COMMUN. MATH. SCI.
Vol. 15, No. 7, pp. 1821–1842

DISCRETE-IN-TIME RANDOM PARTICLE BLOB METHOD FOR THE KELLER–SEGEL EQUATION AND CONVERGENCE ANALYSIS*

HUI HUANG† AND JIAN-GUO LIU‡

Abstract. We establish an error estimate of a discrete-in-time random particle blob method for the Keller–Segel (KS) equation in $\mathbb{R}^d$ $(d\geq 2)$. With a blob size $\varepsilon = N^{-\frac{1}{d(d+1)}}\log(N)$, we prove the convergence rate between the solution to the KS equation and the empirical measure of the random particle method under L^2 norm in probability, where N is the number of the particles.

Keywords. coupling method; concentration inequality; splitting scheme; kernel density estimation; Newtonian aggregation; chemotaxis; Brownian motion; interacting particle system.

AMS subject classifications. 65M75; 65M15; 65M12; 35Q92; 35K55; 60H35.

1. Introduction

Analogously to the random vortex blob method for the Navier–Stokes equations, we propose a discrete-in-time random particle blob method for the Keller–Segel (KS) equation in $\mathbb{R}^d$ $(d\geq 2)$ [19, 25], which reads

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

where ν is a positive constant. 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 from the fundamental solution of Laplace's equation 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

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

*Received: April 3, 2016; accepted (in revised form): April 15, 2017. Communicated by Peter Markowich.
Hui Huang is partially supported by the Alan Mekler Postdoctoral Fellowship in the Department of Mathematics at Simon Fraser University. He also hopes to thank Professor Huai-Yu Jian for his support and encouragement. Jian-Guo Liu is partially supported by KI-Net NSF RNMS (Grant No. 1107444) and NSF DMS (Grant No. 1514826).
†Department of Mathematics, Simon Fraser University, Burnaby, BC, V5A 1S6, Canada (huanghui12@mails.tsinghua.edu.cn).
‡Departments of Physics and Mathematics, Duke University, Durham, NC 27708, USA (jliu@phy.duke.edu).

Furthermore, we take the gradient of the Newtonian potential $\Phi(x)$ as the attractive force $F(x)$. Thus we have

$$F(x)=\nabla\Phi(x)=-\frac{C_* x}{|x|^d}, \quad \forall\, x\in\mathbb{R}^d\backslash\{0\},\ d\geq 2, \tag{1.4}$$

with $C_*=\frac{\Gamma(d/2)}{2\pi^{d/2}}$.

In recent years, there has been a surge of activities focused on properties of the KS equation, both for parabolic-elliptic and parabolic-parabilic systems. In the case of two dimensions, for the parabolic-elliptic model (1.1), a sharp bound on the critical mass, $m_c=8\pi$, was given by Dolbeault and Perthame in [12] through using the logarithmic Hardy–Littlewood–Sobolev inequality. Critical mass means that if the initial mass is less than m_c, the solution will exist globally; otherwise there must be mass concentration [12]. This result was further completed and improved in [5], where the existence of free-energy solutions had been proved. In [3], Bedrossian and Masmoudi proved a local existence and uniqueness of mild solutions for initial measure only satisfying $\max_{x\in\mathbb{R}^2}\mu\{x\}<8\pi$. For the parabolic-parabolic KS model, the global existence was analyzed and the critical mass (which is also 8π) was derived in [8]. There was in-depth analyses for the case of critical mass $m_c=8\pi$ in [4,6]. In space dimension $d\geq 3$, global existence, finite time blow-up and large time asymptotic behavior were studied in [7,26,28]. Last, we refer readers to the review paper [16] or Chapter 5 in [26] for more details.

Since the error estimates obtained later are valid when the solution of the KS equation is regular enough, we assume that

$$0\leq\rho_0\in L^1\cap H^k(\mathbb{R}^d) \text{ with } k>d/2+3, \tag{1.5}$$

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

$$\|\rho\|_{L^\infty(0,T;H^k(\mathbb{R}^d))}\leq C\left(\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}\right), \tag{1.6}$$

where $T>0$ depends on $\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}$. The proof of this result is a standard process and it can be found in [17, Theorem A.1]. As a direct result of the Sobolev imbedding theorem, one has

$$\|\rho\|_{L^\infty(0,T;W^{3,\infty}(\mathbb{R}^d))}\leq C\left(\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}\right), \tag{1.7}$$

for $k>d/2+3$.

Let $\{X_i(0)\}_{i=1}^N$ be N independent, identically distributed (i.i.d) random vectors with common density $\rho_0(x)$. Then we introduce a random particle blob method for the KS equation, and it is given by the following stochastic particle system of N particle paths $\{X_i(t)\}_{i=1}^N$

$$dX_i(t)=\frac{1}{N-1}\sum_{j=1}^N F_\varepsilon\big(X_i(t)-X_j(t)\big)\,dt+\sqrt{2\nu}\,dB_i(t), \quad i=1,\cdots,N, \tag{1.8}$$

where $\{B_i(t)\}_{i=1}^N$ are N independent standard Brownian motions and

$$F_\varepsilon=\psi_\varepsilon * F, \quad \psi_\varepsilon(x)=\varepsilon^{-d}\psi(\varepsilon^{-1}x), \quad \varepsilon>0. \tag{1.9}$$

In this article, we take the cut-off function $0\leq\psi(x)\in C_0^\infty(\mathbb{R}^d)$ as in [17], which satisfies $\psi(x)=\psi(|x|)$ and $\int_{\mathbb{R}^d}\psi(x)\,dx=1$.

To discretize equation (1.8) in time, take a time step $\Delta t>0$ and let $t_n=n\Delta t$. Our approximation $X_i^{(n)}\approx X_i(t_n)$, will satisfy

$$X_i^{(n+1)}=X_i^{(n)}+\Delta t G_N\left(X_i^{(n)}\right)+\sqrt{2\nu\Delta t}N_i^{(n)},\quad i=1,\cdots,N, \tag{1.10}$$

where the $N_i^{(n)}$ are independent standard Gaussian random vectors and

$$G_N\left(X_i^{(n)}\right):=\frac{1}{N-1}\sum_{j=1}^{N}F_\varepsilon\left(X_i^{(n)}-X_j^{(n)}\right). \tag{1.11}$$

Our error estimate is based on the coupling method. To do this, we will construct independent particles $\{Y_i^{(n)}\}_{i=1}^N$ from the mean field equation and prove that they are close to particles $\{X_i^{(n)}\}_{i=1}^N$ with high probability (see Theorem 3.1).

Before the construction of $\{Y_i^{(n)}\}_{i=1}^N$, we recall a simple fact from the probability theory. Suppose $X\in\mathbb{R}^d$ is a random vector with density $\rho(x)$. Let $u(x)$ be a smooth vector field so that the mapping $y=x+\Delta t u(x)$ is a smooth homeomorphism for all Δt small enough. If $Y=X+\Delta t u(X)$, then Y has the density function $\bar\rho(x)$ defined by

$$\bar\rho\big(x+\Delta t u(x)\big)=\det{}^{-1}\big(I+\Delta t Du(x)\big)\rho(x). \tag{1.12}$$

Now, we can construct independent particles $\{Y_i^{(n)}\}_{i=1}^N$ to approximate $\{X_i^{(n)}\}_{i=1}^N$ by following the approach in [13]. Assume that we have independent random variables $\{Y_i^{(0)}\}_{i=1}^N$ with the same density $\rho^{(0)}(x)=\rho_0(x)$. If $G^{(0)}=F_\varepsilon*\rho^{(0)}$ and $Y_i^{(1/2)}=Y_i^{(0)}+\Delta t G^{(0)}\big(Y_i^{(0)}\big)$, then $\{Y_i^{(1/2)}\}_{i=1}^N$ are independent and have common density $\rho^{(1/2)}$ (see equation (1.14)) by the definition (1.12). Furthermore, if $Y_i^{(1)}=Y_i^{(1/2)}+\sqrt{2\nu\Delta t}N_i^{(0)}$, the $N_i^{(0)}$ being the independent standard Gaussian random vectors, then $\{Y_i^{(1)}\}_{i=1}^N$ has density $\rho^{(1)}$ given by equation (1.15). Continuing this process, we have constructed $\{Y_i^{(n)}\}_{i=1}^N$ with common density $\rho^{(n)}$ satisfying the discretized equations:

$$G^{(n)}(x)=F_\varepsilon*\rho^{(n)}(x), \tag{1.13}$$

$$\rho^{(n+1/2)}\big(x+\Delta t G^{(n)}(x)\big)=\det{}^{-1}\big(I+\Delta t DG^{(n)}(x)\big)\rho^{(n)}(x), \tag{1.14}$$

$$\rho^{(n+1)}(x)=H(\sqrt{\nu}\Delta t)\rho^{(n+1/2)}(x). \tag{1.15}$$

Here, the operator $H(s)$ is the solution operator of the heat equation at time s. In addition, the splitting algorithm from equations (1.13)–(1.15) is a splitting scheme with the linear transport approximation, and it has been proved in [18] that $\rho^{(n)}$ converges to $\rho(t_n,x)$ with the rate $C\Delta t$ when F_ε in equation (1.13) is replaced by F.

To approximate $X_i^{(n)}$, we define $Y_i^{(0)}=X_i^{(0)}$ and

$$Y_i^{(n+1)}=Y_i^{(n)}+\Delta t G^{(n)}\big(Y_i^{(n)}\big)+\sqrt{2\nu\Delta t}N_i^{(n)},\quad i=1,\cdots,N, \tag{1.16}$$

where $G^{(n)}\big(Y_i^{(n)}\big)$ is the vector field constructed in equation (1.13). Note that $Y_i^{(n)}$ is independent of $Y_j^{(n)}$ if $i\neq j$, and they share the common density $\rho^{(n)}$.

Now, we define the regularized empirical measure as follows: consider a non-negative function $\varphi(x)=\varphi(|x|)\in C_0^\infty(\mathbb{R}^d)$, which satisfies:

- $\varphi(x)$ is bounded and compactly supported;
- $\int_{\mathbb{R}^d}\varphi(x)\,dx=1$ and $\int_{\mathbb{R}^d}xx^T\varphi(x)\,dx=aI_d$, where $a=\frac{1}{d}\int_{\mathbb{R}^d}x^2\varphi(x)\,dx$ and I_d is the identity matrix.

And for some $\delta>0$, which will be specified later, let $\varphi_\delta(x)=\delta^{-d}\varphi(\delta^{-1}x)$. Then the regularized empirical measure of $\{X_i^{(n)}\}_{i=1}^N$ can be defined as

$$\mu_X^{(n)}(x):=\frac{1}{N}\sum_{i=1}^N\varphi_\delta\big(x-X_i^{(n)}\big), \tag{1.17}$$

and similarly, one has the regularized empirical measure of $\{Y_i^{(n)}\}_{i=1}^N$

$$\mu_Y^{(n)}(x):=\frac{1}{N}\sum_{i=1}^N\varphi_\delta\big(x-Y_i^{(n)}\big). \tag{1.18}$$

The use of such regularized empirical measures above is important in computation. The vortex method was first introduced by Chorin in 1973 [9], 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 *et al.* [14,15], Beale and Majda [1,2]. 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 [13, 22–24] in 1980s. Lastly, we refer to the book [10] for theoretical and practical use of the vortex methods, and also refer to [11] for recent progress on a blob method for the aggregation equation.

For the KS equation, a random particle blob method has been studied in [20], where a rigorous global convergence without probability rate has been obtained. Furthermore, our recent paper [17] studied the time continuous system (1.8), and proved the convergence of particle paths in probability. However, in this paper, we study the fully discretized system (1.10) and prove the convergence of the regularized empirical measure in L^2 space. Since the system (1.8) has a very large size N and the kernel F_ε is singular, the standard convergence analyses of time discretization does not work. To overcome this, in [18], we proposed a splitting method with the linear transport approximation as in equations (1.13)–(1.15), so that we can fully use the regularity of $\rho^{(n)}$. Based on the stability and regularity of the splitting method, we use the coupling method to achieve our objective. In order to realize this approach, this splitting method requires us to take the initial positions as i.i.d. random vectors $\{X_i(0)\}_{i=1}^N$ with the common density ρ_0. On the contrary, the initial positions of the particles were taken on the lattice points $N^{-1/d}\alpha\in\mathbb{R}^d$ with mass $\frac{1}{N}\rho_0\big(N^{-1/d}\alpha\big)$ in [17], where $\alpha=(\alpha_1,\cdots,\alpha_d)\in\mathbb{Z}^d$, and a sampling letter plays a crucial part in the proof. This tool could not be adapted here. Based on the convergence result of random particle trajectories in [17], Liu and Zhang [21] showed the convergence of regularized empirical measures of many particle systems in probability under a Sobolev norm to the corresponding mean field PDE by using a new martingale method.

With defining the regularized empirical measure $\mu_X^{(n)}$, we can state the main result in this paper, which shows that $\mu_X^{(n)}$ converges to the unique solution of KS Equation (1.1) in L^2 norm with high probability.

THEOREM 1.1. *Suppose that* $0 \le \rho_0(x) \in L^1 \cap H^k(\mathbb{R}^d)$ *with* $k > d/2+3$. *Let* $\rho(t,x)$ *be the unique regular solution of the KS Equation* (1.1) *with local existence time* T *and* $\mu_X^{(n)}$ *be the regularized empirical measure as in definition* (1.17) *with* $\delta = N^{-\frac{1}{d(d+2)}}$ *(N sufficiently large). If* $\varepsilon = N^{-\frac{1}{d(d+1)}}\log(N)$ *, then there exists some* T_**,* C_* *depending only on* T *and* $\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}$*, for* $T_*/N \le \Delta t \le C_*$ *and* $(n+1)\Delta t \le T_*$*, such that the following estimate holds*

$$P\left(\max_{0\le n\Delta t\le T_*}\|\rho(t_n,\cdot)-\mu_X^{(n)}(\cdot)\|_2 < C_1\left(N^{-\frac{1}{d(d+2)}}+\Delta t\right)\right) \ge 1 - C_2 N^{-\frac{2}{d(d+2)}},$$

where C_1*,* C_2 *depend only on* T_* *and* $\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}$*.*

To conclude the introduction, we give the outline of this paper. First (Section 2), we provide several lemmas including kernel estimates, the regularity of $\rho^{(n)}$ and a concentration inequality. In Section 3, we consider the independent particles $Y_i^{(n)}$ moving with the velocity field constructed from equations (1.13)–(1.15). And given sufficient smoothing, we prove that these independent particles satisfy Equation (1.10) up to an error. Then we establish the error estimate between the regularized empirical measure $\mu_X^{(n)}$ and the solution to KS Equation (1.1) in L^2 norm in Section 4. Then (Section 5), we obtain the error estimate on the interaction. In Section 6, we extend our results to general regular attractive force F.

2. Preliminaries

Notation: For convenience, we denote the index set of the particles as $I := \{1,\cdots,N\}$ and the number of particles N is assumed to be sufficiently large in the sequel. In this paper, we use $\|\cdot\|_p$ for L^p norm of a function and use $\|\cdot\|_{\ell 2}$ for discrete L^2 norm of a vector, which can be represented as

$$\|(v_i)_{i\in I}\|_{\ell^2} = \left(\frac{1}{N}\sum_{i=1}^N |v_i|^2\right)^{1/2}. \tag{2.1}$$

Moreover, we denote $\hat{f}(\xi)$ as the Fourier transformation of $f(x)$, which is

$$\hat{f}(\xi) := \int_{\mathbb{R}^d} f(x) e^{-2\pi i x\cdot\xi}\,dx. \tag{2.2}$$

The generic constant will be denoted generically by C, even if it is different from line to line.

Firstly, we summarize some useful estimates about the regularized kernel F_ε as in definition (1.9).

LEMMA 2.1 (Kernel estimates).

(i) $F_\varepsilon(0) = 0$*,* $F_\varepsilon(x) = F(x)$ *for any* $|x| \ge \varepsilon$ *and* $|F_\varepsilon(x)| \le |F(x)|$*;*

(ii) $|\partial^\beta F_\varepsilon(x)| \le C_\beta \varepsilon^{1-d-|\beta|}$*,for any* $x \in \mathbb{R}^d$*;*

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

These results can be found in [17, Lemma 2.1-2.2].

The following lemma shows that the algorithm from equations (1.13)–(1.15) is H^k stable.

LEMMA 2.2 ([18, Proposition 3]). *Suppose that the initial density* $0 \le \rho_0(x) \in L^1 \cap H^k(\mathbb{R}^d)$ *with* $k > \frac{d}{2}+1$. *Then there exists some* $C_*, T_1 > 0$ *depending only on*

$\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}$, *such that for the algorithm* (1.13)–(1.15) *with* $\Delta t \le C_*$, *we have*

$$\|\rho^{(n)}\|_{H^k} \le C(T_1, \|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}), \quad \forall\ 0 \le n\Delta t \le T_1. \tag{2.3}$$

The next lemma can be found in [13], which provides us the probability bounds of random vectors in the sequel.

LEMMA 2.3 ([13, Lemma 1]). *Let* $Z_i, \cdots, Z_N$ *be i.i.d. random vectors with* $\mathbb{E}[Z_i]=0$, $\mathbb{E}[Z_i^2] \le g(N)$ *and* $|Z_i| \le C\sqrt{Ng(N)}$. *Then the sample mean* $\bar{Z} = \frac{1}{N}\sum_{i=1}^N Z_i$ *satisfies*

$$\mathbb{E}\left[\exp\left(\sqrt{\frac{N}{g(N)}}|\bar{Z}|\right)\right] \le C, \tag{2.4}$$

and

$$P\left(|\bar{Z}| \ge \frac{C_p\sqrt{g(N)}\log(N)}{\sqrt{N}}\right) \le N^{-p}, \tag{2.5}$$

where C_p *depends only on* C *and* $p>0$.

Applying Lemma 2.3, we obtain the probability bound of the gradient of the regularized kernel F_ε:

LEMMA 2.4. *Let* F_ε *be the regularized kernel as in definition* (1.9) *with* $\varepsilon \ge N^{-\frac{1}{2d}}\log(N)$. *Then there exists a positive constant* C *depending only on* $p>0$, T_1 *and* $\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}$, *such that if* $T_1/N \le \Delta t \le C_*$ (T_1, C_* *are used in Lemma* 2.2), *then*

$$P\left(\max_x \left|\frac{1}{N-1}\sum_{j=1}^N \nabla F_\varepsilon\left(x - Y_j^{(n)}\right)\right| \ge C \textit{ for some } n\Delta t \le T_1\right) \le N^{1-p}. \tag{2.6}$$

Proof. The main idea of this proof is to use the Fourier transform and Markov's inequality. First, let us define

$$L^{(n)}(x) := \frac{1}{N}\sum_{j=1}^N \nabla F_\varepsilon\left(x - Y_j^{(n)}\right) - \int_{\mathbb{R}^d} \nabla F_\varepsilon(x-y)\rho^{(n)}(y)\,dy. \tag{2.7}$$

So in order to give a rough bound on

$$\frac{1}{N-1}\sum_{j=1}^N \nabla F_\varepsilon\left(x - Y_j^{(n)}\right), \tag{2.8}$$

we only need to bound

$$S_n := \max_x |L^{(n)}(x)|, \tag{2.9}$$

since the integral contribution to $L^{(n)}(x)$ is bounded. We shall use the fact that

$$\max_x |L^{(n)}(x)| \le \int_{\mathbb{R}^d} |\widehat{L}^{(n)}(\xi)|\,d\xi, \tag{2.10}$$

with

$$\widehat{L}^{(n)}(\xi) = \frac{1}{N}\sum_{j=1}^{N}\left(e^{-2\pi i\xi\cdot Y_j^{(n)}} - \hat{\rho}^{(n)}(\xi)\right)\widehat{\nabla F_\varepsilon}(\xi). \tag{2.11}$$

If we define

$$L_1(\xi) := \frac{1}{N}\sum_{j=1}^{N}\left(e^{-2\pi i\xi\cdot Y_j^{(n)}} - \hat{\rho}^{(n)}(\xi)\right) =: \frac{1}{N}\sum_{j=1}^{N} Z_j, \tag{2.12}$$

then it is easy to verify that

$$\mathbb{E}(Z_j) = 0, \quad \mathbb{E}(|Z_j|^2) \le C, \quad |Z_j| \le C. \tag{2.13}$$

Thus it follows from Lemma 2.3 that

$$\mathbb{E}\left[\exp\{\sqrt{N}L_1(\xi)\}\right] \le C. \tag{2.14}$$

Set $L_2(\varepsilon) = \int_{\mathbb{R}^d} |\widehat{\nabla F_\varepsilon}(\xi)|\,d\xi$. Then one has

$$\int_{\mathbb{R}^d} \frac{|\widehat{\nabla F_\varepsilon}(\xi)|}{L_2(\varepsilon)}\,d\xi = 1. \tag{2.15}$$

Now we can apply Jensen's inequality since e^x is a convex function:

$$\exp\left\{\int_{\mathbb{R}^d} \sqrt{N}L_1(\xi)\frac{|\widehat{\nabla F_\varepsilon}(\xi)|}{L_2(\varepsilon)}\,d\xi\right\} \le \int_{\mathbb{R}^d} \exp\{\sqrt{N}L_1(\xi)\}\frac{|\widehat{\nabla F_\varepsilon}(\xi)|}{L_2(\varepsilon)}\,d\xi, \tag{2.16}$$

which leads to

$$\begin{aligned}\mathbb{E}\left[\exp\left\{\int_{\mathbb{R}^d} \sqrt{N}L_1(\xi)\frac{|\widehat{\nabla F_\varepsilon}(\xi)|}{L_2(\varepsilon)}\,d\xi\right\}\right] &\le \int_{\mathbb{R}^d} \mathbb{E}\left[\exp\{\sqrt{N}L_1(\xi)\}\right]\frac{|\widehat{\nabla F_\varepsilon}(\xi)|}{L_2(\varepsilon)}\,d\xi \\ &\le C\int_{\mathbb{R}^d} \frac{|\widehat{\nabla F_\varepsilon}(\xi)|}{L_2(\varepsilon)}\,d\xi = C. \end{aligned} \tag{2.17}$$

Hence,

$$\begin{aligned}\mathbb{E}\left[\exp\left\{\frac{\sqrt{N}S_n}{L_2(\varepsilon)}\right\}\right] &\le \mathbb{E}\left[\exp\left\{\frac{\sqrt{N}\int_{\mathbb{R}^d}|\widehat{L}^{(n)}(\xi)|\,d\xi}{L_2(\varepsilon)}\right\}\right] \\ &= \mathbb{E}\left[\exp\left\{\int_{\mathbb{R}^d}\sqrt{N}L_1(\xi)\frac{|\widehat{\nabla F_\varepsilon}(\xi)|}{L_2(\varepsilon)}\,d\xi\right\}\right] \le C. \end{aligned} \tag{2.18}$$

If we use Markov's inequality, it follows that

$$P\left(S_n \ge \frac{C_p\log(N)L_2(\varepsilon)}{\sqrt{N}}\right) \le N^{-p}. \tag{2.19}$$

Notice that

$$L_2(\varepsilon) = \int_{\mathbb{R}^d}|\widehat{\nabla F_\varepsilon}(\xi)|\,d\xi = \int_{\mathbb{R}^d}|\widehat{F}(\xi)||\widehat{\nabla\psi_\varepsilon}(\xi)|\,d\xi \le C\int_{\mathbb{R}^d}|\widehat{\psi}(\varepsilon\xi)|\,d\xi \le C\varepsilon^{-d}, \tag{2.20}$$

since $|\widehat{F}(\xi)| \le \frac{C}{|\xi|}$ and $|\widehat{\nabla\psi_\varepsilon}(\xi)| = |\xi||\widehat{\psi}(\varepsilon\xi)|$. So if we assume that $\varepsilon \ge N^{-\frac{1}{2d}}\log(N)$, then inequalities (2.19) and (2.20) imply that

$$P\left(S_n \ge C_p \log^{1-d}(N)\right) \le N^{-p}. \tag{2.21}$$

By the constriction of time size $T_1/N \le \Delta t$, one has

$$n \le T_1/\Delta t \le N. \tag{2.22}$$

Taking unions of N sets of the events

$$\Gamma_n := \left\{S_n \ge C_p \log^{1-d}(N)\right\}, \tag{2.23}$$

inequality (2.21) implies that

$$P\left(\max_x |L^{(n)}(x)| \ge C_p \log^{1-d}(N) \text{ for some } n\Delta t \le T_1\right) = P(\cup_n \Gamma_n) \le N^{1-p}. \tag{2.24}$$

For N sufficiently large, there exists some Λ such that

$$P\left(\max_x |L^{(n)}(x)| \ge \Lambda \text{ for some } n\Delta t \le T_1\right) \le N^{1-p}. \tag{2.25}$$

and

$$P\left(\max_x \left|\int_{\mathbb{R}^d} \nabla F_\varepsilon(x-y)\rho^{(n)}(y)\,dy\right| \ge \Lambda \text{ for some } n\Delta t \le T_1\right) = 0. \tag{2.26}$$

Hence, it follows from the definition (2.7), that

$$\begin{aligned}
&P\left(\max_x \left|\frac{1}{N-1}\sum_{j=1}^{N} \nabla F_\varepsilon\big(x - Y_j^{(n)}\big)\right| \ge 2\Lambda \text{ for some } n\Delta t \le T_1\right)\\
\le& P\left(\max_x |L^{(n)}(x)| \ge \Lambda \text{ for some } n\Delta t \le T_1\right)\\
&+ P\left(\max_x \left|\int_{\mathbb{R}^d} \nabla F_\varepsilon(x-y)\rho^{(n)}(y)\,dy\right| \ge \Lambda \text{ for some } n\Delta t \le T_1\right)\\
\le& N^{1-p},
\end{aligned} \tag{2.27}$$

which concludes the proof. ∎

3. The error estimate between $X_i^{(n)}$ and $Y_i^{(n)}$

In this section, we will show that $Y_i^{(n)}$ is a good approximation of $X_i^{(n)}$. Actually, it satisfies

$$Y_i^{(n+1)} = Y_i^{(n)} + \Delta t G_N\big(Y_i^{(n)}\big) + \sqrt{2\nu\Delta t}\, N_i^{(n)} + \Delta t\ r_i^{(n)}, \quad i \in I, \tag{3.1}$$

where

$$G_N\big(Y_i^{(n)}\big) := \frac{1}{N-1}\sum_{j=1}^{N} F_\varepsilon\big(Y_i^{(n)} - Y_j^{(n)}\big), \tag{3.2}$$

and

$$r_i^{(n)} := G^{(n)}\big(Y_i^{(n)}\big) - \frac{1}{N-1}\sum_{j=1}^{N} F_\varepsilon\big(Y_i^{(n)} - Y_j^{(n)}\big). \tag{3.3}$$

The following proposition states that the residual vector $r^{(n)} := \big(r_i^{(n)}\big)_{i\in I}$ is small in ℓ_2 norm with high probability.

PROPOSITION 3.1. *Assume that $\varepsilon \geq N^{-1/d}$. Then there exists a positive constant C depending only on $p>1$, T_1 and $\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}$, such that if $T_1/N \leq \Delta t \leq C_*$ (T_1, C_* are used in Lemma 2.2), then*

$$P\left(|r_i^{(n)}| \geq CN^{-1/d}\log(N) \text{ for some } i\in I, n\Delta t\leq T_1\right)\leq N^{2-p}, \tag{3.4}$$

and

$$P\left(\|r^{(n)}\|_{\ell_2} \geq CN^{-1/d}\log(N) \text{ for some } n\Delta t\leq T_1\right)\leq N^{2-p}, \tag{3.5}$$

where $r_i^{(n)}$ is the residual as in definition (3.3).

Proof. First, we are ready to bound

$$r_1^{(n)} = G^{(n)}\big(Y_1^{(n)}\big) - \frac{1}{N-1}\sum_{j=1}^{N} F_\varepsilon\big(Y_1^{(n)} - Y_j^{(n)}\big) = \frac{1}{N-1}\sum_{j=2}^{N} Z_j, \tag{3.6}$$

where $Z_j = G^{(n)}\big(Y_1^{(n)}\big) - F_\varepsilon\big(Y_1^{(n)} - Y_j^{(n)}\big)$ and we have used $F_\varepsilon(0)=0$. Since Y_1 and Y_j are independent, let us consider Y_1 as given and denote $\mathbb{E}'[\cdot] = \mathbb{E}[\cdot|Y_1]$. It is easy to show that $\mathbb{E}'[Z_j]=0$ since

$$\mathbb{E}'\left[F_\varepsilon\big(Y_1^{(n)} - Y_j^{(n)}\big)\right] = \int_{\mathbb{R}^d} F_\varepsilon\big(Y_1^{(n)} - y\big)\rho^{(n)}(y)\,dy = \mathbb{E}'\big[G^{(n)}(Y_1^{(n)})\big]. \tag{3.7}$$

To use Lemma 2.3, we need a bound for the variance

$$\mathbb{E}'\big[|Z_j|^2\big] = \mathbb{E}'\left[|G^{(n)}(Y_1^{(n)}) - F_\varepsilon\big(Y_1^{(n)} - Y_j^{(n)}\big)|^2\right]. \tag{3.8}$$

Since $G^{(n)}$ is bounded and $Y_1^{(n)}$ could be any point, it suffices to bound

$$\mathbb{E}'\big[F_\varepsilon(Y_1^{(n)} - Y_j^{(n)})\big] = \int_{\mathbb{R}^d} F_\varepsilon(Y_1^{(n)} - x)\rho^{(n)}(x)\,dx \leq \|\rho^{(n)}\|_2\|F_\varepsilon\|_2 \leq C\varepsilon^{1-d/2} \tag{3.9}$$

and

$$\mathbb{E}'\big[F_\varepsilon(Y_1^{(n)} - Y_j^{(n)})^2\big] = \int_{\mathbb{R}^d} F_\varepsilon(Y_1^{(n)} - x)^2\rho^{(n)}(x)\,dx \leq \|\rho^{(n)}\|_\infty\|F_\varepsilon\|_2^2 \leq C\varepsilon^{2-d}, \tag{3.10}$$

where we have used $\|\rho^{(n)}\|_\infty \leq C\|\rho^{(n)}\|_{H^k} \leq C(T_1, \|\rho^{(0)}\|_{H^k})$ and $\|F_\varepsilon\|_2 \leq C\varepsilon^{1-d/2}$. Hence one has

$$\mathbb{E}'\big[|Z_j|^2\big] \leq C\varepsilon^{2-d}. \tag{3.11}$$

Under the assumption that $\varepsilon \geq N^{-1/d}$, the hypotheses of Lemma 2.3 are satisfied with $g(N) = CN^{1-2/d}$. In addition, $|Z_i| \leq C\varepsilon^{1-d} \leq CN^{1-1/d} \leq C\sqrt{Ng(N)}$. Hence, we have the probability bound of $r_1^{(n)}$ by Lemma 2.3:

$$P\left(|r_1^{(n)}| \geq C_p N^{-1/d}\log(N)\right) \leq N^{-p}. \tag{3.12}$$

Similarly, the same bound must also apply hold to other $r_i^{(n)}$ with $i=2,\cdots,N$, which leads to

$$P\left(|r_i^{(n)}|\geq C_p N^{-1/d}\log(N) \text{ for some } i\in I, n\Delta t\leq T_1\right)\leq N^{2-p}. \tag{3.13}$$

As a direct result of inequality (3.13), we obtain inequality (3.5) by the definition of ℓ_2 norm. ∎

Recall the definition (1.11) of $G_N(X_i^{(n)})$ and the definition (3.2) of $G_N(Y_i^{(n)})$. We obtain the following proposition of stability.

PROPOSITION 3.2. *Suppose that $\varepsilon\geq N^{-\frac{1}{d(d+1)}}\log(N)$, vectors $X^{(n)}:=\left(X_i^{(n)}\right)_{i\in I}$ and $Y^{(n)}:=\left(Y_i^{(n)}\right)_{i\in I}$ satisfy equations (1.10) and (1.16) respectively. If we denote events*

$$\mathcal{A}:=\left\{\left\|G_N(X^{(n)})-G_N(Y^{(n)})\right\|_{\ell_2}<C\|X^{(n)}-Y^{(n)}\|_{\ell_2} \text{ for any } n\Delta t\leq T_1\right\}, \tag{3.14}$$

$$\mathcal{B}:=\left\{\max_{0\leq n\Delta t\leq T_1}\|X^{(n)}-Y^{(n)}\|_{\ell_2}\leq N^{-\frac{1}{d}}\log^{\frac{3}{2}}(N)\right\}, \tag{3.15}$$

and

$$\mathcal{L}:=\left\{\max_{x}\max_{0\leq n\Delta t\leq T_1}\left|\frac{1}{N-1}\sum_{j=1}^{N}\nabla F_\varepsilon\left(x-Y_j^{(n)}\right)\right|<C\right\}, \tag{3.16}$$

where C depends only on $p>0$, T_1 and $\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}$, then for $T_1/N\leq\Delta t\leq C_*$ (T_1, C_* are used in Lemma 2.2), we have

$$\mathcal{B}\cap\mathcal{L}\subset\mathcal{A}. \tag{3.17}$$

Here the event $\mathcal{A}$ can be seen as the stability result and the event $\mathcal{B}\cap\mathcal{L}$ can be treated as the stability condition.

Proof. First, we split $G_N(X_i^{(n)})-G_N(Y_i^{(n)})$ into two parts:

$$\begin{aligned}
G_N(X_i^{(n)})-G_N(Y_i^{(n)})&=\frac{1}{N-1}\sum_{j=1}^{N}\left(F_\varepsilon(X_i^{(n)}-X_j^{(n)})-F_\varepsilon(Y_i^{(n)}-Y_j^{(n)})\right)\\
&=\frac{1}{N-1}\sum_{j=1}^{N}\left(F_\varepsilon(X_i^{(n)}-X_j^{(n)})-F_\varepsilon(Y_i^{(n)}-X_j^{(n)})\right)\\
&\quad+\frac{1}{N-1}\sum_{j=1}^{N}\left(F_\varepsilon(Y_i^{(n)}-X_j^{(n)})-F_\varepsilon(Y_i^{(n)}-Y_j^{(n)})\right)\\
&=:I_{1,i}+I_{2,i}.
\end{aligned} \tag{3.18}$$

To estimate $I_{1,i}$, one has

$$I_{1,i}=\frac{1}{N-1}\sum_{j=1}^{N}\int_0^1\partial_s F_\varepsilon\left(sX_i^{(n)}+(1-s)Y_i^{(n)}-X_j^{(n)}\right)ds$$

$$
\begin{aligned}
&= \left[\int_0^1 \frac{1}{N-1}\sum_{j=1}^{N}\nabla F_\varepsilon\big(sX_i^{(n)}+(1-s)Y_i^{(n)}-X_j^{(n)}\big)\,ds\right]\big(X_i^{(n)}-Y_i^{(n)}\big)\\
&=: \int_0^1 f_i(s)\,ds(X_i^{(n)}-Y_i^{(n)}). \qquad (3.19)
\end{aligned}
$$

Notice that

$$
|f_i|\le\left|\frac{1}{N-1}\sum_{j=1}^{N}\nabla F_\varepsilon\big(sX_i^{(n)}+(1-s)Y_i^{(n)}-Y_j^{(n)}\big)\right|+\frac{C}{\varepsilon^{d+1}(N-1)}\sum_{j=1}^{N}|X_j^{(n)}-Y_j^{(n)}|. \tag{3.20}
$$

The first term in inequality (3.20) can be bounded under the event $\mathcal{L}$:

$$
\max_i \max_{0\le n\Delta t\le T_1}\left|\frac{1}{N-1}\sum_{j=1}^{N}\nabla F_\varepsilon\big(sX_i^{(n)}+(1-s)Y_i^{(n)}-Y_j^{(n)}\big)\right|<C, \tag{3.21}
$$

for $\varepsilon\ge N^{-\frac{1}{2d}}\log(N)$. To estimate the second term in inequality (3.20), since we assume that $\varepsilon\ge N^{-\frac{1}{d(d+1)}}\log(N)$, using the Hölder inequality, one has

$$
\frac{C}{\varepsilon^{d+1}(N-1)}\sum_{j=1}^{N}|X_j^{(n)}-Y_j^{(n)}|\le C\log^{-\frac{1}{2}-d}(N)\le C, \tag{3.22}
$$

under the event $\mathcal{B}$.

Hence, it follows from inequalities (3.21) and (3.22) that

$$
\max_i \max_{0\le n\Delta t\le T_1}|f_i|<C, \tag{3.23}
$$

under the event $\mathcal{B}\cap\mathcal{L}$, which leads to the following bound of $I_{1,i}$ under the event $\mathcal{B}\cap\mathcal{L}$

$$
\frac{1}{N}\sum_{i=1}^{N}|I_{1,i}|^2<\frac{C}{N}\sum_{i=1}^{N}|X_i^{(n)}-Y_i^{(n)}|^2 \text{ for any } n\Delta t\le T_1. \tag{3.24}
$$

Next, we will estimate $I_{2,i}$:

$$
\begin{aligned}
I_{2,i}&=\frac{1}{N-1}\sum_{j=1}^{N}\nabla F_\varepsilon\big(Y_i^{(n)}-Y_j^{(n)}\big)\big(X_j^{(n)}-Y_j^{(n)}\big)+\frac{1}{N-1}\sum_{j=1}^{N}\nabla^2F_\varepsilon(\zeta_{ij})\big(X_j^{(n)}-Y_j^{(n)}\big)^2\\
&=: g_{1i}+g_{2i}. \qquad (3.25)
\end{aligned}
$$

For g_{1i}, by Young's inequality, under the event $\mathcal{L}$, we conclude that

$$
\|(g_{1i})_{i\in I}\|_{\ell_2}<C\|X^{(n)}-Y^{(n)}\|_{\ell_2} \text{ for any } n\Delta t\le T_1. \tag{3.26}
$$

For g_{2i}, we notice that

$$
|g_{2i}|\le\frac{C}{\varepsilon^{d+1}}\|X^{(n)}-Y^{(n)}\|_{\ell_2}^2, \tag{3.27}
$$

which leads to

$$\|(g_{2i})_{i\in I}\|_{\ell_2} \le \frac{C}{\varepsilon^{d+1}} \|X^{(n)} - Y^{(n)}\|_{\ell_2} \|X^{(n)} - Y^{(n)}\|_{\ell_2} \le C\|X^{(n)} - Y^{(n)}\|_{\ell_2} \tag{3.28}$$

under the event $\mathcal{B}$, since the assumption $\varepsilon \ge N^{-\frac{1}{d(d+1)}} \log(N)$. Hence, it follows from inequalities (3.26) and (3.28) that

$$\|(I_{2,i})_{i\in I}\|_{\ell_2} < C\|X^{(n)} - Y^{(n)}\|_{\ell_2} \text{ for any } n\Delta t \le T_1 \tag{3.29}$$

under the event $\mathcal{B}\cap\mathcal{L}$. Combing inequalities (3.24) and (3.29) and equation (3.18) implies that

$$\mathcal{B}\cap\mathcal{L} \subset \mathcal{A}. \tag{3.30}$$

□

As a direct result of the Proposition 3.1 and Proposition 3.2, we have the following theorem:

THEOREM 3.1. *Under the same assumption as Proposition 3.2, then for $T_1/N \le \Delta t \le C_*$ (T_1, C_* are used in Lemma 2.2) and $n\Delta t \le T_1$, $Y^{(n)}$ is a good approximation of $X^{(n)}$, and the following estimate holds*

$$P\left(\max_{0\le n\Delta t\le T_1} \|X^{(n)} - Y^{(n)}\|_{\ell_2} < C_1 N^{-1/d} \log(N)\right) \ge 1 - C_2 N^{3-p}, \tag{3.31}$$

where C_1, C_2 depend only on $p > 3$, T_1 and $\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}$.

Proof. First, we prove the following inequality by the induction on n.

$$P\left(\max_{0\le k\le n} \|X^{(k)} - Y^{(k)}\|_{\ell_2} \le C_1^{(n)} N^{-1/d} \log(N)\right) \ge 1 - 3nN^{2-p}, \tag{3.32}$$

where

$$C_1^{(n)} := \frac{C_4}{C_3}[(1 + C_3\Delta t)^n - 1] \le \frac{C_4}{C_3}\exp(C_3 n\Delta t) \le \frac{C_4}{C_3}\exp(C_3 T_1), \tag{3.33}$$

with C_3, C_4 given below. It is obvious that inequality (3.32) holds for $n = 0$, since $P\left(\|X^{(0)} - Y^{(0)}\|_{\ell_2} = 0\right) = 1$. Assume that inequality (3.32) holds true up to n. That is, if we denote

$$\mathcal{A}^n := \left\{\max_{0\le k\le n} \|X^{(k)} - Y^{(k)}\|_{\ell_2} \le C_1^{(n)} N^{-1/d} \log(N)\right\},$$

then one has

$$P((\mathcal{A}^n)^c) \le 3nN^{2-p}. \tag{3.34}$$

If we denote the event

$$\mathcal{B} := \left\{\max_{0\le k\le n} \|X^{(k)} - Y^{(k)}\|_{\ell_2} \le N^{-\frac{1}{d}} \log^{\frac{3}{2}}(N)\right\}, \tag{3.35}$$

then we have

$$\mathcal{A}^n \subset \mathcal{B}, \text{ for } N \text{ sufficiently large.} \tag{3.36}$$

In addition, we define the event

$$\mathcal{A}_1^n := \left\{ \left\| G_N(X^{(k)}) - G_N(Y^{(k)}) \right\|_{\ell_2} < C_3 \|X^{(k)} - Y^{(k)}\|_{\ell_2} \text{ for any } k \leq n \right\}, \tag{3.37}$$

where C_3 is independent of n, then it follows from (3.17) in Proposition 3.2 that

$$\mathcal{A}^n \cap \mathcal{L} \subset \mathcal{B} \cap \mathcal{L} \subset \mathcal{A}_1^n. \tag{3.38}$$

Now, we show inequality (3.32) holds for $n+1$. Indeed, notice that

$$X^{(k+1)} - Y^{(k+1)} = X^{(k)} - Y^{(k)} + \Delta t \left(G_N(X^{(k)}) - G_N(Y^{(k)}) \right) - \Delta t r^{(k)}, \tag{3.39}$$

then one concludes that

$$\max_{0\leq k\leq n} \|X^{(k+1)} - Y^{(k+1)}\|_{\ell_2} < (1 + C_3 \Delta t) \max_{0\leq k\leq n} \|X^{(n)} - Y^{(n)}\|_{\ell_2} + \Delta t \max_{0\leq k\leq n} \|r^{(k)}\|_{\ell_2}, \tag{3.40}$$

under the event $\mathcal{A}_1^n$.

Moreover, we define event

$$\mathcal{A}_2^n := \left\{ \|r^{(k)}\|_{\ell_2} < C_4 N^{-1/d} \log(N) \text{ for any } k \leq n \right\}, \tag{3.41}$$

where C_4 is independent of n. Thus we have

$$\begin{aligned} \max_{0\leq k\leq n} \|X^{(k+1)} - Y^{(k+1)}\|_{\ell_2} &< (1 + C_3 \Delta t) \max_{0\leq k\leq n} \|X^{(k)} - Y^{(k)}\|_{\ell_2} + \Delta t \max_{0\leq k\leq n} \|r^{(k)}\|_{\ell_2} \\ &< (C_1^{(n)} + C_1^{(n)} C_3 \Delta t + C_4 \Delta t) N^{-1/d} \log(N) \\ &= C_1^{(n+1)} N^{-1/d} \log(N), \end{aligned} \tag{3.42}$$

under the event $\mathcal{A}^n \cap \mathcal{A}_1^n \cap \mathcal{A}_2^n \supset \mathcal{A}^n \cap \mathcal{B} \cap \mathcal{L} \cap \mathcal{A}_2^n \supset \mathcal{A}^n \cap \mathcal{L} \cap \mathcal{A}_2^n$, where we have used the definition of $C_1^{(n)}$ in (3.33).

Now collecting inequality (3.5) in Proposition 3.1, inequality (2.6) in Lemma 2.4 and the induction assumption (3.34), one has

$$\begin{aligned} &P\left(\max_{0\leq k\leq n} \|X^{(k+1)} - Y^{(k+1)}\|_{\ell_2} \geq C_1^{(n+1)} N^{-1/d} \log(N) \right) \\ &\leq P((\mathcal{A}^n \cap \mathcal{A}_1^n \cap \mathcal{A}_2^n)^c) \leq P((\mathcal{A}^n \cap \mathcal{L} \cap \mathcal{A}_2^n)^c) \\ &\leq P(\mathcal{L}^c) + P((\mathcal{A}^n)^c) + P((\mathcal{A}_2^n)^c) \leq N^{1-p} + 3nN^{2-p} + N^{2-p} \leq 3(n+1)N^{2-p}, \end{aligned} \tag{3.43}$$

which leads to

$$P\left(\max_{0\leq k\leq n+1} \|X^{(k)} - Y^{(k)}\|_{\ell_2} \leq C_1^{(n+1)} N^{-1/d} \log(N) \right) \geq 1 - 3(n+1) N^{2-p}. \tag{3.44}$$

Hence, we finish the proof of inequality (3.32) by induction. Using the fact $n \leq N$ from inequality (2.22) and definition (3.33), inequality (3.32) implies the theorem. ∎

4. Convergence analysis and the proof of Theorem 1.1

In order to prove the error estimate between ρ and $\mu_X^{(n)}$, let us split the error into three parts

$$\|\rho - \mu_X^{(n)}\|_2 \leq \|\rho - \rho^{(n)}\|_2 + \|\rho^{(n)} - \mu_Y^{(n)}\|_2 + \|\mu_Y^{(n)} - \mu_X^{(n)}\|_2. \tag{4.1}$$

Then the idea of the proof of Theorem 1.1 is to obtain the error estimates of those three parts respectively.

4.1. The error estimate between ρ and $\rho^{(n)}$.

LEMMA 4.1. *Under the same assumption as Theorem 1.1, the solution to the splitting algorithm from equations (1.13)–(1.15) is convergent to $\rho(t_n,x)$ in L^2 norm. There exists some $T_* := \min\{T, T_1\}$ such that the following estimate*

$$\max_{0\le n\Delta t\le T_*} \|\rho^{(n)} - \rho(t_n,\cdot)\|_2 \le C(T_*, \|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)})(\Delta t+\varepsilon), \tag{4.2}$$

holds for $\Delta t \le C_*$ and $(n+1)\Delta t \le T_*$ (T_1, C_* are used in Lemma 2.2).

The process of the proof of this lemma is almost the same as the proof of [18, Theorem 1.2]. The only difference is that we consider F_ε here instead of F. Hence we omit the details.

4.2. The error estimate between $\rho^{(n)}$ and $\mu_Y^{(n)}$. Before we establish the error estimate between $\rho^{(n)}$ and $\mu_Y^{(n)}$, let us introduce the following lemma about kernel density estimation.

LEMMA 4.2. *Assume that $\{Y_i^{(n)}\}_{i=1}^N$ are i.i.d. random vectors that we have constructed in definition (1.16), and they share with the common density $\rho^{(n)}$. Let $\mu_Y^{(n)}$ be the regularized empirical measure of $\{Y_i^{(n)}\}_{i=1}^N$ as in definition (1.18), then we have the following mean integrated squared error estimate*

$$\mathbb{E}\left[\|\rho^{(n)} - \mu_Y^{(n)}\|_2^2\right] \le CN^{-1}\delta^{-d} + C\delta^4, \tag{4.3}$$

where C depends only on T_1 (T_1 is used in Lemma 2.2), $\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}$ and $\int_{\mathbb{R}^d} |x|^2\varphi(x)\,dx$.

Proof. This lemma can be found in [27, Chap. 4], and for completeness and precise regularity used in the bound, we give the sketch of the proof here. First, we calculate that

$$\begin{aligned}
\mathbb{E}[\mu_Y^{(n)}] &= \int_{\mathbb{R}^d} \rho^{(n)}(x-y)\frac{1}{\delta^d}\varphi\left(\frac{y}{\delta}\right)dy = \int_{\mathbb{R}^d} \rho^{(n)}(x-\delta z)\varphi(z)\,dz \\
&= \int_{\mathbb{R}^d} \varphi(z)\left(\rho^{(n)}(x) - (\delta z)^T\cdot\nabla\rho^{(n)}(x) + \frac{1}{2}\delta^2 z^T D^2\rho^{(n)}(x)z\right)dz + o(\delta^2) \\
&= \rho^{(n)}(x) - \delta\int_{\mathbb{R}^d} z^T\cdot\nabla\rho^{(n)}(x)\varphi(z)\,dz + \frac{1}{2}\delta^2\int_{\mathbb{R}^d} z^T D^2\rho^{(n)}(x)z\varphi(z)\,dz + o(\delta^2) \\
&= \rho^{(n)}(x) + \frac{1}{2}\delta^2 \mathrm{trace}\left(D^2\rho^{(n)}(x)\int_{\mathbb{R}^d} zz^T\varphi(z)\,dz\right) + o(\delta^2) \\
&= \rho^{(n)}(x) + \frac{1}{2}\delta^2\sum_{i=1}^d \rho^{(n)}_{x_ix_i}(x)\int_{\mathbb{R}^d}|z_i|^2\varphi(z)\,dz + o(\delta^2).
\end{aligned} \tag{4.4}$$

Hence,

$$\|\mathbb{E}[\mu_Y^{(n)}] - \rho^{(n)}\|_2^2 \le C\left(\int_{\mathbb{R}^d}|z|^2\varphi(z)\,dz, \|\rho^{(n)}\|_{H^2}\right)\delta^4. \tag{4.5}$$

Then, we estimate $\mathrm{Var}(\mu_X^{(n)})$ as

$$\mathrm{Var}(\mu_Y^{(n)}) = \frac{1}{N}\left[\frac{1}{\delta^d}\int_{\mathbb{R}^d}\varphi^2(z)\rho^{(n)}(x-\delta z)\,dz - \left(\int_{\mathbb{R}^d}\varphi(z)\rho^{(n)}(x-\delta z)\,dz\right)^2\right]$$

$$=\frac{1}{N\delta^d}\rho^{(n)}(x)\int_{\mathbb{R}^d}\varphi^2(z)\,dz+o(\frac{1}{N\delta^d}). \tag{4.6}$$

Notice that

$$\mathbb{E}\left[|\rho^{(n)}-\mu_Y^{(n)}|^2\right]=\mathrm{Var}(\mu_Y^{(n)})+\left(\mathbb{E}[\mu_Y^{(n)}]-\rho^{(n)}\right)^2. \tag{4.7}$$

Integrating equation (4.7) over x in $\mathbb{R}^d$ and applying inequality (4.5) with equation (4.6), one has

$$\mathbb{E}\left[\|\rho^{(n)}-\mu_Y^{(n)}\|_2^2\right]\le C\frac{1}{N\delta^d}+C\delta^4, \tag{4.8}$$

which concludes the proof of this lemma. □

As a direct result of Lemma 4.2, we can get the distance between $\rho^{(n)}$ and $\mu_Y^{(n)}$. Indeed, one has

$$\begin{aligned}
&P\left(\max_{0\le n\Delta t\le T_*}\|\rho^{(n)}-\mu_Y^{(n)}\|_2\ge N^{-\frac{1}{d(d+2)}}\right)=P\left(\max_{0\le n\Delta t\le T_*}\|\rho^{(n)}-\mu_Y^{(n)}\|_2^2\ge N^{-\frac{2}{d(d+2)}}\right)\\
=&\mathbb{E}\left[1_{\left\{\max_{0\le n\Delta t\le T_*}\|\rho^{(n)}-\mu_Y^{(n)}\|_2^2\ge N^{-\frac{2}{d(d+2)}}\right\}}\right]\\
\le&\mathbb{E}\left[\max_{0\le n\Delta t\le T_*}\|\rho^{(n)}-\mu_Y^{(n)}\|_2^2N^{\frac{2}{d(d+2)}}1_{\left\{\max_{0\le n\Delta t\le T_*}\|\rho^{(n)}-\mu_Y^{(n)}\|_2^2\ge N^{-\frac{2}{d(d+2)}}\right\}}\right]\\
\le&N^{\frac{2}{d(d+2)}}\mathbb{E}\left[\max_{0\le n\Delta t\le T_*}\|\rho^{(n)}-\mu_Y^{(n)}\|_2^2\right].
\end{aligned} \tag{4.9}$$

If we choose $\delta=N^{-\frac{1}{d(d+2)}}$ in estimate (4.3), then one has

$$P\left(\max_{0\le n\Delta t\le T_*}\|\rho^{(n)}-\mu_Y^{(n)}\|_2\ge N^{-\frac{1}{d(d+2)}}\right)\le CN^{-\frac{2}{d(d+2)}}, \tag{4.10}$$

which gives the probability bound of the error between $\rho^{(n)}$ and $\mu_Y^{(n)}$.

4.3. The error estimate between $\mu_X^{(n)}$ and $\mu_Y^{(n)}$. Recall that

$$\mu_X^{(n)}:=\frac{1}{N}\sum_{i=1}^N\varphi_\delta\left(x-X_i^{(n)}\right), \tag{4.11}$$

and

$$\mu_Y^{(n)}:=\frac{1}{N}\sum_{i=1}^N\varphi_\delta\left(x-Y_i^{(n)}\right). \tag{4.12}$$

The L^2 norm of the difference between $\mu_X^{(n)}$ and $\mu_Y^{(n)}$ is given by

$$\|\mu_X^{(n)}-\mu_Y^{(n)}\|_2=\|\widehat{\mu^{(n)}}_X-\widehat{\mu^{(n)}}_Y\|_2\le\frac{1}{N}\sum_{i=1}^N\left\|\widehat{\varphi}(\delta\xi)(e^{-i\xi X_i^{(n)}}-e^{-i\xi Y_i^{(n)}})\right\|_2, \tag{4.13}$$

where we have used the properties of Fourier transformation.

Since $|e^{-i\xi X_i^{(n)}} - e^{-i\xi Y_i^{(n)}}| \le |\xi||X_i^{(n)} - Y_i^{(n)}|$, one has

$$\begin{aligned}\|\mu_X^{(n)} - \mu_Y^{(n)}\|_2 &\le \frac{1}{N}\sum_{i=1}^N |X_i^{(n)} - Y_i^{(n)}| \big\|\widehat{\varphi}(\delta\xi)|\xi|\big\|_2 \\ &\le \big\|\widehat{\varphi}(\delta\xi)|\xi|\big\|_2 \|X^{(n)} - Y^{(n)}\|_{\ell_2}. \end{aligned} \tag{4.14}$$

Notice that

$$\big\|\widehat{\varphi}(\delta\xi)|\xi|\big\|_2^2 = \delta^{-(d+2)} \int_{\mathbb{R}^d} \hat{\varphi}^2(y)|y|^2 dy \le C\delta^{-(d+2)}, \tag{4.15}$$

which leads to

$$\|\mu_X^{(n)} - \mu_Y^{(n)}\|_2 \le C\delta^{-1-d/2} \|X^{(n)} - Y^{(n)}\|_{\ell_2}. \tag{4.16}$$

So applying Theorem 3.1 and choosing $\delta = N^{-\frac{1}{d(d+2)}}$, one concludes that

$$P\left(\max_{0\le n\Delta t\le T_*} \|\mu_X^{(n)} - \mu_Y^{(n)}\|_2 < CN^{-\frac{1}{2d}} \log(N)\right) \ge 1 - CN^{3-p}. \tag{4.17}$$

4.4. The proof of Theorem 1.1. Finally, collecting inequalities (4.2), (4.10) and (4.17), we compute that:

$$\begin{aligned} &P\left(\max_{0\le n\Delta t\le T_*} \|\rho - \mu_X^{(n)}\|_2 \ge 3C\left(N^{-\frac{1}{d(d+2)}} + \Delta t + \varepsilon\right)\right) \\ \le & P\left(\max_{0\le n\Delta t\le T_*} \{\|\rho - \rho^{(n)}\|_2 + \|\rho^{(n)} - \mu_Y^{(n)}\|_2 + \|\mu_Y^{(n)} - \mu_X^{(n)}\|_2\} \ge 3C\left(N^{-\frac{1}{d(d+2)}} + \Delta t + \varepsilon\right)\right) \\ \le & P\left(\max_{0\le n\Delta t\le T_*} \|\rho - \rho^{(n)}\|_2 \ge C(\Delta t + \varepsilon)\right) + P\left(\max_{0\le n\Delta t\le T_*} \|\rho^{(n)} - \mu_Y^{(n)}\|_2 \ge CN^{-\frac{1}{d(d+2)}}\right) \\ &\quad + P\left(\max_{0\le n\Delta t\le T_*} \|\mu_Y^{(n)} - \mu_X^{(n)}\|_2 \ge CN^{-\frac{1}{2d}} \log(N)\right) \\ \le & 0 + CN^{-\frac{2}{d(d+2)}} + CN^{3-p} \le CN^{-\frac{2}{d(d+2)}}, \end{aligned} \tag{4.18}$$

for $p \ge 3 + \frac{2}{d(d+2)}$. Hence Theorem 1.1 has been proved.

5. The error estimate on interaction

THEOREM 5.1. *For $2 \le d \le 3$, suppose that $0 \le \rho_0(x) \in L^1 \cap H^k(\mathbb{R}^d)$ with $k > d/2 + 3$ and F, F_ε satisfy equations (1.4), (1.9) respectively. Let $\rho(t,x)$ be the regular solution of the KS Equation (1.1) with local existence time T and $\{X_i^{(n)}\}_{i=1}^N$ satisfy equation (1.10) (N sufficiently large). If we choose $N^{-\frac{1}{d(d+1)}} \log(N) = \varepsilon$, then there exists some T_*, C_* depending only on T and $\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}$, for $T_*/N \le \Delta t \le C_*$ and $(n+1)\Delta t \le T_*$, such that the following estimate holds*

$$\begin{aligned} P\Bigg(\max_{0\le n\Delta t\le T_*} \max_x &\left| \int_{\mathbb{R}^d} F(x-y)\rho(t_n,y)dy - \frac{1}{N-1}\sum_{j=1}^N F_\varepsilon(x - X_j^{(n)})\right| \\ &< C_1\left(\varepsilon^{1-\frac{d}{2}}\Delta t + \varepsilon^{2-\frac{d}{2}} + N^{-\frac{1}{d(d+1)}}\right)\Bigg) \\ \ge 1 - C_2 N^{3-p},& \end{aligned} \tag{5.1}$$

where C_1, C_2 *depend only on* $p>3$, T_* *and* $\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}$.

Proof. First, we split the error into two parts:

$$\begin{aligned}
&\left|\int_{\mathbb{R}^d} F(x-y)\rho(t_n,y)\,dy - \frac{1}{N-1}\sum_{j=1}^N F_\varepsilon\big(x-X_j^{(n)}\big)\right| \\
&\le \left|\int_{\mathbb{R}^d} F(x-y)\rho(t_n,y)\,dy - \int_{\mathbb{R}^d} F_\varepsilon(x-y)\rho^{(n)}(y)\,dy\right| \\
&\quad + \left|\int_{\mathbb{R}^d} F_\varepsilon(x-y)\rho^{(n)}(y)\,dy - \frac{1}{N-1}\sum_{j=1}^N F_\varepsilon\big(x-X_j^{(n)}\big)\right| =: e_1^n(x)+e_2^n(x). \qquad (5.2)
\end{aligned}$$

To estimate $e_1^n(x)$, we have

$$\begin{aligned}
e_1^n(x) &\le \left|\int_{\mathbb{R}^d}\big(F(x-y)-F_\varepsilon(x-y)\big)\rho(t_n,y)\,dy\right| + \left|\int_{\mathbb{R}^d} F_\varepsilon(x-y)\big(\rho(t_n,y)-\rho^{(n)}(y)\big)\,dy\right| \\
&=: e_{11}^n(x)+e_{12}^n(x). \qquad (5.3)
\end{aligned}$$

A simple computation we know that

$$e_{11}^n(x)\le C\varepsilon, \qquad (5.4)$$

and

$$e_{12}^n(x)\le \|F_\varepsilon\|_2\left\|\rho(t_n,\cdot)-\rho^{(n)}\right\|_2 \le C\varepsilon^{1-\frac{d}{2}}(\Delta t+\varepsilon), \qquad (5.5)$$

which leads to

$$\max_x e_1^n(x)\le C\left(\varepsilon^{1-\frac{d}{2}}\Delta t+\varepsilon^{2-\frac{d}{2}}\right). \qquad (5.6)$$

As for $e_2^n(x)$, we compute

$$\begin{aligned}
e_2^n(x) &\le \left|\int_{\mathbb{R}^d} F_\varepsilon(x-y)\rho^{(n)}(y)\,dy - \frac{1}{N-1}\sum_{j=1}^N F_\varepsilon\big(x-Y_j^{(n)}\big)\right| \\
&\quad + \left|\frac{1}{N-1}\sum_{j=1}^N F_\varepsilon\big(x-Y_j^{(n)}\big) - \frac{1}{N-1}\sum_{j=1}^N F_\varepsilon\big(x-X_j^{(n)}\big)\right| \\
&=: e_{21}^n(x)+e_{22}^n(x). \qquad (5.7)
\end{aligned}$$

To estimate $e_{21}^n(x)$, we follow the same procedure in the proof of Lemma 2.4. Indeed, similar to definition (2.9), we let $S_n := \max_x e_{21}^n(x)$, and we can prove that

$$P\left(S_n \ge \frac{C_p\log(N)L_2(\varepsilon)}{\sqrt{N}}\right)\le N^{-p}, \qquad (5.8)$$

as presented in inequality (2.19), where $L_2(\varepsilon)=\int_{\mathbb{R}^d}|\widehat{F_\varepsilon}(\xi)|d\xi \le C\varepsilon^{1-d}$.

So if we assume that $\varepsilon \ge N^{-\frac{1}{d(d+1)}}\log(N)\ge N^{-\frac{1}{4(d-1)}}\log(N)$, then it implies that

$$P\left(\max_x e_{21}^n(x) < CN^{-\frac{1}{4}}\log^{2-d}(N) \text{ for any } n\Delta t\le T_*\right)\ge 1-N^{1-p}, \qquad (5.9)$$

similar to inequality (2.24).

To estimate $e_{22}^n(x)$, by Lemma 2.1, one has

$$e_{22}^n(x) \le C\varepsilon^{-d}\frac{1}{N-1}\sum_{j=1}^{N}|X_j^{(n)}-Y_j^{(n)}| \le CN^{\frac{1}{d+1}}\log^{-d}(N)\|X^{(n)}-Y^{(n)}\|_{\ell_2}, \tag{5.10}$$

under the assumption that $\varepsilon \ge N^{-\frac{1}{d(d+1)}}\log(N)$. Then we apply Theorem 3.1 and obtain

$$P\left(\max_x e_{22}^n(x) < CN^{-\frac{1}{d(d+1)}}\log^{1-d}(N) \text{ for any } n\Delta t \le T_*\right) \ge 1 - CN^{3-p}. \tag{5.11}$$

Collecting estimates (5.9) and (5.11), we conclude that

$$P\left(\max_x e_2^n(x) < CN^{-\frac{1}{d(d+1)}} \text{ for any } n\Delta t \le T_*\right) \ge 1 - CN^{3-p}. \tag{5.12}$$

Combining estimates (5.6) and (5.12), one has

$$\begin{aligned}
&P\Bigg(\max_{0\le n\Delta t\le T_*}\max_x\left|\int_{\mathbb{R}^d}F(x-y)\rho(t_n,y)dy - \frac{1}{N-1}\sum_{j=1}^{N}F_\varepsilon\left(x-X_j^{(n)}\right)\right| \\
&\qquad \ge 2C\left(\varepsilon^{1-\frac{d}{2}}\Delta t + \varepsilon^{2-\frac{d}{2}} + N^{-\frac{1}{d(d+1)}}\right)\Bigg) \\
&\le P\left(\max_{0\le n\Delta t\le T_*}\max_x\{e_1^n(x)+e_2^n(x)\} \ge 2C\left(\varepsilon^{1-\frac{d}{2}}\Delta t + \varepsilon^{2-\frac{d}{2}} + N^{-\frac{1}{d(d+1)}}\right)\right) \\
&\le P\left(\max_{0\le n\Delta t\le T_*}\max_x e_1^n(x) \ge C\left(\varepsilon^{1-\frac{d}{2}}\Delta t + \varepsilon^{2-\frac{d}{2}}\right)\right) \\
&\quad + P\left(\max_{0\le n\Delta t\le T_*}\max_x e_2^n(x) \ge CN^{-\frac{1}{d(d+1)}}\right) \\
&\le 0 + CN^{3-p} = CN^{3-p}.
\end{aligned} \tag{5.13}$$

Then we conclude the proof. □

6. Extension to general regular attractive force F

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

$$F \in H^k(\mathbb{R}^d) \text{ with } k > \frac{d}{2} + 3. \tag{6.1}$$

We consider the regular solution ρ of the following Fokker–Planck equation

$$\begin{cases} \partial_t\rho = \nu\Delta\rho - \nabla\cdot(\rho F*\rho), & x\in\mathbb{R}^d,\ t>0, \\ \rho(0,x) = \rho_0(x), \end{cases} \tag{6.2}$$

where $0 \le \rho_0 \in H^k(\mathbb{R}^d)$ with $k > \frac{d}{2}+3$. Then ρ has the following regularity for any $T>0$

$$\|\rho\|_{L^\infty(0,T;H^k(\mathbb{R}^d))} \le C\left(T, \|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}, \|F\|_{H^k(\mathbb{R}^d)}\right). \tag{6.3}$$

Since F is non-singular, there is no need to mollify the force F anymore. To be specific, we consider trajectories $\{X_i(t)\}_{i=1}^N$ satisfying SDEs:

$$dX_i(t) = \frac{1}{N-1}\sum_{j=1}^{N} F\big(X_i(t) - X_j(t)\big)\,dt + \sqrt{2\nu}\,dB_i(t), \quad i = 1, \cdots, N, \tag{6.4}$$

where $\{B_i(t)\}_{i=1}^N$ are N independent standard Brownian motions. To discretize this system, one has

$$X_i^{(n+1)} = X_i^{(n)} + \Delta t G_N\big(X_i^{(n)}\big) + \sqrt{2\nu\Delta t}\,N_i^{(n)}, \quad i = 1, \cdots, N, \tag{6.5}$$

where the $N_i^{(n)}$ are independent standard Gaussian random vectors and

$$G_N\big(X_i^{(n)}\big) := \frac{1}{N-1}\sum_{j=1}^{N} F\big(X_i^{(n)} - X_j^{(n)}\big). \tag{6.6}$$

Moreover, the splitting scheme we constructed from equations (1.13)–(1.15) becomes

$$G^{(n)}(x) = F * \rho^{(n)}(x), \tag{6.7}$$

$$\rho^{(n+\frac{1}{2})}\big(x + \Delta t G^{(n)}(x)\big) = \det{}^{-1}\big(I + \Delta t DG^{(n)}(x)\big)\rho^{(n)}(x), \tag{6.8}$$

$$\rho^{(n+1)}(x) = H(\sqrt{\nu}\Delta t)\rho^{(n+\frac{1}{2})}(x). \tag{6.9}$$

To approximate $X_i^{(n)}$, we define $Y_i^{(0)} = X_i^{(0)}$ and

$$Y_i^{(n+1)} = Y_i^{(n)} + \Delta t G^{(n)}\big(Y_i^{(n)}\big) + \sqrt{2\nu\Delta t}\,N_i^{(n)}, \quad i = 1, \cdots, N, \tag{6.10}$$

where $G^{(n)}\big(Y_i^{(n)}\big)$ are the vector fields constructed in definition (6.7). Note that $Y_i^{(n)}$ is independent of $Y_j^{(n)}$ if $i \neq j$, and they share the common density $\rho^{(n)}$.

Then the regularized empirical measure of $\{X_i^{(n)}\}_{i=1}^N$ can be defined as

$$\mu_X^{(n)}(x) := \frac{1}{N}\sum_{i=1}^{N} \varphi_\delta\big(x - X_i^{(n)}\big), \tag{6.11}$$

and similarly, we can define

$$\mu_Y^{(n)}(x) := \frac{1}{N}\sum_{i=1}^{N} \varphi_\delta\big(x - Y_i^{(n)}\big). \tag{6.12}$$

First, similar to Theorem 3.1, we have the following extended result.

THEOREM 6.1. *Suppose that* $0 \le \rho_0(x) \in L^1 \cap H^k(\mathbb{R}^d)$ *and* $F \in H^k(\mathbb{R}^d)$ *with* $k > d/2 + 3$. *Let* $\rho(t,x)$ *be the regular solution of equation* (6.2) *and* $\{X_i^{(n)}\}_{i=1}^N$, $\{Y_i^{(n)}\}_{i=1}^N$ *(N sufficiently large) satisfy equations* (6.5) *and* (6.10) *respectively. Then for any* $T > 0$, *there exists some* C_* *depending only on* T, $\|F\|_{H^k(\mathbb{R}^d)}$ *and* $\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}$, *for* $T/N \le$

$\Delta t \le C_*$, such that $Y^{(n)}$ is a good approximation of $X^{(n)}$, and the following estimate holds

$$P\left(\max_{0\le n\Delta t\le T}\|X^{(n)}-Y^{(n)}\|_{\ell_2}<C_1N^{-\frac{1}{2}}\log(N)\right)\ge 1-C_2N^{3-p}, \tag{6.13}$$

where C_1, C_2 depend only on $p>3$, T, $\|F\|_{H^k(\mathbb{R}^d)}$ and $\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}$.

Proof. As we have done in Proposition 3.1, for

$$r_i^{(n)}=G^{(n)}(Y_i^{(n)})-\frac{1}{N-1}\sum_{j=1}^{N}F(Y_i^{(n)}-Y_j^{(n)}), \tag{6.14}$$

we prove that

$$P\left(\max_{0\le n\Delta t\le T}\|r^{(n)}\|_{\ell_2}\ge CN^{-1/2}\log(N)\right)\le N^{2-p}. \tag{6.15}$$

Then, similar to Proposition 3.2, we have the stability result

$$\left\|G_N(X^{(n)})-G_N(Y^{(n)})\right\|_{\ell_2}<C\|X^{(n)}-Y^{(n)}\|_{\ell_2},\ \text{for any } n\Delta t\le T, \tag{6.16}$$

under the following event

$$\mathcal{B}:=\left\{\max_{0\le n\Delta t\le T}\|X^{(n)}-Y^{(n)}\|_{\ell_2}<N^{-\frac{1}{2}}\log^{\frac{3}{2}}(N)\right\}. \tag{6.17}$$

Finally, following the approach of the proof in Theorem 3.1, inequality (6.13) can be obtained from inequality (6.15). ∎

Next we can extend the result in Theorem 1.1 to the following theorem.

THEOREM 6.2. *Under the same assumption as in Theorem 6.1, let $\mu_X^{(n)}$ be the regularized empirical measure as in definition (6.11) with $\delta=N^{-\frac{1}{2(d+2)}}$ (N sufficiently large). Then for any $T>0$, there exists some C_* depending only on T, $\|F\|_{H^k(\mathbb{R}^d)}$ and $\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}$, for $T/N\le\Delta t\le C_*$ and $(n+1)\Delta t\le T$, such that the following estimate holds*

$$P\left(\max_{0\le n\Delta t\le T}\|\rho(t_n,\cdot)-\mu_X^{(n)}\|_2<C_1(N^{-\frac{1}{2(d+2)}}+\Delta t)\right)\ge 1-C_2N^{-\frac{1}{d+2}},$$

where C_1, C_2 depend only on T, $\|F\|_{H^k(\mathbb{R}^d)}$ and $\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}$.

Proof. We only give the sketch of the proof here, since it is almost the same as we have done in Theorem 1.1. First, we have the error estimate between ρ and $\rho^{(n)}$

$$\max_{0\le n\Delta t\le T}\|\rho^{(n)}-\rho(t_n,\cdot)\|_2\le C\Delta t. \tag{6.18}$$

Next, we obtain the error estimate between $\rho^{(n)}$ and $\mu_Y^{(n)}$

$$P\left(\max_{0\le n\Delta t\le T}\|\rho^{(n)}-\mu_Y^{(n)}\|_2\ge N^{-\frac{1}{2(d+2)}}\right)\le CN^{-\frac{1}{d+2}}. \tag{6.19}$$

Furthermore, we can prove the error estimate between $\mu_X^{(n)}$ and $\mu_Y^{(n)}$

$$P\left(\max_{0\le n\Delta t\le T}\|\mu_X^{(n)}-\mu_Y^{(n)}\|_2<CN^{-\frac{1}{4}}\log(N)\right)\ge 1-CN^{3-p}. \tag{6.20}$$

Collecting inequalities (6.18)–(6.20), we conclude the proof. ∎

Moreover, we can extend the result in Theorem 5.1.

THEOREM 6.3. *For $d \geq 2$, suppose that $0 \leq \rho_0(x) \in L^1 \cap H^k(\mathbb{R}^d)$, $F \in H^k(\mathbb{R}^d)$ with $k > d/2+3$. Let $\rho(t,x)$ be the regular solution of equation (6.2) and $\{X_i^{(n)}\}_{i=1}^N$ satisfy equation (6.5) (N sufficiently large). Then for any $T>0$, there exists some C_* depending only on T, $\|F\|_{H^k(\mathbb{R}^d)}$ and $\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}$, for $T/N \leq \Delta t \leq C_*$ and $(n+1)\Delta t \leq T$, such that the following estimate holds*

$$P\Bigg(\max_{0\leq n\Delta t\leq T}\max_x\left|\int_{\mathbb{R}^d} F(x-y)\rho(t_n,y)dy - \frac{1}{N-1}\sum_{j=1}^N F\big(x-X_j^{(n)}\big)\right| < C_1(N^{-\frac12}\log(N)+\Delta t)\Bigg) \geq 1 - C_2 N^{3-p},$$

where C_1, C_2 depend only on $p>1$, T, $\|F\|_{H^k(\mathbb{R}^d)}$ and $\|\rho_0\|_{L^1\cap H^k(\mathbb{R}^d)}$.

Proof. We split the error into two parts:

$$\begin{aligned}
&\left|\int_{\mathbb{R}^d} F(x-y)\rho(t_n,y)\,dy - \frac{1}{N-1}\sum_{j=1}^N F\big(x-X_j^{(n)}\big)\right| \\
&\leq \left|\int_{\mathbb{R}^d} F(x-y)\rho(t_n,y)\,dy - \int_{\mathbb{R}^d} F(x-y)\rho^{(n)}(y)\,dy\right| \\
&\quad + \left|\int_{\mathbb{R}^d} F(x-y)\rho^{(n)}(y)\,dy - \frac{1}{N-1}\sum_{j=1}^N F\big(x-X_j^{(n)}\big)\right| =: e_1^n(x)+e_2^n(x). \qquad (6.21)
\end{aligned}$$

Following the method in Theorem 5.1, it is easy to prove that

$$\max_x e_1^n(x) \leq C\Delta t, \qquad (6.22)$$

and

$$P\left(\max_{0\leq n\Delta t\leq T}\max_x e_2^n(x) < CN^{-\frac12}\log(N)\right) \geq 1 - CN^{3-p}, \qquad (6.23)$$

which leads to inequality (6.21). ∎

REFERENCES

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

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

[3] J. Bedrossian and N. Masmoudi, *Existence, uniqueness and Lipschitz dependence for Patlak–Keller–Segel and Navier–Stokes in $\mathbb{R}^2$ with measure-valued initial data*, Arch. Ration. Mech. Anal., 214(3):717–801, 2014.

[4] A. Blanchet, E.A. Carlen, and J.A. Carrillo, *Functional inequalities, thick tails and asymptotics for the critical mass Patlak–Keller–Segel model*, Journal of Functional Analysis, 262(5):2142–2230, 2012.

[5] A. Blanchet, J. Dolbeault, and B. Perthame, *Two-dimensional Keller–Segel model: optimal critical mass and qualitative properties of the solutions*, Electronic Journal of Differential Equations, 44:1–32, 2006.

[6] A. Blanchet, N. Masmoudi, J.A. Carrillo de la Plata, et al. *Infinite time aggregation for the critical Patlak–Keller–Segel model in* $\mathbb{R}^2$, Commun. Pure Appl. Math., 61(10):1449–1481, 2008.
[7] V. Calvez, L. Corrias, and M.A. Ebde, *Blow-up, concentration phenomenon and global existence for the Keller–Segel model in high dimension*, Commun. Part. Diff. Eqs., 37(4):561–584, 2012.
[8] V. Calvez, L. Corrias, et al. *The parabolic-parabolic Keller-Segel model in* $\mathbb{R}^2$, Commun. Math. Sci., 6(2):417–447, 2008.
[9] A.J. Chorin, *Numerical study of slightly viscous flow*, Journal of Fluid Mechanics, 57(04):785–796, 1973.
[10] G.-H. Cottet and P.D. Koumoutsakos, *Vortex Methods: Theory and Practice*, Cambridge University Press, 2000.
[11] K. Craig and A. Bertozzi, *A blob method for the aggregation equation*, Math. Comput., 85:1681–1717, 2016.
[12] J. Dolbeault and B. Perthame, *Optimal critical mass in the two dimensional Keller–Segel model in* $\mathbb{R}^2$, Comptes Rendus Mathematique, 339(9):611–616, 2004.
[13] J. Goodman, *Convergence of the random vortex method*, Commun. Pure Appl. Math., 40(2):189–220, 1987.
[14] O. Hald and V.M. Del Prete, *Convergence of vortex methods for Euler's equations*, Math. Comput., 32(143):791–809, 1978.
[15] O.H. Hald, *Convergence of vortex methods for Euler's equations II*, SIAM J. Numer. Anal., 16(5):726–755, 1979.
[16] D. Horstmann et al. *From 1970 until present: the Keller–Segel model in chemotaxis and its consequences*, Jahresbericht der Deutschen Mathematiker-Vereinigung, 105(3):103–165, 2003.
[17] H. Huang and J.-G. Liu, *Error estimate of a random particle blob method for the Keller–Segel equation*, Math. Comput., 86(308):2719–2744, 2017.
[18] H. Huang and J.-G. Liu, *Error estimates of the aggregation-diffusion splitting algorithms for the Keller–Segel equations*, Discrete and Continuous Dynamical Systems - Series B, 21(10):3463C–3478, 2016.
[19] 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.
[20] J.-G. Liu and R. Yang, *A random particle blob method for the Keller–Segel equation and convergence analysis*, Math. Comput., 86:725–745, 2017.
[21] J.-G. Liu and Y. Zhang, *Convergence of diffusion-drift many particle systems in probability under Sobolevnorm*, P. Gonçalves and A.J. Soares (eds.), in From Particle Systems to Partial Differential Equations-III, Springer, 195–223, 2016.
[22] D.-G. Long, *Convergence of the random vortex method in two dimensions*, J. Amer. Math. Soci., 1(4):779–804, 1988.
[23] C. Marchioro and M. Pulvirenti, *Vortex methods in two-dimensional fluid dynamics*, Lecture notes in Physics, 203:1–137, 1984.
[24] 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.
[25] C.S. Patlak, *Random walk with persistence and external bias*, The Bulletin of Mathematical Biophysics, 15(3):311–338, 1953.
[26] B. Perthame, *Transport Equations in Biology*, Springer Science & Business Media, 2007.
[27] M.P. Wand and M.C. Jones, *Kernel Smoothing*, Crc Press, 1994.
[28] M. Winkler, *Aggregation vs. global diffusive behavior in the higher-dimensional Keller–Segel model*, J. Diff. Eqs., 248(12):2889–2905, 2010.