ERROR ESTIMATE OF A RANDOM PARTICLE BLOB METHOD FOR THE KELLER-SEGEL EQUATION

HUI HUANG AND JIAN-GUO LIU

Abstract. We establish an optimal error estimate for a random particle blob method for the Keller-Segel equation in $\mathbb{R}^d$ ($d \geq 2$). With a blob size $\varepsilon = h^\kappa$ ($1/2 < \kappa < 1$), we prove a rate $h|\ln h|$ of convergence in ℓ_h^p ($p > \frac{d}{1-\kappa}$) norm up to a probability $1 - h^{C|\ln h|}$, where h is the initial grid size.

1. Introduction

The vortex method was first introduced by Chorin in 1973 [6], which is one of the most significant computational methods for fluid dynamics and other related fields. The convergence of the vortex method for two and three dimensional inviscid incompressible fluid flows was first proved by Hald [13], Beale and Majda [2, 3]. Then Anderson and Greengard [1] gave a simpler proof for the estimate of the consistency error. When the effect of viscosity is involved, the vortex method is replaced by the so called random vortex method by adding a Brownian motion to every vortex. The convergence analysis of the random vortex method for the Navier-Stokes equation have been given by [11, 19, 20, 23] in 1980s.

Generally speaking, there are two ways to set up the initial data. On one hand, some authors like Marchioro and Pulvirenti [20], Osada [23], Goodman [11] and [17] took the initial positions as independent identically distributed random variables $X_i(0)$ with common density $\rho_0(x)$. Specifically, Goodman proved a rate of convergence for the incompressible Navier-Stokes equation in two dimension of the order $N^{-1/4}\ln N$, where N is the number of vortices used in the computation. However, this Monte Carlo sampling method is very inefficient in the computation. On the other hand, the Chorin's original method assumed that initial positions of the vortices are on the lattice points $hi \in \mathbb{R}^2$ with mass $\rho_0(hi)h^2$. Especially, Long [19] achieved an almost optimal rate of convergence of the order $N^{-1/2}\ln N \sim h|\ln h|$ except an event of probability $h^{C'C}$. And much of his technique will be adapted to this article. A similar probabilistic approach has been used on Vlasov-Poisson system by [5]. Lastly, we refer to the book [7] for theoretical and practical use of

2010 *Mathematics Subject Classification.* Primary 65M75, 65M15, 65M12, 35Q92, 35K55, 60H30.

Key words and phrases. Concentration inequality, Bennett's inequality, Newtonian aggregation, sampling estimate, chemotaxis, Brownian motion, propagation of chaos.

Hui Huang is partially supported by National Natural Science Foundation of China (Grant No: 41390452, 11271118).

The work of Jian-Guo Liu is partially supported by KI-Net NSF RNMS grant No. 1107291 and NSF DMS grant No. 1514826.

the vortex methods, and also refer to [8] for recent progress on a blob method for the aggregation equation.

In this paper, we introduce a random particle blob method for the following classical Keller-Segel (KS) equation [14] in $\mathbb{R}^d$ $(d \geq 2)$:

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

where ν is a positive constant. This model is developed to describe the biological phenomenon chemotaxis. In the context of biological aggregation, $\rho(t,x)$ represents the bacteria density, and $c(t,x)$ represents the chemical substance concentration, which is given by fundamental solution as follows

$$c(t,x) = \begin{cases} C_d \displaystyle\int_{\mathbb{R}^d} \frac{\rho(t,y)}{|x-y|^{d-2}} dy, & \text{if } d \geq 3, \\ -\dfrac{1}{2\pi} \displaystyle\int_{\mathbb{R}^d} \ln|x-y| \rho(t,y) dy, & \text{if } d = 2, \end{cases} \tag{1.2}$$

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 the d-dimensional unit ball. We can recast $c(t,x)$ as $c(t,x) = \Phi * \rho(t,x)$ with Newton potential $\Phi(x)$, which can be represented as

$$\Phi(x) = \begin{cases} \dfrac{C_d}{|x|^{d-2}}, & \text{if } d \geq 3, \\ -\dfrac{1}{2\pi} \ln|x|, & \text{if } d = 2. \end{cases} \tag{1.3}$$

Furthermore, we take the gradient of the Newtonian potential $\Phi(x)$ as the attractive force $F(x)$. Thus we have $F(x) = \nabla \Phi(x) = -\frac{C_* x}{|x|^d}$, $\forall \ x \in \mathbb{R}^d \backslash \{0\}$, $d \geq 2$, where $C_* = \frac{\Gamma(d/2)}{2\pi^{d/2}}$.

Now we consider the KS equation (1.1) under the following assumption

Assumption 1. The initial density $\rho_0(x)$ satisfies

(1) $\rho_0(x)$ has a compact support D with $D \subseteq B(R_0)$;
(2) $0 \leq \rho_0 \in H^k(\mathbb{R}^d)$ for $k \geq \frac{3d}{2} + 1$.

In fact, the above assumption is sufficient for the existence of the unique local solution to (1.1) with the following regularity

$$\|\rho\|_{L^\infty(0,T;H^k(\mathbb{R}^d))} \leq C(\|\rho_0\|_{H^k(\mathbb{R}^d)}), \tag{1.4}$$

$$\|\partial_t \rho\|_{L^\infty(0,T;H^{k-2}(\mathbb{R}^d))} \leq C(\|\rho_0\|_{H^k(\mathbb{R}^d)}), \tag{1.5}$$

where $T > 0$ only depends on $\|\rho_0\|_{H^k(\mathbb{R}^d)}$. The proof of this result is a standard process and it will be given in Appendix A. Denote $T_{\max}$ to be the largest existence time, such that (1.4) and (1.5) are valid for any $0 < T < T_{\max}$. As a direct result of the Sobolev imbedding theorem, one has $\rho(t,x) \in C^{k-d/2-1}(\mathbb{R}^d)$ for any $t \in [0,T]$. We define the drift term

$$G(t,x) := \int_{\mathbb{R}^d} F(x-y) \rho(t,y) dy, \tag{1.6}$$

then $-\Delta G(t,x) = \nabla\rho(t,x)$. By using the Sobolev imbedding theorem, one has

$$\|G\|_{L^\infty\left(0,T;W^{k-\frac{d}{2},\infty}(\mathbb{R}^d)\right)} \le C\|G\|_{L^\infty(0,T;H^{k+1}(\mathbb{R}^d))} \le C(\|\rho_0\|_{H^k(\mathbb{R}^d)}), \tag{1.7}$$

$$\|\partial_t G\|_{L^\infty\left(0,T;W^{k-\frac{d}{2}-2,\infty}(\mathbb{R}^d)\right)} \le C(\|\rho_0\|_{H^k(\mathbb{R}^d)}). \tag{1.8}$$

So $G(t,x)$ is bounded and Lipschitz continuous with respect to x for $k \ge d/2+1$ from (1.7) and the Sobolev imbedding theorem. Thus the following stochastic differential equation (SDE):

$$X(t) = X(0) + \int_0^t \int_{\mathbb{R}^d} F(X(s)-y)\rho(s,y)dyds + \sqrt{2\nu}B(t), \tag{1.9}$$

has a unique strong solution $X(t)$ by a basic theorem of SDE [22, Theorem 5.2.1], where $X(0) = \alpha \in D$ and $B(t)$ is a standard Brownian motion.

If we denote the fundamental solution of the following PDE

$$\begin{cases} u_t = \nu\Delta u - \nabla\cdot(uG), \\ u(0,x) = \delta_\alpha(x), \ \alpha \in D. \end{cases} \tag{1.10}$$

to be $g(t, x \leftarrow 0, \alpha)$, then it is the transition probability density of the diffusion process $X(t)$, i.e., $g(t, x \leftarrow 0, \alpha)$ is the density that a particle reached the position x at time t from position α at time 0. And we have

$$\rho(t,x) = \int_{\mathbb{R}^d} g(t, x \leftarrow 0, \alpha)\rho_0(\alpha)d\alpha. \tag{1.11}$$

See Friedman [10, Theorem 5.4 on P.149].

We take h as the grid size and decompose the domain D into the union of non-overlapping cells $C_i = X_i(0) + [-\frac{h}{2}, \frac{h}{2}]^d$ with center $X_i(0) = hi =: \alpha_i \in D$, i.e. $D \subset \bigcup_{i\in I} C_i$, where $I = \{i\} \subset \mathbb{Z}^d$ is the index set for cells. The total number of cells is given by $N = \sum_{i\in I} \approx \frac{|D|}{h^d}$.

Suppose $X_i(t)$ is the strong solution to (1.9), i.e.

$$X_i(t) = X_i(0) + \int_0^t \int_{\mathbb{R}^d} F\big(X_i(s)-y\big)\rho(s,y)dyds + \sqrt{2\nu}B_i(t), \quad i \in I, \tag{1.12}$$

with the initial data $X_i(0) = \alpha_i = hi$ where $B_i(t)$ are independent standard Brownian motions.

For any test function $\varphi \in C_0^\infty(\mathbb{R}^d)$ and $t \in [0,T]$, we define

$$u(s,\alpha) = \int_{\mathbb{R}^d} \varphi(x) g(t, x \leftarrow s, \alpha)dx, \quad s \in [0,t]. \tag{1.13}$$

Then $u(s,\alpha)$ is the solution to the following backward Kolmogrov equation

$$\begin{cases} \partial_s u = -\nu\Delta u - G\cdot\nabla u, \quad \alpha \in \mathbb{R}^d, \ s \in [0,t], \\ u(t,\alpha) = \varphi(\alpha). \end{cases} \tag{1.14}$$

Following the standard regularity estimate, we have

$$\|u(s,\cdot)\|_{H^{d+1}(\mathbb{R}^d)} \le C_T\|\varphi\|_{H^{d+1}(\mathbb{R}^d)}, \quad s \in [0,t]. \tag{1.15}$$

Moreover, on one hand, we have

$$< \varphi, \rho > = \int_{\mathbb{R}^d} \varphi(x)\rho(x)dx = \int_{\mathbb{R}^d} \varphi(x) \int_D \rho_0(\alpha) g(t, x \leftarrow 0, \alpha) d\alpha dx = \int_D u(0,\alpha)\rho_0(\alpha)d\alpha. \tag{1.16}$$

On the other hand, we define the empirical measure $\mu_N(t) := \sum_{j\in I} \delta(x - X_j(t))\rho_0(\alpha_j)h^d$. And define $\mathbb{E}[\mu_N(t)]$ in the sense of Pettis integral [25], i.e.

$$\begin{aligned} < \varphi, \mathbb{E}[\mu_N(t)] > &= \mathbb{E}[< \varphi, \mu_N(t) >] \\ &= \sum_{j\in I} \int_{\mathbb{R}^d} \varphi(x) g(t, x \leftarrow 0, \alpha_j) dx \rho_0(\alpha_j) h^d \\ &= \sum_{j\in I} u(0, \alpha_j)\rho_0(\alpha_j)h^d. \end{aligned} \tag{1.17}$$

Combining (1.16) and (1.17), and using (2.1) from Lemma 2.3, we conclude that

$$\begin{aligned} |< \varphi, \mathbb{E}[\mu_N(t)] - \rho >| &= \left| \sum_{j\in I} u(0,\alpha_j)\rho_0(\alpha_j)h^d - \int_D u(0,\alpha)\rho_0(\alpha)d\alpha \right| \\ &\leq Ch^{d+1} \|u(0,\cdot)\rho_0\|_{W^{d+1,1}(\mathbb{R}^d)} \\ &\leq Ch^{d+1} \|u(0,\cdot)\|_{H^{d+1}(\mathbb{R}^d)} \leq Ch^{d+1}\|\varphi\|_{H^{d+1}(\mathbb{R}^d)} \end{aligned} \tag{1.18}$$

which leads to

$$\|\mathbb{E}[\mu_N(t)] - \rho\|_{H^{-(d+1)}(\mathbb{R}^d)} \leq Ch^{d+1}. \tag{1.19}$$

Our above error estimate (1.19) is in the weak sense (see [12] for the concept). Recently, the error estimate in the strong sense up to a small probability was obtained by [18]. Therefore, the main task of this article is to establish the error estimate between $X_i(t)$ and $X_{i,\varepsilon}(t)$. Here $X_{i,\varepsilon}(t)$ is the solution to the random particle blob method which we will describe below.

Introducing a random particle blob method for the KS equation as in [19], we have the following system of SDEs

$$X_{i,\varepsilon}(t) = X_{i,\varepsilon}(0) + \int_0^t \sum_{j\in I} F_\varepsilon\big(X_{i,\varepsilon}(s) - X_{j,\varepsilon}(s)\big)\rho_j h^d ds + \sqrt{2\nu}B_i(t), \quad i \in I, \tag{1.20}$$

with the initial data $X_{i,\varepsilon}(0) = \alpha_i = hi$ where

$$\rho_j = \rho_0(\alpha_j), \quad F_\varepsilon = F * \psi_\varepsilon, \quad \psi_\varepsilon(x) = \varepsilon^{-d}\psi(\varepsilon^{-1}x), \quad \varepsilon > 0. \tag{1.21}$$

The choice of the blob function ψ is closely related to the accuracy of our method. Following [16], we choose $\psi(x) \geq 0$, $\psi(x) \in C_0^{2d+2}(\mathbb{R}^d)$,

$$\psi(x) = \begin{cases} C(1 + \cos \pi|x|)^{d+2}, & \text{if } |x| \leq 1, \\ 0, & \text{if } |x| > 1, \end{cases} \tag{1.22}$$

where C is a constant such that $\int_{\mathbb{R}^d} \psi(x)dx = 1$.

For convenience, we will give the following notations for the drift term

$$G(t,x) := F * \rho = \int_{\mathbb{R}^d} F(x-y)\rho(t,y)dy; \tag{1.23}$$

$$G_\varepsilon^h(t,x) := \sum_{j\in I} F_\varepsilon\big(x - X_j(t)\big)\rho_j h^d; \tag{1.24}$$

$$\hat{G}_\varepsilon^h(t,x) := \sum_{j\in I} F_\varepsilon\big(x - X_{j,\varepsilon}(t)\big)\rho_j h^d. \tag{1.25}$$

Define the discrete ℓ_h^p norm of a vector $v = (v_i)_{i\in I}$ such that

$$\|v\|_{\ell_h^p} = \|(v_i)_{i\in I}\|_{\ell_h^p} = \left(\sum_{i\in I} |v_i|^p h^d\right)^{1/p}, \quad p > 1. \tag{1.26}$$

Then we have the following main theorem:

Theorem 1.1. *Suppose the initial density $\rho_0(x)$ satisfies Assumption 1. Let T_{max} be the largest existence time of the regular solution* (1.4), (1.5) *to KS equation* (1.1)*. Assume that $X_h(t) = (X_i(t))_{i\in I}$ is the exact path of* (1.12) *and $X_{h,\varepsilon}(t) = (X_{i,\varepsilon}(t))_{i\in I}$ is the solution to the random particle blob method* (1.20)*. We take $\varepsilon = h^\kappa$ with any $\frac{1}{2} < \kappa < 1$ and $p > \frac{d}{1-\kappa}$, then for all $0 < h \le h_0$ with h_0 sufficiently small, there exist two positive constants C and C' depending on T_{max}, p, d, R_0 and $\|\rho_0\|_{H^k(\mathbb{R}^d)}$, such that the following estimate holds*

$$P\left(\max_{0\le t\le T} ||X_{h,\varepsilon}(t) - X_h(t)||_{\ell_h^p} < \Lambda h|\ln h|\right) \ge 1 - h^{C\Lambda|\ln h|},$$

for any $\Lambda > C'$ and $0 < T < T_{max}$.

Remark 1.2. For the Coulomb interaction case $F = -\nabla\Phi(x)$, the above estimate holds for any $T > 0$, since the regular solution ρ exists globally.

To conclude this introduction, we present the outline of the paper. In Section 2, we give some essential lemmas including kernel, sampling, concentration and far field estimates. In Section 3, we give a proof of the consistency error at the fixed time $t \in [0,T]$. Then we give a stability theorem in Section 4. Next, by using results from Sections 3-4, we conclude the proof of the convergence of the particle path in Section 5. In Appendix A, we give a sketch proof of the regularity $\rho \in L^\infty\big(0,T;H^k(\mathbb{R}^d)\big)$. Lastly, we extend our result to the particle system with regular force in Appendix B.

2. Preliminaries on kernel, sampling, concentration and far field estimates

Notation: The inessential constants will be denoted generically by C, even if it is different from line to line.

Firstly, we summarize some useful estimates about the kernel F_ε in (1.21) and its derivatives.

Lemma 2.1. *(Pointwise estimates)*

(i) $F_\varepsilon(0) = 0$ *and* $F_\varepsilon(x) = F(x)h(\frac{|x|}{\varepsilon})$ *for any* $x \ne 0$, *where* $h(r) = \frac{2\pi^{d/2}}{\Gamma(d/2)}\int_0^r \psi(s)s^{d-1}ds$;
(ii) $F_\varepsilon(x) = F(x)$ *for any* $|x| \ge \varepsilon$ *and* $F_\varepsilon(x) \le \min\{C\frac{|x|}{\varepsilon^d}, |F(x)|\}$;
(iii) $|\partial^\beta F_\varepsilon(x)| \le C_\beta \varepsilon^{1-d-|\beta|}$, *for any* $x \in \mathbb{R}^d$;
(iv) $|\partial^\beta F_\varepsilon(x)| \le C_\beta |x|^{1-d-|\beta|}$, *for any* $|x| \ge \varepsilon$.

The estimates (i) and (ii) have been proved in [16, Lemma 2.1]. For (iii) and (iv), we can follow the argument of [2, Lemma 5.1] by making dimensional changes and using the definition of $F_\varepsilon(x)$ in our paper.

Lemma 2.2. *(Integral estimates)*

(i) $\int_{|x|\le R} |F_\varepsilon(x)|dx \le CR,\ \forall\ \varepsilon < 1$;

(ii) $\|F_\varepsilon\|_{W^{|\beta|,q}(\mathbb{R}^d)} \le C\varepsilon^{d/q+1-d-|\beta|}$, *for* $q > 1$.

Lemma 2.2 is a direct result from Lemma 2.1.

Also, we will need the following sampling lemma, which is essential to our error estimate.

Lemma 2.3. *(*[1, Lemma 2.2]*) Suppose that* $f \in W^{d+1,1}(\mathbb{R}^d)$*, then*

$$\left|\sum_{i\in\mathbb{Z}^d} f(hi)h^d - \int_{\mathbb{R}^d} f(x)dx\right| \le C_d h^{d+1}\|f\|_{W^{d+1,1}(\mathbb{R}^d)}. \tag{2.1}$$

The proof of this lemma is based on the Poisson summation formula, which was given by Anderson and Greengard [1].

Since the initial positions $X_i(0)$ are chosen on the lattice points instead of being chosen randomly, the following lemma is essential to our analysis.

Lemma 2.4. *Let* $X(t,\alpha)$ *be the solution of the following SDE under Assumption 1*

$$X(t;\alpha) = X(0;\alpha) + \int_0^t G\left(s, X(s;\alpha)\right) ds + \sqrt{2\nu}B_\alpha(t),$$

with initial data $X(0;\alpha) = \alpha \in D$ *and* $B_\alpha(t)$ *is the standard Brownian motion. Assume* $\{X_i(t)\}$ *are solutions of the SDEs*

$$X_i(t) = X_i(0) + \int_0^t G\left(s, X_i(s)\right) ds + \sqrt{2\nu}B_i(t), \quad i \in I,$$

with initial data $X_i(0) = \alpha_i = hi \in D$ *and* $\{B_i(t)\}$ *are independent standard Brownian motions. For functions* $f \in W^{d+1,q}(\mathbb{R}^d;\mathbb{R}^{d'})$ *and* $\Gamma \in W_0^{d+1,q'}(\mathbb{R}^d;\mathbb{R}^{d'})$ *with supp* $\Gamma = D$ *and* $1/q + 1/q' = 1$*, we have the following estimate for the quadrature error*

$$\begin{aligned}\max_{0\le t\le T}&\left|\sum_{i\in I} \mathbb{E}\left[f\left(X_i(t)\right)\right]\Gamma(\alpha_i)h^d - \int_D \mathbb{E}\left[f\left(X(t;\alpha)\right)\right]\Gamma(\alpha)d\alpha\right| \\ \le& Ch^{d+1}\|f\|_{W^{d+1,q}(\mathbb{R}^d;\mathbb{R}^{d'})},\end{aligned} \tag{2.2}$$

where C *depends only on* $d,\ d',\ T,\ \|\rho_0\|_{H^k(\mathbb{R}^d)}$ *and* $\|\Gamma\|_{W_0^{d+1,q'}(\mathbb{R}^d;\mathbb{R}^{d'})}$.

Proof. To prove this lemma, for any $t \in [0,T]$, we define

$$u(s,y) = \int_{\mathbb{R}^d} f(x)g(t,x \leftarrow s,y)dx, \quad s \in [0,t]. \tag{2.3}$$

Then, one has

$$u(t,y) = f(y); \quad u(0,y) = \int_{\mathbb{R}^d} f(x)g(t,x\leftarrow 0,y)dx = \mathbb{E}\left[f\left(X(t;y)\right)\right]. \tag{2.4}$$

Thus $u(s,y)$ is the solution to the following backward Kolmogrov equation

$$\begin{cases} \partial_s u = -\nu\Delta u - G\cdot\nabla u, \quad y\in\mathbb{R}^d,\ s\in[0,t], \\ u(t,y) = f(y). \end{cases} \tag{2.5}$$

Following the standard regularity estimate, we have

$$\|u(s,\cdot)\|_{W^{d+1,q}(\mathbb{R}^d)} \le C\|f\|_{W^{d+1,q}(\mathbb{R}^d;\mathbb{R}^{d'})}, \quad s\in[0,t], \tag{2.6}$$

where C depends only on $d,\ d',\ T$ and $\|G\|_{L^\infty(0,T;W^{d+1,\infty}(\mathbb{R}^d))}$.

Notice the fact that $\Gamma(\alpha)$ has support D, and we can use Lemma 2.3, which leads to

$$\begin{aligned} &\left|\sum_{i\in I} u(0,\alpha_i)\Gamma(\alpha_i)h^d - \int_D u(0,\alpha)\Gamma(\alpha)d\alpha\right| \\ \le & Ch^{d+1}\|u(0,\alpha)\Gamma(\alpha)\|_{W^{d+1,1}(\mathbb{R}^d)} \le Ch^{d+1}\|u(0,\alpha)\|_{W^{d+1,q}(\mathbb{R}^d)} \\ \le & Ch^{d+1}\|f\|_{W^{d+1,q}(\mathbb{R}^d;\mathbb{R}^{d'})}, \end{aligned} \tag{2.7}$$

where C depends only on $d,\ d',\ T\ \|\rho_0\|_{H^k(\mathbb{R}^d)}$ and $\|\Gamma\|_{W_0^{d+1,q'}(\mathbb{R}^d;\mathbb{R}^{d'})}$.

Substitute $u(0,\alpha) = \mathbb{E}\left[f\left(X(t;\alpha)\right)\right]$ in (2.7), one has

$$\begin{aligned} &\max_{0\le t\le T}\left|\sum_{i\in I}\mathbb{E}\left[f\left(X(t;\alpha_i)\right)\right]\Gamma(\alpha_i)h^d - \int_D \mathbb{E}\left[f\left(X(t;\alpha)\right)\right]\Gamma(\alpha)d\alpha\right| \\ \le & Ch^{d+1}\|f\|_{W^{d+1,q}(\mathbb{R}^d;\mathbb{R}^{d'})} \end{aligned} \tag{2.8}$$

where C depends only on $d,\ d',\ T,\ \|\rho_0\|_{H^k(\mathbb{R}^d)}$ and $\|\Gamma\|_{W_0^{d+1,q'}(\mathbb{R}^d;\mathbb{R}^{d'})}$. Since $X_i(t)$ and $X(t;\alpha_i)$ have the same distribution, so we have

$$\mathbb{E}\left[f\left(X_i(t)\right)\right] = \mathbb{E}\left[f\left(X(t;\alpha_i)\right)\right], \tag{2.9}$$

which leads to our lemma. □

Next, we introduce the following concentration inequality, which is a reformation of the well-known Bennett's inequality. And it plays an very important role in the sequel analysis.

Lemma 2.5. *Let $\{Y_i\}_{i=1}^n$ be n independent bounded d-dimensional random vectors satisfying*

(i) $\mathbb{E}[Y_i] = 0$ and $|Y_i| \le M$ for all $i = 1,\cdots,n$;

(ii) $\sum_{i=1}^n Var(Y_i) \le V$ with $Var(Y_i) = \mathbb{E}[|Y_i|^2]$.

If $M \le C\frac{\sqrt{V}}{\eta}$ with some positive constant C, then we have

$$P\left(|\sum_{i=1}^n Y_i| \ge \eta\sqrt{V}\right) \le \exp\left(-C'\eta^2\right), \tag{2.10}$$

for all $\eta > 0$, where C' only depends on C and d.

Proof. See Pollard [24, Appendix B] for a proof of the Bennet's inequality in case $d = 1$, which leads to

$$P\left(|\sum_{i=1}^n Y_i| \ge \eta\sqrt{V}\right) \le 2\exp\left[-\frac{1}{2}\eta^2 B(M\eta V^{-\frac{1}{2}})\right], \tag{2.11}$$

where $B(\lambda) = 2\lambda^{-2}[(1+\lambda)\ln(1+\lambda) - \lambda],\ \lambda > 0,\ \lim_{\lambda\to 0^+} B(\lambda) = 1,\ \lim_{\lambda\to +\infty} B(\lambda) = 0$ and $B(\lambda)$ is decreasing in $(0, +\infty)$.

Since $M \le C\frac{\sqrt{V}}{\eta}$ and $B(\lambda)$ is decreasing, one concludes that

$$P\left(|\sum_{i=1}^{n} Y_i| \ge \eta\sqrt{V}\right) \le 2\exp\left[-\frac{1}{2}B(C)\eta^2\right]. \tag{2.12}$$

Denote $S = \sum_{i=1}^{n} Y_i$, then for d dimensional random vector Y_i, we have

$$\begin{aligned} P\left(|S| \ge \eta\sqrt{V}\right) &\le \sum_{j}^{d} P(|S_j| \ge \frac{\eta\sqrt{V}}{\sqrt{d}}) \\ &\le 2d\exp\left[-\frac{1}{2d}B(C)\eta^2\right] \le \exp\left[-C'\eta^2\right], \end{aligned} \tag{2.13}$$

where C' only depending on C and d. □

Lemma 2.6. *For $i, j \in I$, let $M_{ij}^{\ell} = \max_{|y|\le C_0\varepsilon} \max_{|\beta|=l} |\partial^\beta F_\varepsilon(X_i(t) - X_j(t) + y)|$ with some positive constant C_0. We take $\varepsilon \ge h|\ln h|^{\frac{2}{d}}$, then there exist two positive constants C, $C_\sharp$ depending on T, d and $\|\rho_0\|_{H^k(\mathbb{R}^d)}$, such that*

$$\begin{cases} P\left(\sum_{j\in I} M_{ij}^{\ell} h^d \ge \Lambda|\ln\varepsilon|\right) \le h^{C\Lambda|\ln h|}, & \text{for any } i \in I \text{ , if } \ell = 1, \\ P\left(\sum_{j\in I} M_{ij}^{\ell} h^d \ge \Lambda\varepsilon^{-1}\right) \le h^{C\Lambda|\ln h|}, & \text{for any } i \in I \text{ , if } \ell = 2, \end{cases} \tag{2.14}$$

holds true at any fixed time t for any $\Lambda > C_\sharp$.

This lemma can be obtained through the same approach as in [19, Lemma 9].

Lemma 2.7. [19, Lemma 10] *Assume $B(t)$ is a standard Brownian motion in $\mathbb{R}^d$. Then*

$$P\left\{\max_{t\le s\le t+\Delta t} |B(s) - B(t)| \ge b\right\} \le C(\sqrt{\Delta t}/b)\exp(-C'b^2/\Delta t),$$

where $b > 0$ and the positive constants C, C' depend only on d.

Proof. We give the proof of $d = 1$, then the case $d \ge 2$ can be obtained easily. See Freedman [9, P.18], then one has

$$P\left\{\max_{t\le s\le t+\Delta t} |B(s) - B(t)| \ge b\right\} \le 2P\left\{|B(\Delta t)| \ge b\right\} = 4P\left\{|B(1)| \ge b/\sqrt{\Delta t}\right\}. \tag{2.15}$$

Since $B(1) \sim N(0, 1)$, a simple computation leads to our lemma. □

Lastly, we introduce the following far field estimate:

Lemma 2.8. *Assume that $X_i(t)$ is the exact solution to (1.12), for R bigger than the diameter of D, then we have*

$$P(|X_i(t)| \ge R) \le \frac{C}{R^2}, \tag{2.16}$$

where C depends on d, T, R_0 and $\|\rho\|_{H^k(\mathbb{R}^d)}$.

Proof. Recall that $g(t, x \leftarrow 0, \alpha)$ is the solution to the following equation

$$\begin{cases} u_t = \nu\Delta u - \nabla\cdot(uG), \\ u(0,x) = \delta_\alpha(x), \ \ \alpha\in D. \end{cases} \tag{2.17}$$

We denote the second moment estimate of u as $m_2(t) = \int_{\mathbb{R}^d} |x|^2 u(t,x)dx$, then one has

$$\begin{aligned} \frac{dm_2(t)}{dt} &= 2d + 2\int_{\mathbb{R}^d}(x\cdot G)udx \qquad (2.18)\\ &=2d + 2C_*\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\frac{x\cdot(x-y)}{|x-y|^d}\rho(y)u(x)dxdy \\ &\le 2d + 2C_*\int_{\mathbb{R}^d}|x|u(x)\left[\int_{|x-y|\le 1}\frac{\rho(y)}{|x-y|^{d-1}}dy + \int_{|x-y|>1}\frac{\rho(y)}{|x-y|^{d-1}}dy\right]dx \\ &\le 2d + C(\|\rho\|_{L^1},\|\rho\|_{L^\infty})(\|u\|_{L^1}+m_2(t)), \quad t\in(0,T]. \end{aligned}$$

By using Gronwall's inequality, we have

$$m_2(t) \le e^{C_1 T}(m_2(0) + C_2 T), \quad t\in(0,T]. \tag{2.19}$$

Notice that $m_2(0) = \int_{\mathbb{R}^d}|x|^2\delta_\alpha(x)dx \le CR_0^2$, which leads to

$$m_2(t) \le C_1 R_0^2 + C_2, \quad t\in[0,T], \tag{2.20}$$

where D satisfies $D\subseteq B(R_0)$ and C_1, C_2 depend on $d,\ T,\ \|\rho\|_1,\ \|\rho\|_\infty$. Now, we compute $P(|X_i(t)|\ge R)$, and one has

$$P(|X_i(t)|\ge R) = \int_{|x|\ge R} g(t,x\leftarrow 0,\alpha_i)dx \le \frac{m_2(t)}{R^2}. \tag{2.21}$$

Thus, we conclude the proof. □

3. Consistency error at the fixed time

In this section, we will achieve the following consistency estimate result at any fixed time. Recall the definition of $G(t,x)$, $G_\varepsilon^h(t,x)$ in (1.23) and (1.24), then we have the result as below.

Theorem 3.1. *Assume that $X_i(t)$ is the exact path of (1.12). Under the same assumption as in Theorem 1.1, there exist two constants $C, C' > 0$ depending only on $T,\ d,\ R_0$ and $\|\rho_0\|_{H^k(\mathbb{R}^d)}$, such that at any fixed time $t\in[0,T]$, we have*

$$P\left(\max_{i\in I}\left|G_\varepsilon^h(t,X_i(t)) - G(t,X_i(t))\right| < \Lambda h|\ln h|\right) \ge 1 - h^{C\Lambda|\ln h|}, \tag{3.1}$$

for all $\Lambda > C'$.

Proof. For any fixed x and t, we decompose the consistency error into the sampling error, the discretization error and the moment error as follows

$$
\begin{aligned}
(3.2)\qquad |G_\varepsilon^h(t,x)-G(t,x)| &= \left|\sum_{j\in I} F_\varepsilon(x-X_j(t))\rho_j h^d - \int_{\mathbb{R}^d} F(x-y)\rho(t,y)dy\right| \\
&\le \left|\sum_{j\in I} F_\varepsilon(x-X_j(t))\rho_j h^d - \sum_{j\in I}\mathbb{E}[F_\varepsilon(x-X_j(t))]\rho_j h^d\right| \\
&\quad + \left|\sum_{j\in I}\mathbb{E}[F_\varepsilon(x-X_j(t))]\rho_j h^d - \int_{\mathbb{R}^d} F_\varepsilon(x-y)\rho(t,y)dy\right| \\
&\quad + \left|\int_{\mathbb{R}^d} F_\varepsilon(x-y)\rho(t,y)dy - \int_{\mathbb{R}^d} F(x-y)\rho(t,y)dy\right| \\
&=: e_s(t,x)+e_d(t,x)+e_m(t,x).
\end{aligned}
$$

Step 1 For the moment error, it can be proved that

$$
(3.3)\qquad e_m(t,x)\le C_1\varepsilon^2.
$$

Indeed, if we rewrite $e_m(t,x)=\left|\int_{\mathbb{R}^d}[F_\varepsilon(y)-F(y)]\rho(t,x-y)dy\right|$, and from (i) in Lemma 2.1, then one has

$$
\begin{aligned}
(3.4)\qquad e_m(t,x) &= \left|\int_{\mathbb{R}^d}[h(\frac{|y|}{\varepsilon})-1]F(y)\rho(t,x-y)dy\right| \\
&= \varepsilon\left|\int_{\mathbb{R}^d}[h(|z|)-1]F(z)\rho(t,x-\varepsilon z)dz\right| \\
&= \varepsilon\left|\int_{\mathbb{R}^d}[h(|z|)-1]F(z)[\rho(t,x-\varepsilon z)-\rho(t,x)]dz\right| \\
&\le C\varepsilon^2\|\nabla\rho\|_{L^\infty}\int_0^1 r^2|1-h(r)|dr \le C_1\varepsilon^2,
\end{aligned}
$$

where C_1 depending only on d, $\|\rho_0\|_{H^k(\mathbb{R}^d)}$.

Step 2 For the discretization error $e_d(t,x)$, we notice that

$$
\begin{aligned}
(3.5)\qquad \int_{\mathbb{R}^d} F_\varepsilon(x-y)\rho(t,y)dy &= \int_{\mathbb{R}^d} F_\varepsilon(x-y)\left[\int_D g(t,y;0,\alpha)\rho_0(\alpha)d\alpha\right]dy \\
&= \int_D\left[\int_{\mathbb{R}^d} F_\varepsilon(x-y)g(t,y;0,\alpha)dy\right]\rho_0(\alpha)d\alpha \\
&= \int_D \mathbb{E}\left[F_\varepsilon(x-X(t;\alpha))\right]\rho_0(\alpha)d\alpha.
\end{aligned}
$$

By applying Lemma 2.4 with $f(y)=F_\varepsilon(x-y),\Gamma(\alpha)=\rho_0(\alpha)$, we obtain

$$
\begin{aligned}
(3.6)\qquad &\left|\sum_{j\in I}\mathbb{E}[F_\varepsilon(x-X_j(t))]\rho_j h^d - \int_D \mathbb{E}\left[F_\varepsilon(x-X(t;\alpha))\right]\rho_0(\alpha)d\alpha\right| \\
&\le Ch^{d+1}\|F_\varepsilon\|_{W^{d+1,q}(\mathbb{R}^d)} \le C_2 h^{d+1}\varepsilon^{d/q-2d}.
\end{aligned}
$$

It follows from (3.5) and (3.6)that

$$
\begin{aligned}
e_d(t,x) &= \left| \sum_{j\in I} \mathbb{E}[F_\varepsilon(x - X_j(t))]\rho_j h^d - \int_{\mathbb{R}^d} F_\varepsilon(x-y)\rho(t,y)dy \right| \\
&\le C_2 h^{d+1}\varepsilon^{d/q-2d},
\end{aligned} \tag{3.7}
$$

where C_2 only depends on $T,\ d$ and $\|\rho_0\|_{H^k(\mathbb{R}^d)}$.

Step 3 For the sampling error $e_s(t,x)$, we will use Lemma 2.5 to give an estimate of $e_s(t,x) = |\sum_{j\in I} Y_j|$, where

$$
Y_j = \big(F_\varepsilon(x - X_j(t)) - \mathbb{E}[F_\varepsilon(x - X_j(t))]\big)\rho_j h^d. \tag{3.8}
$$

It is obvious that

$$
\mathbb{E}[Y_j] = 0 \text{ and } |Y_j| \le Ch^d\varepsilon^{1-d} =: M, \quad \text{for all } j \in I. \tag{3.9}
$$

Next, we will show that $\sum_{j\in I} \operatorname{Var} Y_j$ is uniformly bounded by some V. Actually,

$$
\begin{aligned}
\sum_{j\in I} \operatorname{Var} Y_j &= \sum_{j\in I} \left\{ \mathbb{E}[|F_\varepsilon(x - X_j(t))|^2] - |\mathbb{E}[F_\varepsilon(x - X_j(t))]|^2 \right\} \rho_j^2 h^{2d} \\
&\le \sum_{j\in I} \mathbb{E}[|F_\varepsilon(x - X_j(t))|^2]\rho_j^2 h^{2d}.
\end{aligned} \tag{3.10}
$$

We apply Lemma 2.4 again with $f(y) = |F_\varepsilon(x-y)|^2 = C|\partial^{d-1}F_\varepsilon(x-y)|, \Gamma(\alpha) = \rho_0(\alpha)^2$, then one has

$$
\begin{aligned}
&\left| \sum_{j\in I} \mathbb{E}[|F_\varepsilon(x - X_j(t))|^2]\rho_j^2 h^{2d} - h^d \int_D \mathbb{E}[|F_\varepsilon(x - X(t;\alpha))|^2]\rho_0(\alpha)^2 d\alpha \right| \\
&\le Ch^d \|F_\varepsilon\|_{W^{2d,q}(\mathbb{R}^d)} \le Ch^{d+1}\varepsilon^{d/q-3d+1}h^d,
\end{aligned} \tag{3.11}
$$

which follows from Lemma 2.2 as we have done in (3.7).

Notice that

$$
\begin{aligned}
&\int_D \mathbb{E}[|F_\varepsilon(x - X(t;\alpha))|^2]\rho_0(\alpha)^2 d\alpha \\
&= \int_D \int_{\mathbb{R}^d} |F_\varepsilon(x-y)|^2 g(t, y \leftarrow 0, \alpha)\rho_0(\alpha)^2 dy d\alpha \\
&= \int_{\mathbb{R}^d} |F_\varepsilon(x-y)|^2 \left\{ \int_D g(t, y \leftarrow 0, \alpha)\rho_0(\alpha)^2 d\alpha \right\} dy,
\end{aligned} \tag{3.12}
$$

where $g(t, y \leftarrow 0, \alpha)$ is the Green's function. Notice that $u(t,x) := \int_D g(t, x \leftarrow 0, \alpha)\rho_0(\alpha)^2 d\alpha$ is the solution of the following equation

$$
\begin{cases}
\partial_t u = \nu\Delta u - \nabla\cdot(u\nabla c), \quad x\in\mathbb{R}^d,\ t>0, \\
-\Delta c = u(t,x), \\
u(0,x) = \rho_0^2(x).
\end{cases} \tag{3.13}
$$

So the L^∞ norm of u are bounded by $\|\rho_0\|_{H^k}$. Therefore, using Lemma 2.2, one has that (3.12) is bounded by

$$C\int_{\mathbb{R}^d}|F_\varepsilon(x-y)|^2dy = C\|F_\varepsilon\|_2^2 \le C\varepsilon^{2-d}. \tag{3.14}$$

Collecting (3.10), (3.11) and (3.14), we have

$$\begin{aligned}\sum_{j\in I}\text{Var } Y_j &\le Ch^{d+1}\varepsilon^{d/q-3d+1}h^d + Ch^d\varepsilon^{2-d} \\ &\le Ch^{d+\frac{q(2-d)}{2q-1}} =: V \quad (\text{by } \varepsilon = h^{\frac{q}{2q-1}}),\end{aligned} \tag{3.15}$$

where the constant C depends only on $T,\ d$ and $\|\rho_0\|_{H^k(\mathbb{R}^d)}$.

For any $C_3>0$, we let $\eta = C_3|\ln h|$ in Lemma 2.5. In order to use Lemma 2.5, we need to verify that $M \le C\frac{\sqrt{V}}{\eta}$, which leads to

$$h^{\frac{3qd-d-2q}{2(2q-1)(d-1)}}(|\ln h|)^{\frac{1}{d-1}} \le \varepsilon. \tag{3.16}$$

Since we choose $\varepsilon = h^{\frac{q}{2q-1}}$ with $q>1$, (3.16) can be verified when h is sufficiently small. Hence , it follows from the concentration inequality (2.10) that

$$P\left(e_s(t,x) \ge C_3|\ln h|\sqrt{V}\right) \le \exp[-C'C_3^2|\ln h|^2] \le h^{C''C_3|\ln h|}, \tag{3.17}$$

for some $C''>0$ depending only on $T,\ d$ and $\|\rho_0\|_{H^k(\mathbb{R}^d)}$.

Step 4 We take $\varepsilon = h^{\frac{q}{2q-1}}$ with any $q>1$ and that h is sufficiently small, then it has

$$C_1\varepsilon^2 + C_2h^{d+1}\varepsilon^{d/q-2d} + C_3|\ln h|\sqrt{V} < C_4h|\ln h|, \tag{3.18}$$

where C_4 is bigger than a positive constant depending only on $T,\ d$ and $\|\rho_0\|_{H^k(\mathbb{R}^d)}$. In summary, at any fixed x and t, we have

$$\begin{aligned}&P(|G_\varepsilon^h(t,x)-G(t,x)| \ge 3C_4h|\ln h|) \\ \le& P(e_m(t,x)+e_d(t,x)+e_s(t,x) \ge 3C_4h|\ln h|) \\ \le& P(e_m(t,x)\ge C_4h|\ln h|) + P(e_d(t,x)\ge C_4h|\ln h|) + P(e_s(t,x)\ge C_4h|\ln h|) \\ \le& 0+0+P\left(e_s(t,x)\ge C_4h|\ln h|\right) \le h^{C'''C_4|\ln h|},\end{aligned} \tag{3.19}$$

with some $C'''>0$.

Step 5 For the lattice points $z_k = hk$ in ball $B(R)$ with $R = h^{-\gamma|\ln h|}$ (γ will be determined later), it follows from inequality (3.19) that

$$\begin{aligned}&P\left(\max_k |G_\varepsilon^h(t,z_k)-G(t,z_k)| \ge 3C_4h|\ln h|\right) \\ \le& \sum_k P(|G_\varepsilon^h(t,z_k)-G(t,z_k)| \ge 3C_4h|\ln h|) \\ \le& C''''h^{-(1+\gamma|\ln h|)d}h^{C'''C_4|\ln h|} = C''''h^{-d+(C'''C_4-\gamma d)|\ln h|},\end{aligned} \tag{3.20}$$

which leads to

$$P\left(\max_k |G_\varepsilon^h(t,z_k)-G(t,z_k)| \ge C_4'h|\ln h|\right) \le h^{CC_4'|\ln h|}, \tag{3.21}$$

with some constant $C > 0$ provided that $C'''C_4 - \gamma d > 0$.

Step 6 For any fixed t, denote the event $U := \{X_i(t) \in B(R)\}$, then we know from Lemma 2.8 that $P(U^c) \leq \frac{C}{R^2} = Ch^{2\gamma|\ln h|}$. Now, we do the estimate under event U, and suppose z_i is the closet lattice point to $X_i(t)$ with $|X_i(t) - z_i| \leq h$.

We compute

$$
\begin{aligned}
(3.22)\quad & \left|G_\varepsilon^h(t, X_i(t)) - G(t, X_i(t))\right| \\
& \leq \left|G_\varepsilon^h(t, X_i(t)) - G_\varepsilon^h(t, z_i)\right| + \left|G_\varepsilon^h(t, z_i) - G(t, z_i)\right| + |G(t, z_i) - G(t, X_i(t))| \\
& =: I_1 + I_2 + I_3.
\end{aligned}
$$

Then $P\left(I_2 \geq C_4' h|\ln h|\right) \leq h^{CC_4'|\ln h|}$ follows from (3.21).

For I_1, we have

$$
\begin{aligned}
(3.23)\quad I_1 &= \left|\sum_{j\in I} F_\varepsilon(X_i(t) - X_j(t))\rho_j h^d - \sum_{j\in I} F_\varepsilon(z_i - X_j(t))\rho_j h^d\right| \\
&= \left|\sum_{j\in I} \nabla F_\varepsilon(X_i(t) - X_j(t) + \xi)\rho_j h^d\right| |X_i(t) - z_i| \leq C\sum_{j\in I} M_{ij}^1 h^d h,
\end{aligned}
$$

by applying mean-value theorem. We take $\varepsilon = h^{\frac{q}{2q-1}}$ with any $q > 1$ and that h is sufficiently small, which make sure $\varepsilon \geq h|\ln h|^{\frac{2}{d}}$. Apply Lemma 2.6 for any $C_5 > CC_\sharp$, one has

$$
(3.24)\qquad P(I_1 \geq C_5 h|\ln\varepsilon|) \leq P(\sum_{j\in I} M_{ij}^1 h^d \geq \frac{C_5}{C}|\ln\varepsilon|) \leq h^{CC_5|\ln h|},
$$

which leads to

$$
(3.25)\qquad P(I_1 \geq C_5 h|\ln h|) \leq h^{CC_5|\ln h|} \quad (\text{by } h \leq \varepsilon),
$$

with some $C > 0$.

For I_3, since $G(t, x)$ is smooth enough, one has

$$
(3.26)\qquad I_3 = |\nabla G(t, z_i + \xi)\|\, z_i - X_i(t)| \leq C_6 h \leq C_6 h|\ln h|,
$$

by using mean-value theorem.

Take $C_7 > \max\{C_4', C_5, C_6\}$, and we collect the estimates of I_1, I_2 and I_3, then one has

$$
\begin{aligned}
(3.27)\qquad & P\left(\max_{i\in I}\left|G_\varepsilon^h(t, X_i(t)) - G(t, X_i(t))\right| \geq 3C_7 h|\ln h|\right) \\
\leq & NP\left(I_1 + I_2 + I_3 \geq 3C_7 h|\ln h|\right) \\
\leq & Nh^{CC_7|\ln h|} + Nh^{CC_7|\ln h|} + 0 \leq h^{CC_7|\ln h|},
\end{aligned}
$$

with some $C > 0$ and that C_7 is bigger than a positive constant depending only on T, d and $\|\rho_0\|_{H^k(\mathbb{R}^d)}$.

Until now, we have only proved that

$$
(3.28)\quad P\left(\left\{\max_{i\in I}\left|G_\varepsilon^h(t, X_i(t)) - G(t, X_i(t))\right| \geq C_7' h|\ln h|\right\} \cap U\right) \leq h^{CC_7'|\ln h|}.
$$

Hence, we have

$$P\left(\max_{i\in I}\left|G_\varepsilon^h(t,X_i(t))-G(t,X_i(t))\right|\geq C_7'h|\ln h|\right)\tag{3.29}$$
$$\leq h^{CC_7'|\ln h|}+P(U^c)\leq h^{CC_7'|\ln h|}+Ch^{2\gamma|\ln h|}\leq h^{CC_7'|\ln h|}.$$

Finally, we concludes the proof of this theorem by using $P(A^c)=1-P(A)$. $\square$

4. Stability estimate

In this section, we will focus on giving a proof of the stability estimate, which can be expressed as follows.

Theorem 4.1. *(Stability) Under the same assumption as in Theorem 1.1. Assume*

$$\max_{0\leq t\leq T}\max_{i\in I}|X_{i,\varepsilon}(t)-X_i(t)|\leq\varepsilon,\tag{4.1}$$

then there exist two positive constants C, C' depending only on T, p, d, R_0 and $\|\rho_0\|_{H^k(\mathbb{R}^d)}$, such that for any $\Lambda>C'$, the following stability estimate holds

$$P\left(\|\hat{G}_\varepsilon^h(t,X_{h,\varepsilon}(t))-G_\varepsilon^h(t,X_h(t))\|_{\ell_h^p}<\Lambda\|X_{h,\varepsilon}(t)-X_h(t)\|_{\ell_h^p},\ \forall\ t\in[0,T]\right)\tag{4.2}$$
$$\geq 1-h^{C\Lambda|\ln h|},$$

where G_ε^h, $\hat{G}_\varepsilon^h$ are defined in (1.24) and (1.25).

Proof. In order to prove (4.2), we divide $[0,T]$ into N' subintervals with length $\Delta t=h^r$ for some $r>2$ and $t_n=nh^r$, $n=0,\ldots,N'$. If we denote the following events

$$A_n:=\left\{\|\hat{G}_\varepsilon^h(t,X_{h,\varepsilon}(t))-G_\varepsilon^h(t,X_h(t))\|_{\ell_h^p}\geq\Lambda\|X_{h,\varepsilon}(t)-X_h(t)\|_{\ell_h^p},\ \exists\ t\in[t_n,t_{n+1}]\right\},\tag{4.3}$$

$$\tilde{A}:=\left\{\|\hat{G}_\varepsilon^h(t,X_{h,\varepsilon}(t))-G_\varepsilon^h(t,X_h(t))\|_{\ell_h^p}\geq\Lambda\|X_{h,\varepsilon}(t)-X_h(t)\|_{\ell_h^p},\ \exists\ t\in[0,T]\right\},\tag{4.4}$$

then, one has

$$P\left(\tilde{A}\right)=P\left(\bigcup_{n=0}^{N'-1}A_n\right).\tag{4.5}$$

So our main idea of this proof is to give the estimate of $P(A_n)$ first.

Directly, we apply Lemma 2.7 and get

$$P\left(\max_n\max_{t_n\leq t\leq t_{n+1}}|X_i(t)-X_i(t_n)|\geq Ch^r+\sqrt{2\nu}h\right)\tag{4.6}$$
$$\leq C'h^{r/2-1}\exp(-C''h^{2-r})\to 0,$$

which leads to

$$P\left(\max_n\max_{t_n\leq t\leq t_{n+1}}|X_i(t)-X_i(t_n)|\geq\varepsilon\right)\leq C'h^{r/2-1}\exp(-C''h^{2-r})\to 0,\tag{4.7}$$

provided that

$$Ch^r+\sqrt{2\nu}h\leq\varepsilon.\tag{4.8}$$

Again, (4.8) can be verified by our choice of $\varepsilon = h^{\frac{q}{2q-1}}$ with h sufficiently small. Actually, (4.7) makes sure that the position $X_i(t)$ for $t \in [t_n, t_{n+1}]$ are close to $X_i(t_n)$.

For $t \in [t_n, t_{n+1}]$, recalling the definition of drift term (1.24) and (1.25), we write

$$
\begin{aligned}
(4.9)\quad & \hat{G}^h_\varepsilon(t, X_{i,\varepsilon}(t)) - G^h_\varepsilon(t, X_i(t)) \\
&= \sum_{j\in I} \left[F_\varepsilon\big(X_{i,\varepsilon}(t) - X_{j,\varepsilon}(t)\big) - F_\varepsilon\big(X_i(t) - X_j(t)\big)\right] \rho_j h^d \\
&= \sum_{j\in I} \nabla F_\varepsilon\big(X_i(t_n) - X_j(t_n) + \xi_{ij}\big) \cdot \big(X_{i,\varepsilon}(t) - X_i(t) + X_j(t) - X_{j,\varepsilon}(t)\big)\rho_j h^d \\
&= \sum_{j\in I} \nabla F_\varepsilon\big(X_i(t_n) - X_j(t_n) + \xi_{ij}\big) \cdot \big(X_{i,\varepsilon}(t) - X_i(t)\big)\rho_j h^d \\
&\quad + \sum_{j\in I} \nabla F_\varepsilon\big(X_i(t_n) - X_j(t_n) + \xi_{ij}\big) \cdot \big(X_j(t) - X_{j,\varepsilon}(t)\big)\rho_j h^d \\
&=: \mathcal{I}_i + \mathcal{J}_i,
\end{aligned}
$$

where the term ξ_{ij} is from mean-value theorem, which may depend on the components $X_{i,\varepsilon}(t)$,$X_{j,\varepsilon}(t)$, $X_i(t)$, $X_j(t)$, $X_i(t_n)$, $X_j(t_n)$. Furthermore, combining (4.1) and (4.7), one has

$$
P(|\xi_{ij}| \geq 4\varepsilon,\ \exists\, t \in [t_n, t_{n+1}]) \leq C' h^{r/2-1} \exp(-C'' h^{2-r}). \tag{4.10}
$$

We will give the estimates of $\mathcal{I}_i$ and $\mathcal{J}_i$ under the event $A := \{\xi_{ij} : |\xi_{ij}| < 4\varepsilon,\ \forall\, t \in [t_n, t_{n+1}]\}$ in the following steps 1-2.

Step 1(Estimate of $\mathcal{I}_i$) In order to do the estimate of $\mathcal{I}_i$, we need to give the uniform bound of $\sum_{j\in I} \nabla F_\varepsilon\big(X_i(t_n) - X_j(t_n) + \xi_{ij}\big)\rho_j h^d$ first. To do that, we are required to prove the uniform bound of $\sum_{j\in I} \nabla F_\varepsilon\big(X_i(t_n) - X_j(t_n)\big)\rho_j h^d$.

We may write

$$
\begin{aligned}
\sum_{j\in I} \nabla F_\varepsilon\big(x - X_j(t_n)\big)\rho_j h^d &= \sum_{j} \mathbb{E}[\nabla F_\varepsilon\big(x - X_j(t_n)\big)]\rho_j h^d \\
&\quad + \sum_{j\in I} \left[\nabla F_\varepsilon\big(x - X_j(t_n)\big) - \mathbb{E}[\nabla F_\varepsilon\big(x - X_j(t_n)\big)]\right] \rho_j h^d \\
&=: I_1 + I_2.
\end{aligned}
$$

For I_1, it can be estimated by Lemma 2.4 with $f(y) = \nabla F_\varepsilon(x-y)$, $\Gamma(\alpha) = \rho_0(\alpha)$ as we have done before

$$
\begin{aligned}
(4.11)\qquad & \left| I_1 - \int_{\mathbb{R}^d} \mathbb{E}[\nabla F_\varepsilon\big(x - X(t_n;\alpha)\big)]\rho_0(\alpha) d\alpha \right| \\
\leq & \|F_\varepsilon\|_{W^{d+2,q}(\mathbb{R}^d)} \leq C h^{d+1} \varepsilon^{d/q-2d-1} \leq C \quad (\text{ by } \varepsilon = h^{\frac{q}{2q-1}}),
\end{aligned}
$$

where C depends on $T,\ d,\ \|\rho_0\|_{H^k(\mathbb{R}^d)}$. On the other hand, we notice that

$$\begin{aligned}(4.12)\qquad &\left|\int_{\mathbb{R}^d}\mathbb{E}[\nabla F_\varepsilon\big(x-X(t_n;\alpha)\big)]\rho_0(\alpha)d\alpha\right|\\ =&\left|\int_{\mathbb{R}^d}F_\varepsilon(x-y)\nabla\rho(t_n,y)dy\right|\\ \leq&\|\nabla\rho\|_{L^\infty}\|F_\varepsilon\|_{L^1(B)}+\|\nabla\rho\|_{L^1}\|F_\varepsilon\|_{L^\infty(\mathbb{R}^d/B)}\leq C,\end{aligned}$$

where $\|\nabla\rho\|_{L^1}\leq C(T,d,R_0,\|\rho_0\|_{H^k(\mathbb{R}^d)})$ has been used and B is the unit ball in $\mathbb{R}^d$. Actually, the proof of the estimate of $\|\nabla\rho\|_{L^1}$ can be done by using the standard semigroup method. We recall the heat semigroup operator $e^{t\Delta}$ defined by

$$e^{t\Delta}\rho:=H(t,x)*\rho, \tag{4.13}$$

where $H(t,x)=\frac{1}{(4\pi t)^{d/2}}e^{-\frac{|x|^2}{4t}}$ is the heat kernel. Then the solution to KS equation (1.1) can be represented as

$$\rho=e^{t\Delta}\rho_0+\int_0^t e^{(t-s)\Delta}(-\nabla\cdot(\rho\nabla c))ds. \tag{4.14}$$

A simple computation leads to

$$\|\nabla\rho\|_{L^1}\leq C\|\nabla\rho_0\|_{L^1}+C(T)\|\nabla\cdot(\rho\nabla c)\|_{L^\infty(0,T;L^1(\mathbb{R}^d))}. \tag{4.15}$$

Further more, one has

$$\begin{aligned}(4.16)\qquad \|\nabla\cdot(\rho\nabla c)\|_{L^1}&=\|\nabla\rho\cdot\nabla c-\rho^2\|_{L^1}\leq\|\nabla\rho\|_{L^2}\|\nabla c\|_{L^2}+\|\rho\|_{L^2}^2\\ &\leq C(d)\|\nabla\rho\|_{L^2}\|\rho\|_{L^{\frac{2d}{d+2}}}+\|\rho\|_{L^2}^2\leq C(d,\|\rho_0\|_{H^k}).\end{aligned}$$

Thus we have $\|\nabla\rho\|_{L^1}\leq C(T,d,R_0,\|\rho_0\|_{H^k(\mathbb{R}^d)})$. Combining (4.11) and (4.12), one concludes that

$$|I_1|\leq C_1(T,\ d,\ R_0,\ \|\rho_0\|_{H^k(\mathbb{R}^d)}). \tag{4.17}$$

To estimate I_2, let $I_2=\sum\limits_{j\in I}Y_j$

$$Y_j=\left[\nabla F_\varepsilon\big(x-X_j(t_n)\big)-\mathbb{E}[\nabla F_\varepsilon\big(x-X_j(t_n)\big)]\right]\rho_j h^d. \tag{4.18}$$

We have $\mathbb{E}[Y_j]=0$, $|Y_j|\leq Ch^d\varepsilon^{-d}\leq C|\ln h|^{-2}:=M$ provided that

$$h|\ln h|^{\frac{2}{d}}\leq\varepsilon. \tag{4.19}$$

Indeed, (4.19) can be verified since we choose $\varepsilon=h^{\frac{q}{2q-1}}$ with $1<q$ and sufficiently small h.

Furthermore,

$$\sum_{j\in I}\mathrm{Var}Y_j\leq\sum_{j\in I}\mathbb{E}\left[|\nabla F_\varepsilon\big(x-X_j(t_n)\big)|^2\right]\rho_j^2h^{2d}. \tag{4.20}$$

We once again apply Lemma 2.4 with $f(y)=|\nabla F_\varepsilon(x-y)|^2=C|\partial^{d+1}F_\varepsilon(x-y)|$, $\Gamma(\alpha)=\rho_0(\alpha)^2$

(4.21)

$$\left|\sum_{j\in I}\mathbb{E}\left[|\nabla F_\varepsilon\big(x-X_j(t_n)\big)|^2\right]\rho_j^2h^{2d}-h^d\int_{\mathbb{R}^d}\mathbb{E}\left[|\nabla F_\varepsilon\big(x-X(t_n;\alpha)\big)|^2\right]\rho_0(\alpha)^2d\alpha\right|$$

$$\leq Ch^dh^{d+1}\varepsilon^{d/q-3d-1}\leq C|\ln h|^{-2}\quad(\text{ by }\varepsilon=h^{\frac{q}{2q-1}}),$$

where the constant C depends only on $T,\ d$ and $\|\rho_0\|_{H^k(\mathbb{R}^d)}$.

On the other hand, as we have done in (3.12) and (3.14), then we have

$$\left|h^d\int_{\mathbb{R}^d}\mathbb{E}\left[|\nabla F_\varepsilon\big(x-X(t_n;\alpha)\big)|^2\right]\rho_0(\alpha)^2d\alpha\right|\le Ch^d\varepsilon^{-d}\le C|\ln h|^{-2}. \tag{4.22}$$

Hence, one has $\sum_j \mathrm{Var}Y_j\le C|\ln h|^{-2}=:V$.

For any $C_2>0$, we choose $\eta=C_2|\ln h|$ in Lemma 2.5. It is easy to check that

$$M=C|\ln h|^{-2}\le C\frac{\sqrt{V}}{\eta}. \tag{4.23}$$

Thus we can use Lemma 2.5 now, for any $C_2>0$

$$\begin{aligned}&P\left(|I_2|\ge C_2|\ln h|\sqrt{C}|\ln h|^{-1}\right)\\ =&P\left(|I_2|\ge C_2\sqrt{C}\right)\le\exp\left\{-C'C_2^2|\ln h|^2\right\}\le h^{C''C_2|\ln h|}.\end{aligned} \tag{4.24}$$

We take $C_3>C_1+C_2\sqrt{C}$, thus we have

$$P\left(\left|\sum_{j\in I}\nabla F_\varepsilon\big(x-X_j(t_n)\big)\rho_jh^d\right|\ge 2C_3\right)\le P\left(|I_1|+|I_2|\ge 2C_3\right)\le h^{C'''C_3|\ln h|}, \tag{4.25}$$

with some $C'''>0$. Hence, at the fixed time t_n,

$$\left|\sum_{j\in I}\nabla F_\varepsilon\big(X_i(t_n)-X_j(t_n)\big)\rho_jh^d\right|<C_3',$$

except for an event of probability less than $h^{CC_3'|\ln h|}$ with C_3' bigger than a positive constant depending only on $T,\ d,\ R_0$ and $\|\rho_0\|_{H^k(\mathbb{R}^d)}$.

Notice that

$$\begin{aligned}&\left|\sum_{j\in I}\left[\nabla F_\varepsilon\big(X_i(t_n)-X_j(t_n)+\xi_{ij}\big)-\nabla F_\varepsilon\big(X_i(t_n)-X_j(t_n)\big)\right]\rho_jh^d\right|\\ \le&\varepsilon C''''\sum_{j\in I}M_{ij}^2h^d.\end{aligned} \tag{4.26}$$

So one has

$$\begin{aligned}&P\left(\left|\sum_{j\in I}\nabla F_\varepsilon\big(X_i(t_n)-X_j(t_n)+\xi_{ij}\big)\rho_jh^d\right|\ge 2C_3'\right)\\ \le&P\left(\left|\sum_{j\in I}\nabla F_\varepsilon\big(X_i(t_n)-X_j(t_n)\big)\rho_jh^d\right|+\varepsilon C''''\sum_{j\in I}M_{ij}^2h^d\ge 2C_3'\right)\\ \le&P\left(\left|\sum_{j\in I}\nabla F_\varepsilon\big(X_i(t_n)-X_j(t_n)\big)\rho_jh^d\right|\ge C_3'\right)+P\left(\varepsilon C''''\sum_{j\in I}M_{ij}^2h^d\ge C_3'\right)\\ \le&h^{CC_3'|\ln h|},\end{aligned} \tag{4.27}$$

where Lemma 2.6 have been used in the last inequality since we can choose $C_3' > C''''C_\sharp$.

Recall $\mathcal{I}_i = \sum_{j\in I} \nabla F_\varepsilon\big(X_i(t_n)-X_j(t_n)+\xi_{ij}\big)\cdot\big(X_{i,\varepsilon}(t)-X_i(t)\big)\rho_j h^d$, hence it follows from (4.27) that

$$P\left(\|(\mathcal{I}_i)_{i\in I}\|_{\ell_h^p} \geq C_3''\|X_{h,\varepsilon}(t)-X_h(t)\|_{\ell_h^p},\ \exists\, t\in[t_n,t_{n+1}]\right) \leq h^{CC_3''|\ln h|}, \tag{4.28}$$

with some $C>0$ and C_3'' bigger than a positive constant depending only on $T,\ d,\ R_0$ and $\|\rho_0\|_{H^k(\mathbb{R}^d)}$.

Step 2(Estimate of $\mathcal{J}_i$) At the fixed time t_n, let $Z_i\in\varepsilon\cdot\mathbb{Z}^d$ be the closest lattice point to $X_i(t_n)$. If there is more than one lattice point closest to $X_i(t_n)$, then we chose an arbitrary one. We write

$$\begin{aligned}\mathcal{J}_i =& \sum_{j\in I}\nabla F_\varepsilon(Z_i-Z_j)\cdot e_j\rho_j h^d\\ &+\sum_{j\in I}[\nabla F_\varepsilon\big(X_i(t_n)-X_j(t_n)+\xi_{ij}\big)-\nabla F_\varepsilon(Z_i-Z_j)]\cdot e_j\rho_j h^d\\ =:& \mathcal{J}_{1i}+\mathcal{J}_{2i},\end{aligned} \tag{4.29}$$

where $e_j = X_j(t)-X_{j,\varepsilon}(t)$. For each $z_k=\varepsilon k,\ k\in\mathbb{Z}^d$, we define f_k to be the average of all $e_j\rho_j h^d$ where $X_j(t_n)$ is in the square $Q_k = z_k+[-\frac{\varepsilon}{2},\frac{\varepsilon}{2}]^d$. Namely,

$$f_k = \varepsilon^{-d}\sum_{X_j(t_n)\in Q_k} e_j\rho_j h^d, \tag{4.30}$$

with convention of $f_k=0$ if Q_k contains none of the $X_j(t_n)$. Then one has

$$\|(f_k)_{k\in\mathbb{Z}^d}\|_{\ell_\varepsilon^p} \leq C_p\|(e_j\rho_j)_{j\in I}\|_{\ell_h^p}, \tag{4.31}$$

$$\begin{aligned}&P\left(\|(\mathcal{J}_{1i})_{i\in I}\|_{\ell_h^p}\geq C_4\|(\sum_k\nabla F_\varepsilon(z_{k'}-z_k)\cdot f_k\varepsilon^d)_{k'\in\mathbb{Z}^d}\|_{\ell_\varepsilon^p},\ \exists\, t\in[t_n,t_{n+1}]\right)\\ &\leq h^{CC_4|\ln h|},\end{aligned} \tag{4.32}$$

for some $C>0$ and C_4 is bigger than a positive constant depending on $T,\ p,\ d$ and $\|\rho_0\|_{H^k(\mathbb{R}^d)}$. The derivation of these two results can be achieved by the argument in [19, P.797]. In addition, it follows from Beale [3, P.47-48] that

$$\|(\sum_k\nabla F_\varepsilon(z_{k'}-z_k)\cdot f_k\varepsilon^d)_{k'\in\mathbb{Z}^d}\|_{\ell_\varepsilon^p}\leq C\|(f_k)_{k\in\mathbb{Z}^d}\|_{\ell_\varepsilon^p}, \tag{4.33}$$

which implies

$$P\left(\|(\mathcal{J}_{1i})_{i\in I}\|_{\ell_h^p}\geq C_4\|(e_j\rho_j)_{j\in I}\|_{\ell_h^p},\ \exists\, t\in[t_n,t_{n+1}]\right)\leq h^{CC_4|\ln h|}. \tag{4.34}$$

For $\mathcal{J}_{2i}$, we use mean-value theorem again

$$\mathcal{J}_{2i} = \sum_{j\in I}[\nabla^2 F_\varepsilon\big(X_i(t_n)-X_j(t_n)+\xi_{ij}+\xi_{ij}'\big)\cdot\xi_{ij}'']\cdot e_j\rho_j h^d, \tag{4.35}$$

where $\xi''_{ij} = \xi_{ij} + (X_i(t_n) - Z_i) - (X_j(t_n) - Z_j)$. Since $|\xi_{ij}| \le \varepsilon$, $|\xi'_{ij}| \le |\xi''_{ij}| \le 3\varepsilon$, then one has

$$|\mathcal{J}_{2i}| \le \sum_{j\in I} 3M_{ij}^2 \varepsilon |e_j \rho_j| h^d. \tag{4.36}$$

Applying the discrete version of Young's inequality, we conclude that

$$\|(\mathcal{J}_{2i})_{i\in I}\|_{\ell_h^p} \le 3\varepsilon \sum_{j\in I} M_{ij}^2 h^d \|(e_j\rho_j)_{j\in I}\|_{\ell_h^p}. \tag{4.37}$$

By Lemma 2.6 with $C_0 = 4$, one has

$$P\left(\|(\mathcal{J}_{2i})_{i\in I}\|_{\ell_h^p} \ge C_5 \|(e_j\rho_j)_{j\in I}\|_{\ell_h^p},\ \exists\, t \in [t_n, t_{n+1}]\right) \le h^{CC_5|\ln h|}, \tag{4.38}$$

with any $C_5 > 3C_\sharp$.

Recall (4.34) and (4.37), then we have

$$P\left(\|(\mathcal{J}_{i})_{i\in I}\|_{\ell_h^p} \ge 2C_6 \|(e_j\rho_j)_{j\in I}\|_{\ell_h^p},\ \exists\, t \in [t_n, t_{n+1}]\right) \le h^{CC_6|\ln h|}, \tag{4.39}$$

for any $C_6 > C_4 + C_5$.

Step 3 Collecting the estimate of $\mathcal{I}_i$ (4.28), the estimate of $\mathcal{J}_i$ (4.39) and the definition of event A_n (4.3), one concludes that

$$P(A_n \cap A) \le h^{C\Lambda|\ln h|}, \quad n = 0, \cdots, N'-1, \tag{4.40}$$

for some $C > 0$ and Λ bigger than a positive constant depending on T, p, d, R_0 and $\|\rho_0\|_{H^k(\mathbb{R}^d)}$. Since (4.5) and (4.40), we have

(4.41)
$$P\left(\tilde{A} \cap A\right) = P\left(\bigcup_{n=0}^{N'-1} (A_n \cap A)\right) \le N' h^{C\Lambda|\ln h|} = C' h^{-r} h^{C\Lambda|\ln h|} \le h^{C\Lambda|\ln h|},$$

for some $C > 0$. Finally, we have

(4.42)
$$P\left(\tilde{A}\right) \le P\left(\tilde{A} \cap A\right) + P(A^c) \le h^{C\Lambda|\ln h|} + C' h^{r/2-1} \exp(-C'' h^{2-r}) \le h^{C\Lambda|\ln h|},$$

for some $C > 0$ and Λ is bigger than a positive constant depending on T, p, d, R_0 and $\|\rho_0\|_{H^k(\mathbb{R}^d)}$. Until now, the proof of the stability theorem can be completed, since $P\left(\tilde{A}^c\right) = 1 - P\left(\tilde{A}\right)$. □

5. The convergence analysis and the proof of Theorem 1.1

In order to prove the convergence of particle paths, we need to extend the consistency error to all time. It can be obtained by combining the consistency estimates for a finite number of times $0 = t_0 < t_1 < \cdots < t_{N'} = T$ where $\Delta t = h^r$ with $r > 2$. We denote the following events

$$A_1^n : \left\{ \max_{i\in I} \left| G_\varepsilon^h(t_n, X_i(t_n)) - G(t_n, X_i(t_n)) \right| < \Lambda_1 h |\ln h| \right\}; \tag{5.1}$$

$$A_2 : \left\{ \max_{n} \max_{t_n \le t \le t_{n+1}} |X_i(t) - X_i(t_n)| < C(h^r + \nu^{1/2} h) \right\}; \tag{5.2}$$

(5.3)
$$A_3 : \left\{ \|\hat{G}^h_\varepsilon(t, X_{h,\varepsilon}(t)) - G^h_\varepsilon(t, X_h(t))\|_{\ell^p_h} < \Lambda_3 \|X_{h,\varepsilon}(t) - X_h(t)\|_{\ell^p_h},\ \forall\, t \in [0,T] \right\},$$

with that Λ_1, Λ_3 are bigger than a constant depending only on T, p, d, R_0 and $\|\rho_0\|_{H^k(\mathbb{R}^d)}$. For all $t \in [t_n, t_{n+1}]$, under the event $A_1^n \cap A_2$, we obtain

$$\begin{aligned}
&\|G^h_\varepsilon(t, X_h(t)) - G(t, X_h(t))\|_{\ell^p_h} \\
\le& \|G^h_\varepsilon(t, X_h(t)) - G^h_\varepsilon(t_n, X_h(t_n))\|_{\ell^p_h} \\
&+ \|G^h_\varepsilon(t_n, X_h(t_n)) - G(t_n, X_h(t_n))\|_{\ell^p_h} + \|G(t_n, X_h(t_n)) - G(t, X_h(t))\|_{\ell^p_h} \\
<& C\|X_h(t) - X_h(t_n)\|_{\ell^p_h} + Ch^r + \Lambda_1 h|\ln h| \\
<& (C + \Lambda_1) h |\ln h|,
\end{aligned} \tag{5.4}$$

by the estimate $|X_i(t) - X_i(t_n)| \le C(h^r + \nu^{1/2} h)$ and the fact that $G(t,x)$ has bounded derivatives. Therefore

$$\max_{0\le t\le T} \|G^h_\varepsilon(t, X_h(t)) - G(t, X_h(t))\|_{\ell^p_h} < (C + \Lambda_1) h|\ln h|, \tag{5.5}$$

under the event $\bigcap_{n=0}^{N'} A_1^n \cap A_2$.

The convergence can be proved by the same argument as in [2, 3]. Denote $e_i(t) = X_{i,\varepsilon}(t) - X_i(t)$ and vector $e(t) = (e_i)_{i\in I} = X_{h,\varepsilon}(t) - X_h(t)$. One has

$$\frac{de_i}{dt} = \hat{G}^h_\varepsilon(t, X_{i,\varepsilon}(t)) - G(t, X_i(t)) \tag{5.6}$$

and the differential inequality

(5.7)
$$\begin{aligned}
\|\frac{de}{dt}\|_{\ell^p_h} \le& \|\hat{G}^h_\varepsilon(t, X_{h,\varepsilon}(t)) - G^h_\varepsilon(t, X_h(t))\|_{\ell^p_h} + \|G^h_\varepsilon(t, X_h(t)) - G(t, X_h(t))\|_{\ell^p_h} \\
<& \Lambda_3 \|e(t)\|_{\ell^p_h} + (C + \Lambda_1) h|\ln h|,
\end{aligned}$$

under the event $\bigcap_{n=0}^{N'} A_1^n \cap A_2 \cap A_3$ by the stability Theorem 4.1 and the consistency estimate (5.5). It follows from (5.7) and the fact $\frac{d\|e\|_{\ell^p_h}}{dt} \le \|\frac{de}{dt}\|_{\ell^p_h}$, by using Gronwall's inequality with $e(0) = 0$ that

$$\max_{0\le t\le T} \|e(t)\|_{\ell^p_h} < C(T, \Lambda_1, \Lambda_3) h|\ln h| = \Lambda h|\ln h|, \tag{5.8}$$

under the event $\bigcap_{n=0}^{N'} A_1^n \cap A_2 \cap A_3$. Here we denote $\Lambda := C(T, \Lambda_1, \Lambda_3)$.

To complete the proof, we need to justify the stability condition: $|e_i(t)| \le \varepsilon$ for all i and $0 \le t \le T$ under the event $\bigcap_{n=0}^{N'} A_1^n \cap A_2 \cap A_3$. Since $h^d \max_{i\in I} |e_i(t)|^p \le (\|e(t)\|_{\ell^p_h})^p$, one has

$$\max_{i\in I} |e_i(t)| \le h^{-d/p} \|e(t)\|_{\ell^p_h} < Ch^{1-d/p} |\ln h| < \frac{\varepsilon}{2}, \quad \text{for } 0 \le t \le T, \tag{5.9}$$

by choosing $p > \frac{d(2q-1)}{q-1}$, $\varepsilon = h^{\frac{q}{2q-1}}$ with $q > 1$, and h small enough. Hence, $\max_{i\in I}|e_i|$ can hardly reach ε. From the discussion above, we have

$$\begin{aligned}(5.10)\qquad & P\left(\max_{0\le t\le T}\|X_{h,\varepsilon}(t)-X_h(t)\|_{L_h^p}\ge \Lambda h|\ln h|\right)\\ \le & P\left(\bigcup_{n=0}^{N'-1}(A_1^n)^c\cup A_2^c\cup A_3^c\right)\le \sum_{n=0}^{N'-1}P((A_1^n)^c)+P(A_2^c)+P(A_3^c)\\ \le & Ch^{-r}h^{C\Lambda_1|\ln h|}+C'h^{r/2-1}\exp(-C''h^{2-r})+h^{C\Lambda_3|\ln h|}\le h^{C\Lambda|\ln h|},\end{aligned}$$

where we have used (3.1) in Theorem 3.1, (4.2) in Theorem 4.1 and (4.7). Finally, we denote $\kappa = \frac{q}{2q-1}$, then the proof has been completed.

Appendices

Appendix A. Proof of $\rho \in H^k(\mathbb{R}^d)$ with initial data $\rho_0 \in L^1 \cap H^k(\mathbb{R}^d)$.

Theorem A.1. *Assume that the initial data ρ_0 satisfies*

$$0 \le \rho_0 \in L^1 \cap H^k(\mathbb{R}^d) \text{ with } k > \frac{d}{2}, \tag{A.1}$$

then the KS system (1.1) has a local solution with the following regularity

(A.2)
$$\|\rho\|_{L^\infty(0,T;H^k(\mathbb{R}^d))} \le C(\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}), \quad \|\rho\|_{L^2(0,T;H^{k+1}(\mathbb{R}^d))} \le C(\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}),$$

where $T > 0$ only depends on $\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}$.

Proof. We give a sketch of the proof here. By weak Young's inequality [15, P.107], one has

$$\|\nabla c\|_2 \le C\||x|^{-(d-1)}\|_{\frac{d}{d-1},w}\|\rho\|_{\frac{2d}{d+2}} \le C\|\rho\|_{\frac{2d}{d+2}}. \tag{A.3}$$

To estimate $\|\rho\|_{\frac{2d}{d+2}}$, we multiply (1.1) by $d\rho^{d-1}$ and integrate over $\mathbb{R}^d$, which leads to

$$\frac{d}{dt}\|\rho\|_d^d + \frac{4(d-1)}{d}\|\nabla\rho^{\frac{d}{2}}\|_2^2 \le (d-1)\|\rho\|_{d+1}^{d+1}. \tag{A.4}$$

Let us recall the Gagliardo-Nirenberg inequality [21, P.176, (2.3.50)]:

$$\|\rho\|_q \le C\|\nabla\rho\|_p^\theta\|\rho\|_r^{1-\theta}, \tag{A.5}$$

where $1 \le p, r \le \infty$, $0 \le \theta \le 1$, and $\frac{1}{q} = \theta(\frac{1}{p} - \frac{1}{d}) + \frac{1-\theta}{r}$. We choose $q = \frac{2(d+1)}{d}$ and $p = r = 2$, then one has

$$\|\rho\|_{d+1}^{d+1} = \|\rho^{\frac{d}{2}}\|_{\frac{2(d+1)}{d}}^{\frac{2(d+1)}{d}} \le C\|\nabla\rho^{\frac{d}{2}}\|_2\|\rho^{\frac{d}{2}}\|_2^{\frac{d+2}{d}}. \tag{A.6}$$

Hence by using the Young's inequality, we obtain

$$\frac{d}{dt}\|\rho\|_d^d \le C(\|\rho\|_d^d)^{\frac{d+2}{d}}. \tag{A.7}$$

Solving the above ordinary differential inequality, we know there exists there exists a $T_1 > 0$ depending on $\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}$, such that

$$\|\rho\|_{L^\infty(0,T_1;L^d(\mathbb{R}^d))} \le C(\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}). \tag{A.8}$$

Further more, as showed in [4], we know the mass conservation holds true

$$\|\rho\|_1 = \|\rho_0\|_1 \tag{A.9}$$

Hence by applying the interpolation inequality ($1 \leq \frac{2d}{d+2} < d$), we know

$$\|\rho\|_{\frac{2d}{d+2}} \leq C(\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}), \tag{A.10}$$

which leads to

$$\|\nabla c\|_2 \leq C(\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}), \tag{A.11}$$

for $0 < t \leq T_1$.

A simple computation of system (1.1) shows that, for $0 \leq |s| \leq k$,

$$\frac{d}{dt}\|D^s\rho\|_2^2 + \frac{1}{2}\|D^s\nabla\rho\|_2^2 \leq \frac{1}{2}\|D^s(\rho\nabla c)\|_2^2. \tag{A.12}$$

Using the Leibniz formula and Sobolev imbedding theorem, one concludes

$$\|D^s(\rho\nabla c)\|_2 \leq \|\rho\|_\infty\|\nabla c\|_{H^k} + \|\rho\|_{H^k}\|\nabla c\|_\infty \leq C\|\rho\|_{H^k}\|\nabla c\|_{H^k}. \tag{A.13}$$

Recall the fact that

$$\|\nabla c\|_{H^k} \leq \|\rho\|_{H^k} + \|\nabla c\|_2 \leq \|\rho\|_{H^k} + C(\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}), \tag{A.14}$$

which leads to

$$\frac{d}{dt}\|\rho\|_{H^k}^2 + \frac{1}{2}\|\rho\|_{H^{k+1}}^2 \leq C(\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)})(1 + \|\rho\|_{H^k}^2)^2. \tag{A.15}$$

Solving the above ordinary differential inequality, there exists a $0 < T \leq T_1$ depending on $\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}$, such that

$$\|\rho\|_{L^\infty(0,T;H^k(\mathbb{R}^d))} \leq C(\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}), \tag{A.16}$$

and

$$\|\rho\|_{L^2(0,T;H^{k+1}(\mathbb{R}^d))} \leq C(\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}), \tag{A.17}$$

which concludes our proof. □

Appendix B. Extension to general regular attractive force F

In this section, we will extend our result to the particle system with interacting function F regular enough, which satisfies

$$F \in H^{d+1}(\mathbb{R}^d). \tag{B.1}$$

We consider the regular solution ρ of the following PDE:

$$\begin{cases} \partial_t\rho = \nu\Delta\rho - \nabla\cdot(\rho F * \rho), & x \in \mathbb{R}^d,\ t > 0, \\ \rho(0,x) = \rho_0(x), \end{cases} \tag{B.2}$$

with ρ_0 has a compact support D with $D \subseteq B(R_0)$ and $0 \leq \rho_0 \in H^k(\mathbb{R}^d)$ with $k \geq d+1$. Then ρ has the following regularity for any $T > 0$

$$\|\rho\|_{L^\infty(0,T;H^{d+1}(\mathbb{R}^d))} \leq C\left(T, \|\rho_0\|_{H^{d+1}(\mathbb{R}^d)}, \|F\|_{H^{d+1}(\mathbb{R}^d)}\right) \tag{B.3}$$

and

$$\begin{aligned} &\|G\|_{L^\infty(0,T;W^{d+1,\infty}(\mathbb{R}^d))} \\ =&\|F * \rho\|_{L^\infty(0,T;W^{d+1,\infty}(\mathbb{R}^d))} \leq C\left(T, \|\rho_0\|_{H^{d+1}(\mathbb{R}^d)}, \|F\|_{H^{d+1}(\mathbb{R}^d)}\right). \end{aligned} \tag{B.4}$$

Again we suppose the self-consistent process $X_i(t)$ satisfying

$$X_i(t) = X_i(0) + \int_0^t \int_{\mathbb{R}^d} F\big(X_i(s) - y\big)\rho(s,y)dyds + \sqrt{2\nu}B_i(t), \quad i \in I, \tag{B.5}$$

with the initial data $X_i(0) = \alpha_i$.

Since F is regular enough, there is no need to mollify the force F anymore. To be specific, we consider trajectories $\{\hat{X}_i(t)\}_{i\in I}$ satisfying the SDEs:

$$\hat{X}_i(t) = \hat{X}_i(0) + \int_0^t \sum_{j\in I} F\big(\hat{X}_i(s) - \hat{X}_j(s)\big)\rho_j h^d ds + \sqrt{2\nu}B_i(t), \quad i \in I, \tag{B.6}$$

with initial data $\hat{X}_i(0) = \alpha_i$. And we denote

$$G_1^h(t,x) := \sum_{j\in I} F\big(x - X_j(t)\big)\rho_j h^d; \tag{B.7}$$

$$\hat{G}^h(t,x) := \sum_{j\in I} F\big(x - \hat{X}_j(t)\big)\rho_j h^d. \tag{B.8}$$

The extended result can be described in the following theorem

Theorem B.1. *Suppose the initial density $\rho_0(x)$ has a compact support D with $D \subseteq B(R_0)$ and $0 \le \rho_0 \in H^k(\mathbb{R}^d)$ with $k \ge d+1$. For the attractive force F satisfying* (B.1)*, ρ is the global regular solution to* (B.2)*. Assume that $X_h(t) = (X_i(t))_{i\in I}$ is the exact path of* (B.5) *and $\hat{X}_h(t) = \Big(\hat{X}_i(t)\Big)_{i\in I}$ is the solution to the particle system* (B.6)*. There exists two positive constants C and C' depending on T, p, d, R_0, $\|F\|_{H^{d+1}(\mathbb{R}^d)}$ and $\|\rho_0\|_{H^k(\mathbb{R}^d)}$ such that the following estimate holds*

$$P\left(\max_{0\le t\le T}\Big\|\hat{X}_h(t) - X_h(t)\Big\|_{\ell_h^p} < \Lambda h|\ln h|\right) \ge 1 - h^{C\Lambda|\ln h|},$$

for any $\Lambda > C'$, $p \ge 1$ and $T > 0$.

The idea of the proof of Theorem B.1 can be done as before, which is the consistency and stability implying convergence.

Like we have done in Section 3, the consistency can be proved.

Theorem B.2. *Under the same assumption as Theorem B.1, there exists two constants $C, C' > 0$ depending only on T, d, R_0, $\|F\|_{H^{d+1}(\mathbb{R}^d)}$ and $\|\rho_0\|_{H^k(\mathbb{R}^d)}$ such that at any fixed time $t \in [0,T]$, we have*

$$P\left(\max_{i\in I}\big|G_1^h(t, X_i(t)) - G(t, X_i(t))\big| < \Lambda h|\ln h|\right) \ge 1 - h^{C\Lambda|\ln h|}, \tag{B.9}$$

for all $\Lambda > C'$, where $G = F * \rho$ and G_1^h is defined (B.7)*.*

Proof. The proof is almost the same as the proof of Theorem 3.1. In this case, we have

$$\begin{aligned}|G_1^h(t,x)-G(t,x)| &\le \left|\sum_{j\in I} F(x-X_j(t))\rho_j h^d - \sum_{j\in I}\mathbb{E}[F(x-X_j(t))]\rho_j h^d\right| \\ &\quad + \left|\sum_{j\in I}\mathbb{E}[F(x-X_j(t))]\rho_j h^d - \int_{\mathbb{R}^d} F(x-y)\rho(t,y)dy\right| \\ &=: e_s(t,x)+e_d(t,x).\end{aligned} \tag{B.10}$$

and one can prove that

$$e_d(t,x)\le Ch^{d+1}; \quad P(e_s(t,x)\ge \Lambda h|\ln h|)\le h^{C\Lambda|\ln h|}. \tag{B.11}$$

Then this theorem can be proved similarly. □

As we have done in Section 4, we have the following stability result:

Theorem B.3. *Under the same assumption as Theorem B.1, there exists a constant $C>0$ depending only on T, p, d, R_0, $\|F\|_{H^{d+1}(\mathbb{R}^d)}$ and $\|\rho_0\|_{H^k(\mathbb{R}^d)}$ such that*

$$\|\hat{G}^h(t,\hat{X}_h(t))-G_1^h(t,X_h(t))\|_{\ell_h^p}\le C\|\hat{X}_h(t)-X_h(t)\|_{\ell_h^p},\ \forall\, t\in[0,T], \tag{B.12}$$

where G_1^h, $\hat{G}^h$ are defined in (B.7) *and* (B.8)*.*

Proof. Instead of using Lemma 2.6, we have that $M_{i,j}^1=|\nabla F(X_i(t)-X_j(t)+y)|\le C$ for any $t\in[0,T]$, which leads to

$$\sum_{j\in I} M_{i,j}^1 h^d\le C,\ \forall\, t\in[0,T]. \tag{B.13}$$

In addition, one has

$$\begin{aligned}&\hat{G}^h(t,\hat{X}_i(t))-G_1^h(t,X_i(t)) \\ =&\sum_{j\in I}\Big[F\big(\hat{X}_i(t)-\hat{X}_j(t)\big)-F\big(X_i(t)-X_j(t)\big)\Big]\rho_j h^d \\ =&\sum_{j\in I}\nabla F\big(X_i(t)-X_j(t)+\xi_{ij}\big)\cdot\big(\hat{X}_i(t)-X_i(t)+X_j(t)-\hat{X}_j(t)\big)\rho_j h^d \\ =&\sum_{j\in I}\nabla F\big(X_i(t)-X_j(t)+\xi_{ij}\big)\cdot\big(\hat{X}_i(t)-X_i(t)\big)\rho_j h^d \\ &+\sum_{j\in I}\nabla F\big(X_i(t)-X_j(t)+\xi_{ij}\big)\cdot\big(X_j(t)-\hat{X}_j(t)\big)\rho_j h^d \\ =:&\,\mathcal{I}_i+\mathcal{J}_i.\end{aligned} \tag{B.14}$$

Hence, we have

$$|\mathcal{I}_i|\le C|\hat{X}_i(t)-X_i(t)|; \quad |\mathcal{J}_i|\le\sum_{j\in I}M_{i,j}^1|X_j(t)-\hat{X}_j(t)|\rho_j h^d. \tag{B.15}$$

which concludes the proof. □

Finally, combining Theorem B.2 and Theorem B.3, we can get Theorem B.1 as we have done in Section 5.

Acknowledgments

Hui Huang wishes to acknowledge a scholarship from China Scholarship Council and thank Professor Huai-Yu Jian for his support and encouragement.

References

[1] C. Anderson and C. Greengard. On vortex methods. *SIAM journal on numerical analysis*, 22(3):413–440, 1985.
[2] J. T. Beale and A. Majda. Vortex methods. I: Convergence in three dimensions. *Mathematics of Computation*, 39(159):1–27, 1982.
[3] J. T. Beale and A. Majda. Vortex methods. II: Higher order accuracy in two and and three dimensions. *Mathematics of Computation*, 39(159):29–52, 1982.
[4] S. Bian and J.-G. Liu. Dynamic and steady states for multi-dimensional keller-segel model with diffusion exponent $m > 0$. *Communications in Mathematical Physics*, 323(3):1017–1070, 2013.
[5] N. Boers and P. Pickl. On mean field limits for dynamical systems. *Journal of Statistical Physics*, pages 1–16, 2015.
[6] A. J. Chorin. Numerical study of slightly viscous flow. *Journal of fluid mechanics*, 57(04):785–796, 1973.
[7] G.-H. Cottet and P. D. Koumoutsakos. *Vortex methods: theory and practice*. Cambridge university press, 2000.
[8] K. Craig and A. L. Bertozzi. A blob method for the aggregation equation. *Mathematics of Computation*, 85:1681–1717, 2016.
[9] D. Freedman. *Brownian motion and diffusion*. Springer Science & Business Media, 2012.
[10] A. Friedman. *Stochastic differential equations and applications*, volume 1. Academic press, 2014.
[11] J. Goodman. Convergence of the random vortex method. *Communications on Pure and Applied Mathematics*, 40(2):189–220, 1987.
[12] C. Graham and D. Talay. *Stochastic Simulation and Monte Carlo Methods: Mathematical Foundations of Stochastic Simulation*, volume 68. Springer Science & Business Media, 2013.
[13] O. H. Hald. Convergence of vortex methods for euler's equations. ii. *SIAM Journal on Numerical Analysis*, 16(5):726–755, 1979.
[14] E. F. Keller and L. A. Segel. Initiation of slime mold aggregation viewed as an instability. *Journal of Theoretical Biology*, 26(3):399–415, 1970.
[15] E. H. Lieb and M. Loss. *Analysis*. American Mathematical Society, 2001.
[16] J.-G. Liu and R. Yang. Propagation of chaos for the keller-segel equation with a logarithmic cut-off. preprint.
[17] J.-G. Liu and R. Yang. A random particle blob method for the keller-segel equation and convergence analysis. *Mathematics of Computation*, 2016. in press.
[18] J.-G. Liu and Z. Yuan. Convergence of diffusion-drift many particle systems in probability under sobolev norm. In P. Gonçalves and A. J. Soares, editors, *From Particle Systems to Partial Differential Equations-III*. Springer, 2016.
[19] D.-G. Long. Convergence of the random vortex method in two dimensions. *Journal of the American Mathematical Society*, 1(4):779–804, 1988.
[20] C. Marchioro and M. Pulvirenti. Vortex methods in two-dimensional fluid dynamics. *Lecture notes in physics*, 203:1–137, 1984.
[21] V. Maz'ya. *Sobolev spaces*. Springer, 2013.
[22] B. Oksendal. *Stochastic differential equations: an introduction with applications*. Springer Science & Business Media, 2013.
[23] H. Osada. Propagation of chaos for the two dimensional navier-stokes equation. *Proceedings of the Japan Academy. Series A Mathematical sciences*, 62(1):8–11, 1986.
[24] D. Pollard. *Convergence of stochastic processes*. Springer Science & Business Media, 2012.
[25] M. Talagrand. *Pettis integral and measure theory*. American Mathematical Society, 1984.

Department of Mathematical Sciences, Tsinghua University, Beijing, 100084, People's Republic of China and Departments of Physics and Mathematics, Duke University, Durham, NC 27708, USA.
E-mail address: huanghui12@mails.tsinghua.edu.cn

Departments of Physics and Mathematics, Duke University, Durham, NC 27708, USA.
E-mail address: jliu@phy.duke.edu