Learning interacting particle systems: diffusion parameter estimation for aggregation equations

Hui Huang*, Jian-Guo Liu†, Jianfeng Lu‡

February 8, 2018

Abstract

In this article, we study the parameter estimation of interacting particle systems subject to the Newtonian aggregation. Specifically, we construct an estimator $\widehat{\nu}$ with partial observed data to approximate the diffusion parameter ν, and the estimation error is achieved. Furthermore, we extend this result to general aggregation equations with a bounded Lipschitz interaction field.

Keywords: Inverse problem, parameter identification of agent based mode, mean-field limit, data assimilation, concentration inequality, discrete observation.

1 Introduction

Parameter estimation of (stochastic) dynamical systems has ubiquitous applications in many areas in science and technology. This is often known as data assimilation (see the recent book [32] for a mathematical introduction) in particular in the context of numerical weather forecast. In such problems, a physical model of the form of a dynamical system is proposed with a few unknown parameters, (partial) observations of the evolution of the system are then used to determine those parameters and to give predictions of the dynamics in the future.

In this work, we are interested in a particular class of physical systems that can be modeled by interacting particle systems. This means that the dynamics of the system is determined by interactions between agents (particles) together with some intrinsic or extrinsic random effects. Such systems are widely used to establish different mathematical models describing collective behaviors of organisms and social aggregations, for instance flocks of birds [23], aggregation of bacteria [3], schools of fish [22], swarms formed by insects [4], opinion dynamics [35] and robotics and space missions [29]. We study the parameter estimation of such interacting particle systems with partial observed data.

More precisely, the microscopic agent-based model investigated here describes the evolution of positions of N agents, denoted by $\{X_i^t\} \subset \mathbb{R}^d$, $i = 1, \ldots, N$, whose evolution is governed by a system of stochastic differential equations (SDEs) of the type

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

where F models some pairwise interaction between the agents and B_i^t are independent realizations of Brownian motions which count for extrinsic random perturbation of the agent positions. In such systems, the agents are assumed to be identical, so that the noise level ν is the same for each agent. In this work, we assume that the interaction kernel F between

*Department of Mathematics, Simon Fraser University, Burnaby, BC, Canada. Email: hha101@sfu.ca

†Department of Mathematics and Department of Physics, Duke University, Durham, NC, USA. Email: jliu@phy.duke.edu

‡Department of Mathematics, Department of Physics and Department of Chemistry, Duke University, Durham, NC, USA. Email: jianfeng@math.duke.edu

agents are known, while the noise level ν is to be determined. More specifically, we will focus on the case when the interaction between agents are given by Newtonian type interaction for dimension $d \geq 2$, or more precisely, a regularized Newtonian interaction, to be specified below. Suppose we observe or track the trajectories of K agents on the time interval $[0, T]$, where $1 \ll K \ll N$, the question we address in this work is *how to estimate ν and to quantify the error of the estimator.*

A more general situation one may consider is the problem for which the interaction kernel is also to be determined, this will be left for future works. We note the recent work [7] which considers learning the interaction kernel for a deterministic interacting particle systems through a variational approach.

Observe that the scaling of (1) is chosen such that we are in the mean-field regime, as the interaction strength decreases as $1/N$ as the number of agents $N \to \infty$. It is thus expected that in the limit $N \to \infty$, the system can be well described by a mean-field dynamics, which can be described as the following nonlinear partial differential equation (PDE)

$$\partial_t \rho = \nu \Delta \rho - \nabla \cdot (\rho F * \rho), \quad x \in \mathbb{R}^d, \ t > 0, \tag{2a}$$

$$\rho(x, 0) = \rho_0(x), \tag{2b}$$

where the noise level $\nu > 0$ enters the PDE system as a diffusion parameter. In particular, here the interaction kernel is chosen as Newtonian:

$$F(x) = \mp \frac{C_* x}{|x|^d}, \quad \forall\, x \in \mathbb{R}^d \backslash \{0\}, \ d \geq 2, \tag{3}$$

with $C_* = \frac{\Gamma(d/2)}{2\pi^{d/2}}$. Here the sign $\mp$ indicates that the interaction between individuals can either be attraction or repulsion. Specifically, when the mechanism of interaction is attraction, the mean field equation (2) becomes the parabolic-elliptic Keller-Segel equation [30, 36], which is a prototypical model for chemotaxis and has been used in many related modeling scenarios.

While it would be intriguing to study the parameter identification problem for the particle system (1) with the Newtonian interaction (3), such microscopic system is however ill-posed, as shown by the recent deep result by Fournier and Jourdain [18, Proposition 4]: For any $N \geq 2$ and $T > 0$, denote $\{X_i(t); t \in [0, T], i = 1, \ldots, N\}$ the solution to (1) with F given in (3), then

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

i.e., the singularity cannot be avoided in any finite time with a positive probability and thus the particle system is not well-defined. As a result, on the microscopic level, one has to use a regularized interaction kernel for (1). We will in this work consider the regularized kernel F^N:

$$F^N = F * \psi_N, \quad \psi_N(x) = N^{d\delta} \psi(N^\delta x), \tag{4}$$

where δ the cut-off index and $0 \leq \psi(x) \in C_0^\infty(\mathbb{R}^d)$ is a cut-off function, which satisfies $\psi(x) = \psi(|x|)$ and $\int_{\mathbb{R}^d} \psi(x)\, dx = 1$.

Then we have the regularized stochastic particle system $\{X_i^t\}_{i=1}^N$ satisfying

$$dX_i^t = \frac{1}{N-1} \sum_{j \neq i}^{N} F^N\left(X_i^t - X_j^t\right) dt + \sqrt{2\nu}\, dB_i^t, \quad i = 1, \cdots, N, \tag{5}$$

where the initial data $\{X_i^0\}_{i=1}^N$ are independently, identically distributed (i.i.d.) with common density function ρ_0. Since the regularized kernel is Lipschitz for fixed N, the system above has a unique global strong solution. The corresponding aggregation equation has the form

$$\partial_t \rho = \nu \Delta \rho - \nabla \cdot (\rho F^N * \rho), \quad x \in \mathbb{R}^d, \ t > 0, \tag{6a}$$

$$\rho(x, 0) = \rho_0(x). \tag{6b}$$

The analysis of the scaling limits of interacting particle system (5) is usually called the *mean-field limits*, which pass limits from microscopic discrete particle systems to macroscopic continuum models. Classical results for globally Lipschitz forces was obtained by Braun and Hepp [9] and Dobrushin [14]. Then Bolley, Cañizo and Carrilo [6] presented an extension of the classical theory to the particle system with only locally Lipschitz interacting force. The last few years have seen great progress in mean-field limits for singular forces by treating them with an N-dependent cut-off. In particular, the mean-field limit for the Keller-Segel model has been rigorously proved in [19, 24, 25]. And the deterministic particle method for aggregation equations can be found in [10, 12]. For a general overview of this topic we refer readers to [11, 20, 28, 38].

Considering the parameter estimation problem for diffusion processes, there is a huge literature in statistics and econometrics, often related to the estimation of volatility in financial models. A complete literature review is beyond our scope and we refer the readers to the book [37] for an overview. To make the scenario more realistic, instead of assuming the availability of some trajectories $\{X_i^t\}_{i=1}^N$ for all time $t \in [0,T]$, we consider the case that trajectories are only observed at discrete time snapshots during the time interval. Diffusion parameter estimation problems based on discrete observations have been discussed by many authors [1, 2, 5, 13, 15, 16, 27, 31, 39]. However, to our knowledge, no previous work has been done for diffusion estimation in the context of interaction particle systems. Specifically, there are a few differences between our work with these works: 1) We consider parameter estimation of a interacting particle system, however authors mentioned above studied a single diffusion process. 2) Our estimator (8) concerns the information of interacting particles, but they only investigated one trajectory and take the expectation value of this stochastic process. 3) In our setting, the interacting force F is singular, while the drift function is assumed to be regular enough in usual statistics literature as mentioned earlier. Our main result, given below in Theorem 1.1 after we make precise the estimator $\widehat{\nu}$, quantifies the estimation error of the proposed estimator.

Take a time step $\Delta t > 0$ and let $t_n := n\Delta t$ and $M := \frac{T}{\Delta t}$ (we assume that $\frac{T}{\Delta t}$ is an integer). Denote $X_i^{(n)} := X_i^{t_n}$ as the solution to (5) at time t_n. Namely, one has

$$
\begin{aligned}
X_i^{(n+1)} - X_i^{(n)} &= \int_{t_n}^{t_{n+1}} \frac{1}{N-1}\sum_{j\neq i}^{N} F^N\left(X_i^s - X_j^s\right) ds + \sqrt{2\nu}(B_i^{t_{n+1}} - B_i^{t_n}) \\
&= \int_{t_n}^{t_{n+1}} \frac{1}{N-1}\sum_{j\neq i}^{N} F^N\left(X_i^s - X_j^s\right) ds + \sqrt{2\nu\Delta t}\mathcal{N}_i^{(n)}, \qquad (7)
\end{aligned}
$$

where $\mathcal{N}_i^{(n)} \sim \mathcal{N}(0,1)^d$, i.e. the standard Gaussian distribution in dimension d.

Then we are ready to define our estimator for the diffusion parameter

$$
\widehat{\nu} := \frac{1}{6dKT}\sum_{i=1}^{K}\sum_{n=0}^{M-1}\left|X_i^{(n+1)} - X_i^{(n)}\right|^2, \qquad (8)
$$

where $1 \ll K \ll N$, which means we only have partial observations.

Our main result quantifies the estimation error of the proposed estimator (8), which is summarized as below

Theorem 1.1. *Suppose the initial data $0 \le \rho_0(x) \in L^1 \cap L^\infty(\mathbb{R}^d)$ and let $\rho(x,t)$ be the regular solution of the aggregation equation (6) with local existence time T such that $\rho \in L^\infty(0,T;L^1\cap L^\infty(\mathbb{R}^d))$. Take a time step $\Delta t > 0$ and let $t_n := n\Delta t$ and $M := \frac{T}{\Delta t}$. Assume $\{X_i^{(n)}\}_{i=1,n=0}^{K,M}$ be the K $(1 \ll K \ll N)$ sample trajectories satisfying (5) with the cut-off index $0 < \delta < \frac{1}{d}$ at time t_n. For any $\alpha > 0$, there exists some constants $C_\alpha > 0$ and $N_0 > 0$ depending only on ν, α, T and $\|\rho_0\|_{L^1\cap L^\infty(\mathbb{R}^d)}$, such that for $N \ge N_0$, the estimator $\hat{\nu}$ defined in (8) is an approximation of ν, and the following estimate holds*

$$
\mathbb{P}\Big(|\widehat{\nu} - \nu| \le C_\alpha \Delta t(1 + N^{-2\delta}\log^2(N)) + \gamma\nu\Big) \ge 1 - N^{-\alpha} - 2e^{-\frac{dKM\gamma^2}{8}}, \qquad (9)
$$

with all $\gamma \in (0,1)$. In particular, if we choose $\Delta t = \gamma = K^{\delta-\frac{1}{2}}$, it follows from (9) that

$$
\mathbb{P}\left(|\widehat{\nu} - \nu| \le C_\alpha K^{\delta-\frac{1}{2}}\right) \ge 1 - K^{-\alpha}. \qquad (10)
$$

To prove the estimate error, we split the estimator $\widehat{\nu}$ into three parts:

$$
\begin{aligned}
\hat{\nu} &= \frac{1}{6dKT}\sum_{i=1}^{K}\sum_{n=0}^{M-1}\left|X_i^{(n+1)}-X_i^{(n)}\right|^2 \\
&\le \frac{1}{2dKT}\sum_{i=1}^{K}\sum_{n=0}^{M-1}\left|X_i^{(n+1)}-X_i^{(n)}-\int_{t_n}^{t_{n+1}}\frac{1}{N-1}\sum_{j\neq i}^{N}F^N(X_i^s-X_j^s)ds\right|^2 \\
&\quad+\frac{1}{2dKT}\sum_{i=1}^{K}\sum_{n=0}^{M-1}\left|\int_{t_n}^{t_{n+1}}\left(\frac{1}{N-1}\sum_{j\neq i}^{N}F^N\left(X_i^s-X_j^s\right)-\int_{\mathbb{R}^d}F^N(X_i^s-y)\rho(y,s)dy\right)ds\right|^2 \\
&\quad+\frac{1}{2dKT}\sum_{i=1}^{K}\sum_{n=0}^{M-1}\left|\int_{t_n}^{t_{n+1}}\int_{\mathbb{R}^d}F^N(X_i^s-y)\rho(y,s)dyds\right|^2 \\
&=: \mathcal{I}_1+\mathcal{I}_2+\mathcal{I}_3,
\end{aligned}
\tag{11}
$$

then one has the error estimate

$$
|\widehat{\nu}-\nu|\le|\mathcal{I}_1-\nu|+|\mathcal{I}_2|+|\mathcal{I}_3|. \tag{12}
$$

In the sequel, we will see that the estimate of $|\mathcal{I}_3|$ is a direct result from the property of regularized kernel F^N. And the estimate of $|\mathcal{I}_2|$ is an estimate of interaction, which follows from the mean-field limit theory. As for the estimate of $|\mathcal{I}_1-\nu|$, it can be deduced from a concentration inequality of chi-squared distribution.

The work is organized as follows: In the next section, we will give a rigorous proof of the mean-filed limit for aggregation equations with Newtonian potential. Base on this, we also obtain an error estimate on interaction. Section 3 is devoted to prove that our estimator $\widehat{\nu}$ is a good approximation of ν and the convergence rate between them is achieved. Then in Section 4 we further extend our result to the case where the aggregation equation has a bounded Lipschitz interacting force.

2 Mean-field limit and estimate on interaction

In this section we will prove the mean-filed limit for particle system (5). Namely, given the solution ρ to the mean-field equation (6), we construct a mean-field trajectories $\{Y_i^t\}_{i=1}^N$ from (6), then we prove the closeness between $X_t=(X_1^t,\cdots,X_N^t)$ and $Y_t=(Y_1^t,\cdots,Y_N^t)$. To do this, we shall consider again a Newtonian system with noise, however, this time not subject to the pair interaction but under the influence of the external mean field $F^N*\rho(x,t)$

$$
dY_i^t=\int_{\mathbb{R}^d}F^N\left(Y_i^t-y\right)\rho(y,t)dy\,dt+\sqrt{2\nu}\,dB_i^t,\quad i=1,\cdots,N, \tag{13}
$$

here we let $\{Y_i^t\}_{i=1}^N$ has the same initial condition as $\{X_i^t\}_{i=1}^N$ (i.i.d. with common density ρ_0). Since the particles are subject to an external field, the independence is conserved. Therefore the $\{Y_i^t\}_{i=1}^N$ are distributed i.i.d. according to the common probability density ρ. We remark that the aggregation equation (6) is Kolmogorov's forward equation for any solution of (13), and in particular their probability distribution ρ solves (6).

2.1 Preliminaries

Notations: The generic constant will be denoted generically by C, even if it is different from line to line. For a vector $X_t=(X_1^t,\ldots,X_N^t)$, we denote

$$
\|X_t\|_\infty := \sup_{i=1,\cdots,N}|X_i^t|. \tag{14}
$$

Since error estimates obtained later are valid when the solution of PDE (6) is regular enough, we assume that

$$
0\le\rho_0\in L^1\cap L^\infty(\mathbb{R}^d), \tag{15}
$$

then equation (6) has a unique local solution with the following regularity

$$\|\rho\|_{L^\infty(0,T;L^1\cap L^\infty(\mathbb{R}^d))} \le C\left(\|\rho_0\|_{L^1\cap L^\infty(\mathbb{R}^d)}\right) =: C_{\rho_0}, \tag{16}$$

where $T > 0$ depends only on $\|\rho_0\|_{L^1\cap L^\infty(\mathbb{R}^d)}$. The proof of this result is a standard process and it can be found in [17, Proposition 3.2].

Let us recall some estimates of the regularized kernel F^N defined in (4):

Lemma 2.1. *([24, Lemma 2.1])*

(i) $F^N(0) = 0$, $F^N(x) = F(x)$ *for any* $|x| \ge N^{-\delta}$ *and* $|F^N(x)| \le |F(x)|$;

(ii) $|\partial^\beta F^N(x)| \le CN^{(d+|\beta|-1)\delta}$, *for any* $x \in \mathbb{R}^d$;

(iii) $\|F^N\|_2 \le CN^{(\frac{d}{2}-1)\delta}$.

Next we define a cut-off function L^N, which will provide the local Lipschitz bound for F^N.

Definition 2.1. Let

$$L^N(x) = \begin{cases} \dfrac{6^d}{|x|^d}, & \text{if } |x| \ge 6N^{-\delta}, \\ N^{d\delta}, & \text{else }, \end{cases} \tag{17}$$

and $L^N : \mathbb{R}^{6N} \to \mathbb{R}^N$ be defined by $(L^N(X_t))_i := \frac{1}{N-1}\sum\limits_{i\ne j} L^N(X_i^t - X_j^t)$. Furthermore, we define $\overline{L}^N(Y_t)$ by $(\overline{L}^N(Y_t))_i := \int_{\mathbb{R}^d} L^N(Y_i^t - x)\rho(x,t)dx$.

Denote

$$(F^N(Y^t))_i := \frac{1}{N-1}\sum_{j\ne i} F^N(Y_i^t - Y_j^t), \tag{18}$$

then we have the local Lipschitz continuity of F^N:

Lemma 2.2. *[26, Lemma 2.3] If* $\|X_t - Y_t\|_\infty \le 2N^{-\delta}$, *then it holds that*

$$\|F^N(X_t) - F^N(Y_t)\|_\infty \le C\|L^N(Y_t)\|_\infty \|X_t - Y_t\|_\infty, \tag{19}$$

for some $C > 0$ *independent of* N.

The following observations of F^N and L^N turn out to be very helpful in the sequel:

Lemma 2.3. *[26, Lemma 2.4] Let* $L^N(x)$ *be defined in Definition 2.1 and* $\rho \in L^1\cap L^\infty(\mathbb{R}^3)$. *Then there exists a constant* $C > 0$ *independent of* N *such that*

$$\|L^N * \rho\|_\infty \le C\log(N)(\|\rho\|_1 + \|\rho\|_\infty), \quad \|(L^N)^2 * \rho\|_\infty \le CN^{d\delta}(\|\rho\|_1 + \|\rho\|_\infty); \tag{20}$$

and

$$\|F^N * \rho\|_\infty \le C(\|\rho\|_1 + \|\rho\|_\infty), \quad \|\nabla F^N * \rho\|_\infty \le C\log(N)(\|\rho\|_1 + \|\rho\|_\infty). \tag{21}$$

Also, we need the following concentration inequality to provide us the probability bounds of random variables:

Lemma 2.4. *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 for any* $\alpha > 0$, *the sample mean* $\bar{Z} = \frac{1}{N}\sum_{i=1}^N Z_i$ *satisfies*

$$\mathbb{P}\left(|\bar{Z}| \ge \frac{C_\alpha\sqrt{g(N)}\log(N)}{\sqrt{N}}\right) \le N^{-\alpha}, \tag{22}$$

where C_α *depends only on* C *and* α.

The proof can be seen in [21, Lemma 1], which is a direct result of the Taylor expansion and the Markov's inequality.

2.2 Mean-field limit for the aggregation equation with Newtonian potential

In this section, we obtain the maximal distance between the exact microscopic dynamics (5) and the approximate mean-field dynamics (13). Denote

$$(\overline{F}^N(Y_t))_i := \int_{\mathbb{R}^d} F^N(Y_i^t - x)\rho(x,t)dx, \tag{23}$$

then we can introduce the following lemma of consistency at fixed time:

Lemma 2.5. *At any fixed time $t \in [0,T]$, suppose that Y_t satisfies the mean-field dynamics (13) with i.i.d initial data sharing the common density ρ_0 satisfying (15). Assume that F^N and $\overline{F}^N$ are defined in (18) and (23) respectively, L^N and $\overline{L}^N$ are showed in Definition 2.1. For any $\alpha > 0$ and $0 < \delta \le \frac{1}{d}$, there exist a constant $C_{1,\alpha} > 0$ depending only on α, T and C_{ρ_0} such that*

$$\mathbb{P}\left(\left\|F^N(Y_t) - \overline{F}^N(Y_t)\right\|_\infty \ge C_{1,\alpha} N^{\frac{\delta(d-2)-1}{2}} \log(N)\right) \le N^{-\alpha}, \tag{24}$$

and

$$\mathbb{P}\left(\left\|L^N(Y_t) - \overline{L}^N(Y_t)\right\|_\infty \ge C_{1,\alpha} N^{\frac{d\delta-1}{2}} \log(N)\right) \le N^{-\alpha}. \tag{25}$$

Proof. We can prove this lemma by using Lemma 2.4. Due to the exchangeability of the particles, we are ready to bound

$$(F^N(Y_t))_1 - (\overline{F}^N(Y_t))_1 = \frac{1}{N-1}\sum_{j=2}^N F^N(Y_1^t - Y_j^t) - \int_{\mathbb{R}^3} F^N(Y_1^t - x)\rho(x,t)dx = \frac{1}{N-1}\sum_{j=2}^N Z_j, \tag{26}$$

where

$$Z_j := F^N(Y_1^t - Y_j^t) - \int_{\mathbb{R}^d} F^N(Y_1^t - x)\rho(x,t)dx.$$

Since Y_1^t and Y_j^t are independent when $j \ne 1$ and $F^N(0) = 0$, let us consider Y_1^t as given and denote $\mathbb{E}'[\cdot] = \mathbb{E}[\cdot|Y_1^t]$. It is easy to show that $\mathbb{E}'[Z_j] = 0$ since

$$\mathbb{E}'\left[F^N(Y_1^t - Y_j^t)\right] = \int_{\mathbb{R}^d} F^N(Y_1^t - x)\rho(x,t)dx.$$

To use Lemma 2.4, we need a bound for the variance

$$\mathbb{E}'\left[|Z_j|^2\right] = \mathbb{E}'\left[\left|F^N(Y_1^t - Y_j^t) - \int_{\mathbb{R}^3} F^N(Y_1^t - x)\rho(x,t)dx\right|^2\right]. \tag{27}$$

Since it follows from Lemma 2.3 that

$$\int_{\mathbb{R}^3} F^N(Y_1^t - x)\rho(x,t)dx \le C(\|\rho\|_1 + \|\rho\|_\infty), \tag{28}$$

it suffices to bound

$$\mathbb{E}'\left[F^N(Y_1^t - Y_j^t)\right] = \int_{\mathbb{R}^d} F^N(Y_1^t - x)\rho(x,t)dx \le C(\|\rho\|_1 + \|\rho\|_\infty) \le C(T, C_{\rho_0}), \tag{29}$$

and

$$\mathbb{E}'\left[F^N(Y_1^t - Y_j^t)^2\right] = \int_{\mathbb{R}^d} F^N(Y_1^t - x)^2\rho(x,t)dx \le \|\rho\|_\infty \|F^N\|_2^2 \le C(T, C_{\rho_0}) N^{\delta(d-2)}, \tag{30}$$

where we have used $\|F^N\|_2 \le C\varepsilon^{\delta(\frac{d}{2}-1)}$ in Lemma 2.1 (iii). Hence one has

$$\mathbb{E}'\left[|Z_j|^2\right] \le CN^{\delta(d-2)}. \tag{31}$$

So the hypotheses of Lemma 2.4 are satisfied with $g(N) = CN^{\delta(d-2)}$. In addition, it follows from (ii) in Lemma 2.1 that $|Z_j| \le CN^{\delta(d-1)} \le C\sqrt{Ng(N)}$. Hence, using Lemma 2.4, we have the probability bound

$$\mathbb{P}\left(\left|(F^N(Y_t))_1 - (\overline{F}^N(Y_t))_1\right| \ge C(\alpha, T, C_{\rho_0})N^{\frac{\delta(d-2)-1}{2}}\log(N)\right) \le N^{-\alpha}. \tag{32}$$

Similarly, the same bound must also apply hold to other term with $i = 2, \cdots, N$, which leads to

$$\mathbb{P}\left(\left\|F^N(Y_t) - \overline{F}^N(Y_t)\right\|_\infty \ge C(\alpha, T, C_{\rho_0})N^{\frac{\delta(d-2)-1}{2}}\log(N)\right) \le N^{1-\alpha}. \tag{33}$$

Let $C_{1,\alpha}$ be the constant in (33), we conclude (24).

To prove (25), we follow the same procedure above

$$(L^N(Y_t))_1 - (\overline{L}^N(Y_t))_1 = \frac{1}{N-1}\sum_{j=2}^N L^N(Y_1^t - Y_j^t) - \int_{\mathbb{R}^d} L^N(Y_1^t - x)\rho(x,t)dx = \frac{1}{N-1}\sum_{j=2}^N Z_j, \tag{34}$$

where

$$Z_j = L^N(Y_1^t - Y_j^t) - \int_{\mathbb{R}^3} L^N(Y_1^t - x)\rho(x,t)dx.$$

It is easy to show that $\mathbb{E}'[Z_j] = 0$. To use Lemma 2.4, we need a bound for the variance. One computes that

$$\mathbb{E}'\left[L^N(Y_1^t - Y_j^t)\right] = \int_{\mathbb{R}^d} L^N(Y_1^t - x)\rho(x,t)dx \le C\log(N)(\|\rho\|_1 + \|\rho\|_\infty) \le C(T, C_{\rho_0})\log(N), \tag{35}$$

and

$$\mathbb{E}'\left[L^N(Y_1^t - Y_j^t)^2\right] = \int_{\mathbb{R}^d} L^N(Y_1^t - x)^2\rho(x,t)dx \le CN^{d\delta}(\|\rho\|_1 + \|\rho\|_\infty) \le C(T, C_{\rho_0})N^{d\delta}, \tag{36}$$

where we have used the estimates of L^N in Lemma 2.3. Hence one has

$$\mathbb{E}'\left[|Z_j|^2\right] \le CN^{d\delta}. \tag{37}$$

So the hypotheses of Lemma 2.4 are satisfied with $g(N) = CN^{d\delta}$. In addition, it follows from Definition 2.1 that $|Z_j| \le CN^{d\delta} \le C\sqrt{Ng(N)}$. Hence, we have the probability bound

$$\mathbb{P}\left(\left|(L^N(Y_t))_1 - (\overline{L}^N(Y_t))_1\right| \ge C(\alpha, T, C_{\rho_0})N^{\frac{d\delta-1}{2}}\log(N)\right) \le N^{-\alpha}. \tag{38}$$

by Lemma 2.4, which leads to

$$\mathbb{P}\left(\left\|L^N(Y_t) - \overline{L}^N(Y_t)\right\|_\infty \ge C(\alpha, T, C_{\rho_0})N^{\frac{d\delta-1}{2}}\log(N)\right) \le N^{1-\alpha}. \tag{39}$$

Thus, (25) follows from (39).

□

Next we improve the consistency error to all time through the same procedure in [26, Proposition 3.1]. To do this, we need the following lemma:

Lemma 2.6. *Assume that the time step size $\Delta t = t_{n+1} - t_n = N^{-\frac{\beta}{d}}$ for $\beta > 2$ and Y_t satisfies the mean-field dynamics (13). There exists some constant $C_B > 0$ depending only on ν, T and C_{f_0}, such that it holds*

$$\mathbb{P}\left(\sup_n \sup_{t\in[t_n, t_{n+1}]} \|Y_t - Y_{t_n}\|_\infty \ge C_B N^{-\frac{\beta+2}{2d}}\right) \le C_B N^{\frac{2-\beta}{2d}}\exp(-C_B N^{\frac{\beta-2}{d}}). \tag{40}$$

Proof. Notice that for $t \in [t_n, t_{n+1}]$

$$\begin{aligned} Y_t - Y_{t_n} &= \int_{t_n}^{t} \int_{\mathbb{R}^d} F^N(Y_i^s - y)\rho(y,s)dy\,ds + \sqrt{2\nu\Delta t}(B^t - B^{t_n}) \\ &=: I_1(t) + I_2(t). \end{aligned} \tag{41}$$

It follows from Lemma 2.3 that

$$\sup_{t_n \le t \le t_{n+1}} \|I_1(t)\|_\infty \le C\Delta t \le CN^{-\frac{\beta}{d}}, \tag{42}$$

where C depending only on T and C_{f_0}. To estimate $I_2(t)$, recall a basic property of the Brownian motion [25, Lemma 2.7]:

$$\mathbb{P}\left(\sup_{t \le s \le t+\Delta t} \|B^s - B^t\|_\infty \ge b\right) \le C_1(\sqrt{\Delta t}/b)\exp(-C_2 b^2/\Delta t), \tag{43}$$

where C_1 and C_2 depend only on d. Choosing $b = N^{-\frac{1}{d}}$ in (43), it leads to

$$\mathbb{P}\left(\sup_{t_n \le t \le t_{n+1}} \|I_2(t)\|_\infty \ge N^{-\frac{\beta+2}{2d}}\right) \le C_1 N^{\frac{2-\beta}{2d}} \exp(-C_2 N^{\frac{\beta-2}{d}}). \tag{44}$$

Collecting (42) and (44), it yields that

$$\mathbb{P}\left(\sup_n \sup_{t \in [t_n, t_{n+1}]} \|Y_t - Y_{t_n}\|_\infty \ge CN^{-\frac{\beta+2}{2d}}\right) \le C_1 N^{\frac{2-\beta}{2d}} \exp(-C_2 N^{\frac{\beta-2}{d}}), \tag{45}$$

for some $\beta > 2$, which concludes the proof. $\square$

Now we can prove the consistency error in all time.

Proposition 2.1. *(Consistency) Let Y_t satisfies the mean-field dynamics (13) with i.i.d initial data sharing the common density ρ_0 satisfying (15). Assume that F^N and $\overline{F}^N$ be defined in (18) and (23) respectively. For any $\alpha > 0$ and $0 < \delta \le \frac{1}{d}$, there exist a constant $C_{2,\alpha} > 0$ depending only on depends on ν, α, T and C_{ρ_0} such that*

$$\mathbb{P}\left(\sup_{t \in [0,T]} \left\|F^N(Y_t) - \overline{F}^N(Y_t)\right\|_\infty \ge C_{2,\alpha} N^{\frac{\delta(d-2)-1}{2}} \log(N)\right) \le N^{-\alpha}, \tag{46}$$

and

$$\mathbb{P}\left(\sup_{t \in [0,T]} \left\|L^N(Y_t) - \overline{L}^N(Y_t)\right\|_\infty \ge C_{2,\alpha} N^{\frac{d\delta-1}{2}} \log(N)\right) \le N^{-\alpha}. \tag{47}$$

Proof. Denote events:

$$\mathcal{H} := \left\{\sup_n \sup_{t \in [t_n, t_{n+1}]} \|Y_t - Y_{t_n}\|_\infty \le C_B N^{-\frac{\beta+2}{2d}}\right\}, \tag{48}$$

and

$$\mathcal{C}_{t_n} := \left\{\left\|F^N(Y_{t_n}) - \overline{F}^N(Y_{t_n})\right\|_\infty \ge C_{1,\alpha} N^{\frac{\delta(d-2)-1}{2}} \log(N)\right\}, \tag{49}$$

where C_B and $C_{1,\alpha}$ are used in Lemma 2.5 and Lemma 2.6 respectively. According to the Lemma 2.5 and Lemma 2.6, one has

$$\mathbb{P}(\mathcal{C}_{t_n}^c) \le N^{-\alpha}, \quad \mathbb{P}(\mathcal{H}^c) \le C_B N^{\frac{2-\beta}{2d}} \exp(-C_B N^{\frac{\beta-2}{d}}). \tag{50}$$

for any $\alpha > 0$ and $\beta > 2$.

Furthermore, we denote

$$\mathcal{B}_{t_n} := \left\{ \left\| L^N(Y_{t_n}) - \overline{L}^N(Y_{t_n}) \right\|_\infty \le C_{1,\alpha} N^{\frac{d\delta-1}{2}} \log(N) \right\}, \tag{51}$$

then under the event $\mathcal{B}_{t_n}$, it holds that

$$\|L^N(Y_{t_n})\|_\infty \le \|\overline{L}^N(Y_{t_n})\|_\infty + C_{1,\alpha} N^{\frac{d\delta-1}{2}} \log(N) \le C(\alpha, T, C_{f_0}) \log(N). \tag{52}$$

and $\mathbb{P}(\mathcal{B}_{t_n}^c) \le N^{-\alpha}$ by Lemma 2.5.

For all $t \in [t_n, t_{n+1}]$, under the event $\mathcal{B}_{t_n} \cap \mathcal{C}_{t_n} \cap \mathcal{H}$, we obtain

$$\begin{aligned}
&\left\| F^N(Y_t) - \overline{F}^N(Y_t) \right\|_\infty \\
\le &\left\| F^N(Y_t) - F^N(Y_{t_n}) \right\|_\infty + \left\| F^N(Y_{t_n}) - \overline{F}^N(Y_{t_n}) \right\|_\infty + \left\| \overline{F}^N(Y_{t_n}) - \overline{F}^N(Y_t) \right\|_\infty \\
\le &C \|L^N(Y_{t_n})\|_\infty \|Y_t - Y_{t_n}\|_\infty + C_{1,\alpha} N^{\frac{\delta(d-2)-1}{2}} \log(N) + C \log(N) \|Y_t - Y_{t_n}\|_\infty \\
\le &C(\alpha, T, C_{f_0}) \log(N) N^{-\frac{\beta+2}{2d}} + C_{1,\alpha} N^{\frac{\delta(d-2)-1}{2}} \log(N) \\
\le &C(\alpha, T, C_{f_0}) N^{\frac{\delta(d-2)-1}{2}} \log(N), \quad (\beta > (d-2)(1-d\delta))
\end{aligned}$$

where in the second inequality we have used the local Lipschitz bound of F^N

$$\left\| F^N(Y_t) - F^N(Y_{t_n}) \right\|_\infty \le C \|L^N(Y_{t_n})\|_\infty \|Y_t - Y_{t_n}\|_\infty, \tag{53}$$

under the event $\mathcal{H}$ (see in Lemma 2.2). It yields that

$$\sup_{t\in[0,T]} \left\| F^N(Y_t) - \overline{F}^N(Y_t) \right\|_\infty \le C(\nu, \alpha, T, C_{f_0}) N^{\frac{\delta(d-2)-1}{2}} \log(N), \tag{54}$$

holds under the event $\bigcap_{n=0}^{M-1} \mathcal{B}_{t_n} \cap \mathcal{C}_{t_n} \cap \mathcal{H}$. Therefore

$$\begin{aligned}
&\mathbb{P}\left(\sup_{t\in[0,T]} \left\| F^N(Y_t) - \overline{F}^N(Y_t) \right\|_\infty \ge C(\alpha, T, C_{f_0}) N^{\frac{\delta(d-2)-1}{2}} \log(N) \right) \\
\le &\sum_{n=0}^{M} P(\mathcal{B}_{t_n}^c) + \sum_{n=0}^{M-1} P(\mathcal{C}_{t_n}^c) + P(\mathcal{H}^c) \\
\le &TN^{-\frac{d\alpha-\beta}{d}} + TN^{-\frac{d\alpha-\beta}{d}} + C_B N^{\frac{2-\beta}{2d}} \exp(-C_B N^{\frac{\beta-2}{d}}) \le N^{-\alpha'}.
\end{aligned} \tag{55}$$

Denote $C_{2,\alpha'}$ to be the constant $C(\nu, \alpha, T, C_{f_0})$ in (55). Since $\alpha > 0$ is arbitrary and so is α', hence (46) holds true. The proof of (47) can be done similarly. □

In order to prove the convergence, we still need the stability result which states:

Proposition 2.2. *(Stability) Assume that trajectories $X_t = (X_t^i)_{i=1,\cdots,N}$, $Y_t = (Y_t^i)_{i=1,\cdots,N}$ satisfy (5) and (13) respectively with the initial data $X_0 = Y_0$, which is i.i.d. sharing the common density ρ_0 satisfying (15). Let events $\mathcal{B}_{t_n}$ and $\mathcal{H}$ be defined in (51) and (48) respectively, F^N be defined in (18). Denote events:*

$$\mathcal{A} := \left\{ \sup_{t\in[0,T]} \|X_t - Y_t\|_\infty < N^{-\delta} \right\}, \tag{56}$$

and

$$\mathcal{S}(\Lambda) := \left\{ \|F^N(X_t) - F^N(Y_t)\|_\infty \le \Lambda \log(N) \|X_t - Y_t\|_\infty + \Lambda \log(N) N^{-\frac{\beta+2}{2d}}, \ \forall\, t \in [0,T] \right\}.$$

For any $\alpha > 0$, there exists some $C_{3,\alpha} > 0$ depending only on ν, α, T and C_{ρ_0} such that

$$\bigcap_{n=0}^{M} \mathcal{B}_{t_n} \cap \mathcal{A} \cap \mathcal{H} \subset \mathcal{S}(C_{3,\alpha}).$$

Here the event $\mathcal{S}(C_{3,\alpha})$ can be seen as the stability result and the events $\mathcal{B}_{t_n}$, $\mathcal{A}$ and $\mathcal{H}$ can be treated as the stability conditions.

Proof. First, we split $\mathcal{S}(\Lambda)$ into the union of non-overlapping sets $\{\mathcal{S}_n\}_{n=1}^{N'}(\Lambda)$, where

$$\begin{aligned}\mathcal{S}_n(\Lambda) := \Big\{ &\|F^N(X_t) - F^N(Y_t)\|_\infty \le \Lambda \log(N) \|X_t - Y_t\|_\infty \\ &+ \Lambda \log(N) N^{-\frac{\beta+2}{2d}}, \ \forall\, t \in [t_n, t_{n+1}] \Big\}.\end{aligned}$$

Notice that for any $t \in [t_n, t_{n+1}]$, under the event $\mathcal{A} \cap \mathcal{H}$, one has

$$\begin{aligned}\sup_{t\in[t_n,t_{n+1}]} \|X_t - Y_{t_n}\|_\infty &\le \sup_{t\in[t_n,t_{n+1}]} \|X_t - Y_t\|_\infty + \sup_{t\in[t_n,t_{n+1}]} \|Y_t - Y_{t_n}\|_\infty \\ &\le N^{-\delta} + C_B N^{-\frac{\beta+2}{2d}} \le 2N^{-\delta}, \quad (\beta > 2d\delta - 2)\end{aligned}$$

and

$$\sup_{t\in[t_n,t_{n+1}]} \|Y_t - Y_{t_n}\|_\infty < C_B N^{-\frac{\beta+2}{2d}} < N^{-\delta}. \quad (\beta > 2d\delta - 2)$$

Then applying the local Lipschitz bound of F^N (see in Lemma 2.2) leads to

$$\begin{aligned}\|F^N(X_t) - F^N(Y_t)\|_\infty &\le \|F^N(X_t) - F^N(Y_{t_n})\|_\infty + \|F^N(Y_{t_n}) - F^N(Y_t)\|_\infty \\ &\le C\|L^N(Y_{t_n})\|_\infty \left(\|X_t - Y_{t_n}\|_\infty + \|Y_{t_n} - Y_t\|_\infty\right) \\ &\le C\|L^N(Y_{t_n})\|_\infty \|X_t - Y_t\|_\infty + 2C\|L^N(Y_{t_n})\|_\infty \|Y_{t_n} - Y_t\|_\infty\end{aligned}$$

under the event $\mathcal{A} \cap \mathcal{H}$.

Furthermore, under the event $\mathcal{B}_{t_n}$, it follows from (52) that

$$\|L^N(Y_{t_n})\|_\infty \le C \log(N), \tag{57}$$

Hence, for all $t \in [t_n, t_{n+1}]$ one has

$$\begin{aligned}\|F^N(X_t) - F^N(Y_t)\|_\infty &\le C(\nu, \alpha, T, C_{\rho_0}) \log(N) \|X_t - Y_t\|_\infty \\ &+ C(\nu, \alpha, T, C_{\rho_0}) \log(N) N^{-\frac{\beta+2}{2d}},\end{aligned}$$

under event $\mathcal{A} \cap \mathcal{H} \cap \mathcal{B}_{t_n}$. Denote the $C(\nu, \alpha, T, C_{\rho_0})$ in the above as $C_{3,\alpha}$. This implies $\mathcal{A} \cap \mathcal{H} \cap \mathcal{B}_{t_n} \subset \mathcal{S}_n(C_{3,\alpha})$, which yields

$$\bigcap_{n=0}^{M-1} \mathcal{B}_{t_n} \cap \mathcal{H} \cap \mathcal{A} \subset \mathcal{S}(C_{3,\alpha}).$$

Thus, the proposition has been proved.

□

Our next theorem states that the N-particle trajectory X_t starting from X_0 (i.i.d. with common density ρ_0) remains close to the mean-field trajectory Y_t with the same initial configuration $Y_0 = X_0$. More precisely, we prove that the measure of the set where the maximal distance $\sup_{t\in[0,T]} \|X_t - Y_t\|_\infty$ on $[0, T]$ excedes $N^{-\delta}$ decreases exponentially with the number of particles N, as N grows to infinity:

Theorem 2.1. *(Convergence) Assume that trajectories* $X_t = (X_t^i)_{i=1,\cdots,N}$, $Y_t = (Y_t^i)_{i=1,\cdots,N}$ *satisfy* (5) *and* (13) *respectively with the initial data* $X_0 = Y_0$, *which is i.i.d. sharing the common density* ρ_0 *satisfying* (15)*. Then for any* $\alpha > 0$, *there exist some constant* $N_0 > 0$ *depending only on* ν, α, T *and* C_{ρ_0}, *such that for* $N \geq N_0$, *the following estimate holds with the cut-off index* $0 < \delta < \frac{1}{d}$

$$\mathbb{P}\left(\sup_{t\in[0,T]} \|X_t - Y_t\|_\infty \leq N^{-\delta}\right) \geq 1 - N^{-\alpha}.$$

Proof. We can prove the convergence result by using the consistency from Proposition 2.1 and the stability from Proposition 2.2. Denote the event

$$\mathcal{C} := \left\{ \sup_{t\in[0,T]} \left\| F^N(Y_t) - \overline{F}^N(Y_t) \right\|_\infty \leq C_{2,\alpha} N^{\frac{\delta(d-2)-1}{2}} \log(N) \right\}. \tag{58}$$

Consider the quantity $e(t)$ defined as

$$e(t) := \|X_t - Y_t\|_\infty . \tag{59}$$

Computing under the event $\mathcal{C} \cap \mathcal{S}(C_{3,\alpha})$ and using the fact $\frac{d\|x\|_\infty}{dt} \leq \|\frac{dx}{dt}\|_\infty$, one has

$$\begin{aligned}
\frac{de(t)}{dt} &\leq \left\| F^N(X_t) - \overline{F}^N(Y_t) \right\|_\infty \\
&\leq \left\| F^N(X_t) - F^N(Y_t) \right\|_\infty + \left\| F^N(Y_t) - \overline{F}^N(Y_t) \right\|_\infty \\
&\leq C_{3,\alpha} \log(N) \|X_t - Y_t\|_\infty + C_{3,\alpha} \log(N) N^{-\frac{\beta+2}{2d}} + C_{2,\alpha} N^{\frac{\delta(d-2)-1}{2}} \log(N) \\
&\leq C(\nu, \alpha, T, C_{\rho_0}) \log(N) e(t) + C(\nu, \alpha, T, C_{\rho_0}) N^{\frac{\delta(d-2)-1}{2}} \log(N).
\end{aligned} \tag{60}$$

Using Gronwall's inequality with $e(0) = 0$, it follows from (60) that

$$\sup_{t\in[0,T]} e(t) \leq CTN^{\frac{\delta(d-2)-1}{2}} \log(N) e^{CT\log(N)}, \tag{61}$$

Where C depends only on ν, α, T and C_{ρ_0}.

We denote the event

$$\mathcal{M} := \left\{ \sup_{t\in[0,T]} e(t) \leq CTN^{\frac{\delta(d-2)-1}{2}} \log(N) e^{CT\log(N)} \right\}. \tag{62}$$

then it follows from Proposition 2.2 that

$$\mathcal{C} \cap \bigcap_{n=0}^{M-1} \mathcal{B}_{t_n} \cap \mathcal{H} \cap \mathcal{A} \subset \mathcal{C} \cap \mathcal{S}(C_{3,\alpha}) \subset \mathcal{M}. \tag{63}$$

Notice that for $0 < \delta < \frac{1}{3}$ there exists some N_0 depending only on ν, α, T and C_{ρ_0}, such that for $N \geq N_0$

$$\begin{aligned}
&\sup_{t\in[0,T]} \|X_t - Y_t\|_\infty \leq \sup_{t\in[0,T]} e(t) \\
\leq &CTN^{\frac{\delta(d-2)-1}{2}} \log(N) e^{CT\log(N)} \leq \frac{1}{2} N^{-\delta} < N^{-\delta}.
\end{aligned}$$

Since $\sup_{t\in[0,T]} \|X_t - Y_t\|_\infty$ is a continuous function and it vanishes at $t = 0$, it can never reach $N^{-\delta}$. So the condition $\mathcal{A}$ defined in (56) has never been used. The above argument is a standard *a-priori* estimate in PDE analysis, which has been used in [21, 25, 33, 34]. Thus it follows from (63) that

$$\mathcal{C} \cap \bigcap_{n=0}^{M-1} \mathcal{B}_{t_n} \cap \mathcal{H} \subset \mathcal{M}, \tag{64}$$

which concludes that

$$\mathbb{P}\left(\sup_{t\in[0,T]} e(t) \geq N^{-\delta}\right) \leq \mathbb{P}(\mathcal{M}^c) \leq \sum_{n=0}^{M-1} \mathbb{P}(\mathcal{B}^c_{t_n}) + \mathbb{P}(\mathcal{H}^c) + \mathbb{P}(\mathcal{C}^c)$$
$$\leq TN^{\frac{\beta}{d}-\alpha} + C_B N^{\frac{2-\beta}{2d}} \exp(-C_B N^{\frac{\beta-2}{d}}) + N^{-\alpha} \leq N^{-\alpha'},$$

by using Proposition 2.1, Lemma 2.6 and Lemma 2.5. Since $\alpha > 0$ is arbitrary and so is α', we have proved Theorem 2.1. □

2.3 The error estimate on interaction

Using Theorem 2.1, we obtain the error estimate on interaction:

Theorem 2.2. *Under the same assumption as Theorem 2.1, let $\rho(x,t)$ be the regular solution to the aggregation equation (6) up to time T such that $\rho \in L^\infty(0,T;L^1\cap L^\infty(\mathbb{R}^d))$. Assume that $\{X^t_i\}_{i=1}^N$ satisfy the particle system (5) and F^N satisfies (4). Then for any $\alpha > 0$ there exists some constants $C_{4,\alpha} > 0$, $N_0 > 0$ depending only on ν, α, T and C_{ρ_0} such that for $N \geq N_0$ the following estimate holds with the cut-off index $0 < \delta < \frac{1}{3}$*

$$\mathbb{P}\Bigg(\sup_{0\leq t\leq T}\sup_{x}\left|\int_{\mathbb{R}^d} F^N(x-y)\rho(y,t)dy - \frac{1}{N}\sum_{j=1}^N F^N(x-X^t_j)\right|$$
$$\leq C_{4,\alpha} N^{-\delta}\log(N)\Bigg) \geq 1 - N^{-\alpha},$$

Proof. Let us denote

$$e^t_2(x) := \left|\int_{\mathbb{R}^d} F^N(x-y)\rho(y,t)dy - \frac{1}{N}\sum_{j=1}^N F^N(x-X^t_j)\right|,$$

then one split it into two parts:

$$\begin{aligned} e^t_2(x) \leq &\left|\int_{\mathbb{R}^d} F^N(x-y)\rho(y,t)dy - \frac{1}{N}\sum_{j=1}^N F^N(x-Y^t_j)\right| \\ &+ \left|\frac{1}{N}\sum_{j=1}^N F^N(x-Y^t_j) - \frac{1}{N}\sum_{j=1}^N F^N(x-X^t_j)\right| \\ &=: e^t_{21}(x) + e^t_{22}(x). \end{aligned}$$

To estimate $e^t_{22}(x)$, we can directly use the result from Theorem 2.1. Let us recall the event

$$\mathcal{A} = \left\{\sup_{t\in[0,T]} \|X_t - Y_t\|_\infty \leq N^{-\delta}\right\}, \tag{65}$$

then it follows from Theorem 2.1 that

$$\mathbb{P}(\mathcal{A}^c) \leq N^{-\alpha}. \tag{66}$$

Furthermore, we define event

$$\mathcal{B}_1 := \left\{\sup_{t\in[0,T]} \left\|L^N(x-Y_t) - \overline{L}^N(x-Y_t)\right\|_\infty \leq C_{2,\alpha} N^{\frac{d\delta-1}{2}}\log(N)\right\}, \tag{67}$$

and according to Proposition 2.1 one has that

$$\mathbb{P}(\mathcal{B}_1^c) \leq N^{-\alpha}. \tag{68}$$

Hence under the event $\mathcal{A}\cap\mathcal{B}_1$, it follows from Lemma 2.2 and Lemma 2.3 (i) that

$$\begin{aligned} e_{22}^t(x) &\le C\|L^N(x-Y_t)\|_\infty\|X_t-Y_t\|_\infty \\ &\le C\left(\|\overline{L}^N(x-Y_t)\|_\infty + C_{2,\alpha}N^{\frac{d\delta-1}{2}}\log(N)\right)N^{-\delta} \\ &\le C\log(N)N^{-\delta}, \end{aligned}$$

which implies that

$$\mathbb{P}\left(\sup_{0\le t\le T}\sup_x e_{22}^t(x)\le C\log(N)N^{-\delta}\right)\ge 1-N^{-\alpha}, \tag{69}$$

where C depends only on ν, α, T and C_{ρ_0}.

To estimate $e_{21}^t(x)$, we will use by using Lemma 2.4. Let us define

$$\frac{1}{N}\sum_{j=1}^N\left(F^N(x-Y_j^t)-\int_{\mathbb{R}^d}F^N(x-y)\rho(y,t)\,dy\right)=:\frac{1}{N}\sum_{j=1}^N Z_j^t, \tag{70}$$

where

$$Z_j^t = F^N(x-Y_j^t)-\int_{\mathbb{R}^d}F^N(x-y)\rho(y,t)\,dy. \tag{71}$$

And it is obvious that $\mathbb{E}[Z_j^t]=0$. Since $\{Y_i^t\}_{i=1}^N$ are i.i.d., we know $\{Z_i^t\}_{i=1}^N$ are also i.i.d.. To use Lemma 2.4, we need a bound for the variance

$$\mathbb{E}\big[|Z_j^t|^2\big]=\mathbb{E}\left[\left|F^N(x-Y_j^t)-\int_{\mathbb{R}^d}F^N(x-y)\rho(y,t)\,dy\right|^2\right]. \tag{72}$$

Since it follows from Lemma 2.3 that

$$\mathbb{E}[F^N(x-Y_j^t)]=\int_{\mathbb{R}^3}F^N(x-y)\rho(y,t)dy\le C(\|\rho\|_1+\|\rho\|_\infty), \tag{73}$$

it suffices to bound

$$\mathbb{E}\big[F^N(x-Y_j^t)^2\big]=\int_{\mathbb{R}^d}F^N(x-y)^2\rho(y,t)dy\le\|\rho\|_\infty\|F^N\|_2^2\le C(T,C_{\rho_0})N^{\delta(d-2)}, \tag{74}$$

where we have used $\|F^N\|_2\le C\varepsilon^{\delta(\frac{d}{2}-1)}$ in Lemma 2.1 (iii). Hence one has

$$\mathbb{E}\big[|Z_j^t|^2\big]\le CN^{\delta(d-2)}. \tag{75}$$

So the hypotheses of Lemma 2.4 are satisfied with $g(N)=CN^{\delta(d-2)}$. In addition, it follows from (ii) in Lemma 2.1 that $|Z_j^t|\le CN^{\delta(d-1)}\le C\sqrt{Ng(N)}$. Hence, using Lemma 2.4, we have the probability bound at any fix time $t\in[0,T]$

$$\mathbb{P}\left(\left|\int_{\mathbb{R}^d}F^N(x-y)\rho(y,t)dy-\frac{1}{N}\sum_{j=1}^N F^N(x-Y_j^t)\right|\ge CN^{\frac{\delta(d-2)-1}{2}}\log(N)\right)\le N^{-\alpha}, \tag{76}$$

where C depends only on ν, α, T and C_{ρ_0}. Then following the same procedure as in Proposition 2.1, we can improve the estimate (76) to all time, which is

$$\mathbb{P}\left(\sup_{0\le t\le T}\sup_x e_{21}^t(x)\le CN^{\frac{\delta(d-2)-1}{2}}\log(N)\right)\ge 1-N^{-\alpha}. \tag{77}$$

Collecting (77) and (69), it yields that

$$\mathbb{P}\left(\sup_{0\le t\le T}\sup_x e_2^t(x)\le CN^{-\delta}\log(N)\right)\ge 1-N^{1-\alpha}. \tag{78}$$

where C depends only on ν, α, T and C_{ρ_0}, which concludes that

$$\mathbb{P}\Bigg(\sup_{0\le t\le T}\sup_x\left|\int_{\mathbb{R}^d}F^N(x-y)\rho(y,t)dy-\frac{1}{N}\sum_{j=1}^N F^N(x-X_j^t)\right| \\ \le CN^{-\delta}\log(N)\Bigg)\ge 1-N^{-\alpha},$$

where C depends only on ν, α, T and C_{ρ_0}, which finishes the proof. ☐

3 Parameter estimation and the proof Theorem 1.1

In this section, we obtain the diffusion parameter estimation and prove our main Theorem 1.1. Let us define

$$\nu_{K,N} := \frac{1}{2dKT}\sum_{i=1}^K\sum_{n=0}^{M-1}\left|X_i^{(n+1)}-X_i^{(n)}-\int_{t_n}^{t_{n+1}}\frac{1}{N-1}\sum_{j\neq i}^N F^N\left(X_i^s-X_j^s\right)ds\right|^2. \tag{79}$$

and recall (11), one concludes that

$$|\widehat{\nu}-\nu|\le|\nu_{K,N}-\nu|+|\mathcal{I}_2|+|\mathcal{I}_3|, \tag{80}$$

where

$$\mathcal{I}_2=\frac{1}{2dKT}\sum_{i=1}^K\sum_{n=0}^{M-1}\left|\int_{t_n}^{t_{n+1}}\left(\frac{1}{N-1}\sum_{j\neq i}^N F^N\left(X_i^s-X_j^s\right)-\int_{\mathbb{R}^d}F^N(X_i^s-y)\rho(y,s)dy\right)ds\right|^2, \tag{81}$$

and

$$\mathcal{I}_3=\frac{1}{2dKT}\sum_{i=1}^K\sum_{n=0}^{M-1}\left|\int_{t_n}^{t_{n+1}}\int_{\mathbb{R}^d}F^N(X_i^s-y)\rho(y,s)\,dyds\right|^2. \tag{82}$$

According to Lemma 2.3, one has

$$|\mathcal{I}_3|\le C\Delta t, \tag{83}$$

where C depends only on T and C_{ρ_0}. Then it follows from Theorem 2.2 that

$$\mathbb{P}\left(|\mathcal{I}_2|\le C\Delta tN^{-2\delta}\log^2(N)\right)\ge 1-N^{-\alpha}, \tag{84}$$

where C depends only on ν, α, T and C_{ρ_0}. It is left to estimate the error between $\nu_{K,N}$ and ν, which can be done by using the concentration property of χ^2 random variable.

Proposition 3.1. *Under the assumption as in Theorem 1.1. Suppose that $\nu_{K,N}$ satisfies (79), then the following estimate holds*

$$\mathbb{P}\left(|\nu_{K,N}-\nu|>\gamma\nu\right)\le 2e^{-\frac{dKM\gamma^2}{8}},\quad \textit{for all } \gamma\in(0,1). \tag{85}$$

Proof. Recall that

$$X_i^{(n+1)}=X_i^{(n)}+\int_{t_n}^{t_{n+1}}\frac{1}{N-1}\sum_{j\neq i}^N F^N\left(X_i^s-X_j^s\right)ds+\sqrt{2\nu\Delta t}\mathcal{N}_i^{(n)},\quad i=1,\cdots,K, \tag{86}$$

then we know

$$\frac{X_i^{(n+1)}-X_i^{(n)}-\int_{t_n}^{t_{n+1}}\frac{1}{N-1}\sum_{j\neq i}^N F^N\left(X_i^s-X_j^s\right)ds}{\sqrt{2\nu\Delta t}}\sim\mathcal{N}(0,1)^d. \tag{87}$$

Notice that the random variable

$$S := \frac{1}{2\nu\Delta t}\sum_{i=1}^{K}\sum_{n=0}^{M-1}\left|X_i^{(n+1)} - X_i^{(n)} - \int_{t_n}^{t_{n+1}} \frac{1}{N-1}\sum_{j\neq i}^{N} F^N\big(X_i^s - X_j^s\big)ds\right|^2$$

is distributed according to the chi-squared distribution with dNM degrees of freedom. This is usually denoted as

$$S \;\sim\; \chi^2(dKM). \tag{88}$$

Recall a simple fact from probability theory, we know $\mathbb{E}[S] = dKM$ and

$$\mathrm{Var}[S] = \mathbb{E}\left[(S - dKM)^2\right] = 2dKM. \tag{89}$$

Recall that the estimate of ν is given by

$$\nu_{K,N} = \frac{1}{2dKT}\sum_{i=1}^{K}\sum_{n=0}^{M-1}\left|X_i^{(n+1)} - X_i^{(n)} - \int_{t_n}^{t_{n+1}} \frac{1}{N-1}\sum_{j\neq i}^{N} F^N\big(X_i^s - X_j^s\big)ds\right|^2, \tag{90}$$

which leads to

$$\mathbb{E}\left[\left(\frac{\nu_{K,N}}{\nu} - 1\right)^2\right] = \frac{2}{dKM}. \tag{91}$$

Hence we have

$$\mathbb{E}\left[(\nu_{K,N} - \nu)^2\right] = \frac{2\nu^2}{dKM}. \tag{92}$$

Also by the concentration of χ^2 variable, we have the following two sided tail bound

$$\mathbb{P}\left(\left|\frac{S}{dKM} - 1\right| > \gamma\right) \le 2e^{-\frac{dKM\gamma^2}{8}}, \quad \text{for all } \gamma \in (0,1), \tag{93}$$

which is a direct result from the Bernstein's inequality as the form showed in [8, Corollary 2.11]. And it leads to

$$\mathbb{P}\left(|\nu_{K,N} - \nu| > \gamma\nu\right) \le 2e^{-\frac{dKM\gamma^2}{8}}, \quad \text{for all } \gamma \in (0,1). \tag{94}$$

Hence it concludes the proof. □

Collecting estimates (94), (84) and (83), one has

$$\mathbb{P}\left(|\widehat{\nu} - \nu| \le C\Delta t(1 + N^{-2\delta}\log^2(N)) + \gamma\nu\right) \ge 1 - N^{\alpha} - 2e^{-\frac{dKM\gamma^2}{8}}, \tag{95}$$

for all $\gamma \in (0,1)$. Hence Theorem 1.1 has been proved.

4 Extension to regular interacting kernel $F \in W^{1,\infty}(\mathbb{R}^d)$

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

$$F \in W^{1,\infty}(\mathbb{R}^d). \tag{96}$$

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:

$$dX_i^t = \frac{1}{N-1}\sum_{j\neq i}^{N} F\big(X_i^t - X_j^t\big)\,dt + \sqrt{2\nu}\,dB_i^t, \quad i = 1,\cdots,N, \tag{97}$$

where the initial data $\{X_i^0\}_{i=1}^N$ are i.i.d. sharing the common density $\rho_0 \in L^1 \cap L^\infty(\mathbb{R}^d)$. Then the solution ρ to the mean field equation:

$$\partial_t \rho = \nu \Delta \rho - \nabla \cdot (\rho F * \rho), \quad x \in \mathbb{R}^d,\ t > 0, \tag{98a}$$

$$\rho(x, 0) = \rho_0(x), \tag{98b}$$

has the following regularity for any $T > 0$

$$\|\rho\|_{L^\infty(0,T;L^1\cap L^\infty(\mathbb{R}^d))} \le C\left(T, \|\rho_0\|_{L^1\cap L^\infty(\mathbb{R}^d)}, \|F\|_{W^{1,\infty}(\mathbb{R}^d)}\right) =: C_{F,\rho_0}. \tag{99}$$

Take a time step $\Delta t > 0$ and let $t_n := n\Delta t$ and $M := \frac{T}{\Delta t}$ (we assume that $\frac{T}{\Delta t}$ is an integer). Denote $X_i^{(n)} := X_i^{t_n} = X_i^{n\Delta t}$ as the solution to (97) at time t_n. Namely, one has

$$X_i^{(n+1)} - X_i^{(n)} = \int_{t_n}^{t_{n+1}} \frac{1}{N-1} \sum_{j\neq i}^{N} F\left(X_i^s - X_j^s\right) ds + \sqrt{2\nu\Delta t}\mathcal{N}_i^{(n)}, \tag{100}$$

where $\mathcal{N}_i^{(n)} \sim \mathcal{N}(0, 1)^d$, i.e. the standard Gaussian distribution in dimension d.

Then we are ready to define our estimator for the diffusion parameter as before

$$\widehat{\nu} := \frac{1}{6dKT} \sum_{i=1}^{K} \sum_{n=0}^{M-1} \left|X_i^{(n+1)} - X_i^{(n)}\right|^2, \tag{101}$$

where $1 \ll K \ll N$, which means we only have partial observations.

The extended result can be described in the following theorem.

Theorem 4.1. *Suppose that $F(x) \in W^{1,\infty}(\mathbb{R}^d)$ and $0 \le \rho_0(x) \in L^1 \cap L^\infty(\mathbb{R}^d)$. For any $T > 0$, take a time step $\Delta t > 0$ and define $t_n := n\Delta t$ and $M := \frac{T}{\Delta t}$. Let $\{X_i^{(n)}\}_{i=1,n=0}^{K,M}$ be the sample trajectories satisfying (97) at time t_n. Then there exists some $C_\alpha > 0$ and $N_0 > 0$ depending only on ν, α, T, $\|F\|_{W^{1,\infty}(\mathbb{R}^d)}$ and $\|\rho_0\|_{L^1\cap L^\infty(\mathbb{R}^d)}$, such that for $N \ge N_0$, the estimator $\widehat{\nu}$ defined in (101) is an approximation of ν, and the following estimate holds*

$$\mathbb{P}\left(|\widehat{\nu} - \nu| \le C_\alpha \Delta t(1 + N^{-1}\log^2(N)) + \nu\gamma\right) \ge 1 - N^{-\alpha} - 2e^{-\frac{dKM\gamma^2}{8}}, \tag{102}$$

for all $\gamma \in (0, 1)$. In particular, if we choose $\Delta t = \gamma = K^{-\frac{1}{4}}$, it follows from (102) that

$$\mathbb{P}\left(|\widehat{\nu} - \nu| \le C_\alpha K^{-\frac{1}{4}}\right) \ge 1 - K^{-\alpha}. \tag{103}$$

Proof. First, let us we split the estimator $\widehat{\nu}$ into three parts:

$$\begin{aligned}
\hat{\nu} &= \frac{1}{6dKT} \sum_{i=1}^{K} \sum_{n=0}^{M-1} |X_i^{(n+1)} - X_i^{(n)}|^2 \\
&\le \frac{1}{2dKT} \sum_{i=1}^{K} \sum_{n=0}^{M-1} \left| X_i^{(n+1)} - X_i^{(n)} - \int_{t_n}^{t_{n+1}} \frac{1}{N-1} \sum_{j\neq i}^{N} F\left(X_i^s - X_j^s\right) ds \right|^2 \\
&\quad + \frac{1}{2dKT} \sum_{i=1}^{K} \sum_{n=0}^{M-1} \left| \int_{t_n}^{t_{n+1}} \left(\frac{1}{N-1} \sum_{j\neq i}^{N} F\left(X_i^s - X_j^s\right) - \int_{\mathbb{R}^d} F(X_i^s - y)\rho(y, s)\, dy \right) ds \right|^2 \\
&\quad + \frac{1}{2dKT} \sum_{i=1}^{K} \sum_{n=0}^{M-1} \left| \int_{t_n}^{t_{n+1}} \int_{\mathbb{R}^d} F(X_i^s - y)\rho(y, s)\, dyds \right|^2 \\
&=: \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3.
\end{aligned} \tag{104}$$

According to Lemma 2.3, one has

$$|\mathcal{I}_3| \le C\Delta t, \tag{105}$$

where C depends only on T, $\|F\|_{W^{1,\infty}(\mathbb{R}^d)}$ and $\|\rho_0\|_{L^1\cap L^\infty(\mathbb{R}^d)}$. It follows from Proposition 3.1 that

$$\mathbb{P}\left(|\mathcal{I}_1-\nu|>\gamma\nu\right)\le 2e^{-\frac{dKM\gamma^2}{8}},\quad \text{for all } \gamma\in(0,1). \tag{106}$$

Now it is left to get the estimate of $\mathcal{I}_2$. The main idea behind the proof is also to construct a mean-field dynamic system $\{Y_i^t\}_{i=1}^N$ without interaction:

$$dY_i^t=\int_{\mathbb{R}^d}F\big(Y_i^t-y\big)\rho(y,t)dy\,dt+\sqrt{2\nu}\,dB_i^t,\quad i=1,\cdots,N, \tag{107}$$

here again we let $\{Y_i^t\}_{i=1}^N$ has the same initial condition as $\{X_i^t\}_{i=1}^N$ (i.i.d. with common density ρ_0) . Consider the quantity $e(t)$ defined as

$$e(t):=\|X_t-Y_t\|_\infty. \tag{108}$$

Following the same procedure as in Lemma 2.5 and Proposition 2.1, one can prove that there exists some $C_{1,\alpha}$ depending only on ν, α, T, $\|F\|_{W^{1,\infty}(\mathbb{R}^d)}$ and $\|\rho_0\|_{L^1\cap L^\infty(\mathbb{R}^d)}$ such that

$$\mathbb{P}\left(\left\|\frac{1}{N-1}\sum_{j\neq i}^N F\big(Y_i^t-Y_j^t\big)-\int_{\mathbb{R}^d}F\big(Y_i^t-y\big)\rho(y,t)dy\right\|_\infty\ge C_{1,\alpha}N^{-\frac12}\log(N)\right)\le N^{-\alpha}. \tag{109}$$

We denote the event

$$\mathcal{C}:=\left\{\left\|\frac{1}{N-1}\sum_{j\neq i}^N F\big(Y_i^t-Y_j^t\big)-\int_{\mathbb{R}^d}F\big(Y_i^t-y\big)\rho(y,t)dy\right\|_\infty\le C_{1,\alpha}N^{-\frac12}\log(N)\right\}, \tag{110}$$

Then using the fact $\frac{d\|x\|_\infty}{dt}\le\|\frac{dx}{dt}\|_\infty$, one concludes that under the event $\mathcal{C}$

$$\begin{aligned}\frac{de(t)}{dt}&\le\left\|\frac{1}{N-1}\sum_{j\neq i}^N F\big(X_i^t-X_j^t\big)-\int_{\mathbb{R}^d}F\big(Y_i^t-y\big)\rho(y,t)dy\right\|_\infty\\&\le\left\|\frac{1}{N-1}\sum_{j\neq i}^N F\big(X_i^t-X_j^t\big)-\frac{1}{N-1}\sum_{j\neq i}^N F\big(Y_i^t-Y_j^t\big)\right\|_\infty\\&\quad+\left\|\frac{1}{N-1}\sum_{j\neq i}^N F\big(Y_i^t-Y_j^t\big)-\int_{\mathbb{R}^d}F\big(Y_i^t-y\big)\rho(y,t)dy\right\|_\infty\\&\le C\|X_t-Y_t\|_\infty+CN^{-\frac12}\log(N),\end{aligned} \tag{111}$$

which leads to

$$\sup_{t\in[0,T]}\|X_t-Y_t\|_\infty\le CN^{-\frac12}\log(N), \tag{112}$$

where C depends only on ν, α, T, $\|F\|_{W^{1,\infty}(\mathbb{R}^d)}$ and $\|\rho_0\|_{L^1\cap L^\infty(\mathbb{R}^d)}$. Based on this mean-field limit result, we can prove error estimate on interaction as in Theorem 2.2. It is easy to compute that

$$\begin{aligned}&\left|\int_{\mathbb{R}^d}F(x-y)\rho(y,t)dy-\frac1N\sum_{j=1}^N F(x-X_j^t)\right|\\\le&\left|\int_{\mathbb{R}^d}F(x-y)\rho(y,t)dy-\frac1N\sum_{j=1}^N F(x-Y_j^t)\right|\\&+\left|\frac{1}{N-1}\sum_{j\neq i}^N F(x-Y_j^t)-\frac1N\sum_{j=1}^N F(x-X_j^t)\right|\\=:&e_1^t(x)+e_2^t(x).\end{aligned}$$

Using (112) implies that under the event $\mathcal{C}$

$$\sup_{0\le t\le T}\sup_x e_2^t(x) \le CN^{-\frac{1}{2}}\log(N), \tag{113}$$

where C depends only on ν, α, T, $\|F\|_{W^{1,\infty}(\mathbb{R}^d)}$ and $\|\rho_0\|_{L^1\cap L^\infty(\mathbb{R}^d)}$. As for $e_1^t(x)$, following the procedure in the proof of (77), it yields that

$$\mathbb{P}\left(\sup_{0\le t\le T}\sup_x e_1^t(x) \le CN^{-\frac{1}{2}}\log(N)\right) \ge 1 - N^{-\alpha}. \tag{114}$$

Combining (113) and (114), it leads to

$$\mathbb{P}\Bigg(\sup_{0\le t\le T}\sup_x \left|\int_{\mathbb{R}^d} F(x-y)\rho(y,t)dy - \frac{1}{N}\sum_{j=1}^N F(x-X_j^t)\right| \\ \le CN^{-\frac{1}{2}}\log(N)\Bigg) \le 1 - N^{-\alpha},$$

which leads to

$$\mathbb{P}\left(|\mathcal{I}_2| \le C\Delta t N^{-1}\log^2(N)\right) \ge 1 - N^{-\alpha}, \tag{115}$$

where C depends only on ν, α, T, $\|F\|_{W^{1,\infty}(\mathbb{R}^d)}$ and $\|\rho_0\|_{L^1\cap L^\infty(\mathbb{R}^d)}$.

Collecting (106), (115) and (105), we obtain our result

$$\mathbb{P}\left(|\widehat{\nu}-\nu| \le C\Delta t(1 + N^{-1}\log^2(N)) + \nu\gamma\right) \ge 1 - N^{-\alpha} - 2e^{-\frac{dKM\gamma^2}{8}}, \quad \text{for all } \gamma\in(0,1),$$

where C depends only on ν, α, T, $\|F\|_{W^{1,\infty}(\mathbb{R}^d)}$ and $\|\rho_0\|_{L^1\cap L^\infty(\mathbb{R}^d)}$. $\square$

References

[1] Y. Aït-Sahalia, P. A. Mykland, et al. Estimators of diffusions with randomly spaced discrete observations: a general theory. *The Annals of Statistics*, 32(5):2186–2222, 2004.

[2] F. M. Bandi and G. Moloche. On the functional estimation of multivariate diffusion processes. *Econometric Theory*, pages 1–51, 2017.

[3] N. Bellomo, A. Bellouquid, Y. Tao, and M. Winkler. Toward a mathematical theory of Keller-Segel models of pattern formation in biological tissues. *Mathematical Models and Methods in Applied Sciences*, 25(09):1663–1763, 2015.

[4] A. J. Bernoff and C. M. Topaz. Nonlocal aggregation models: A primer of swarm equilibria. *SIAM Review*, 55(4):709–747, 2013.

[5] B. M. Bibby and M. Sørensen. Martingale estimation functions for discretely observed diffusion processes. *Bernoulli*, pages 17–39, 1995.

[6] F. Bolley, J. A. Canizo, and J. A. Carrillo. Stochastic mean-field limit: non-Lipschitz forces and swarming. *Mathematical Models and Methods in Applied Sciences*, 21(11):2179–2210, 2011.

[7] M. Bongini, M. Fornasier, M. Hansen, and M. Maggioni. Inferring interaction rules from observations of evolutive systems I: The variational approach. *Math. Model. Methods Appl. Sci.*, 27:909–951, 2017.

[8] S. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities: A nonasymptotic theory of independence*. Oxford university press, 2013.

[9] W. Braun and K. Hepp. The Vlasov dynamics and its fluctuations in the $\frac{1}{N}$ limit of interacting classical particles. *Communications in mathematical physics*, 56(2):101–113, 1977.

[10] J. A. Carrillo, K. Craig, and F. S. Patacchini. A blob method for diffusion. *arXiv preprint arXiv:1709.09195*, 2017.

[11] J. A. Carrillo, M. Fornasier, G. Toscani, and F. Vecil. Particle, kinetic, and hydrodynamic models of swarming. *Mathematical modeling of collective behavior in socio-economic and life sciences*, pages 297–336, 2010.

[12] K. Craig and A. Bertozzi. A blob method for the aggregation equation. *Mathematics of computation*, 85(300):1681–1717, 2016.

[13] D. Dacunha-Castelle and D. Florens-Zmirou. Estimation of the coefficients of a diffusion from discrete observations. *Stochastics: An International Journal of Probability and Stochastic Processes*, 19(4):263–284, 1986.

[14] R. L. Dobrushin. Vlasov equations. *Functional Analysis and Its Applications*, 13(2):115–123, 1979.

[15] G. Dohnal. On estimating the diffusion coefficient. *Journal of Applied Probability*, 24(1):105–114, 1987.

[16] J. Fan and C. Zhang. A reexamination of diffusion estimators with applications to financial model validation. *Journal of the American Statistical Association*, 98(461):118–134, 2003.

[17] R. C. Fetecau, H. Huang, and W. Sun. Propagation of chaos for aggregation equations with newtonian potential over bounded domains. preprint.

[18] N. Fournier and B. Jourdain. Stochastic particle approximation of the Keller-Segel equation and two-dimensional generalization of Bessel processes. *The Annals of Applied Probability*, 27(5):2807–2861, 2017.

[19] A. García-Cañizares and P. Pickl. Microscopic derivation of the Keller-Segel equation in the sub-critical regime. *arXiv preprint arXiv:1703.04376*, 2017.

[20] F. Golse. On the dynamics of large particle systems in the mean field limit. In *Macroscopic and Large Scale Phenomena: Coarse Graining, Mean Field Limits and Ergodicity*, pages 1–144. Springer, 2016.

[21] J. Goodman. Convergence of the random vortex method. *Communications on Pure and Applied Mathematics*, 40(2):189–220, 1987.

[22] C. K. Hemelrijk and H. Hildenbrandt. Self-organized shape and frontal density of fish schools. *Ethology*, 114(3):245–254, 2008.

[23] H. Hildenbrandt, C. Carere, and C. K. Hemelrijk. Self-organized aerial displays of thousands of starlings: a model. *Behavioral Ecology*, 21(6):1349–1359, 2010.

[24] H. Huang and J.-G. Liu. Discrete-in-time random particle blob method for the Keller-Segel equation and convergence analysis. *Communication in Mathematical Sciences*, 15(7):1821–1842, 2017.

[25] H. Huang and J.-G. Liu. Error estimate of a random particle blob method for the Keller-Segel equation. *Mathematics of Computation*, 86:2719–2744, 2017.

[26] H. Huang, J.-G. Liu, and P. Pickl. On the mean field limit for the Vlasov-Poisson-Fokker-Planck system. preprint.

[27] M. Huzak. Approximate maximum likelihood estimation of diffusion parameters from discrete observations. *arXiv preprint arXiv:1607.06699*, 2016.

[28] P.-E. Jabin. A review of the mean field limits for Vlasov equations. *Kinet. Relat. Models*, 7(4):661–711, 2014.

[29] M. Ji and M. Egerstedt. Distributed coordination control of multiagent systems while preserving connectedness. *IEEE Transactions on Robotics*, 23(4):693–703, 2007.

[30] 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.

[31] M. Kessler. Estimation of an ergodic diffusion from discrete observations. *Scandinavian Journal of Statistics*, 24(2):211–229, 1997.

[32] K. Law, A. Stuart, and K. Zygalakis. *Data Assimilation*. Springer, 2015.

[33] D. Lazarovici and P. Pickl. A mean field limit for the Vlasov-Poisson system. *Archive for Rational Mechanics and Analysis*, 225:1201–1231, 2017.

[34] D.-G. Long. Convergence of the random vortex method in two dimensions. *Journal of the American Mathematical Society*, 1(4):779–804, 1988.

[35] S. Motsch and E. Tadmor. Heterophilious dynamics enhances consensus. *SIAM review*, 56(4):577–621, 2014.

[36] C. S. Patlak. Random walk with persistence and external bias. *The bulletin of mathematical biophysics*, 15(3):311–338, 1953.

[37] B. P. Rao. *Statistical inference for diffusion type processes*. Wiley, 2010.

[38] H. Spohn. *Dynamics of charged particles and their radiation field*. Cambridge university press, 2004.

[39] N. Yoshida. Estimation for diffusion processes from discrete observation. *Journal of Multivariate Analysis*, 41(2):220–242, 1992.