qquad \int \varphi df_k(x) \xrightarrow{k\to\infty} \int \varphi df(x) \ \ \text{for any } \varphi \in C_b(E),$$

where $C_b(E)$ is the space of continuous and bounded functions.

Now recalling a recent result of the authors [15], we give some sufficient conditions of the initial data for the existence and uniqueness of the global weak solution in $L^\infty\big(0, T; L^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d, (1 + |x|)dx)\big)$ to (1.1).

Assumption 1. *1. $\rho_0(x) \in L^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d, (1 + |x|)dx)$, $\int_{\mathbb{R}^d} \rho_0(x)dx = 1$;*

2.

$$\|\rho_0\|_{L^{\frac{d}{2}}(\mathbb{R}^d)} < \begin{cases} 8\pi\nu & \text{if } d=2, \\ \dfrac{8\nu S_d}{d} & \text{if } d\geq 3, \end{cases} \tag{2.1}$$

where $S_d = \frac{d(d-2)}{4} 2^{2/d} \pi^{1+1/d} \Gamma\left(\frac{d+1}{2}\right)^{-2/d}$*, which is the best constant in the Sobolev inequality [14, pp.202].*

We use the following definition of the weak solution to (1.1).

Definition 1. (weak solution) Let the initial data $\rho_0(x) \in L^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d, (1+|x|)dx)$ and $T>0$, we shall say that $\rho(t,x)$,

$$\rho(t,x) \in L^\infty\big(0,T; L^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d, (1+|x|)dx)\big) \cap L^2(0,T; H^1(\mathbb{R}^d)),$$

$$\partial_t \rho \in L^2\big(0,T; H^{-1}(\mathbb{R}^d)\big),$$

is a weak solution to (1.1) with the initial data $\rho_0(x)$ if it satisfies:

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

$$\begin{aligned} \int_{\mathbb{R}^d} \rho(t,x)\varphi(x)dx - \int_{\mathbb{R}^d} \rho_0(x)\varphi(x)dx + \nu \int_0^t \int_{\mathbb{R}^d} \nabla\varphi(x)\cdot\nabla\rho(s,x)dxds \\ = \int_0^t \int_{\mathbb{R}^d} \rho(s,x)\big(\int_{\mathbb{R}^d} F(x-y)\rho(s,y)dy\big)\cdot\nabla\varphi(x)dxds, \end{aligned} \tag{2.2}$$

   where the interacting force $F(x) = \nabla\Phi(x) = -\frac{C^* x}{|x|^d}$, $\forall\ x \in \mathbb{R}^d\backslash\{0\},\ d\geq 2$, where $C^* = |\mathbb{S}^{d-1}|^{-1}$, $|\mathbb{S}^{d-1}| = \dfrac{2\pi^{d/2}}{\Gamma(d/2)}$ and $\Phi(x)$ is the Newtonian potential, which can be represented as

$$\Phi(x) = \begin{cases} \dfrac{C_d}{|x|^{d-2}} & \text{if } d\geq 3, \\ -\dfrac{1}{2\pi}\ln|x| & \text{if } d=2, \end{cases} \tag{2.3}$$

   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 the d-dimensional unit ball.

2. c is the chemical substance concentration associated with ρ and given by

$$c(t,x) = \Phi * \rho(t,x). \tag{2.4}$$

The Assumption 1 is sufficient for the existence of global weak solution to (1.1), see [1, 2]. Recently, the existence, uniqueness and Dobrushins type stability of the weak solution in the space $L^\infty\big(0,T;L^\infty(\mathbb{R}^d)\cap L^1(\mathbb{R}^d,(1+|x|)dx)\big)$ was established in [15, Theorem 1.1 and Theorem 1.2], which is summarized in the following theorem.

Theorem 2.1. *Consider the KS equation* (1.1)*, we have*

(i) Assume that $\rho_0(x)$ satisfies Assumption 1, then for any $T>0$, there exists a unique weak solution $\rho(t,x)$ to (1.1) *with the initial data ρ_0.*

(ii) Assume that $\rho_0^1(x)$, $\rho_0^2(x)$ both satisfy the Assumption 1. For any $T>0$, let ρ^1,ρ^2 be two weak solutions to (1.1) *with the initial conditions $\rho_0^1(x)$ and $\rho_0^2(x)$ respectively, then there exists two constants C (depending only on $\|\rho^1\|_{L^\infty(0,T;L^\infty\cap L^1(\mathbb{R}^d))}$ and $\|\rho^2\|_{L^\infty(0,T;L^\infty\cap L^1(\mathbb{R}^d))}$) and C_T (depending only on T) such that*

$$\sup_{t\in[0,T]}\mathcal{W}_1(\rho_t^1,\rho_t^2)\le C_T\max\Big\{\mathcal{W}_1(\rho_0^1,\ \rho_0^2),\{\mathcal{W}_1(\rho_0^1,\ \rho_0^2)\}^{\exp(-CT)}\Big\}.$$

Now we introduce the mean-field nonlinear stochastic process for the particle model (1.2) (see Proposition 3.2), and it's density satisfies the KS equation (see Proposition 3.1).

Definition 2. Let $\rho_0\in L^1\cap L^\infty(\mathbb{R}^d)$ and B_t be a Brownian motion. Chose a probability measure g_0 on $\mathbb{R}_+\times\mathbb{R}^d$ such that $\rho_0(x)dx=\int_{\mathbb{R}_+}mg_0(dm,dx)$. Let (M,X_0) be a g_0-distributed random variable that is independent of B_t. We say that a pair $(X_t,\rho(t,x))$, where X_t is a stochastic process and $\rho\in L^\infty\big(0,T;L^\infty\cap L^1(\mathbb{R}^d)\big)$, is a strong solution to the nonlinear stochastic differential equation (1.3), if for all $t\ge 0$,

$$X_t=X_0+\int_0^t\int_{\mathbb{R}^d}F(X_s-y)\rho(s,y)dyds+\sqrt{2\nu}B_t,$$

where $\rho(t,x)dx=\int_{\mathbb{R}_+}mg_t(dm,dx), g_t(m,x)=\mathcal{L}(M,X_t)$.

Theorem 2.2. *Suppose we are given a probability measure g_0 with support contained in $(0,\bar{M}]\times\mathbb{R}^d$, with g_0 satisfying the consistency condition $\rho_0(x)dx=\int_0^{\bar{M}}mg_0(dm,dx)$ and ρ_0 satisfying Assumption 1. Then there exists a unique strong solution $(X_t,\rho(t,x))$ to* (1.3) *associated to the g_0-distributed initial random variable (M,X_0).*

Below we give a example of constructing the g_0-distributed initial random variable (M,X_0).

Remark 2.1. Suppose $\rho_0(x) \in L^1 \cap L^\infty(D)$, where $D \subset \mathbb{R}^d$ is a bounded domain and it is the support of ρ_0. Suppose $|\partial D| = 0$. Let $X_0 : \Omega \to D$ be a random variable with density $\frac{\chi_D}{|D|}$, $M = |D|\rho_0(X_0)$ and $g_0(m, x) = \mathcal{L}(M, X_0)$. Then $\int_0^{\bar{M}} m g_0(dm, dx) = \rho_0(x)dx$ and $0 < M \le \bar{M}$ a.s., where $\bar{M} = |D|\|\rho_0\|_{L^\infty}$.

Proof. For any $\varphi \in C_b(\mathbb{R}^d)$, take $\psi(m, x) = m\varphi \in C_b([0, \bar{M}] \times \mathbb{R}^d)$, one has

$$\mathbb{E}[\psi(M, X_0)] = \int_{\mathbb{R}^d} |D|\rho_0(x)\varphi(x)\frac{\chi_D}{|D|}dx = \int_{\mathbb{R}^d} \rho_0(x)\varphi(x)dx. \tag{2.5}$$

On the other hand, by $g_0(m, x) = \mathcal{L}(M, X_0)$,

$$\mathbb{E}[\psi(M, X_0)] = \int_{\mathbb{R}^d} \int_0^{\bar{M}} m\varphi(x)g_0(dm, dx). \tag{2.6}$$

Combining (2.5) and (2.6) together and using the Fubini theorem, we get

$$\int_{\mathbb{R}^d} \rho_0(x)\varphi(x)dx = \int_{\mathbb{R}^d} \varphi(x) \int_0^{\bar{M}} m g_0(dm, dx). \tag{2.7}$$

Since φ is arbitrary, we obtain $\int_0^{\bar{M}} m g_0(dm, dx) = \rho_0(x)dx$.

Finally, from the fact that $|\partial D| = 0$, one has $0 < M \le \bar{M}$ a.s.. □

Before giving the main Theorem 2.3 of this paper, we describe how to regularize the Newtonian potential with a blob function.

Lemma 2.1. *[15, Lemma 2.1] Suppose $J(x) \in C^2(\mathbb{R}^d)$, supp $J(x) \in B(0, 1)$, $J(x) = J(|x|)$ and $J(x) \ge 0$. Let $J_\varepsilon(x) = \frac{1}{\varepsilon^d}J(\frac{x}{\varepsilon})$. Let $\Phi_\varepsilon(x) = J_\varepsilon * \Phi(x)$ 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 $F_\varepsilon(x) = F(x)g(\frac{|x|}{\varepsilon})$ for any $x \ne 0$, where $g(r) = \frac{1}{C^*}\int_0^r J(s)s^{d-1}ds$, $C^* = \dfrac{\Gamma(d/2)}{2\pi^{d/2}}$, $d \ge 2$;

(ii) $|F_\varepsilon(x)| \le \min\{\frac{C|x|}{\varepsilon^d}, |F(x)|\}$ and $|\nabla F_\varepsilon(x)| \le \frac{C}{\varepsilon^d}$.

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

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

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

Theorem 2.3. *Let $\rho(t,x)$ be the unique weak solution to* (1.1) *with the initial density ρ_0 satisfying Assumption 1 and $\{X_t^{1,\varepsilon},\cdots,X_t^{N,\varepsilon}\}$ be the unique strong solution to* (1.2) *associated to the independent and identically distributed (i.i.d.) initial random variables $\{(M_i,X_0^i)\}_{i=1}^N$. The initial data have common distribution $g_0(m,x)$ satisfying $\rho_0(x)dx = \int_0^{\bar{M}} mg_0(dm,dx)$ and $0<M_i\le \bar{M}$ a.s., where $\bar{M}$ is a constant. Denote $f_t(dx):=\rho(t,x)dx$, then there exists a subsequence of $\{X_t^{1,\varepsilon},\cdots,X_t^{N,\varepsilon}\}$ (which we have taken without relabeling N) and regularized parameters $\varepsilon(N)\sim(\ln N)^{-\frac{1}{d}}\to 0$ as $N\to\infty$ such that for any $t>0$,*

$$\frac{1}{N}\sum_{i=1}^N M_i\delta_{X_t^{i,\varepsilon(N)}} \to f_t(x) \qquad a.s.\ as\ N\to\infty. \tag{2.8}$$

Remark 2.2. $\frac{1}{N}\sum\limits_{i=1}^N M_i\delta_{X_t^{i,\varepsilon(N)}}$ are $\mathcal{M}_+(\mathbb{R}^d;\frac{1}{N}\sum\limits_{i=1}^N M_i)$-valued random variables, where $\mathcal{M}_+(\mathbb{R}^d;a)$ denotes the space of positive Radon measures on $\mathbb{R}^d$ where the total variation equals a. Since $\{M_i\}_{i=1}^N$ are i.i.d. and

$$\mathbb{E}[M_i] = \int_{(0,\bar{M}]\times\mathbb{R}^d} mg_0(dm,dx) = \int_{\mathbb{R}^d}\rho_0(x)dx = 1, \tag{2.9}$$

by the strong law of large numbers, we have

$$\frac{1}{N}\sum_{i=1}^N M_i \to 1, \qquad \text{a.s. as } N\to\infty. \tag{2.10}$$

3 Convergence proof

3.1 Preliminaries

Before giving the convergence proof of Theorem 2.3, we recall two lemmas stated in [15] and prove a result on the regularity of the regularized drift term $\int_{\mathbb{R}^d} F_\varepsilon(X_s-y)\rho(s,y)dy$ of (1.3).

Lemma 3.1. *[15, Lemma 2.2] For any function $\rho(x)\in L^\infty\cap L^1(\mathbb{R}^d)$, there exists a constant C such that for all $0\le\varepsilon'\le\varepsilon$,*

(i) $\int_{\mathbb{R}^d}|\rho(y)F_\varepsilon(x-y)|dy \le C\,\|\rho\|_{L^\infty\cap L^1}$.

(ii) $\int_{\mathbb{R}^d}|\rho(y)||F_\varepsilon(x-y)-F_\varepsilon(x'-y)|dy \le C\,\omega(|x-x'|)\,\|\rho\|_{L^\infty\cap L^1}$*, where*

$$\omega(r) = \begin{cases} 1 & \text{if } r\ge 1,\\ r(1-\ln r) & \text{if } 0<r<1. \end{cases} \tag{3.1}$$

(iii) $\int_{\mathbb{R}^d} |\rho(y)||F_\varepsilon(x-y) - F_{\varepsilon'}(x-y)|dy \le C\,\|\rho\|_{L^\infty}\,\varepsilon$.

The following lemma is a Gronwall type inequality with a logarithmic singularity.

Lemma 3.2. *[15, Lemma 2.4] Assume that there exists a family of nonnegative continuous functions $\{\alpha_\varepsilon(t)\}_{\varepsilon>0}$ satisfying*

$$\alpha_\varepsilon(t) \le C\int_0^t \alpha_\varepsilon(s)[1-\ln \alpha_\varepsilon(s)]ds + C\varepsilon T \quad \text{for } t\in[0,T],$$

where C is a constant. Then there exists two constants C_T and $\varepsilon_0(T) > 0$ such that if $\varepsilon < \varepsilon_0(T)$,

$$\sup_{t\in[0,T]} \alpha_\varepsilon(t) \le C_T \varepsilon^{\exp(-CT)} < 1. \tag{3.2}$$

Lemma 3.3. *Let (M, X^i) $(i = 1, 2)$ be two random variables with distributions $g^i(m, x)$, $0 < M \le \bar{M}$ a.s., where $\bar{M}$ is a constant. Suppose that there exists $\rho^i \in L^\infty \cap L^1(\mathbb{R}^d)$ such that $\rho^i(x)dx = \int_0^{\bar{M}} m g^i(dm, dx)$. For any $0 \le \varepsilon' \le \varepsilon$, define*

$$I := \int_{\mathbb{R}^d} MF_\varepsilon(X^1-y)\rho^1(y)dy - \int_{\mathbb{R}^d} MF_{\varepsilon'}(X^2-y)\rho^2(y)dy.$$

Then there exists a constant C depending on $\bar{M}$, $\|\rho^1\|_{L^\infty\cap L^1}$ and $\|\rho^2\|_{L^\infty\cap L^1}$ such that

$$\mathbb{E}\big[|I|\big] \le C\big(\varepsilon + \omega(\mathbb{E}[M|X^1-X^2|])\big). \tag{3.3}$$

Proof. A direct computation shows that

$$\begin{aligned} |I| \;\le\; & \int_{\mathbb{R}^d} M|F_\varepsilon(X^1-y) - F_\varepsilon(X^2-y)|\rho^1(y)dy + \int_{\mathbb{R}^d} M|F_\varepsilon(X^2-y) - F_{\varepsilon'}(X^2-y)|\rho^1(y)dy \\ & + \int_{\mathbb{R}^d} M|F_{\varepsilon'}(X^2-y)\rho^1(y) - F_{\varepsilon'}(X^2-y)\rho^2(y)|dy =: I_1 + I_2 + I_3. \end{aligned} \tag{3.4}$$

By (ii) and (iii) in Lemma 3.1, one has

$$I_1 \le C\,\|\rho^1(y)\|_{L^\infty\cap L^1(\mathbb{R}^d)} M\omega\big(|X^1-X^2|\big). \tag{3.5}$$

$$I_2 \le C\,\|\rho^1(y)\|_{L^\infty(\mathbb{R}^d)}\varepsilon. \tag{3.6}$$

Suppose (M, X^1) and (N, Y^1) are i.i.d., and (M, X^2) and (N, Y^2) are i.i.d., we have

$$\begin{aligned} \mathbb{E}[I_3] &= \mathbb{E}\big[\int_{\mathbb{R}^d} M|F_{\varepsilon'}(X^2-y)\rho^1(y) - F_{\varepsilon'}(X^2-y)\rho^2(y)|dy\big] \\ &= \mathbb{E}_x\mathbb{E}_y\big[MN|F_{\varepsilon'}(X^2-Y^1) - F_{\varepsilon'}(X^2-Y^2)|\big] \\ &= \mathbb{E}_y\big[\int_{\mathbb{R}^d} N|\big(F_{\varepsilon'}(x-Y^1) - F_{\varepsilon'}(x-Y^2)\big)\rho^2(x)|dx\big] \\ &\le C\mathbb{E}[N\omega\big(|Y^1-Y^2|\big)\big] = C\mathbb{E}[M\omega\big(|X^1-X^2|\big)]. \end{aligned} \tag{3.7}$$

By taking the expectation of (3.4) and combining (3.5), (3.6) and (3.7), we obtain that

$$\mathbb{E}[|I|] \quad \leq \quad C\varepsilon + C\mathbb{E}[M\omega\big(|X^1 - X^2|\big)]. \tag{3.8}$$

Using the facts that $x\omega(r) \leq \omega(xr)$ for any $0 < x \leq 1$, that $M \leq \bar{M}$, and that $\omega(r)$ is concave, we then obtain that

$$\begin{aligned}\mathbb{E}[|I|] &\leq C\varepsilon + C\mathbb{E}\big[\frac{M}{\bar{M}}\omega\big(|X^1 - X^2|\big)\big] \leq C\varepsilon + C\mathbb{E}\big[\omega\big(\frac{M}{\bar{M}}|X^1 - X^2|\big)\big] \\ &\leq C\varepsilon + C\omega\big(\frac{1}{\bar{M}}\mathbb{E}\big[M|X^1 - X^2|\big]\big).\end{aligned} \tag{3.9}$$

Notice that for any given constant C_1, there exists a constant C_2 depending on C_1 such that $\omega(C_1 x) \leq C_2\omega(x)$, which allows us to simplify (3.9) to read

$$\mathbb{E}[|I|] \leq C\varepsilon + C\omega\big(\mathbb{E}[M|X^1 - X^2|]\big), \tag{3.10}$$

which finishes the proof. □

3.2 Convergence of the paths

First, we prove the well-posedness of the nonlinear stochastic differential equation corresponding to the KS equation.

Proof of Theorem 2.2: Let (M, X_0) be g_0-distributed and independent of the Brownian motion B_t. For any fixed $\varepsilon > 0$ and initial distribution $g_0^\varepsilon(m, x) = \mathcal{L}(M, X_0)$, the regularized equation

$$X_t^\varepsilon = X_0 + \int_0^t \int_{\mathbb{R}^d} F_\varepsilon(X_s^\varepsilon - y)\rho_s^\varepsilon(y)dyds + \sqrt{2\nu}B_t, \tag{3.11}$$

is well-posed [15, see Theorem 2.1], where $g_t^\varepsilon(m, x) = \mathcal{L}(M, X_t^\varepsilon)$, $\rho^\varepsilon(t, x)dx = \int_0^{\bar{M}} m g_t^\varepsilon(dm, dx)$. We denote the unique solution as $(X_t^\varepsilon, \rho^\varepsilon(t, x))$. From the Itô formula, we know that $\rho^\varepsilon(t, x)$ is the unique classical solution to the following regularized KS equation,

$$\begin{cases} \partial_t\rho^\varepsilon = \nu\triangle\rho^\varepsilon - \nabla\cdot[\rho^\varepsilon\nabla c^\varepsilon], \quad x \in \mathbb{R}^d, \quad t > 0, \\ -\triangle c^\varepsilon = J_\varepsilon * \rho^\varepsilon(t, x), \\ \rho^\varepsilon(t, x)_{t=0} = \rho_0(x). \end{cases} \tag{3.12}$$

In [15, Theorem 2.2], we obtained that there exists a constant C depending on T, $\|\rho_0\|_{L^\infty(\mathbb{R}^d)\cap L^1(\mathbb{R}^d,(1+|x|)dx)}$ and data in (2.1) such that

$$\|\rho^\varepsilon\|_{L^\infty(0,T;L^1(\mathbb{R}^d))} = 1, \ \|\rho^\varepsilon\|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} \leq C, \quad \int_{\mathbb{R}^d} |x|\rho^\varepsilon(t, x)dx \leq C; \tag{3.13}$$

and

$$\int_0^T \|\nabla\rho^\varepsilon\|^2_{L^2(\mathbb{R}^d)}dt \le C, \quad \int_0^T \|\partial_t\rho^\varepsilon\|^2_{H^{-1}(\mathbb{R}^d)}dt \le C. \tag{3.14}$$

We divide the rest proof into four steps to show that (i) $\{MX_t^\varepsilon\}_{\varepsilon>0}$ is a Cauchy sequence, (ii) there exists a subsequence of ρ^ε such that the limit of this subsequence is the unique weak solution to the KS equation (1.1) and satisfies the consistent quality (3.18), (iii) there exists subsequence of $(X_t^\varepsilon, \rho^\varepsilon)$ such that the limit of this subsequence is a strong solution to the corresponding nonlinear stochastic equation (1.3), and (iv) the strong solution to (1.3) is unique.

Step 1 For any $0<\varepsilon'<\varepsilon$, let $(X_t^\varepsilon, \rho^\varepsilon)$ and $(X_t^{\varepsilon'}, \rho^{\varepsilon'})$ be the unique strong solutions to (3.11) with the same initial data (M, X_0). We first prove that

$$\lim_{\varepsilon\to0}\mathbb{E}\big[\sup_{t\in[0,T]} M|X_t^\varepsilon - X_t^{\varepsilon'}|\big] = 0. \tag{3.15}$$

By the fact that $M(X_t^\varepsilon - X_t^{\varepsilon'}) = \int_0^t\int_{\mathbb{R}^d} MF_\varepsilon(X_s^\varepsilon - y)\rho_s^\varepsilon(y)dyds - \int_0^t\int_{\mathbb{R}^d} MF_{\varepsilon'}(X_s^{\varepsilon'} - y)\rho_s^{\varepsilon'}(y)dyds$, Lemma 3.3, the fact that ω is nondecreasing, and the uniform estimates of $\rho_s^\varepsilon(y)$, we assert that there exists a constant C independent of ε and ε' such that

$$\begin{aligned}
&\mathbb{E}\big[\sup_{t\in[0,T]} M|X_t^\varepsilon - X_t^{\varepsilon'}|\big]\\
&\quad\le \int_0^T \mathbb{E}\big[M|\int_{\mathbb{R}^d} F_\varepsilon(X_s^\varepsilon - y)\rho_s^\varepsilon(y)dy - \int_{\mathbb{R}^d} F_{\varepsilon'}(X_s^{\varepsilon'} - y)\rho_s^{\varepsilon'}(y)dy|\big]ds\\
&\quad\le C\varepsilon T + C\int_0^T \omega\big(\mathbb{E}\big[M|X_s^\varepsilon - X_s^{\varepsilon'}|\big]\big)ds \le C\varepsilon T + C\int_0^T \omega\big(\mathbb{E}\big[\sup_{\tau\in[0,s]} M|X_\tau^\varepsilon - X_\tau^{\varepsilon'}|\big]\big)ds.
\end{aligned}$$

By Lemma 3.2, we achieve (3.15) immediately.

Step 2 From (3.15), there exists a subsequence of MX_t^ε (without relabeling) and a limiting point MX_t such that

$$\mathbb{E}\big[\sup_{t\in[0,T]} M|X_t^\varepsilon - X_t|\big] \to 0 \ \text{ as } \ \varepsilon\to0. \tag{3.16}$$

On the other hand, based on the uniform estimates (3.13) and (3.14), and using the Lions-Aubin lemma [17, Lemma 10.4.], there exists a subsequence of ρ^ε (without relabeling) such that for any ball B_R,

$$\rho^\varepsilon \to \rho \text{ in } L^2\big(0,T;L^2(B_R)\big) \text{ as } \varepsilon\to0. \tag{3.17}$$

Since ρ^ε is a weak solution to (3.12), taking the limit $\varepsilon\to0$ concludes that $\rho(t,x)$ is a weak solution to (1.1) with the following regularities:

i) $\rho \in L^\infty\big(0,T; L^\infty(\mathbb{R}^d)\cap L^1(\mathbb{R}^d,(1+|x|)dx)\big)$;

ii) $\rho \in L^2\big(0,T;H^1(\mathbb{R}^d)\big)$, $\partial_t\rho \in L^2\big(0,T;H^{-1}(\mathbb{R}^d)\big)$.

By Theorem 2.1, $\rho(t,x)$ is the unique weak solution to (1.1). We claim that

$$\rho(t,x)dx = \int_0^{\bar M} m g_t(dm,dx), \quad \text{where} \;\; g_t(m,x) = \mathcal{L}(M,X_t). \tag{3.18}$$

Indeed for any $t>0$ and $\varphi \in BL(\mathbb{R}^d)$ (here we denote $BL(\mathbb{R}^d)$ as the space of bounded Lipschitz continuous functions), one has

$$M|\varphi(X_t^\varepsilon) - \varphi(X_t)| \le CM|X_t^\varepsilon - X_t|. \tag{3.19}$$

Thus $\mathbb{E}[M|\varphi(X_t^\varepsilon)-\varphi(X_t)|] \le C\mathbb{E}\big[\sup_{t\in[0,T]} M|X_t^\varepsilon - X_t|\big]$, and then when $\varepsilon\to 0$,

$$\begin{aligned}\mathbb{E}[M\varphi(X_t^\varepsilon)] \quad &= \quad \int_{(0,\bar M]\times\mathbb{R}^d} m\varphi(x) g_t^\varepsilon(dm,dx)\\ &\longrightarrow \;\; \mathbb{E}[M\varphi(X_t)] = \int_{(0,\bar M]\times\mathbb{R}^d} m\varphi(x) g_t(dm,dx).\end{aligned} \tag{3.20}$$

By the portemanteau theorem [12, pp.254, Theorem 13.16], $\int_0^{\bar M} m g_t^\varepsilon(dm,\cdot)$ narrowly converges to $\int_0^{\bar M} m g_t(dm,\cdot)$.

On the other hand, for any $t>0$, by (3.17) we have

$$\int_{\mathbb{R}^d}\int_0^{\bar M} m\varphi(x) g_t^\varepsilon(dm,dx) = \int_{\mathbb{R}^d}\varphi(x)\rho_t^\varepsilon(x)dx \longrightarrow \int_{\mathbb{R}^d}\varphi(x)\rho_t(x)dx. \tag{3.21}$$

Combining (3.20) and (3.21) together, we obtain (3.18).

Step 3 Now we show that

$$X_t = X_0 + \int_0^t\int_{\mathbb{R}^d} F(X_s-y)\rho_s(y)dyds + \sqrt{2\nu}B_t \quad \text{a.s..} \tag{3.22}$$

Using Lemma 3.3 and taking $\varepsilon' = 0$, we obtain that for any $s\in[0,T]$,

$$\mathbb{E}\big[\int_{\mathbb{R}^d} M|F_\varepsilon(X_s^\varepsilon - y)\rho_s^\varepsilon(y) - F(X_s-y)\rho_s(y)|dy\big] \le C\varepsilon + C\omega\big(\mathbb{E}\big[M|X_s^\varepsilon - X_s|\big]\big). \tag{3.23}$$

Combine (3.23) and (3.16), one has

$$\begin{aligned}&\mathbb{E}\big[\big|\int_0^t\int_{\mathbb{R}^d} MF_\varepsilon(X_s^\varepsilon-y)\rho_s^\varepsilon(y)dyds - \int_0^t\int_{\mathbb{R}^d} MF(X_s-y)\rho_s(y)dyds\big|\big]\\ &\le CT\varepsilon + CT\omega\big(\mathbb{E}\big[\sup_{s\in[0,T]} M|X_s^\varepsilon - X_s|\big]\big) \to 0 \qquad \text{as } \varepsilon\to 0.\end{aligned}$$

Thus there exists a subsequence (without relabeling) such that

$$\int_0^t \int_{\mathbb{R}^d} MF_\varepsilon(X_s^\varepsilon - y)\rho_s^\varepsilon(y)dyds \to \int_0^t \int_{\mathbb{R}^d} MF(X_s - y)\rho_s(y)dyds \quad \text{a.s. as } \varepsilon \to 0. \tag{3.24}$$

Taking the limit $\varepsilon \to 0$ in (3.11), one has

$$MX_t = MX_0 + M\int_0^t \int_{\mathbb{R}^d} F(X_s - y)\rho_s(y)dyds + M\sqrt{2\nu}B_t \quad \text{a.s.}. \tag{3.25}$$

Since (M, X_0) is g_0-distributed and g_0 has support $(0, \bar{M}] \times \mathbb{R}^d$, then $0 < M \le \bar{M}$ a.s.. Hence the above equation implies that (X_t, ρ) is a strong solution to (1.3).

Step 4 Suppose (X_t^1, ρ^1) and (X_t^2, ρ^2) are two strong solutions to (1.3) with the same initial random variable (M, X_0). Then ρ^1 and ρ^2 both are weak solutions to the KS equation with the same initial data ρ_0 and this will be proved by Proposition 3.1 (i). Since the weak solution to the KS equation is unique, one has $\rho^1 = \rho^2$. Using Lemma 3.3 and taking $\varepsilon' = \varepsilon = 0$, one has

$$\mathbb{E}[M|X_t^1 - X_t^2|] \le \omega\big(\mathbb{E}[M|X_t^1 - X_t^2|]\big).$$

Combining the initial date $X_0^1 = X_0^2 = X_0$ and the fact $0 < M \le \bar{M}$ a.s., we obtain $X_t^1 = X_t^2$ a.s., i.e. the strong solution to (1.3) is unique. □

Proposition 3.1. *Let ρ_0 satisfy Assumption 1 and B_t be a Brownian motion. Chose a probability measure g_0 on $(0, \bar{M}] \times \mathbb{R}^d$ such that $\rho_0(x)dx = \int_{(0,\bar{M}]} mg_0(dm, dx)$. Let (M, X_0) be a g_0-distributed random variable and it is independent of B_t. The relationship between the strong solution to* (1.3) *and the weak solution to* (1.1) *holds as follows:*

(i) If $(X_t, \rho(t, x))$ is a strong solution to (1.3) *associated to the initial random variable (M, X_0), then $\rho(t, x)$ is a weak solution to* (1.1) *with the initial data $\rho_0(x)$.*

(ii) If ρ is a weak solution to (1.1) *with the initial data $\rho_0(x)$, then there exists a unique stochastic process X_t such that $(X_t, \rho(t, x))$ is the unique strong solution to* (1.3) *associated to the initial random variable (M, X_0).*

Proof. Let $(X_t, \rho(t, x))$ be a strong solution to (1.3). Then X_t is an Itô process [20, Definition 4.1.1], for any $\varphi(x) \in C_b^2(\mathbb{R}^d)$, Itô formula states that

$$\begin{aligned}\varphi(X_t) &= \varphi(X_0) + \int_0^t \int_{\mathbb{R}^d} \nabla\varphi(X_s)F(X_s - y)\rho_s(y)dyds \\ &\quad + \sqrt{2\nu}\int_0^t \nabla\varphi(X_s)dB_s + \nu\int_0^t \triangle\varphi(X_s)ds. \end{aligned}\tag{3.26}$$

Multiplying (3.26) by M and taking the expectation, one has

$$\begin{aligned}\int_{\mathbb{R}_+\times\mathbb{R}^d} m\varphi(x)g_t(dm,dx) &= \int_{\mathbb{R}_+\times\mathbb{R}^d} m\varphi(x)g_0(dm,dx) + \nu\int_0^t\int_{\mathbb{R}_+\times\mathbb{R}^d}\triangle\varphi(x)mg_s(dm,dx)ds \\ &\quad + \int_0^t\int_{\mathbb{R}_+\times\mathbb{R}^{2d}} m\nabla\varphi(x)F(x-y)\rho_s(y)dyg_s(dm,dx)ds.\end{aligned}$$

Since $\rho(t,x)dx = \int_{\mathbb{R}_+} mg_t(dm,dx)$, one knows that ρ is a weak solution to (1.1) with the initial data $\rho_0(x)$.

To prove (ii), we first consider the following linear Fokker-Planck equation:

$$\begin{cases}\partial_t\rho = \nu\triangle\rho - \nabla\cdot[V_g(t,x)\rho],\ x\in\mathbb{R}^d,\ t>0,\\ \rho(t,x)_{t=0} = \rho_0(x),\end{cases} \tag{3.27}$$

where $V_g(t,x) = \int_{\mathbb{R}^d} F(x-y)g(t,y)dy$ and $g(t,x)\in L^\infty\big(0,T;L^\infty\cap L^1(\mathbb{R}^d)\big)$ is a given function. Recall the proof of [15, Proposition 2.3], (3.27) has a unique weak solution.

Now let $V_\rho(t,x) = \int_{\mathbb{R}^d} F(x-y)\rho(t,y)dy$, where ρ is a weak solution to (1.1). Hence ρ is also a weak solution to (3.27) associated with V_ρ and the initial data ρ_0. Since V_ρ is log-Lipschitz continuous, repeating the proof of the well-posedness for the nonlinear stochastic equation (1.3), one can show that the stochastic equation

$$X_t = X_0 + \int_0^t V_\rho\big(s,X_s\big)ds + \sqrt{2\nu}B_t, \tag{3.28}$$

has a unique strong solution X_t. Denote $\tilde{g}_t(m,x) = \mathcal{L}(M,X_t)$, one also can obtain that there exists a density $\tilde{\rho}$ such that $\int_{\mathbb{R}_+} m\tilde{g}_t(dm,dx) = \tilde{\rho}(x)dx$. By the Itô formula, $\tilde{\rho}$ is a weak solution to (3.27) associated with V_ρ and the initial data ρ_0. By the uniqueness of (3.27) we obtain $\tilde{\rho}=\rho$, i.e. (X_t,ρ) is a strong solution to (1.3) associated to (M,X_0).

Hence combining the uniqueness of (1.3), we have proven that if $\rho(t,x)$ is a weak solution to (1.1) with the initial data ρ_0, then there exists a unique stochastic process X_t such that $(X_t,\rho(t,x))$ is the unique strong solution to (1.3) associated to the initial random variable (M,X_0). □

Now let us write down the main estimate on the difference between the stochastic process $(M_i,X_t^{i,\varepsilon})$ for the particle method (1.2) and (M_i,X_t^i) for the mean-field nonlinear stochastic equation (1.3).

Proposition 3.2. *Let* $\{X_t^{1,\varepsilon},\cdots,X_t^{N,\varepsilon}\}$ *be the unique strong solution to* (1.2) *associated to the i.i.d. initial random variables* $\{(M_i,X_0^i)\}_{i=1}^N$, $0<M_i\le\bar{M}$ *a.s. and independent Brownian motions* $\{B_t^i\}_{i=1}^N$. *Let* $\{(M_i,X_t^i)\}_{i=1}^N$ *be the unique strong solutions to* (1.3) *with the*

same initial data $\{(M_i, X_0^i)\}_{i=1}^N$ and $\{B_t^i\}_{i=1}^N$ as (1.2). Then $\{(M_i, X_t^{i,\varepsilon})\}_{i=1}^N$ are exchangeable, $\{(M_i, X_t^i)\}_{i=1}^N$ are i.i.d. and there exists regularized parameters $\varepsilon(N) \sim (\ln N)^{-\frac{1}{d}} \to 0$ as $N \to \infty$ such that

$$\lim_{N\to\infty} \mathbb{E}\big[\sup_{t\in[0,T]} M_i|X_t^{i,\varepsilon(N)} - X_t^i|\big] = 0 \text{ for any } 1 \le i \le N. \tag{3.29}$$

Proof. For the i.i.d initial data $\{(M_i, X_0^i)\}_{i=1}^N$ and Brownian motions $\{B_t^i\}_{i=1}^N$, consider the following equation

$$\bar{X}_t^{i,\varepsilon} = X_0^i + \int_0^t \int_{(0,\bar{M}]\times\mathbb{R}^d} mF_\varepsilon(\bar{X}_s^{i,\varepsilon} - y)\bar{g}_s(dm, dy)ds + \sqrt{2\nu}B_t^i, \quad i = 1, \cdots, N \tag{3.30}$$

where $\bar{g}_t(m, x) = \mathcal{L}(M_i, \bar{X}_t^{i,\varepsilon})$, $\int_{\mathbb{R}_+} m\bar{g}_t(dm, dx) = \bar{\rho}(x)dx$. In the proof of Theorem 2.2, we know that (3.30) has a unique solution $\{(M_i, \bar{X}_t^{i,\varepsilon})\}_{i=1}^N$, and they are i.i.d.. This equation serves as a link between (1.2) and (1.3), and it will be verified by the following three steps.

Step 1 We prove that

$$\mathbb{E}\big[\sup_{t\in[0,T]} M_i|X_t^{i,\varepsilon} - \bar{X}_t^{i,\varepsilon}|\big] \le \frac{C_T}{\sqrt{N-1}\varepsilon^{d-1}} \exp(\frac{C_T}{\varepsilon^d}), \tag{3.31}$$

which gives a relationship between (3.30) and (1.2).

Since $\big|\nabla F_\varepsilon(x)\big| \le \frac{C}{\varepsilon^d}$ by Lemma 2.1, one has

$$\begin{aligned} M_i|X_t^{i,\varepsilon} - \bar{X}_t^{i,\varepsilon}| &\le \int_0^t \frac{1}{N-1}\sum_{j\neq i}^N \big|M_iM_jF_\varepsilon(X_s^{i,\varepsilon} - X_s^{j,\varepsilon}) - M_iM_jF_\varepsilon(\bar{X}_s^{i,\varepsilon} - \bar{X}_s^{j,\varepsilon}) \\ &\quad + M_iM_jF_\varepsilon(\bar{X}_s^{i,\varepsilon} - \bar{X}_s^{j,\varepsilon}) - M_i\int_{(0,\bar{M}]\times\mathbb{R}^d} mF_\varepsilon(\bar{X}_s^{i,\varepsilon} - y)\bar{g}_s(dm, dy)\big|ds \\ &\le \int_0^t \frac{1}{N-1}\sum_{j\neq i}^N \Big[\frac{CM_i}{\varepsilon^d}|X_s^{i,\varepsilon} - \bar{X}_s^{i,\varepsilon}| + \frac{CM_j}{\varepsilon^d}|X_s^{j,\varepsilon} - \bar{X}_s^{j,\varepsilon}| + |A_j^i|\Big]ds, \end{aligned} \tag{3.32}$$

where

$$A_j^i := M_iM_jF_\varepsilon(\bar{X}_s^{i,\varepsilon} - \bar{X}_s^{j,\varepsilon}) - M_i\int_{(0,\bar{M}]\times\mathbb{R}^d} mF_\varepsilon(\bar{X}_s^{i,\varepsilon} - y)\bar{g}_s(dm, dy). \tag{3.33}$$

Since $\{(M_i, X_t^{i,\varepsilon})\}_{i=1}^N$ are exchangeable random variables and $\{(M_i, \bar{X}_t^{i,\varepsilon})\}_{i=1}^N$ are i.i.d. random variables, we have the following exchangeability property:

$$E\big[M_i|X_s^{i,\varepsilon} - \bar{X}_s^{i,\varepsilon}|\big] = E\big[M_j|X_s^{j,\varepsilon} - \bar{X}_s^{j,\varepsilon}|\big] \text{ for any } 1 \le i, j \le N.$$

After taking the expectation of (3.32), by the exchangeability property, one has

$$\mathbb{E}\big[\sup_{t\in[0,T]} M_i|X_t^{i,\varepsilon}-\bar{X}_t^{i,\varepsilon}|\big] \leq \int_0^T \mathbb{E}\big[\big|\frac{1}{N-1}\sum_{j\neq i}^N A_j^i\big|\big]ds + \frac{C}{\varepsilon^d}\int_0^T \mathbb{E}\big[M_i|X_s^{i,\varepsilon}-\bar{X}_s^{i,\varepsilon}|\big]ds.$$

Applying the Gronwall's lemma, one deduces that

$$\begin{aligned}\mathbb{E}\big[\sup_{t\in[0,T]} M_i|X_t^{i,\varepsilon}-\bar{X}_t^{i,\varepsilon}|\big] &\leq \exp(\frac{CT}{\varepsilon^d})\int_0^T \mathbb{E}\big[\big|\frac{1}{N-1}\sum_{j\neq i}^N A_j^i\big|\big]ds \\ &\leq \exp(\frac{CT}{\varepsilon^d})\int_0^T \big\{\mathbb{E}\big[\big|\frac{1}{N-1}\sum_{j\neq i}^N A_j^i|^2\big]\big\}^{\frac{1}{2}}ds. \end{aligned} \tag{3.34}$$

Because $\{(M_i,\bar{X}_t^{i,\varepsilon})\}_{i=1}^N$ are i.i.d. random variables, when $j\neq k$,

$$\mathbb{E}\big[A_j^iA_k^i\big]=0.$$

Hence

$$\mathbb{E}\big[\big|\frac{1}{N-1}\sum_{j\neq i}^N A_j^i\big|^2\big]=\frac{1}{(N-1)^2}\mathbb{E}\big[\sum_{j,k\neq i}^N A_j^iA_k^i\big]\leq\frac{\mathbb{E}\big[(A_2^1)^2\big]}{N-1}. \tag{3.35}$$

Since $\{(M_i,\bar{X}_t^{i,\varepsilon})\}_{i=1}^N$ are i.i.d., we have

$$\begin{aligned}\mathbb{E}\big[(A_2^1)^2\big] &= \mathbb{E}\big[|M_1M_2F_\varepsilon(\bar{X}_s^{1,\varepsilon}-\bar{X}_s^{2,\varepsilon})-M_1\int_{(0,\bar{M}]\times\mathbb{R}^d} mF_\varepsilon(\bar{X}_s^{1,\varepsilon}-y)d\bar{g}_s(dm,dy)|^2\big] \\ &\leq 2\mathbb{E}\big[M_1^2M_2^2F_\varepsilon^2(\bar{X}_s^{1,\varepsilon}-\bar{X}_s^{2,\varepsilon})+\big(M_1\int_{(0,\bar{M}]\times\mathbb{R}^d} mF_\varepsilon(\bar{X}_s^{1,\varepsilon}-y)d\bar{g}_s(dm,dy)\big)^2\big] \\ &\leq 2\mathbb{E}\big[M_1^2M_2^2F_\varepsilon^2(\bar{X}_s^{1,\varepsilon}-\bar{X}_s^{2,\varepsilon})+M_1^2\int_{(0,\bar{M}]\times\mathbb{R}^d} m^2F_\varepsilon^2(\bar{X}_s^{1,\varepsilon}-y)d\bar{g}_s(dm,dy)\big] \\ &= 4\mathbb{E}\big[M_1^2M_2^2F_\varepsilon^2(\bar{X}_s^{1,\varepsilon}-\bar{X}_s^{2,\varepsilon})\big]. \end{aligned} \tag{3.36}$$

For all $\varepsilon>0$, using $\big|F_\varepsilon(x)\big|\leq\min\{\frac{C|x|}{\varepsilon^d},\frac{C}{|x|^{d-1}}\}$ from Lemma 2.1, we obtain that

$$\begin{aligned}&\mathbb{E}\big[M_1^2M_2^2F_\varepsilon^2(\bar{X}_s^{1,\varepsilon}-\bar{X}_s^{2,\varepsilon})\big] \\ &\quad\leq C\int_{|x-y|\leq\varepsilon}\frac{|x-y|^2}{\varepsilon^{2d}}df_s^{1,\varepsilon}(x)df_s^{2,\varepsilon}(y)+C\int_{|x-y|>\varepsilon}\frac{1}{|x-y|^{2(d-1)}}df_s^{1,\varepsilon}(x)df_s^{2,\varepsilon}(y) \\ &\quad\leq\frac{C}{\varepsilon^{2(d-1)}}, \end{aligned} \tag{3.37}$$

where $f_s^{i,\varepsilon}(\cdot)=\int_0^{\bar{M}} m\bar{g}_s(dm,\cdot)$, $i=1,2$. Plugging (3.37) into (3.36), we have

$$\mathbb{E}\big[(A_2^1)^2\big]\leq\frac{C}{\varepsilon^{2(d-1)}}. \tag{3.38}$$

Combining (3.34), (3.35) and (3.38) together yields (3.31).

Step 2 We deduce the relationship between (3.30) and (1.3), i. e. there exists C_T and $\varepsilon_0(T) > 0$ such that if $\varepsilon < \varepsilon_0(T)$,

$$\mathbb{E}\big[\sup_{t\in[0,T]} M_i|\bar{X}_t^{i,\varepsilon} - X_t^i|\big] \le C_T\varepsilon^{\exp(-CT)} \text{ for any } 1 \le i \le N. \tag{3.39}$$

From (3.30) and (1.3) directly, one has

$$M_i|\bar{X}_t^{i,\varepsilon} - X_t^i| \le \int_0^t \int_{\mathbb{R}^d} \big|M_i F_\varepsilon(\bar{X}_s^{i,\varepsilon} - y)\bar{\rho}_s(y) - M_i F(X_s^i - y)\rho_s(y)\big| dyds.$$

Taking the expectation and using Lemma 3.3, we have

$$\begin{aligned}\mathbb{E}\big[\sup_{t\in[0,T]} M_i|\bar{X}_t^{i,\varepsilon} - X_t^i|\big] &\le \int_0^T \mathbb{E}\big[\int_{\mathbb{R}^d} |M_i F_\varepsilon(\bar{X}_s^{i,\varepsilon} - y)\bar{\rho}_s(y) - M_i F(X_s^i - y)\rho_s(y)|dy\big] ds \\ &\le C\varepsilon T + C\int_0^T \mathbb{E}\big[\sup_{\tau\in[0,s]} M_i\omega(|\bar{X}_\tau^{i,\varepsilon} - X_\tau^i|)\big] ds. \end{aligned} \tag{3.40}$$

By Lemma 3.2, we achieve (3.39).

Step 3 Combining (3.31) and (3.39), one has

$$\begin{aligned}\mathbb{E}\big[\sup_{t\in[0,T]} M_i|X_t^{i,\varepsilon} - X_t^i|\big] &\le \mathbb{E}\big[\sup_{t\in[0,T]} M_i|X_t^{i,\varepsilon} - \bar{X}_t^{i,\varepsilon}|\big] + \mathbb{E}\big[\sup_{t\in[0,T]} M_i|\bar{X}_t^{i,\varepsilon} - X_t^i|\big] \\ &\le \frac{C_T}{\sqrt{N-1}\varepsilon^{d-1}} \exp\big(\frac{C_T}{\varepsilon^d}\big) + C_T\varepsilon^{\exp(-CT)}. \end{aligned} \tag{3.41}$$

We choose $\varepsilon = \varepsilon(N) = \lambda(\ln N)^{-\frac{1}{d}} \to 0$ as $N \to \infty$ in (3.41), where λ is a large enough positive constant, and conclude

$$\mathbb{E}\big[\sup_{t\in[0,T]} M_i|X_t^{i,\varepsilon(N)} - X_t^i|\big] \le \frac{C_T N^{\frac{C_T}{\lambda^d}} (\ln N)^{\frac{d-1}{d}}}{\lambda^{d-1}\sqrt{N-1}} + C_T\varepsilon^{\exp(-CT)} \to 0 \text{ as } N \to \infty,$$

which completes the proof of Proposition 3.2. □

Denote $G_t^\varepsilon(m_1, x_1, \cdots, m_N, x_N)$ as the joint distribution of $\{(M_i, X_t^{i,\varepsilon})\}_{i=1}^N$. Below we will use the result of Proposition 3.2 to show the convergence of the j-th order marginal distribution of G_t^ε in $\mathcal{W}_1$ distance.

Corollary 3.1. *(Propagation of chaos) Let* $\{(M_i, X_t^{i,\varepsilon})\}_{i=1}^N$ *and* $\{(M_i, X_t^i)\}_{i=1}^N$ *be defined as in Proposition 3.2. Define* $f_t^{(j),\varepsilon} = \int_{(0,\bar{M}]^N\times\mathbb{R}^{(N-j)d}} m_1, \cdots, m_j G_t^\varepsilon(dm_1, \cdots, dm_N, \cdot, dx_{j+1}, \cdots, dx_N)$, $j \ge 1$. *Let* g_t *be the common time marginal of* $\big\{(M_i, X_t^i)\big\}_{i=1}^N$ *and* $f_t(dx) := \int_0^{\bar{M}} mg_t(dm, dx)$. *Then there exists regularized parameters* $\varepsilon(N) \sim (\ln N)^{-\frac{1}{d}} \to 0$ *as* $N \to \infty$ *such that*

$$\sup_{t\in[0,T]} \mathcal{W}_1(f_t^{(j),\varepsilon(N)}, f_t^{\otimes j}) \to 0 \quad \text{as } N \to \infty. \tag{3.42}$$

Proof. Denote $\tilde{F}_t(m_1,\cdots,m_N,x_1,\cdots,x_N,\hat{x}_1,\cdots,\hat{x}_j)$ as the joint distribution of $\big(M_1,\cdots,M_N,X_t^{1,\varepsilon},\cdots,X_t^{N,\varepsilon},X_t^1,\cdots,X_t^j\big)$, then one has the following facts:

$$f_t^{\otimes j} = \int_{(0,\bar{M}]^N\times\mathbb{R}^{Nd}} m_1,\cdots,m_j\tilde{F}_t(dm_1,\cdots,dm_N,dx_1,\cdots,dx_N,\cdot);$$
$$f_t^{(j),\varepsilon} = \int_{(0,\bar{M}]^N\times\mathbb{R}^{Nd}} m_1,\cdots,m_j\tilde{F}_t(dm_1,\cdots,dm_N,\cdot,dx_{j+1},\cdots,dx_N,d\hat{x}_1,\cdots,d\hat{x}_j);$$
$$\int_{(0,\bar{M}]^N\times\mathbb{R}^{(N-j)d}} m_1,\cdots,m_j\tilde{F}_t(dm_1,\cdots,dm_N,\cdot,dx_{j+1},\cdots,dx_N,\cdot)\in\Lambda(f_t^{(j),\varepsilon},f_t^{\otimes j}).$$

By (2.9), we also have

$$\int_{\mathbb{R}^{jd}} df_t^{(j),\varepsilon} = \int_{\mathbb{R}^{jd}} df_t^{\otimes j} = \mathbb{E}[M_1\cdots M_j] = \big(\mathbb{E}[M_1]\big)^j = 1.$$

Applying (3.29), we obtain

$$\begin{aligned}
&\sup_{t\in[0,T]}\mathcal{W}_1(f_t^{(j),\varepsilon(N)},f_t^{\otimes j}) \qquad (3.43)\\
&\le \sup_{t\in[0,T]}\int_{\mathbb{R}^{2jd}}\big(|x_1-\hat{x}_1|+\cdots+|x_j-\hat{x}_j|\big)\int_{(0,\bar{M}]^N\times\mathbb{R}^{(N-j)d}} m_1,\cdots,m_j\tilde{F}_t(dm_1,\cdots,d\hat{x}_j)\\
&\le j\bar{M}^{j-1}\sup_{t\in[0,T]}\int_{(0,\bar{M}]^N\times\mathbb{R}^{(N+j)d}} m_1|x_1-\hat{x}_1|\tilde{F}_t(dm_1,\cdots,dm_N,dx_1,\cdots,dx_N,d\hat{x}_1,\cdots,d\hat{x}_j)\\
&\le j\bar{M}^{j-1}\mathbb{E}_{m_1,\cdots,m_N,x_1,\cdots,x_N,\hat{x}_1,\cdots,\hat{x}_N}\big[\sup_{t\in[0,T]} M_1|X_t^{1,\varepsilon(N)}-X_t^1|\big]\to 0 \quad \text{as } N\to\infty.
\end{aligned}$$

□

3.3 Convergence of the random blob method

Lemma 3.4. *Let (E,d) be a Polish space and $\mathcal{W}_d$ be the Wasserstein distance on $\mathcal{P}(E)$ associated with the metric d. If $x\in E$, y is a random variable on E, and we denote $f=\mathcal{L}(y)$, then $\mathcal{W}_d(f,\delta_x)=\mathbb{E}[d(y,x)]$.*

Proof. Step 1: For any $f,g\in\mathcal{P}(E)$, we show that if $\pi\in\Lambda(f,\ g)$, then $\operatorname{supp}\pi\subset\operatorname{supp}f\times\operatorname{supp}g$.

For any Borel set $A\subset E$,

$$\pi(\operatorname{supp}f\times A) \;=\; \pi(E\times A)-\pi((\operatorname{supp}f)^c\times A)=g(A)-\pi((\operatorname{supp}f)^c\times A).$$

On the other hand, $\pi((\operatorname{supp}f)^c\times A)\le\pi((\operatorname{supp}f)^c\times E)=f((\operatorname{supp}f)^c)=0$. Hence $\pi(\operatorname{supp}f\times A)=g(A)$. Taking $A=\operatorname{supp}g$, one has $\pi(\operatorname{supp}f\times\operatorname{supp}g)=g(\operatorname{supp}g)=1$, i.e. $\operatorname{supp}\pi\subset\operatorname{supp}f\times\operatorname{supp}g$.

Step 2: If $g = \delta_x$, by Step 1,

$$\int_{E\times E} d(y_1, y_2)\pi(dy_1, dy_2) = \int_{E\times\{x\}} d(y_1, y_2)\pi(dy_1, dy_2)$$
$$= \int_{E\times\{x\}} d(y_1, y_2)(f \otimes \delta_x)(dy_1, dy_2) = \int_E d(y_1, x) f(dy_1).$$

Hence

$$\begin{aligned} \mathcal{W}_d(f, \delta_x) &= \inf_{\pi\in\Lambda(f,\delta_x)} \int_{E\times E} d(y_1, y_2)\pi(dy_1, dy_2) \\ &= \int_E d(y_1, x) f(dy_1) = \mathbb{E}[d(y, x)]. \end{aligned}$$

□

Remark 3.1. Let E be a topological space, $\{(M_i, X^i)\}_{i=1}^N$ be N random variables on $(0, \bar{M}]\times E$ and $\mu^N = \frac{1}{N}\sum_{i=1}^{N} M_i\delta_{X^i} \in \mathcal{M}_+(E)$. Introduce $\hat{G}^N = \mathcal{L}(\mu^N) \in \mathcal{P}\big(\mathcal{M}_+(E)\big)$ and $G^N = \mathcal{L}(M_1, X^1, \cdots, M_N, X^N) \in \mathcal{P}\big(((0, \bar{M}] \times E)^N\big)$, then

$$\forall\ \Phi \in C_b(\mathcal{M}_+(E)),\quad \langle \hat{G}^N, \Phi\rangle = \int_{((0,\bar{M}]\times E)^N} \Phi(\mu^N) G^N(dm_1, dx_2, \cdots, dm_N, dx_N). \tag{3.44}$$

In fact, we can construct a measurable map

$$T : ((0, \bar{M}] \times E)^N \to \mathcal{M}_+(E), \qquad (m_1, x_1, \cdots, m_N, x_N) \mapsto \frac{1}{N}\sum_{i=1}^{N} m_i\delta_{x^i}.$$

Therefore T transports G^N onto $\hat{G}^N$. We shall write $\hat{G}^N = T\sharp G^N$ and say that $\hat{G}^N$ is the push-forward of G^N by T, or equivalently $\forall\ \Phi \in C_b(\mathcal{M}_+(E))$,

$$\int_{((0,\bar{M}]\times E)^N} (\Phi \circ T) G^N(dm_1, dx_2, \cdots, dm_N, dx_N) = \int_{\mathcal{M}_+(E)} \Phi(Y)\hat{G}^N(dY),$$

i.e. (3.44) is satisfied.

Proposition 3.3. *Let (E, d) be a locally compact Polish space, $\{(M_i, X^i)\}_{i=1}^N$ be N random variables on $(0, \bar{M}] \times E$ and f be a positive Radon measure on E. Then the following two conditions are equivalent,*

- *(i) The positive Radon measure $\mu^N := \frac{1}{N}\sum_{i=1}^{N} M_i\delta_{X^i}$ converges in law to the constant random variable f as $N \to \infty$.*
- *(ii) $\forall\ \varphi \in C_b(E)$, $\mathbb{E}[|\langle \mu^N - f, \varphi\rangle|] \to 0$ as $N \to \infty$.*

Proof. First, we show $(ii) \Rightarrow (i)$:

Since E is a locally compact Polish space, then there is a dense sequence $(\varphi_n)_{n\in\mathbb{N}}$ in $C_0(E)$. One can define the weak-$*$ distance [22, page 98],

$$\text{for any } g_1, g_2 \in \mathcal{M}_+(E), \qquad d_1(g_1, g_2) := \sum_{n\in\mathbb{N}} \frac{1}{2^n}(1 \wedge |\langle g_1 - g_2, \varphi_n\rangle|),$$

then $(\mathcal{M}_+(E), d_1)$ is a Polish space [10, Section 15.7]. For any $\alpha, \beta \in \mathcal{P}(\mathcal{M}_+(E))$, define a Wasserstein distance on $\mathcal{P}(\mathcal{M}_+(E))$ as

$$\mathcal{W}_{d_1}(\alpha, \beta) := \inf_{\pi\in\Lambda(\alpha,\beta)} \Big\{ \int_{\mathcal{M}_+(E)\times\mathcal{M}_+(E)} d_1(g_1, g_2)\pi(dg_1, dg_2) \Big\}.$$

Then by Lemma 3.4 and Remark 3.1, one has

$$\begin{aligned} \mathcal{W}_{d_1}(\hat{G}^N, \delta_f) &= \int_{\mathcal{M}_+(E)} d_1(g_1, f)\hat{G}^N(dg_1) \\ &= \int_{((0,\bar{M}]\times E)^N} d_1(\mu^N, f)G^N(dm_1, dx_2, \cdots, dm_N, dx_N) \\ &= \sum_{n\in\mathbb{N}} \frac{1}{2^n}\mathbb{E}\big[(1 \wedge |\langle \mu^N - f, \varphi_n\rangle|)\big]. \end{aligned}$$

For any $\epsilon > 0$, there exists K_ϵ such that

$$\sum_{n>K_\epsilon} \frac{1}{2^n}\mathbb{E}\big[(1 \wedge |\langle \mu^N - f, \varphi_n\rangle|)\big] \leq \frac{\epsilon}{2}. \tag{3.45}$$

By (ii), there exists N' such that when $N \geq N'$,

$$\sum_{1\leq n\leq K_\epsilon} \frac{1}{2^n}\mathbb{E}\big[(1 \wedge |\langle \mu^N - f, \varphi_n\rangle|)\big] \leq \frac{\epsilon}{2}. \tag{3.46}$$

Combining (3.45) and (3.46) together, we have

$$\mathcal{W}_{d_1}(\hat{G}^N, \delta_f) \to 0 \ \text{ as } N \to \infty.$$

Using Proposition 2.1, $\hat{G}^N$ narrowly converges to δ_f as $N \to \infty$, which implies that (i) holds.

Now we prove $(i) \Rightarrow (ii)$:

For any $\mu \in \mathcal{M}_+(E)$ and $\varphi \in C_b(E)$, define a functional

$$\Gamma : \mathcal{M}_+(E) \to \mathbb{R}, \quad \mu \mapsto \Gamma(\mu) := |\langle \mu - f, \varphi\rangle| + |\langle f, \varphi\rangle|, \tag{3.47}$$

and notice $\Gamma \in C_b(\mathcal{M}_+(E))$. By (i), we obtain

$$\mathbb{E}[\Gamma(\mu^N)] = \mathbb{E}[|\langle \mu^N - f, \varphi\rangle|] + |\langle f, \varphi\rangle| \to \mathbb{E}[\Gamma(f)] = |\langle f, \varphi\rangle| \text{ as } N \to \infty. \tag{3.48}$$

Thus $\mathbb{E}[|\langle \mu^N - f, \varphi\rangle|] \to 0$ as $N \to \infty$. $\square$

Proof of Theorem 2.3: Let $\{(M_i, X_t^{i,\varepsilon})\}_{i=1}^N$ and $\{(M_i, X_t^i)\}_{i=1}^N$ be defined as in Proposition 3.2, $\mathcal{L}(M_i, X_0^i) = g_0$ and $\rho_0(x)dx = \int_0^{\bar{M}} m g_0(dm, dx)$. Let g_t be the common time marginal of $\{(M_i, X_t^i)\}_{i=1}^N$. From Proposition 3.1, we know that $f_t(dx) := \rho(t,x)dx = \int_0^{\bar{M}} m g_t(dm, dx)$, where ρ is the unique weak solution to (1.1) with the initial data ρ_0. By the exchangeability of $\{(M_i, X_t^{i,\varepsilon})\}_{i=1}^N$, for any $\varphi \in C_b(\mathbb{R}^d)$,

$$
\begin{aligned}
&\mathbb{E}[|\langle \frac{1}{N}\sum_{i=1}^N M_i \delta_{X_t^{i,\varepsilon}} - f, \varphi\rangle|^2] = \mathbb{E}[|\frac{1}{N}\sum_{i=1}^N M_i \varphi(X_t^{i,\varepsilon}) - \langle f, \varphi\rangle|^2] \qquad (3.49)\\
&= \frac{1}{N}\mathbb{E}\big[M_1^2 \varphi(X_t^{1,\varepsilon})^2\big] + \frac{N-1}{N}\mathbb{E}\big[M_1\varphi(X_t^{1,\varepsilon})M_2\varphi(X_t^{2,\varepsilon})\big] - 2\langle f,\varphi\rangle \mathbb{E}\big[M_1\varphi(X_t^{1,\varepsilon})\big] + \langle f,\varphi\rangle^2.
\end{aligned}
$$

Recall that $G_t^\varepsilon(m_1, x_1, \cdots, m_N, x_N)$ is the joint distribution of $\big(M_1, X_t^{1,\varepsilon}, \cdots, M_N, X_t^{N,\varepsilon}\big)$ in Corollary 3.1. Taking $j = 2$ in (3.42) and by Proposition 2.1, there exists regularized parameters $\varepsilon(N) \sim (\ln N)^{-\frac{1}{d}} \to 0$ as $N \to \infty$ such that

$$
\begin{aligned}
&\mathbb{E}\big[M_1\varphi(X_t^{1,\varepsilon(N)})M_2\varphi(X_t^{2,\varepsilon(N)})\big] = \int_{((0,\bar{M}]\times\mathbb{R}^d)^N} m_1 m_2 \varphi(x_1)\varphi(x_2) G_t^\varepsilon(dm_1, dx_1, \cdots, dm_N, dx_N)\\
&\quad = \int_{\mathbb{R}^{2d}} \varphi(x_1)\varphi(x_2) f_t^{(2),\varepsilon(N)}(dx_1, dx_2) \to \Big(\int_{\mathbb{R}^d} \varphi(x) f_t(dx)\Big)^2 \quad \text{as } N\to\infty. \qquad (3.50)
\end{aligned}
$$

Similarly taking $j = 1$ in (3.42), one has

$$
\mathbb{E}\big[M_1\varphi(X_t^{1,\varepsilon(N)})\big] \to \int_{\mathbb{R}^d} \varphi(x) f_t(dx) \quad \text{as } N \to \infty. \qquad (3.51)
$$

Taking the limit $N \to \infty$ in (3.49), thanks to (3.50) and (3.51), we have

$$
\mathbb{E}[|\langle \frac{1}{N}\sum_{i=1}^N M_i \delta_{X_t^{i,\varepsilon(N)}} - f, \varphi\rangle|^2] \to 0 \quad \text{as } N \to \infty. \qquad (3.52)
$$

From (3.52), one also has

$$
\begin{aligned}
\mathbb{E}[|\langle \frac{1}{N}\sum_{i=1}^N M_i \delta_{X_t^{i,\varepsilon(N)}} - f, \varphi\rangle|] &\le \big\{\mathbb{E}[\langle \frac{1}{N}\sum_{i=1}^N M_i \delta_{X_t^{i,\varepsilon(N)}} - f, \varphi\rangle^2]\big\}^{\frac{1}{2}}\\
&\to 0 \quad \text{as } N \to \infty. \qquad (3.53)
\end{aligned}
$$

From Proposition 3.3, we have

$$
\frac{1}{N}\sum_{i=1}^N M_i \delta_{X_t^{i,\varepsilon(N)}} \to f \quad \text{in law as } N \to \infty.
$$

Since f is a constant random variable, then $\frac{1}{N}\sum_{i=1}^N M_i \delta_{X_t^{i,\varepsilon(N)}}$ converges to f in probability [11, Lemma 3.7]. Thus there exists a subsequence converges a.s. [11, Lemma 3.2] and we complete the proof. □

4 Practical algorithms and convergence results

First we approximate the initial density ρ_0. Take h as a grid size and decompose the domain D into the union of nonoverlapping cells $C_\alpha = X_\alpha + [-\frac{h}{2}, \frac{h}{2}]^d$ with center $X_\alpha = h\alpha$, i.e. $D \subset \bigcup_{\alpha \in I} C_\alpha$, where $I = \{\alpha\} \subset \mathbb{Z}^d$ is the index set for cells. The total number of cells is given by $M = \sum_{\alpha \in I} \approx \frac{|\Omega|}{h^d}$.

Define $M_\alpha = M \int_{C_\alpha} \rho_0(x)dx$, then $M_\alpha \le CA =: \bar{M}$, where $A = \|\rho_0(x)\|_{L^\infty(\mathbb{R}^d)}$. Take a function φ satisfying supp $\varphi(x) \subset B(0,1)$, $\varphi(x) \ge 0$ and $\int_{B(0,1)} \varphi(x)dx = 1$. Approximate φ as $\varphi_h(x) = \frac{1}{h^d}\varphi(\frac{x}{h})$. Then approximate ρ_0 as

$$\rho_{0,h} = \frac{1}{M}\sum_{\alpha \in I} M_\alpha \varphi_h(x - X_\alpha), \tag{4.1}$$

We can derive that there exists a constant C such that

$$\|\rho_{0,h}\|_{L^1} = \|\rho_0\|_{L^1}, \quad \|\rho_{0,h}\|_{L^\infty} \le C\|\rho_0\|_{L^\infty}, \quad \int |x|\rho_{0,h}dx \le C\int |x|\rho_0 dx,$$

and

$$\mathcal{W}_1(\rho_{0,h}, \rho_0) \le Ch.$$

To sample the density $\rho_{0,h}$ in (4.1), we construct the random variables $\{(M_i^h, X_0^{i,h})\}_{i=1}^N$ as follows:

(i) Pick $i = \alpha \in I$ with equal probability $\frac{1}{M}$.

(ii) Construct N i.i.d. random variables $\{\xi_i^h\}_{i=1}^N$ with common density $\varphi_h(x)$. Let

$$M_i^h = M_\alpha, \quad X_0^{i,h} = X_\alpha + \xi_i^h. \tag{4.2}$$

From the above construction, it is easy to verify that $(M_i^h, X_0^{i,h})$ are i.i.d. with common distribution $g_0^h(m,x)$, where

$$g_0^h(dm, dx) = \frac{1}{M}\sum_{\alpha \in I} \delta_{M_\alpha} \otimes \varphi_h(x - X_\alpha)dx.$$

Remark 4.1. M_i^h are i.i.d. with the common marginal distribution $R_0(dm) = \frac{1}{M}\sum_{\alpha \in I}\delta_{M_\alpha}$; $X_0^{i,h}$ are i.i.d. with the common marginal density $H_0(x) = \frac{1}{M}\sum_{\alpha \in I}\varphi_h(x - X_\alpha)$; $\mathbb{E}(M_i^h) = \int_0^{\bar{M}} mR_0(dm) = \frac{1}{M}\sum_{\alpha \in I} M_\alpha = 1$ and

$$\int_0^{\bar{M}} m g_0^h(dm, \cdot) = \rho_{0,h}(x)dx. \tag{4.3}$$

Remark 4.2. In practical computation, we take $N = M$ and order all the cells $\alpha \in I$, and denote the order as i. The convergence analysis for this method should be similar to Theorem 2.3.

From Theorem 2.1 (ii), we have the following corollary.

Corollary 4.1. *Let $\rho_h(t,x)$ and $\rho(t,x)$ be the unique weak solutions to the KS equation* (1.1) *with the initial data $\rho_{0,h}$ and ρ_0 respectively. Then there exists three constants $C_0(T)$, C (depending on $\|\rho^1\|_{L^\infty(0,T;L^\infty\cap L^1(\mathbb{R}^d))}$ and $\|\rho^2\|_{L^\infty(0,T;L^\infty\cap L^1(\mathbb{R}^d))}$) and C_T such that if $\mathcal{W}_1(\rho_{0,h},\rho_0) < Ch < C_0(T)$,*

$$\sup_{t\in[0,T]} \mathcal{W}_1(\rho_{t,h},\rho_t) \le C_T \mathcal{W}_1(\rho_{0,h},\rho_0)^{\exp(-CT)} \le C_T h^{\exp(-CT)}. \tag{4.4}$$

Theorem 4.1. *Let $\rho(t,x)$ be the unique weak solution to the KS equation* (1.1) *with the initial data ρ_0 and $\{X_t^{1,\varepsilon,h},\cdots,X_t^{N,\varepsilon,h}\}$ be the unique strong solution to* (1.2) *with the initial data $\{(M_i^h, X_0^{i,h})\}_{i=1}^N$ given by* (4.2)*. Then there exists a subsequence of $\{X_t^{1,\varepsilon,h},\cdots,X_t^{N,\varepsilon,h}\}$ (without relabeling N) and regularized parameters $\varepsilon(N)\sim(\ln N)^{-\frac{1}{d}}\to 0$ as $N\to\infty$ such that*

$$\lim_{h\to 0}\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^N M_i\varphi(X_t^{i,\varepsilon(N),h}) = \int_{\mathbb{R}^d}\varphi(x)\rho_t dx \text{ for any } \varphi\in C_b(\mathbb{R}^d). \tag{4.5}$$

Proof. Fix $h>0$. Let $\rho_h(t,x)$ be the unique weak solutions to the KS equation (1.1) with the initial data $\rho_{0,h}$. From (3.53), we obtain that for any $\psi\in BL(\mathbb{R})$ there exists regularized parameters $\varepsilon(N)\sim(\ln N)^{-\frac{1}{d}}\to 0$ as $N\to\infty$ such that

$$\begin{aligned}
&\mathbb{E}\big[\psi\big(\frac{1}{N}\sum_{i=1}^N M_i\varphi(X_t^{i,\varepsilon(N),h})\big) - \psi\big(\int_{\mathbb{R}^d}\varphi(x)\rho_t^h dx\big)\big] \\
&\quad\le C\mathbb{E}\big[\big|\frac{1}{N}\sum_{i=1}^N M_i\varphi(X_t^{i,\varepsilon(N),h}) - \int_{\mathbb{R}^d}\varphi(x)\rho_t^h dx\big|\big]\to 0 \quad \text{as } N\to\infty.
\end{aligned} \tag{4.6}$$

By the portemanteau theorem, (4.6) means that the law of $\frac{1}{N}\sum_{i=1}^N M_i\varphi(X_t^{i,\varepsilon(N),h})$ narrowly converges to the law of $\int_{\mathbb{R}^d}\varphi(x)\rho_t^h dx$ as $N\to\infty$. Thus there exists a subsequence of $\{X_t^{1,\varepsilon,h},\cdots,X_t^{N,\varepsilon,h}\}$ (without relabeling N) such that

$$\frac{1}{N}\sum_{i=1}^N M_i\varphi(X_t^{i,\varepsilon(N),h}) \to \int_{\mathbb{R}^d}\varphi(x)\rho_t^h dx \quad \text{a.s. as } N\to\infty.$$

From Corollary 4.1 and Proposition 2.1, one has

$$\int_{\mathbb{R}^d} \varphi(x)\rho_t^h dx \to \int_{\mathbb{R}^d} \varphi(x)\rho_t dx \quad \text{as} \quad h \to 0.$$

Combining the above two results together, we finish the proof. □

Acknowledgments The authors would like to thank Greg Herschlag for many helpful suggestions. The research of J.-G. L. was partially supported by KI-Net NSF RNMS Grant No. 1107291.

References

[1] S. Bian and J.-G. Liu, Dynamic and steady states for multi-dimensional KS equations with diffusion exponent $m > 0$, Comm. Math. Phy., **323** (2013), 1017-1070.

[2] A. Blanchet, J. Dolbeault and B. Perthame, Two-dimensional KS equations: optimal critical mass and qualitative properties of the solutions, Eletron. J. Differ. Equ., **44** (2006), 1-33.

[3] A. J. Chorin, Numerical study of slightly viscous flow, J. Fluid. Mech., **57**, (1973) 785-796.

[4] G. H. Cottet and P. Koumoutsakos, Vortex Methods: Theory and Practice, Cambridge University Press, 2000.

[5] K. Craig and A. Bertozzi, A blob method for the aggregation equation, to appear in Math. Comp.

[6] R. Durrett, Probability: Theory and Examples, Third edition, 2007.

[7] N. Fournier, M. Hauray and S. Mischler, Propagation of chaos for the 2D viscous vortex model, J. Eur. Math. Soc., **16**, (2014) 1423-1466.

[8] J. Goodman, Convergence of the random vortex method, Comm. Pure Appl. Math., **40**, (1987) 189-220.

[9] M. Hauray and S. Mischler, On Kac's chaos and related problems, J. Funct. Anal., **266**, (2014) 6055-6157.

[10] O. Kallenberg, Random Measures, 3th edition, Akademie-Verlag, Berlin, 1983.

[11] O. Kallenberg, Foundations of Modern Probability, Probability and its Applications, Springer-Verlag New York, 1997.

[12] A. Klenke, Probability Theory: A Comprehensive Course, Second edition, Springer, 2014.

[13] P. Koumoutsakos and A. Leonard, High-resolution simulations of the flow around an impulsively started cylinder using vortex methods, J. Fluid Mech., **96**, (1995) 1-38.

[14] E. H. Lieb and M. Loss, Analysis, 2nd edition, Grad. Stud. Math., **14**, American Mathematical Society, Providence, RI, 2001.

[15] J.-G. Liu and R. Yang, Propagation of chaos for Keller-Segel equation, preprint.

[16] D. G. Long, Convergence of the random vortex method in two dimensions, J. Amer. Math. Soc, **1**, (1988) 779-804.

[17] A. J. Majda and A. L. Bertozzi, Vorticity and Incompressible Flow, Cambridge Texts in Applied Mathematics, Cambridge University Press, 2003.

[18] C. Marchioro and M. Pulvirenti, Vortex Methods in Two-Dimensional Fluid Dynamics, Springer-verlag Berlin Heidelberg, New York Tokyo, 1984.

[19] S. Mischler and C. Mouhot, Kac's program in kinetic theory, Invent math, **193**, (2013) 1-147.

[20] B. Øksendal, Stochastic Differential Equations: An Introduction with Applications, Sixth Edition, Springer-Verlag, Berlin, 2003.

[21] A.-S. Sznitman, Topics in propagation of chaos, In Ecole d'Eté de Probabilités de Saint-Flour XIX-1989, Lecture Notes in Math. 1464. Springer, Berlin, (1991).

[22] C. Villani, Optimal Transport, Old and New, Grundlehren Math. Wiss. 338, Springer, Berlin, 2009.