Convergence of Diffusion-Drift Many Particle Systems in Probability under a Sobolev Norm

Jian-Guo Liu and Yuan Zhang

Abstract In this paper we develop a new martingale method to show the convergence of the regularized empirical measure of many particle systems in probability under a Sobolev norm to the corresponding mean field PDE. Our method works well for the simple case of Fokker Planck equation and we can estimate a lower bound of the rate of convergence. This method can be generalized to more complicated systems with interactions.

Key words: Many particle system, martingale method, energy-dissipation inequality.

1 Introduction

We are considering the stochastic processes $\{X_i(t)\}_{i=1}^N$ in $\mathbb{R}^d$ of the following SDE

$$dX_i(t) = \mathbf{F}\left(X_i(t),t\right)dt + dB_i(t), \quad i = 1,2,\cdots,N \tag{1}$$

with an initial condition $X_i(0)$ and a sequence of independent d-dimensional standard Brownian motions $\{B_i(t)\}_{i=1}^N$. We show the system converges to the following Fokker-Planck equation

$$\frac{\partial \rho}{\partial t}(x,t) = \frac{1}{2}\Delta\rho(x,t) - \nabla\cdot(\rho(x,t)\mathbf{F}(x,t)) \tag{2}$$

with boundary condition $\rho(x,0)\in L^2(R^d)$, in the $L^\infty(L^2)\cap L^2(H^1)$ norm and in probability, and we estimate the convergence rate. The motivation of developing this kind of estimates is hoping that it can be adapted in the analysis of the propagation of chaos and the mean field limit for some interacting many particle systems

$$dX_i(t) = \frac{1}{N}\sum_{j\neq i}^N \mathbf{F}_0\big(X_i(t)-X_j(t)\big)dt + \sigma dB_t^i, \quad i = 1,\cdots,N\,. \tag{3}$$

For the simplified model (1) and (2), we show in Theorem 1 that the regularized empirical measure:

$$\rho_{\varepsilon,N}(x,t) = \frac{1}{N}\sum_{i=1}^N \varphi_\varepsilon\left(x - X_i(t)\right), \quad \varphi_\varepsilon(x) = \frac{1}{\varepsilon^d}\varphi(\frac{x}{\varepsilon}), \quad \varepsilon = N^{-1/3d} \tag{4}$$

Jian-Guo Liu
Dept. of Physics and Dept. of Mathematics, Duke U., Box 90320 Durham, NC, USA 27708-0320, e-mail: jliu@math.duke.edu

Yuan Zhang
Dept. of Mathematics, UCLA, Box 951555 Los Angeles, CA, USA 90095-1555,
e-mail: yuanzhang@math.ucla.edu

MSC 2010 subject classifications: 35Q70, 65M75.
The research was partially supported by KI-Net NSF RNMS Grant No. 1107291, and NSF DMS Grant No. 1514826

has a rate of convergence to the mean field solution ρ in the Sobolev space $L^\infty\big(0,T;L^2(\mathbb{R}^d)\big)\cap L^2\big(0,T;H^1(\mathbb{R}^d)\big)$, for any $T>0$,

$$P\Big(\sup_{s\in[0,T]}\big(\|(\rho-\rho_{\varepsilon,N})(\cdot,s)\|^2+\int_0^T\|\nabla(\rho-\rho_{\varepsilon,N})(\cdot,s)\|^2ds\big) \geq C_T\big(\|(\rho-\rho_{\varepsilon,N})(\cdot,0)\|^2+N^{-1/6d}\big)\Big)\leq C_TN^{-1/6d} \tag{5}$$

for some function C_T that depends only on T, $||\rho_0||_{L^2}$ and the Lipschitz constant of $\mathbf{F}$. Thus the free energy-dissipation inequality for the difference between the regularized empirical measure $\rho_{\varepsilon,N}$ and the mean field density ρ holds with high probability. The second term in (5) is the dissipation term and it is contributed from the Brownian motion. In the usual coupling method, the contribution of the Brownian motion is removed in the estimation and hence the dissipation term is lost. When the interaction kernel $\mathbf{F}(x)$ is Lipschitz continuous, the propagation of chaos can be directly justified by the McKean's coupling method [13, 17]. In some physically important systems such as two dimensional Navier-Stokes equation [6, 7, 11, 12, 15] and the Keller-Segel equation [16, 10], the interaction is given by the gradient or the curl of the Newtonian potential. The free energy-dissipation inequality plays a curial role in analyze such kind of systems, particularly, mathematical justifications of the propagation of chaos and mean field limit.

Recently, we have shown that the new martingale method developed in this paper can be generalized to prove the convergence under first order Sobolev norm when interactions are introduced in the particle system. Let $\{X_i(t)\}_{i=1}^N$ be the N-particles system defined in (3), where $\mathbf{F}_0$ is a bounded function and Lipschitz continuous against x with Lipschitz constant L_F for all $t\geq 0$, and the deterministic PDE model be as follows:

$$\begin{cases}\dfrac{\partial\rho}{\partial t}(x,t)=\dfrac{1}{2}\Delta\rho-\nabla\cdot(\rho\mathbf{F}(x,t))\\ \mathbf{F}(x,t)=\displaystyle\int_{R^d}\mathbf{F}_0(x-y)\rho(y,t)dy.\end{cases} \tag{6}$$

To show the regularized empirical measure converges to the solution of the PDE above, we first introduce an intermediate self-consistent system $\{\hat{X}_i(t)\}_{i=1}^N$, which is defined by the following SDE

$$d\hat{X}_i(t)=\mathbf{F}\left(\hat{X}_i(t),t\right)dt+dB_i(t),\quad i=1,2,\cdots,N \tag{7}$$

where

$$\mathbf{F}(x,t)=\int_{R^d}\mathbf{F}_0(x-y)\rho(y,t)dy. \tag{8}$$

According to previous study, we are able to control the distance between $\hat{X}_i(t)$ and $X_i(t)$. And we can use the similar martingale method as in this paper to control the distance between the self-consistent system $\{\hat{X}_i(t)\}_{i=1}^N$ and the deterministic PDE model. The details of the proof will be presented in a separate paper.

The use of the regularized empirical measure is important in computation and the regularized kernel φ is known as a blob function in the vortex method. Pioneered by Chorin in 1973 [2], the random vertex blob method is one of the most successful computational methods for fluid dynamics and other related fields. The success of the method is exemplified by the accurate computation of flow past a cylinder at the Reynolds numbers up to 9500 in the 1990s [9]. The convergence analysis for the random vortex method for the Navier-Stokes equation is given by [7, 11, 12] in the 1980s. We refer to the book [3] for theoretical and practical use of vortex methods, refer to Goodman [7] and Long [11] for the convergence analysis of the random vortex method to the Navier-Stokes equation. We also hoped that the estimation (5) can be adapted to do numerical analysis.

2 Convergence in $L^\infty(L^2)\cap L^2(H^1)$ norm

As described in Introduction, in this section we will show that the regularized empirical measure of the many particle system defined in (1) converge to the solution of the Fokker Planck equation (2) in probability under $L^\infty(L^2)\cap L^2(H^1)$ norm. In this paper, we will denote the $L^2(\mathbb{R}^d)$ norm as $\|\cdot\|$. To show this, we need to make some assumption of the initial state to make sure that the particles are not too close to each other. In the following work, we show that our convergence result holds under either of the following 2 assumptions:

Assumption 1: There is some constant $C < \infty$ independent to N, such that for any N and $i \neq j$, $m, n \leq N$, and $\delta_{i,j} = X_i(0) - X_j(0)$, we always have $|\delta_{i,j}| \geq 2CN^{-1/d}$.

Assumption 2: $(X_1(0), \cdots, X_N(0))$ has a joint distribution such that there is some constant $C < \infty$ independent to N, such that for any N and $i \neq j$, $i, j \leq N$, $\delta_{i,j}$ has a density function $f_{i,j,N}$ and it satisfies $\|f_{i,j,N}\|_\infty \leq C$.

Remark: A special case of Assumption 2 is that $X_1(0), X_2(0), \cdots, X_N(0)$ are i.i.d. with density $\rho(x,0)$, since that for any i, j, $X_i(0) - X_j(0)$ has density function

$$f_{i,j,N}(a) = \int_{\mathbb{R}^d} \rho(x,0)\rho(x+a,0)\,dx \leq \|\rho(x,0)\|^2 < \infty.$$

For any N, we consider the regularized empirical measure:

$$\rho_{\varepsilon,N}(x,t) = \frac{1}{N}\sum_{i=1}^{N} \varphi_\varepsilon\left(x - X_i(t)\right). \tag{9}$$

To show that $\rho_{\varepsilon,N}(x,t)$ converges to ρ, we have theorem as follows:

Theorem 1. *Let $X_i(t), i = 1, \cdots, N$ be solutions of stochastic differential equation (1) with initial data $X_i(0)$ satisfing either Assumption 1 or Assumption 2 and $\rho_{\varepsilon,N}$ be the constructed regularized empirical measure (9) with regularized parameter $\varepsilon_N = N^{-1/3d}$. Let ρ be the solution of the corresponding mead field equation (2) with initial density $\rho_0 \in L^2(\mathbb{R}^d)$. Then, there is a positive function $c(t), t > 0$ (will be specified in (83)) dependent only on t, φ, and $||\rho_0||$, such that*

$$\begin{aligned} P\Big(\sup_{s\in[0,t]} \big(\|(\rho - \rho_{\varepsilon_N,N})(\cdot,s)\|^2 + \int_0^s \|\nabla(\rho - \rho_{\varepsilon_N,N})(\cdot,h)\|^2\,dh\big) \\ < 2e^{C_0 t}\big(\|(\rho - \rho_{\varepsilon_N,N})(\cdot,0)\|^2 + c(t)N^{-1/6d}\big)\Big) \geq 1 - c(t)N^{-1/6d}. \end{aligned} \tag{10}$$

where $C_0 = 2dL_F$ and L_F is the Lipschitz constant of $\mathbf{F}$.

The proof of this theorem is divided into one proposition (Proposition 1) and four lemmas (Lemma 2.1-2.4). We will give the proof of this theorem after the proof of these preliminary results.

Proposition 1. *For the difference between the PDE density ρ and the empirical measure $\rho_{\varepsilon,N}$, we have*

$$\begin{aligned} \|(\rho - \rho_{\varepsilon,N})(\cdot,t)\|^2 = \|(\rho - \rho_{\varepsilon,N})(\cdot,0)\|^2 - \int_0^t \|\nabla(\rho - \rho_{\varepsilon,N})(\cdot,s)\|^2\,ds \\ - \int_0^t \int_{\mathbb{R}^d} \nabla\cdot\mathbf{F}(x,s)\left((\rho - \rho_{\varepsilon,N})(x,s)\right)^2 dxds \\ + Res(t) + \tilde{M}_t + M_t + \frac{t}{N}\|\nabla\varphi_\varepsilon\|_2^2 \end{aligned} \tag{11}$$

where M_t is defined by $M_t = \sum_{i=1}^N M_t^i$ with

$$M_t^i = \frac{2}{N}\int_0^t \int_{\mathbb{R}^d} \rho(x,s)\nabla\varphi_\varepsilon\left(x - X_i(s)\right)dx\cdot dB_i(s), \tag{12}$$

and

$$\tilde{M}_t = \sum_{n=1}^{N} \tilde{M}_t^i \tag{13}$$

with $\tilde{M}_t^i$ equals to

$$\frac{2}{N^2}\int_0^t \int_{\mathbb{R}^d} \varphi_\varepsilon(x)\left(\sum_{j=1}^{i-1}\nabla\varphi_\varepsilon\left(x + B_i(s) - B_j(s)\right) - \sum_{j=i+1}^{N}\nabla\varphi_\varepsilon\left(x + B_j(s) - B_i(s)\right)\right)dx\cdot dB_i(s) \tag{14}$$

and

$$Res(t) = 2\int_0^t \int_{\mathbb{R}^d} \frac{1}{N}\sum_{i=1}^{N} \varphi_\varepsilon\left(x - X_i(s)\right)\left(\mathbf{F}(x,s) - \mathbf{F}\left(X_i(s),s\right)\right)\cdot\nabla(\rho-\rho_{\varepsilon,N})(x,s)\,dxds. \tag{15}$$

Proof. To prove the proposition, first note that for any ε and N,

$$\|(\rho-\rho_{\varepsilon,N})(\cdot,t)\|^2 = \|\rho(\cdot,t)\|^2 - 2\int_{\mathbb{R}^d}\rho(x,t)\rho_{\varepsilon,N}(x,t)\,dx + \|\rho_{\varepsilon,N}(\cdot,t)\|^2.$$

First for the deterministic part of $\|\rho(\cdot,t)\|^2$, we have

$$\begin{aligned}\int_{\mathbb{R}^d}\rho(x,t)^2\,dx &= \int_{\mathbb{R}^d}\rho(x,0)^2\,dx + \int_0^t\int_{\mathbb{R}^d}\rho(x,s)\left(\Delta\rho(x,s) - 2\nabla\cdot(\rho\mathbf{F})(x,s)\right)dxds\\ &= \int_{\mathbb{R}^d}\rho(x,0)^2\,dx - \int_0^t\|\nabla\rho\|^2\,ds + 2\int_0^t\int_{\mathbb{R}^d}\rho(x,s)\mathbf{F}(x,s)\cdot\nabla\rho(x,s)\,dxds.\end{aligned} \tag{16}$$

Then for the second part which equals to

$$-\frac{2}{N}\int_{\mathbb{R}^d}\rho(x,t)\sum_{i=1}^{N}\varphi_\varepsilon\left(x - X_i(t)\right)dx,$$

note that for each i, by Ito's formula, we have

$$\begin{aligned}\rho(x,t)\varphi_\varepsilon\left(x-X_i(t)\right) =&\rho(x,0)\varphi_\varepsilon\left(x-X_n(0)\right) + \int_0^t\frac{\partial\rho(x,s)}{\partial t}\varphi_\varepsilon\left(x-X_i(s)\right)\,ds\\ &-\int_0^t\rho(x,s)\nabla\varphi_\varepsilon\left(x-X_i(s)\right)\cdot\mathbf{F}\left(X_i(s),s\right)\,ds\\ &-\int_0^t\rho(x,s)\nabla\varphi_\varepsilon\left(x-X_i(s)\right)\cdot dB_i(s)\\ &+\frac{1}{2}\int_0^t\rho(x,s)\Delta\varphi_\varepsilon\left(x-X_i(s)\right)\,ds.\end{aligned} \tag{17}$$

Note that for the second term of the sum above, according to the definition of the Fokker-Planck's PDE,

$$\int_0^t\frac{\partial\rho(x,s)}{\partial t}\varphi_\varepsilon\left(x-X_i(s)\right)\,ds = \int_0^t\left(\frac{1}{2}\Delta\rho(x,s) - \nabla\cdot\left(\rho(x,s)\mathbf{F}(x,s)\right)\right)\varphi_\varepsilon\left(x-X_i(s)\right)\,ds.$$

Then integrate it over $x\in R^d$, we have

$$\begin{aligned}&\int_0^t\int_{\mathbb{R}^d}\frac{\partial\rho(x,s)}{\partial t}\varphi_\varepsilon\left(x-X_i(s)\right)dxds\\ &\qquad=-\frac{1}{2}\int_0^t\int_{\mathbb{R}^d}\nabla\rho(x,s)\cdot\nabla\varphi_\varepsilon\left(x-X_i(s)\right)dxds\\ &\qquad\quad+\int_0^t\int_{\mathbb{R}^d}\rho(x,s)\mathbf{F}(x,s)\cdot\nabla\varphi_\varepsilon\left(x-X_i(s)\right)\,dxds.\end{aligned} \tag{18}$$

Remark: Here we can also apply Ito's formula on the integration itself and it is more rigorous since we do not need to change the order of integrations. However, since all the calculations are the same, we will use the followings notations for simplicity.

Then integrating the third term in (17) over $x\in R^d$ we have by divergence theorem that

$$\begin{aligned}&-\int_0^t\int_{\mathbb{R}^d}\rho(x,s)\nabla\varphi_\varepsilon\left(x-X_i(s)\right)\cdot\mathbf{F}\left(X_i(s),s\right)dxds\\ &\qquad=\int_0^t\int_{\mathbb{R}^d}\varphi_\varepsilon\left(x-X_i(s)\right)\mathbf{F}(X_i(s),s)\cdot\nabla\rho(x,s)\,dxds.\end{aligned} \tag{19}$$

Combining (17), (18) and (19) we have

$$
\begin{aligned}
&\int_{\mathbb{R}^d} \rho(x,t)\varphi_\varepsilon(x-X_i(t))\,dx \\
&\qquad = \int_{\mathbb{R}^d} \rho(x,0)\varphi_\varepsilon(x-X_i(0))\,dx \\
&\qquad\quad - \int_0^t \int_{\mathbb{R}^d} \nabla\rho(x,s)\cdot\nabla\varphi_\varepsilon(x-X_i(s))\,dxds \\
&\qquad\quad + \int_0^t \int_{\mathbb{R}^d} \rho(x,s)\mathbf{F}(x,s)\cdot\nabla\varphi_\varepsilon(x-X_i(s))\,dxds \\
&\qquad\quad + \int_0^t \int_{\mathbb{R}^d} \varphi_\varepsilon(x-X_i(s))\,\mathbf{F}(X_i(s),s)\cdot\nabla\rho(x,s)dxds - \frac{N}{2}M_t^i
\end{aligned}
\tag{20}
$$

where M_t^i is

$$
M_t^i = \frac{2}{N}\int_0^t \int_{\mathbb{R}^d} \rho(x,s)\nabla\varphi_\varepsilon(x-X_i(s))\,dx\cdot dB_i(s).
$$

Summing up over $i = 1,2,\cdots,N$ we have

$$
\begin{aligned}
&-2\int_{\mathbb{R}^d} \rho_{\varepsilon,N}(x,t)\rho(x,t)\,dx \\
&\qquad = -2\int_{\mathbb{R}^d} \rho_{\varepsilon,N}(x,0)\rho(x,0)\,dx + 2\int_0^t \int_{\mathbb{R}^d} \nabla\rho(x,s)\cdot\nabla\rho_{\varepsilon,N}(x,s)\,dxds \\
&\qquad\quad - 2\int_0^t \int_{\mathbb{R}^d} \rho(x,s)\mathbf{F}(x,s)\cdot\nabla\rho_{\varepsilon,N}(x,s)\,dxds \\
&\qquad\quad - 2\int_0^t \int_{\mathbb{R}^d} \frac{1}{N}\sum_{i=1}^N \varphi_\varepsilon(x-X_i(s))\,\mathbf{F}(X_i(s),s)\cdot\nabla\rho(x,s)\,dxds \\
&\qquad\quad + M_t
\end{aligned}
\tag{21}
$$

where M_t is the first martingale term in (11) defined in (12). I.e.,

$$
M_t = 2\int_0^t \int_{\mathbb{R}^d} \rho(x,s)\frac{1}{N}\sum_{i=1}^N \nabla\varphi_\varepsilon(x-X_i(s))\,dx\cdot dB_i(s).
$$

Lastly, for the part of $\int_{\mathbb{R}^d} \rho_{\varepsilon,N}(x,t)\rho_{\varepsilon,N}(x,t)\,dx$ which equals to

$$
\begin{aligned}
&\frac{1}{N^2}\sum_{i=1}^N \int_{\mathbb{R}^d} \varphi_\varepsilon(x-X_i(t))^2\,dx + \frac{2}{N^2}\sum_{i<j}\int_{\mathbb{R}^d} \varphi_\varepsilon(x-X_j(t))\,\varphi_\varepsilon(x-X_i(t))\,dx \\
&\qquad = \frac{1}{N}\|\varphi_\varepsilon\|_2^2 + \frac{2}{N^2}\sum_{i<j}\int_{\mathbb{R}^d} \varphi_\varepsilon(x-X_j(t))\,\varphi_\varepsilon(x-X_i(t))\,dx.
\end{aligned}
\tag{22}
$$

And for each $i < j$,

$$
\int_{\mathbb{R}^d} \varphi_\varepsilon(x-X_j(t))\,\varphi_\varepsilon(x-X_i(t))\,dx = \int_{\mathbb{R}^d} \varphi_\varepsilon(x)\,\varphi_\varepsilon(x+X_j(t)-X_i(t))\,dx.
$$

Then we can again apply the Ito's formula on $\varphi_\varepsilon(x+X_j(t)-X_i(t))$:

$$
\begin{aligned}
&\varphi_\varepsilon\big(x+X_j(0)-X_i(0)\big) + \int_0^t \Delta\varphi_\varepsilon(x+X_j(s)-X_i(s))\,ds \\
&\qquad + \int_0^t \nabla\varphi_\varepsilon(x+X_j(s)-X_i(s))\cdot(\mathbf{F}(X_j(s),s)-\mathbf{F}(X_i(s),s))\,ds \\
&\qquad + \int_0^t \nabla\varphi_\varepsilon(x+X_j(s)-X_i(s))\cdot\big(dB_j(t)-dB_i(t)\big).
\end{aligned}
\tag{23}
$$

Integrating the first and second terms over x, we have

$$\int_{\mathbb{R}^d} \varphi_\varepsilon(x)\varphi_\varepsilon\left(x+X_j(0)-X_i(0)\right)dx = \int_{\mathbb{R}^d} \varphi_\varepsilon\left(x-X_j(0)\right)\varphi_\varepsilon\left(x-X_i(0)\right)dx$$

and

$$\int_0^t \int_{\mathbb{R}^d} \varphi_\varepsilon(x)\,\Delta\varphi_\varepsilon\left(x+X_j(s)-X_i(s)\right)dxds = -\int_0^t \int_{\mathbb{R}^d} \nabla\varphi_\varepsilon\left(x-X_j(s)\right)\cdot\nabla\varphi_\varepsilon\left(x-X_i(s)\right)dxds.$$

Moreover for the third term, we have

$$\begin{aligned}\int_0^t \int_{\mathbb{R}^d} &\varphi_\varepsilon(x)\nabla\varphi_\varepsilon\left(x+X_j(s)-X_i(s)\right)\cdot\mathbf{F}\left(X_j(s),s\right)dxdt \\ &= \int_0^t \int_{\mathbb{R}^d} \varphi_\varepsilon\left(x-X_j(s)\right)F\left(X_j(s),s\right)\cdot\nabla\varphi_\varepsilon\left(x-X_i(s)\right)dxds\end{aligned}$$

and

$$\begin{aligned}\int_0^t \int_{\mathbb{R}^d} &\varphi_\varepsilon(x)\nabla\varphi_\varepsilon\left(x+X_j(s)-X_i(s)\right)\cdot\mathbf{F}\left(X_i(s),s\right)dxdt \\ &= -\int_0^t \int_{\mathbb{R}^d} \varphi_\varepsilon\left(x-X_i(s)\right)F\left(X_i(s),s\right)\cdot\nabla\varphi_\varepsilon\left(x-X_j(s)\right)dxds.\end{aligned}$$

So after we sum up over all m,n, we have

$$\begin{aligned}\|\rho_{\varepsilon,N}(\cdot,t)\|^2 =& \|\rho_{\varepsilon,N}(\cdot,0)\|^2 - \int_0^t \|\nabla\rho_{\varepsilon,N}(\cdot,s)\|^2 ds \\ &+2\int_0^t \int_{\mathbb{R}^d} \left[\frac{1}{N}\sum_{i=1}^N \varphi_\varepsilon\left(x-X_i(s)\right)F\left(X_i(s),s\right)\right]\cdot\nabla\rho_{\varepsilon,N}(x,s)\,dxds \\ &+\tilde{M}_t + \frac{t}{N}\|\nabla\varphi_\varepsilon\|_2^2\end{aligned} \tag{24}$$

where $\tilde{M}_t$ is the second martingale term in (11), which is defined in (13) and (14). I.e.,

$$\tilde{M}_t = \sum_{i=1}^N \tilde{M}_t^i$$

with $\tilde{M}_t^i$ equals to

$$\frac{2}{N^2}\int_0^t \int_{\mathbb{R}^d} \varphi_\varepsilon(x)\left(\sum_{j=1}^{i-1}\nabla\varphi_\varepsilon\left(x+B_i(s)-B_j(s)\right) - \sum_{j=i+1}^{N}\nabla\varphi_\varepsilon\left(x+B_j(s)-B_i(s)\right)\right)dx\cdot dB_i(s).$$

Then combine (16), (21) and (24) we have

$$\begin{aligned}\|(\rho-\rho_{\varepsilon,N})(\cdot,t)\|^2 =& \|(\rho-\rho_{\varepsilon,N})(\cdot,0)\|^2 - \int_0^t \|\nabla(\rho-\rho_{\varepsilon,N})(\cdot,s)\|^2 ds \\ &+2\int_0^t \int_{\mathbb{R}^d} \left(\rho(x,s)\mathbf{F}(x,s) - \frac{1}{N}\sum_{i=1}^N \varphi_\varepsilon\left(x-X_i(s)\right)\mathbf{F}\left(X_i(s),s\right)\right)\cdot\nabla(\rho-\rho_{\varepsilon,N})(x,s)\,dxds \\ &+\tilde{M}_t + M_t + \frac{t}{N}\|\nabla\varphi_\varepsilon\|_2^2.\end{aligned} \tag{25}$$

Then plus and minus the term of $-\frac{1}{N}\sum_{i=1}^N \varphi_\varepsilon\left(x-X_i(s)\right)\mathbf{F}(x,s)\cdot\nabla(\rho-\rho_{\varepsilon,N})(x,s)$ we have

$$\begin{aligned}\|(\rho-\rho_{\varepsilon,N})(\cdot,t)\|^2 &= \|(\rho-\rho_{\varepsilon,N})(\cdot,0)\|^2 - \int_0^t \|\nabla(\rho-\rho_{\varepsilon,N})(\cdot,s)\|^2\,ds \\ &+ 2\int_0^t\int_{\mathbb{R}^d}(\rho-\rho_{\varepsilon,N})(x,s)\mathbf{F}(x,s)\cdot\nabla(\rho-\rho_{\varepsilon,N})(x,s)\,dxds \\ &+ 2\int_0^t\int_{\mathbb{R}^d}\frac{1}{N}\sum_{i=1}^N\varphi_\varepsilon\left(x-X_i(s)\right)\left(\mathbf{F}(x,s)-\mathbf{F}\left(X_i(s),s\right)\right)\cdot\nabla(\rho-\rho_{\varepsilon,N})(x,s)\,dxds \\ &+ \tilde{M}_t + M_t + \frac{t}{N}\|\nabla\varphi_\varepsilon\|_2^2.\end{aligned} \tag{26}$$

Applying Green's identity on the third term of equation (26) and recalling the definition of $\mathrm{Res}(t)$ in (15) we have

$$\begin{aligned}\|(\rho-\rho_{\varepsilon,N})(\cdot,t)\|^2 = \|(\rho-\rho_{\varepsilon,N})(\cdot,0)\|^2 &- \int_0^t \|\nabla(\rho-\rho_{\varepsilon,N})(\cdot,s)\|^2\,ds \\ &- \int_0^t\int_{\mathbb{R}^d}\nabla\cdot\mathbf{F}(x,s)\left((\rho-\rho_{\varepsilon,N})(x,s)\right)^2\,dxds \\ &+ \mathrm{Res}(t) + \tilde{M}_t + M_t + \frac{t}{N}\|\nabla\varphi_\varepsilon\|_2^2.\end{aligned} \tag{27}$$

Thus, the proof of Proposition 1 is complete.

Lemma 1. *For all $t \geq 0$, we have the second moment control*

$$E\left(M_t^2\right) \leq \frac{4}{N\varepsilon^{d+2}}\|\nabla\varphi\|^2\int_0^t \|\rho(x,s)\|^2 ds. \tag{28}$$

Proof. Here and in Lemma 2, we will use the natural filtration $\mathscr{F}_t^N$, which is generated by the Brownian motions $B_1(t),\cdots B_N(t)$. Note that $M_t = \sum_{i=1}^N M_t^i$ where

$$M_t^i = \frac{2}{N}\int_0^t\int_{\mathbb{R}^d}\rho(x,s)\nabla\varphi_\varepsilon\left(x-X_i(s)\right)dx\cdot dB_i(s) = \sum_{k=1}^d M_t^{i,k},$$

and

$$M_t^{i,k} = \frac{2}{N}\int_0^t\int_{\mathbb{R}^d}\rho(x,s)\frac{\partial\varphi_\varepsilon\left(x-X_i(s)\right)}{\partial x_k}dx\,dB_i^{(k)}(s).$$

The $B_i^{(k)}(s)$ in the equation above is the kth coordinate of the Brownian motion $B_i(t)$ and it is itself a one dimension Brownian motion and a square integrable martingale under filtration $\mathscr{F}_t^N$ noting that $B_i^{(k)}(s)$ is independent to $B_j^{(h)}(s)$ for all $h \neq k$, or $i \neq j$. For each i and k we have the integrand

$$Y_{i,k}(s) = \int_{\mathbb{R}^d}\rho(x,s)\frac{\partial\varphi_\varepsilon\left(x-X_i(s)\right)}{\partial x_k}dx$$

continuous and adapted to filtration $\mathscr{F}_s^N$. Moreover

$$|Y_{i,k}(s)| = \left|\int_{\mathbb{R}^d}\rho(x,s)\frac{\partial\varphi_\varepsilon\left(x-X_i(s)\right)}{\partial x_k}dx\right| \leq \left\|\frac{\partial\varphi_\varepsilon}{\partial x_k}\right\| \times \|\rho(\cdot,s)\| < \infty.$$

Thus by Theorem 5.2.3 in [5], for all $i \in \{1,2,\cdots,N\}$ and $k \in \{1,2,\cdots,d\}$, $M_t^{i,k}$ is a square integrable martingale with

$$E[(M_t^{i,k})^2] = \frac{4}{N^2}E\left(\int_0^t Y_{i,k}(s)^2 ds\right) \leq \frac{4}{N^2}\left\|\frac{\partial\varphi_\varepsilon}{\partial x_k}\right\|^2\int_0^t \|\rho(\cdot,s)\|^2 ds. \tag{29}$$

And for all $(i,k) \neq (j,h)$ we have that

$$\begin{aligned}\langle M_t^{i,k}, M_t^{j,h}\rangle &= \langle Y_{i,k}\cdot B_i^{(k)}(t), Y_{i,h}\cdot B_j^{(h)}(t)\rangle \\ &= \int_0^t Y_{i,k}(s)\cdot Y_{j,h}(s)\, d\langle B_i^{(k)}(t), B_j^{(h)}(t)\rangle \\ &= \int_0^t Y_{i,k}(s)\cdot Y_{j,h}(s)\, d0 = 0,\end{aligned} \tag{30}$$

since $\langle B_i^{(k)}(t), B_j^{(h)}(t)\rangle \equiv 0$ for two independent Brownian motions, where $\langle X_t, Y_t\rangle$ is the quadratic covariance between the two processes X_t and Y_t, defined by

$$\langle X_t, Y_t\rangle = \frac{1}{2}(\langle X_t + Y_t\rangle - \langle X_t\rangle - \langle Y_t\rangle).$$

Noting that $M_t^{i,k}$ and $M_t^{j,h}$ are both square integrable martingales, (37) implies that

$$E\left(M_t^{i,k} M_t^{j,h}\right) \equiv 0. \tag{31}$$

Combining (29) and (31) immediately gives us

$$E[(M_t^i)^2] = \sum_{k=1}^d E[(M_t^{i,k})^2] \le \frac{4}{N^2}\|\nabla\varphi_\varepsilon\|^2 \int_0^t \|\rho(\cdot,s)\|^2\, ds.$$

and

$$E(M_t^i M_t^j) = 0$$

which implies that

$$E\big((M_t)^2\big) = \sum_{i=1}^N E[(M_t^i)^2] \le \frac{4}{N}\|\nabla\varphi_\varepsilon\|^2 \int_0^t \|\rho(\cdot,s)\|^2\, ds = \frac{4}{N\varepsilon^{d+2}}\|\nabla\varphi\|^2 \int_0^t \|\rho(\cdot,s)\|^2\, ds. \tag{32}$$

Lemma 2. *For all $t \ge 0$, we have the second moment control*

$$E\big((\tilde{M}_t)^2\big) \le \frac{4}{N\varepsilon^{2d+2}}\|\varphi\|^2\|\nabla\varphi\|^2 t. \tag{33}$$

Proof. Again note that $\tilde{M}_t = \sum_{i=1}^N \tilde{M}_t^i$ with

$$\tilde{M}_t^i = \sum_{k=1}^d \tilde{M}_t^{i,k}$$

where

$$\tilde{M}_t^{i,k} = \int_0^t Z_{i,k}(s) dB_i^{(k)}(s)$$

and

$$\begin{aligned}Z_{i,k}(t) =& \frac{2}{N^2}\int_{\mathbb{R}^d} \varphi_\varepsilon(x) \sum_{j=1}^{i-1} \frac{\partial\varphi_\varepsilon\,(x + B_i(s) - B_j(s))}{\partial x_k} dx \\ & - \frac{2}{N^2}\int_{\mathbb{R}^d} \varphi_\varepsilon(x) \sum_{j=i+1}^{N} \frac{\partial\varphi_\varepsilon\,(x + B_j(s) - B_i(s))}{\partial x_k} dx.\end{aligned} \tag{34}$$

It is easy to see that the integrand $Z_{i,k}(t)$ is continuous and adapted to $\mathscr{F}_t^N$ and that

$$|Z_{i,k}(t)| \le \frac{2}{N}\|\varphi_\varepsilon\| \cdot \left\|\frac{\partial\varphi_\varepsilon}{\partial x_k}\right\| \tag{35}$$

Then again accordion to Theorem 5.2.3 in [5] we have for all $i \in \{1, 2, \cdots, N\}$ and $k \in \{1, 2, \cdots, d\}$, $M_t^{i,k}$ is a square integrable martingale with that

$$E[(\tilde{M}_t^{i,k})^2] = E\left(\int_0^t Z_{i,k}(s)^2 ds\right) \leq \frac{4t}{N^2}\|\varphi_\varepsilon\|^2 \cdot \left\|\frac{\partial \varphi_\varepsilon}{\partial x_k}\right\|^2 \tag{36}$$

and that for all $(i,k) \neq (j,h)$ we have that

$$\begin{aligned}\langle \tilde{M}_t^{i,k}, \tilde{M}_t^{j,h}\rangle &= \langle Z_{i,k} \cdot B_i^{(k)}(t), Z_{i,h} \cdot B_j^{(h)}(t)\rangle \\ &= \int_0^t Z_{i,k}(s) \cdot Z_{j,h}(s)\, d\langle B_i^{(k)}(t), B_j^{(h)}(t)\rangle \\ &= \int_0^t Z_{i,k}(s) \cdot Z_{j,h}(s)\, d0 = 0,\end{aligned} \tag{37}$$

which implies that

$$E\left(\tilde{M}_t^{i,k}\tilde{M}_t^{j,h}\right) \equiv 0. \tag{38}$$

Thus we immediately have

$$E[(\tilde{M}_t^i)^2] \leq \frac{4t}{N^2}\|\varphi_\varepsilon\|^2 \cdot \|\nabla\varphi_\varepsilon\|^2$$

and $E\left(\tilde{M}_t^i \tilde{M}_t^j\right) = 0$ for all $i \neq j$. Thus

$$E\left((\tilde{M}_t)^2\right) = \sum_{i=1}^N E\left((\tilde{M}_t^i)^2\right) \leq \frac{4t}{N}\|\varphi_\varepsilon\|^2\|\nabla\varphi_\varepsilon\|^2 = \frac{4t}{N\varepsilon^{2d+2}}\|\varphi\|^2\|\nabla\varphi\|^2. \tag{39}$$

With the martingale parts M_t and $\tilde{M}_t$ both controlled, our last step is to bound the “residue” part $\mathrm{Res}(t)$. First, we again apply Cauchy Schwarz inequality and have

$$\begin{aligned}|\mathrm{Res}(t)| \leq &\frac{1}{2}\int_0^t \|\nabla(\rho - \rho_{\varepsilon,N})(\cdot,s)\|^2 ds \\ &+ \frac{2}{N^2}\int_0^t \int_{\mathbb{R}^d}\left|\sum_{i=1}^N \varphi_\varepsilon(x - X_i(s))(\mathbf{F}(x,s) - \mathbf{F}(X_i(s),s))\right|^2 dxds.\end{aligned} \tag{40}$$

It is easy to see that we can rewrite the integrand in the second term as follows:

$$\left|\sum_{i=1}^N \varphi_\varepsilon(x - X_i(s))(\mathbf{F}(x,s) - \mathbf{F}(X_i(s),s))\right|^2 = \sum_{i,j\leq N} R_{i,j}(x,s)$$

where

$$R_{i,j}(x,s) = \varphi_\varepsilon(x - X_i(s))\,\varphi_\varepsilon(x - X_j(s))\,(\mathbf{F}(x,s) - \mathbf{F}(X_i(s),s)) \cdot (\mathbf{F}(x,s) - \mathbf{F}(X_j(s),s)).$$

Note that for any $i,j \leq N$, $R_{i,j}(x,s) = 0$ when $|X_j(s) - X_i(s)| > 2\varepsilon$. And when $|X_i(s) - X_j(s)| \leq 2\varepsilon$, noting that F is Lipschitz continuous with the Lipschitz constant less than or equal to L_F,

$$\left|R_{i,j}(x,s)\right| \leq L_F^2\varepsilon^2\left|\varphi_\varepsilon(x - X_i(s))\,\varphi_\varepsilon(x - X_j(s))\right|.$$

Thus for all $i,j \leq N$, we have the spatial integral

$$\int_{\mathbb{R}^d}|R_{i,j}(x,s)|\,dx \leq \varepsilon^{2-d}L_F^2\|\varphi\|^2\,\mathbb{1}_{|X_i(s)-X_j(s)|\leq 2\varepsilon}. \tag{41}$$

Combining (40) and (41) we have

$$|\mathrm{Res}(t)| \leq \frac{1}{2}\int_0^t \|\nabla(\rho - \rho_{\varepsilon,N})(\cdot,s)\|^2\,ds + R^*(t) \tag{42}$$

where

$$R^*(t) = \frac{2L_F^2\|\varphi\|^2}{N^2\varepsilon^{d-2}} \sum_{i,j\le N} \int_0^t \mathbb{1}_{|X_i(s)-X_j(s)|\le 2\varepsilon}\, ds. \tag{43}$$

By definition, $R^*(t)$ is monotonically increasing over t. So if we want to control $\sup_{s\le t}|\mathrm{Res}(s)|$, it is sufficient to control $R^*(t)$. Noting that $\mathbb{1}_{|X_i-X_i|\le 2\varepsilon}\equiv 1$, we can take expectation on (43) and have

$$E[R^*(t)] = \frac{2L_F^2\|\varphi\|^2 t}{N\varepsilon^{d-2}} + \frac{2L_F^2\|\varphi\|^2}{N^2\varepsilon^{d-2}} \sum_{i,j\le N: i\ne j} E\left(\int_0^t \mathbb{1}_{|X_i(s)-X_j(s)|\le 2\varepsilon}\, ds\right). \tag{44}$$

Noting that $X_i(s)-X_j(s)$ is continuous and adaptable to $\mathscr{F}_t^N$ (which implies progressive), $\mathbb{1}_{|X_i(s)-X_j(s)|\le 2\varepsilon}\times \mathbb{1}_{0\le s\le t}$ is measurable on $[0,t]\times\Omega$ and bounded and thus integrable. By Fubini's Theorem,

$$E[R^*(t)] = \frac{2L_F^2\|\varphi\|^2 t}{N\varepsilon^{d-2}} + \frac{2L_F^2\|\varphi\|^2}{N^2\varepsilon^{d-2}} \sum_{i,j\le N: i\ne j} \int_0^t P\big(|X_i(s)-X_j(s)|\le 2\varepsilon\big)ds. \tag{45}$$

At this point, we have reduced the problem of controlling $|\mathrm{Res}(s)|$ to controlling the upper bounds for the probabilities of $P(|X_i(s)-X_j(s)|\le 2\varepsilon)$, $i\ne j\le N$. We show this under Assumption 1 and 2 respectively. However, since the proofs are similar, we will only show the proof of the more complicated case under Assumption 1. One can proof the same result under Assumption 2 followings exact the same steps but the calculations are easier since the sums will be replaced by integrals in that case.

Under Assumption 1, for $j=1,\cdots,N$, let

$$E_j(t) = \frac{1}{N}\sum_{i\le N: i\ne j} \int_0^t P(|X_i(s)-X_j(s)|\le 2\varepsilon)\, ds \tag{46}$$

which implies that

$$E[R^*(t)] = \frac{2L_F^2\|\varphi\|^2 t}{N\varepsilon^{d-2}} + \frac{2L_F^2\|\varphi\|^2}{\varepsilon^{d-2}} \left(\frac{1}{N}\sum_{j=1}^N E_j(t)\right).$$

We have the following Lemma.

Lemma 3. *Under Assumption 1, for any $t\ge 0$, there exist some constant $C_1(t)$ and $C_2(t)$ depends only on t such that*

$$E_j(t) \le C_1(t)\varepsilon^{d-1} + C_2(t)\frac{1}{N\varepsilon} \tag{47}$$

for all $j=1,2,\cdots,N$, when ε is sufficiently small.

Proof. For any N and j. Fix $i\ne j$, $i\le N$, and let $\{\Omega, \mathscr{F}_t^{i,j}, P\}$ be our probability measure space where , $\mathscr{F}_t^{i,j}$ is the natural filtration generated by $B_{i,j}^*(t) = [B_i(t), B_j(t)]$, which is a $2d$-dimensional Brownian motion. Let $\theta_{i,j}(s) = -\big(F(X_i(s),s), F(X_j(s),s)\big)$ be the integrand and consider the adapted measurable process

$$\Gamma_s = \int_0^s \theta_{i,j}(h)\cdot dB_{i,j}^*(h). \tag{48}$$

Note that for any $s\ge 0$,

$$|\theta_{i,j}(s)|^2 \le 2d\,\|F\|_\infty^2. \tag{49}$$

Thus the Novikov condition (see page 198 of [14] for details) is satisfied, i.e.,

$$E\left[\exp\left(\frac{1}{2}\int_0^s |\theta_{i,j}(h)|^2\, dh\right)\right] \le \exp(sd\,\|F\|_\infty^2) < \infty,$$

by Girsanov Theorem (see Theorem 3.5.1 of [14]) we can define a probability measure Q in our probability space with Radon-Nikodym derivative

$$\frac{dQ_{i,j}}{dP}\bigg|\mathscr{F}_t = \mathscr{E}_t = \exp\left[\Gamma_t - \frac{1}{2}\int_0^s |\theta_{i,j}(h)|^2\, dh\right]. \tag{50}$$

Then we have

$$\begin{bmatrix} X_i(t)-X_i(0) \\ X_j(t)-X_j(0) \end{bmatrix} = \begin{bmatrix} B_i(t)+\int_0^t F(X_i(s),s)ds \\ B_j(t)+\int_0^t F(X_j(s),s)ds \end{bmatrix} = B^*_{i,j}(t) - \langle \Gamma, B^*_{i,j}\rangle_t$$

is a standard $2d$-dimensional Brownian motion under probability measure $Q_{i,j}$, where $\langle \Gamma, B^*_{i,j}\rangle_t$ is again the quadratic covariance between Γ_t and $B^*_{i,j}(t)$. Thus by Radon-Nikodym Theorem we have

$$\int_{|X_i(t)-X_j(t)|\leq 2\varepsilon} \mathscr{E}_t dP = P\left(|B_i(t)-B_j(t)+\delta_{i,j}| \leq 2\varepsilon\right). \tag{51}$$

Moreover,

$$P\left(|X_i(t)-X_j(t)|\leq 2\varepsilon\right) \leq P(\mathscr{E}_t < \varepsilon^{1/2}) + P\left(|X_i(t)-X_j(t)| \leq 2\varepsilon \cap \mathscr{E}_t \geq \varepsilon^{1/2}\right)$$

and for the first part we have,

$$P(\mathscr{E}_t < \varepsilon^{1/2}) \leq P\left(\exp\left[\int_0^t (-\theta_{i,j}(s))\cdot dB^*_{i,j}(s)\right] > \varepsilon^{-1/2}\exp(-td\,\|F\|_\infty^2)\right). \tag{52}$$

To control the right had side of the inequality above, we consider the L^{4d} norm:

$$E\left[\left(\exp\left[\int_0^t (-\theta_{i,j}(s))\cdot dB^*_{i,j}(s)\right]\right)^{4d}\right] = E\left(\exp\left[\int_0^t (-4d\,\theta_{i,j}(s))\cdot dB^*_{i,j}(s)\right]\right) \tag{53}$$

and note that again by Girsanov Theorem,

$$\mathscr{E}'_t = \exp\left[\int_0^t (-4d\theta_{i,j}(s))\cdot dB^*_{i,j}(s)\right]\exp\left(-8d^2\int_0^t |\theta_{i,j}(s)|^2 ds\right)$$

is again a Radon-Nikodym derivative. Thus we have

$$E(\mathscr{E}'_t) = 1$$

which combining with (49), implies

$$E\left(\exp\left[\int_0^t (-4d\theta_{i,j}(s))\cdot dB^*_{i,j}(s)\right]\right) \leq \exp(16d^3t\,\|F\|_\infty^2) < \infty. \tag{54}$$

Combining (52), (54) and Chebyshev's Inequality gives us

$$P(\mathscr{E}_t < \varepsilon^{1/2}) \leq \varepsilon^{2d}\exp\left((4d^2+16d^3)t\,\|F\|_\infty^2\right). \tag{55}$$

Then for the second part, according to (51) we have

$$\int_{|X_i(t)-X_j(t)|\leq 2\varepsilon \cap \mathscr{E}_t \geq \varepsilon^{1/2}} \mathscr{E}_t dP \leq P\left(|B_i(t)-B_j(t)+\delta_{i,j}| \leq 2\varepsilon\right).$$

and thus

$$\begin{aligned} &P\left(|X_i(t)-X_j(t)| \leq 2\varepsilon \cap \mathscr{E}_t \geq \varepsilon^{1/2}\right) \\ &\quad \leq \varepsilon^{-1/2} P\left(|B_i(t)-B_j(t)+\delta_{i,j}| \leq 2\varepsilon\right). \end{aligned} \tag{56}$$

Combining the two inequalities above, we have

$$\begin{aligned} P\left(|X_i(t)-X_j(t)|\leq 2\varepsilon\right) \leq\, &\varepsilon^{2d}\exp\left((4d^2+16d^3)t\,\|F\|_\infty^2\right) \\ &+ \varepsilon^{-1/2} P\left(|B_i(t)-B_j(t)+\delta_{i,j}| \leq 2\varepsilon\right) \end{aligned} \tag{57}$$

for any $t \geq 0$. Integrating (57) on $[0,t]$ and averaging over all $i \neq j$, $i \leq N$, we have

$$
\begin{aligned}
E_j(t) \leq & \frac{\varepsilon^{2d}}{(4d^2+16d^3)\,\|F\|_\infty^2} \exp\left((4d^2+16d^3)t\,\|F\|_\infty^2\right) \\
& + \frac{1}{N\varepsilon^{1/2}} \sum_{i:i\neq j, i\leq N} \int_0^t P\left(|B_i(s)-B_j(s)+\delta_{i,j}| \leq 2\varepsilon\right) ds.
\end{aligned} \tag{58}
$$

According to (58) to proof this lemma it is sufficient to have the following lemma for standard Brownian motions:

Lemma 4. *Under Assumption 1, for any $t \geq 0$, there is some constant $C_1^*(t)$ and $C_2^*(t)$ such that*

$$
\frac{1}{N} \sum_{i:i\neq j, i\leq N} \left[\int_0^t P\left(|B_i(s)-B_j(s)+\delta_{i,j}| \leq 2\varepsilon\right) ds \right] \leq C_1^*(t)\varepsilon^d + C_2^*(t)\frac{1}{N}. \tag{59}
$$

Proof. We first note that for any s and m,n, $B_i(s)-B_j(s)+\delta_{i,j}$ has a d-dimensional normal distribution with mean $\delta_{i,j}$ and variance $2s$. So we have

$$
\begin{aligned}
& \int_0^t P\left(|B_i(s)-B_j(s)+\delta_{i,j}| \leq 2\varepsilon\right) ds \\
& = \int_0^t \int_{|x|\leq 2\varepsilon} \frac{1}{(4\pi s)^{d/2}} \exp\left(-\frac{|\delta_{i,j}-x|^2}{4s}\right) dxds \\
& = \int_{|x|\leq 2\varepsilon} \int_0^t \frac{1}{(4\pi s)^{d/2}} \exp\left(-\frac{|\delta_{i,j}-x|^2}{4s}\right) dsdx.
\end{aligned} \tag{60}
$$

To deal with equation (60), we need to separate the case of $d=1$, $d=2$ and $d \geq 3$.

Case 1: $d=1$. In this case we simply use the bound

$$
\begin{aligned}
& \int_0^t P\left(|B_i(s)-B_j(s)+\delta_{i,j}| \leq 2\varepsilon\right) ds \\
& \leq \int_{-2\varepsilon}^{2\varepsilon} \int_0^t s^{-1/2}\, dsdx = 8\varepsilon\sqrt{t}.
\end{aligned}
$$

Averaging over m gives us the desired result.

Case 2: $d=2$. In this case we have

$$
\begin{aligned}
& \int_0^t P\left(|B_i(s)-B_j(s)+\delta_{i,j}| \leq 2\varepsilon\right) ds \\
& = \int_{|x|\leq 2\varepsilon} \int_0^t \frac{1}{4\pi s} \exp\left(-\frac{|\delta_{i,j}-x|^2}{4s}\right) dsdx.
\end{aligned}
$$

If $\delta_{i,j} \geq 1$, then for all $\varepsilon < 1/4$ and $x < 2\varepsilon$ we have

$$
\begin{aligned}
& \int_{|x|\leq 2\varepsilon} \int_0^t \frac{1}{4\pi s} \exp\left(-\frac{|\delta_{i,j}-x|^2}{4s}\right) dsdx \\
& \leq \int_{|x|\leq 2\varepsilon} \int_0^t \frac{1}{s} \exp\left(-\frac{1}{16s}\right) dsdx \\
& \leq 16\varepsilon^2 \int_0^t \frac{1}{s} \exp\left(-\frac{1}{16s}\right) ds
\end{aligned} \tag{61}
$$

When $\delta_{i,j} < 1$ taking $h = \frac{|\delta_{i,j}-x|^2}{4s}$, we have

$$\begin{aligned}
&\int_0^t P\left(|B_i(s)-B_j(s)+\delta_{i,j}|\le 2\varepsilon\right) ds\\
&=\int_{|x|\le 2\varepsilon}\int_0^t \frac{1}{4\pi s}\exp\left(-\frac{|\delta_{i,j}-x|^2}{4s}\right) dsdx\\
&=\int_{|x|\le 2\varepsilon}\int_{|\delta_{i,j}-x|^2/4t}^{\infty}\frac{1}{4\pi h}\exp(-h)\,dhdx.
\end{aligned}$$

Note that $h^{-1}\exp(-h)<h^{-1}$ and $h^{-1}\exp(-h)\le\exp(-h)$ when $h\ge 1$. We have

$$\begin{aligned}
\int_{|\delta_{i,j}-x|^2/4t}^{\infty} h^{-1}\exp(-h)\,dh &\le \int_{|\delta_{i,j}-x|^2/4t}^{1} h^{-1}dh+\int_1^\infty e^{-h}dh\\
&\le 2|\log(|\delta_{i,j}-x|)|+|\log t|+1+\log 4.
\end{aligned}\tag{62}$$

Moreover, let $\delta_1=CN^{-1/2}$, where C is the constant in Assumption 1, $\delta_2=\delta_1+4\varepsilon$ and $M=[\delta_2^{-1}]+1$. For all $k=0,1,\cdots,M$ consider the following sets

$$A_k:=\left\{i: k\delta_2\le|\delta_{i,j}|<(k+1)\delta_2\right\}.\tag{63}$$

By definition, it is easy to see that when N is large and ε is small

$$\bigcup_{k=0}^{M} A_k\supset\left\{i:|\delta_{i,j}|<1\right\}.\tag{64}$$

If we first look at A_0, according to Assumption 1, the little balls $\{N(\delta_{i,j},\delta)\}_{i\le N,i\ne j}$ (where $N(x,y)$ is the neighborhood of x with radius y) have no intersections with each other. And for all $i\in A_0$,

$$N(\delta_{i,j},\delta_1)\subset N(0,\delta_1+\delta_2)$$

This immediately implies that

$$\mathrm{card}(A_0)\le\left(\frac{\delta_1+\delta_2}{\delta_1}\right)^2=\left(2+\frac{4\varepsilon}{C}N^{1/2}\right)^2\le 8+\frac{32\varepsilon^2}{C^2}N,$$

since the sum of areas of disjoint disks with radius δ_1 in A_0 cannot be larger than the area of A_0 itself. Thus we have

$$\frac{1}{N}\sum_{i\in A_0}\int_0^t P\left(|B_i(s)-B_j(s)+\delta_{i,j}|\le 2\varepsilon\right)ds\le\frac{t}{N}\mathrm{card}(A_0)\le\frac{8t}{N}+\frac{32\varepsilon^2 t}{C^2}\tag{65}$$

Similarly, for each $k\ge 1$ and $i\in A_k$,

$$N(\delta_{i,j},\delta_1)\subset\left\{y:(k-1)\delta_2\le|x|<(k+2)\delta_2\right\}$$

which implies that

$$\mathrm{card}(A_k)\le\frac{[(k+2)^2-(k-1)^2]\delta_2^2}{\delta_1^2}\le 9k\left(1+\frac{4\varepsilon N^{1/2}}{C}\right)^2.$$

Noting that for all $i\in A_k$ and $|x|\le 2\varepsilon$

$$|\log(|\delta_{i,j}-x|)|\le\max\{\log 2,|\log(|k\delta_2-2\varepsilon|)|\}\le\log 2+|\log(k\delta_2)|.$$

Thus according to (62) and the inequality above

$$
\begin{aligned}
&\frac{1}{N}\sum_{i\in A_k}\int_0^t P\left(|B_i(s)-B_j(s)+\delta_{i,j}|\le 2\varepsilon\right)ds\\
&\le \frac{1}{N}\sum_{i\in A_k}\left[\int_{|x|\le 2\varepsilon}|\log t|+\log 4+1+2|\log(|\delta_{i,j}-x|)|\,dx\right]\\
&\le \frac{1}{N}\sum_{i\in A_k}\left[\int_{|x|\le 2\varepsilon}|\log t|+\log 16+1+2|\log(k\delta_2)|\,dx\right]\\
&\le \left[\frac{1}{N}\sum_{i\in A_k}16(|\log t|+\log 16+1)\varepsilon^2\right]+\frac{288\varepsilon^2 k\left(1+\frac{4\varepsilon N^{1/2}}{C}\right)^2}{N}|\log(k\delta_2)|\\
&=\left[\frac{1}{N}\sum_{i\in A_k}16(|\log t|+\log 16+1)\varepsilon^2\right]+\frac{288\varepsilon^2}{C^2}\delta_2\left[k\delta_2|\log(k\delta_2)|\right]
\end{aligned}
\tag{66}
$$

Summing over $k=0,1,\cdots,M$ we have

$$
\begin{aligned}
&\frac{1}{N}\sum_{i:\delta_{i,j}<1}\int_0^t P\left(|B_i(s)-B_j(s)+\delta_{i,j}|\le 2\varepsilon\right)ds\\
&\le \frac{8t}{N}+\frac{32\varepsilon^2 t}{C^2}+16(|\log t|+\log 16+1)\varepsilon^2+\frac{288\varepsilon^2}{C^2}\sum_{k=1}^{[\delta_2^{-1}]+1}\delta_2\left[k\delta_2|\log(k\delta_2)|\right]
\end{aligned}
\tag{67}
$$

Note that the last term in the inequality above is a Riemann sum of function $x|\log x|$ and the fact that $x|\log x|\le\max\{\log 2, e^{-1}\}<1$ on $[0,2]$.

$$
\sum_{k=1}^{[\delta_2^{-1}]+1}\delta_2\left[k\delta_2|\log(k\delta_2)|\right]\le\int_0^2 dt=2.
$$

So we have

$$
\begin{aligned}
&\frac{1}{N}\sum_{i:\delta_{i,j}<1}\int_0^t P\left(|B_i(s)-B_j(s)+\delta_{i,j}|\le 2\varepsilon\right)ds\\
&\le \frac{8t}{N}+\frac{32\varepsilon^2 t}{C^2}+16(|\log t|+\log 16+1)\varepsilon^2+\frac{576\varepsilon^2}{C^2}
\end{aligned}
\tag{68}
$$

Combining (68) and (61), and letting

$$
\begin{aligned}
&C_1^*(t):=16\int_0^t\frac{1}{s}\exp\left(-1/16s\right)ds+\frac{32t}{C^2}+16(|\log t|+\log 16+1)+\frac{576}{C^2}\\
&C_2^*(t):=8t
\end{aligned}
\tag{69}
$$

we finally get

$$
\frac{1}{N}\sum_{i:i\neq j,i\le N}\int_0^t P\left(|B_i(s)-B_j(s)+\delta_{i,j}|\le 2\varepsilon\right)ds\le C_1^*(t)\varepsilon^d+C_2^*(t)\frac{1}{N}
$$

when $d=2$, and the proof for case 2 is complete.

Case 3: $d\ge 3$**.** The proof in this case is similar but simpler than the case of $d=2$. Again we have

$$
\begin{aligned}
&\int_0^t P\left(|B_i(s)-B_j(s)+\delta_{i,j}|\le 2\varepsilon\right)ds\\
&=\int_{|x|\le 2\varepsilon}\int_0^t\frac{1}{(4\pi s)^{d/2}}\exp\left(-\frac{|\delta_{i,j}-x|^2}{4s}\right)dsdx.
\end{aligned}
$$

If $\delta_{i,j}\ge 1$, then for all $\varepsilon<1/4$ and $|x|<2\varepsilon$ we have

$$\begin{aligned}
&\int_{|x|\le 2\varepsilon}\int_0^t \frac{1}{(4\pi s)^{d/2}} \exp\left(-\frac{|\delta_{i,j}-x|^2}{4s}\right) dsdx \\
&\le \int_{|x|\le 2\varepsilon}\int_0^t \frac{1}{s^{d/2}} \exp\left(-\frac{1}{16s}\right) dsdx \\
&\le 2^{2d}\varepsilon^d \int_0^t \frac{1}{s^{d/2}} \exp\left(-\frac{1}{16s}\right) ds
\end{aligned} \tag{70}$$

When $\delta_{i,j} < 1$ taking $h = \frac{|\delta_{i,j}-x|^2}{4s}$, we have

$$\begin{aligned}
&\int_0^t P\left(|B_i(s) - B_j(s) + \delta_{i,j}| \le 2\varepsilon\right) ds \\
&= \int_{|x|\le 2\varepsilon}\int_0^t \frac{1}{(4\pi s)^{d/2}} \exp\left(-\frac{|\delta_{i,j}-x|^2}{4s}\right) dsdx \\
&< C_d \int_{|x|\le 2\varepsilon} |\delta_{i,j}-x|^{-d+2} dx.
\end{aligned} \tag{71}$$

where constant

$$C_d := 2^{4d}\int_0^\infty h^{-2+d/2}\exp(-h)dh.$$

Then again we can define $\delta_1 = CN^{-1/d}$, where C is the constant in Assumption 1, $\delta_2 = \delta_1 + 4\varepsilon$ and $M = [\delta_2^{-1}] + 1$. For all $k = 0, 1, \cdots, M$ consider the following sets

$$A_k := \left\{i : k\delta_2 \le |\delta_{i,j}| < (k+1)\delta_2\right\}. \tag{72}$$

such that

$$\bigcup_{k=0}^{M} A_k \supset \left\{i : |\delta_{i,j}| < 1\right\}.$$

Then similarly, we have

$$\text{card}(A_0) \le \left(\frac{\delta_1+\delta_2}{\delta_1}\right)^d = \left(2 + \frac{4\varepsilon}{C}N^{1/d}\right)^d \le 2^{2d-1} + \frac{2^{3d-1}\varepsilon^d}{C^d}N$$

and

$$\text{card}(A_k) \le \frac{[(k+2)^d - (k-1)^d]\delta_2^d}{\delta_1^d} \le 3^d k^{d-1}\left(1 + \frac{4\varepsilon N^{1/d}}{C}\right)^d.$$

Thus

$$\frac{1}{N}\sum_{i\in A_0}\int_0^t P\left(|B_i(s) - B_j(s) + \delta_{i,j}| \le 2\varepsilon\right) ds \le \frac{t}{N}\text{card}(A_0) \le \frac{2^{2d-1}t}{N} + \frac{2^{3d-1}t\varepsilon^d}{C^d} \tag{73}$$

and

$$\begin{aligned}
&\frac{1}{N}\sum_{i\in A_k}\int_0^t P\left(|B_i(s) - B_j(s) + \delta_{i,j}| \le 2\varepsilon\right) ds \\
&\le \frac{C_d}{N}\sum_{m\in A_k}\int_{|x|\le 2\varepsilon} |\delta_{i,j}-x|^{-d+2} dx \\
&\le \frac{C_d}{N} 3^d k^{d-1}\left(1 + \frac{4\varepsilon N^{1/d}}{C}\right)^d (2^{2d}\varepsilon^d) \times \left(2^d[k(CN^{-1/d}+4\varepsilon)]^{-d+2}\right)
\end{aligned} \tag{74}$$

Summing over $k = 0, 1, \cdots, M$,

$$
\begin{aligned}
&\frac{1}{N}\sum_{i:\delta_{i,j}<1}\int_0^t P\left(|B_i(s)-B_j(s)+\delta_{i,j}|\le 2\varepsilon\right)ds\\
&\le \frac{2^{2d-1}t}{N}+\frac{2^{3d-1}t\varepsilon^d}{C^d}+C_d\left(\frac{24}{C}\right)^d\varepsilon^d\sum_{k=1}^{[\delta_2^{-1}]+1}[(k\delta_2)\delta_2].
\end{aligned}
\tag{75}
$$

Again for the last term we have

$$
\sum_{k=1}^{[\delta_2^{-1}]+1}[(k\delta_2)\delta_2]\le\int_0^2 t\,dt=2.
$$

Thus

$$
\begin{aligned}
&\frac{1}{N}\sum_{i:\delta_{i,j}<1}\int_0^t P\left(|B_i(s)-B_j(s)+\delta_{i,j}|\le 2\varepsilon\right)ds\\
&\le \frac{2^{2d-1}t}{N}+\frac{2^{3d-1}t\varepsilon^d}{C^d}+2C_d\left(\frac{24}{C}\right)^d\varepsilon^d
\end{aligned}
\tag{76}
$$

Then combining (70) and (76), and letting

$$
\begin{aligned}
&C_1^*(t):=2^{2d}\int_0^t\frac{1}{s^{d/2}}\exp\left(-\frac{1}{16s}\right)ds+\frac{2^{3d-1}t}{C^d}+2C_d\left(\frac{24}{C}\right)^d\\
&C_2^*(t):=2^{2d-1}t
\end{aligned}
\tag{77}
$$

We complete the proof of case 3.

With the Lemma 2.4 proved, the proof of Lemma 2.3 is complete.

Plugging the result of this lemma into (45),

$$
E[R^*(t)]\le\frac{2\|\varphi\|^2t}{N\varepsilon^{d-2}}+2\|\varphi\|^2C_1(t)\varepsilon+\frac{2\|\varphi\|^2}{N\varepsilon^{d-1}}C_2(t).
\tag{78}
$$

Proof of Theorem 1: Combining (28), (33) and (78), noting that $\varepsilon_N=N^{-1/3d}\to 0$, then for any $t\ge 0$, letting

$$
\begin{aligned}
&\delta_1(t):=\left(4N^{-1/2d}\|\nabla\varphi\|^2\int_0^t\|\rho(x,s)\|^2ds\right)^{1/3}\\
&\delta_2(t):=\left(4N^{-1/2d}\|\nabla\varphi\|^2\|\varphi\|^2\right)^{1/3}\\
&\delta_3(t):=\left(2N^{-1/3d}\|\varphi\|^2(t+C_1(t)+C_2(t))\right)^{1/2}
\end{aligned}
$$

by Doob and Chebyshev inequality, inequality (78) and the fact that R^* is nonnegative, it is easy to see that as $N\to\infty$ the probability of the following events:

$$
\begin{aligned}
&P(A_N):=P\big(\sup_{s\le t}|M_t|<\delta_1(t)\big)\ge 1-\delta_1(t)\\
&P(B_N):=P\big(\sup_{s\le t}|\tilde{M}_t|<\delta_2(t)\big)\ge 1-\delta_2(t)\\
&P(C_N):=P\big(R^*(t)<\delta_3(t)\big)\ge 1-\delta_3(t).
\end{aligned}
\tag{79}
$$

Note that for the constant terms in Proposition 1, we have

$$
\frac{t}{N}\|\nabla\varphi_{\varepsilon_N}\|_2^2=\frac{t\varepsilon_N^{d+2}}{N}\|\nabla\varphi\|^2<tN^{-1/3d}
$$

when N is sufficiently large. And under event C_N, for any $s\le t$, by (42)

$$
\begin{aligned}
\mathrm{Res}(s) &\le \frac{1}{2}\int_0^s \|\nabla(\rho-\rho_{\varepsilon_N,N})(\cdot,h)\|^2\,dh + R^*(s)\\
&\le \frac{1}{2}\int_0^s \|\nabla(\rho-\rho_{\varepsilon_N,N})(\cdot,h))\|^2\,dh + R^*(t)\\
&\le \frac{1}{2}\int_0^s \|\nabla(\rho-\rho_{\varepsilon_N,N})(\cdot,h)\|^2\,dh + \delta_3(t),
\end{aligned}
$$

which implies that

$$
\sup_{s\le t}\left(\mathrm{Res}(s) - \frac{1}{2}\int_0^s \|\nabla(\rho-\rho_{\varepsilon_N,N})(\cdot,h)\|^2\,dh\right) \le \delta_3(t). \tag{80}
$$

So under the event

$$
P(A_N\cap B_N\cap C_N) \ge 1-\delta_1(t)-\delta_2(t)-\delta_3(t)
$$

we have that for all $s\le t$

$$
\begin{aligned}
\|(\rho-\rho_{\varepsilon,N})(\cdot,s)\|^2 \le& \|(\rho-\rho_{\varepsilon,N})(\cdot,0)\|^2 - \frac{1}{2}\int_0^s \|\nabla(\rho-\rho_{\varepsilon_N,N})(\cdot,h)\|^2\,dh\\
&-\int_0^s\int_{\mathbb{R}^d} \nabla\cdot\mathbf{F}(x,h)\left((\rho-\rho_{\varepsilon,N})(x,h)\right)^2\,dxdh\\
&+\delta_1(t)+\delta_2(t)+\delta_3(t)+N^{-1/3d}\,t
\end{aligned} \tag{81}
$$

which implies

$$
\begin{aligned}
&\sup_{s\in[0,t]}\Big(2\|(\rho-\rho_{\varepsilon,N})(\cdot,s)\|^2 + \int_0^s \|\nabla(\rho-\rho_{\varepsilon_N,N})(\cdot,h)\|^2\,dh\Big)\\
&\qquad\le 2\|(\rho-\rho_{\varepsilon,N})(\cdot,0)\|^2 - 2\int_0^s\int_{\mathbb{R}^d} \nabla\cdot\mathbf{F}(x,h)\left(\rho(x,h)-\rho_{\varepsilon,N}(x,h)\right)^2\,dxdh\\
&\qquad+2\big(\delta_1(t)+\delta_2(t)+\delta_3(t)+N^{-1/3d}t\big)\,.
\end{aligned} \tag{82}
$$

Since F is Lipschitz continuous from $R^d\to R^d$, it is also differentiable almost everywhere by Rademacher's theorem, see Theorem 3.1.6 of [8]. Thus for any x such that F is differentiable, we have $||\nabla\cdot\mathbf{F}||_{L^\infty}\le dL_F$, which implies that $||\rho(\cdot,s)||^2\le e^{C_0 s}||\rho_0||^2$ where C_0 is the constant defend in the statement of Theorem 1. I.e., $C_0=2dL_F\ge 2||\nabla\cdot\mathbf{F}||_{L^\infty}$. Let

$$
\begin{aligned}
c(t) = \Big(\frac{4e^{C_0 t}}{C_0}\|\nabla\varphi\|^2\|\rho_0\|^2\Big)^{1/3} + \big(4\|\nabla\varphi\|^2\|\varphi\|^2\big)^{1/3}\\
+\big(2\|\varphi\|^2[t+C_1(t)+C_2(t)]\big)^{1/2}+t\,.
\end{aligned} \tag{83}
$$

Gronwall's inequality finishes the proof of Theorem 1.

3 Discussion about Higher Order Sobolev Norm

In the previous discussions, we proved the convergence under H^1 norm. We hope what this method can be generalized to prove the convergence in the higher order Sobolev norms. However, currently we are only able to prove the simple case when $\mathbf{F}\equiv 0$. I.e., the limit PDE is now the heat equation

$$
\frac{\partial\rho}{\partial t}(x,t) = \frac{1}{2}\Delta\rho(x,t)
$$

and the paths of the particles are i.i.d. standard Brownian motions $B_i(t)$, $i=1,2,\cdots,N$ and

$$
\rho_{\varepsilon,N}(x,t) = \frac{1}{N}\sum_{i=1}^N \varphi_\varepsilon\left(x-B_i(t)\right). \tag{84}
$$

Again for any t we consider

$$\|D^\alpha(\rho-\rho_{\varepsilon,N})(\cdot,t)\|^2=\|D^\alpha\rho(\cdot,t)\|^2-2\int_{\mathbb{R}^d}D^\alpha\rho(x,t)\,D^\alpha\rho_{\varepsilon,N}(x,t)dx+\|D^\alpha\rho_{\varepsilon,N}(\cdot,t)\|^2. \tag{85}$$

For the first term we have

$$\begin{aligned}\|D^\alpha\rho(\cdot,t)\|^2&=\|D^\alpha\rho(\cdot,0)\|^2+\int_{\mathbb{R}^d}\int_0^t 2D^\alpha\rho(x,s)\,D^\alpha\frac{\partial\rho}{\partial t}(x,s)dsdx\\&=\|D^\alpha\rho(\cdot,0)\|^2-\int_0^t\|\nabla D^\alpha\rho(\cdot,s)\|^2ds.\end{aligned} \tag{86}$$

Then for the second term we have

$$\int_{\mathbb{R}^d}D^\alpha\rho(x,t)\,D^\alpha\rho_{\varepsilon,N}(x,t)dx=\frac{1}{N}\sum_{i=1}^N\int_{\mathbb{R}^d}D^\alpha\rho(x,t)\,D^\alpha\varphi_\varepsilon\left(x-B_i(t)\right)dx$$

and for each $i\le N$, and for all $x\in R^d$, we have by Ito's formula,

$$\begin{aligned}D^\alpha\rho(x,t)\,D^\alpha\varphi_\varepsilon\left(x-B_i(t)\right)=&\,D^\alpha\rho(x,0)\,D^\alpha\varphi_\varepsilon\left(x-B_i(0)\right)\\&+\int_0^t D^\alpha\frac{\partial\rho}{\partial t}(x,s)\,D^\alpha\varphi_\varepsilon\left(x-B_i(s)\right)ds\\&-\int_0^t D^\alpha\rho(x,s)\nabla D^\alpha\varphi_\varepsilon\left(x-B_i(s)\right)\cdot dB_i(s)\\&+\frac{1}{2}\int_0^t D^\alpha\rho(x,s)\Delta D^\alpha\varphi_\varepsilon\left(x-B_i(s)\right)ds.\end{aligned} \tag{87}$$

integrating over R^d and use integration by parts, we have

$$\begin{aligned}\int_{\mathbb{R}^d}D^\alpha\rho(x,t)\,D^\alpha\varphi_\varepsilon\left(x-B_i(t)\right)dx=&\int_{\mathbb{R}^d}D^\alpha\rho(x,0)\,D^\alpha\varphi_\varepsilon\left(x-B_i(0)\right)dx\\&+\int_0^t\int_{\mathbb{R}^d}D^\alpha\frac{\partial\rho}{\partial t}(x,s)\,D^\alpha\varphi_\varepsilon\left(x-B_i(t)\right)dxds\\&-\frac{1}{2}\int_0^t\int_{\mathbb{R}^d}\nabla D^\alpha\rho(x,s)\cdot\nabla D^\alpha\varphi_\varepsilon\left(x-B_i(t)\right)dxds\\&-NM_t^i\end{aligned}$$

where

$$M_t^i=\frac{1}{N}\int_0^t\int_{\mathbb{R}^d}\left[D^\alpha\rho(x,s)\,\nabla D^\alpha\varphi_\varepsilon\left(x-B_i(s)\right)\right]dx\cdot dB_i(s)$$

Averaging over $i=1,2,\cdots,N$ we have

$$\begin{aligned}\int_{\mathbb{R}^d}D^\alpha\rho(x,t)\,D^\alpha\rho_{\varepsilon,N}\left(x-B_n(t)\right)dx=&\int_{\mathbb{R}^d}D^\alpha\rho(x,0)\,D^\alpha\rho_{\varepsilon,N}\left(x-B_n(0)\right)dx\\&-\int_0^t\int_{\mathbb{R}^d}\nabla D^\alpha\rho(x,s)\cdot\nabla D^\alpha\rho_{\varepsilon,N}\left(x-B_n(t)\right)\,dxds\\&-M_t\end{aligned} \tag{88}$$

where $M_t=\sum_{i=1}^N M_t^i$. And by the same argument in Lemma 2.1, we have

$$E\left((M_t)^2\right)\le\frac{1}{N\varepsilon^{d+2|\alpha|+2}}\|\nabla D^\alpha\varphi\|^2\int_0^t\|D^\alpha\rho(x,s)\|^2ds \tag{89}$$

Finally for the last term $\|D^\alpha\rho_{\varepsilon,N}(x,t)\|^2$, we have

$$\|D^\alpha \rho_{\varepsilon,N}(x,t)\|^2 = \frac{1}{N}\|D^\alpha \varphi_\varepsilon\| + \frac{2}{N^2}\sum_{j>i}\int_{\mathbb{R}^d} D^\alpha \varphi_\varepsilon\left(x - B_j(t)\right) D^\alpha \varphi_\varepsilon\left(x - B_i(t)\right) dx. \tag{90}$$

Then for each $j > i$, by change of variables equals to

$$\int_{\mathbb{R}^d} D^\alpha \varphi_\varepsilon\left(x - B_j(t)\right) D^\alpha \varphi_\varepsilon\left(x - B_i(t)\right) dx = \int_{\mathbb{R}^d} D^\alpha \varphi_\varepsilon\left(x\right) D^\alpha \varphi_\varepsilon\left(x + B_j(t) - B_i(t)\right) dx.$$

Again, we use Ito's formula, for any $j > i$ and $x \in R^d$,

$$\begin{aligned} D^\alpha \varphi_\varepsilon\left(x + B_j(t) - B_i(t)\right) =& D^\alpha \varphi_\varepsilon\left(x + B_j(0) - B_i(0)\right) \\ &+ \int_0^t \nabla D^\alpha \varphi_\varepsilon\left(x + B_j(s) - B_i(s)\right)\cdot dB_j(s) \\ &- \int_0^t \nabla D^\alpha \varphi_\varepsilon\left(x + B_j(s) - B_i(s)\right)\cdot dB_i(s) \\ &+ \int_0^t \Delta D^\alpha \varphi_\varepsilon\left(x + B_j(s) - B_i(s)\right) ds. \end{aligned} \tag{91}$$

Integrating over R^d we have $\int_{\mathbb{R}^d} D^\alpha \varphi_\varepsilon\big(x - B_j(t)\big) D^\alpha \varphi_\varepsilon\left(x - B_i(t)\right) dx$ equals to

$$\begin{aligned} &\int_{\mathbb{R}^d} D^\alpha \varphi_\varepsilon\big(x - B_j(0)\big) D^\alpha \varphi_\varepsilon\left(x - B_i(0)\right) dx \\ &\quad - \int_0^t \int_{\mathbb{R}^d} \nabla D^\alpha \varphi_\varepsilon\left(x - B_j(s)\right) \nabla D^\alpha \varphi_\varepsilon\left(x - B_i(s)\right) dxds \\ &\quad + \int_0^t \int_{\mathbb{R}^d} D^\alpha \varphi_\varepsilon\left(x\right) \nabla D^\alpha \varphi_\varepsilon\left(x + B_j(s) - B_i(s)\right) dx \cdot dB_j(s) \\ &\quad - \int_0^t \int_{\mathbb{R}^d} D^\alpha \varphi_\varepsilon\left(x\right) \nabla D^\alpha \varphi_\varepsilon\left(x + B_j(s) - B_i(s)\right) dx \cdot dB_i(s). \end{aligned}$$

Then summing over all $j > i$, we have

$$\|D^\alpha \rho_{\varepsilon,N}(\cdot,t)\|^2 = \|D^\alpha \rho_{\varepsilon,N}(\cdot,0)\|^2 - \int_0^t \|\nabla D^\alpha \rho_{\varepsilon,N}(\cdot,s)\|^2 ds + \tilde{M}_t + \frac{t}{N}\|\nabla D^\alpha \varphi_\varepsilon\|^2 \tag{92}$$

where

$$\tilde{M}_t = \sum_{i=1}^N \tilde{M}_t^i,$$

with

$$\begin{aligned} \tilde{M}_t^i =& \frac{2}{N^2}\int_0^t \int_{\mathbb{R}^d}\left[\sum_{j=1}^{i-1} D^\alpha \varphi_\varepsilon\left(x\right) \nabla D^\alpha \varphi_\varepsilon\left(x + B_i(s) - B_j(s)\right)\right] dx \cdot dB_i(s) \\ &- \frac{2}{N^2}\int_0^t \int_{\mathbb{R}^d}\left[\sum_{j=i+1}^{N} D^\alpha \varphi_\varepsilon\left(x\right) \nabla D^\alpha \varphi_\varepsilon\left(x + B_j(s) - B_i(s)\right)\right] dx \cdot dB_i(s). \end{aligned}$$

Then according to the same argument of Lemma 2.2, we have

$$E\big((\tilde{M}_t)^2\big) \le \frac{4t}{N\varepsilon^{2d+4|\alpha|+2}}\|D^\alpha \varphi\|^2 \|\nabla D^\alpha \varphi\|^2. \tag{93}$$

Combining (86), (88) and (92) we have

$$\begin{aligned} \|D^\alpha(\rho - \rho_{\varepsilon,N})(\cdot,t)\|^2 =& \|D^\alpha(\rho - \rho_{\varepsilon,N})(x,0)\|^2 - \int_0^t \|\nabla D^\alpha(\rho - \rho_{\varepsilon,N})(\cdot,s)\|^2 ds \\ &+ 2M_t + \tilde{M}_t + \frac{t}{N}\|\nabla D^\alpha \varphi_\varepsilon\|^2. \end{aligned} \tag{94}$$

With inequality (89) and (93), again we let $\varepsilon_N = N^{-1/3(d+2|\alpha|)} \to 0$ as $N \to \infty$. then for any $t \geq 0$, letting

$$\delta_1^*(t) := \Big(4N^{-1/2(d+2|\alpha|)}\|\nabla D^\alpha \varphi\|^2 \int_0^t \|D^\alpha \rho(x,s)\|^2 ds\Big)^{1/3}$$
$$\delta_2^*(t) := \Big(4N^{-1/2(d+2|\alpha|)}\|\nabla D^\alpha \varphi\|^2 \|D^\alpha \varphi\|^2\Big)^{1/3}$$

by Doob and Chebyshev inequality, it is easy to see that as $N \to \infty$ the probability of the following events:

$$P(A_N^*) := P(\sup_{s\leq t}|M_t| < \delta_1(t)) \geq 1 - \delta_1(t)$$
$$P(B_N^*) := P(\sup_{s\leq t}|\tilde{M}_t| < \delta_2(t)) \geq 1 - \delta_2(t)$$

Noting that for the the constant term, we have

$$\frac{t}{N}\|\nabla D^\alpha \varphi_{\varepsilon_N}\|_2^2 < N^{-1/3(d+2|\alpha|)}t$$

when N is sufficiently large, so under the event

$$P(A_N^* \cap B_N^*) \geq 1 - \delta_1^*(t) - \delta_2^*(t)$$

we have that for all $s \leq t$

$$\begin{aligned}\|D^\alpha(\rho - \rho_{\varepsilon,N})(\cdot,s)\|^2 \leq & \|D^\alpha(\rho - \rho_{\varepsilon,N})(\cdot,0)\|^2 - \int_0^s \|\nabla D^\alpha(\rho - \rho_{\varepsilon,N})(\cdot,h)\|^2 dh \\ & + \delta_1^*(t) + \delta_2^*(t) + N^{-1/3(d+2|\alpha|)}t.\end{aligned} \tag{95}$$

Let

$$c_\alpha(t) = \big(4t\|\nabla D^\alpha \varphi\|^2 \|D^\alpha \rho_0\|^2\big)^{1/3} + \big(4\|\nabla D^\alpha \varphi\|^2\|D^\alpha \varphi\|^2\big)^{1/3} + t \tag{96}$$

In above we have used the fact that $\|D^\alpha \rho(\cdot,s)\| \leq \|D^\alpha \rho_0\|$. Then for $\varepsilon_N = N^{-1/3(d+2|\alpha|)}$ that for and any $t \geq 0$, and the $c_\alpha(t)$ defined above that depends only on t and $\rho(x,s)$, we have

$$\begin{aligned}P\Big(&\sup_{s\leq t}\big(\|D^\alpha(\rho - \rho_{\varepsilon_N,N})(\cdot,s)\|^2 + \int_0^s \|\nabla D^\alpha(\rho - \rho_{\varepsilon_N,N})\|^2 dh\big) \\ &< \|D^\alpha(\rho - \rho_{\varepsilon_N,N})(\cdot,0)\|^2 + c_\alpha(t)N^{-1/6(d+2|\alpha|)}\Big) \geq 1 - c_\alpha(t)N^{-1/6(d+2|\alpha|)}.\end{aligned} \tag{97}$$

when N is sufficiently large. So we have

Theorem 2. *Let ρ be the solution of the heat equation with initial density $\rho_0 \in H^k(\mathbb{R}^d)$, $B_i(t)$ be independent copies of Brownian Motions with initial data $B_i(0)$, and $\rho_{\varepsilon_N,N}$ be the constructed regularized empirical measure defined in (84), with regularized parameter $\varepsilon_N = N^{-1/3(d+2k)}$. For any $t \geq 0$, $|\alpha| \leq k$, we have*

$$\begin{aligned}P\Big(&\sup_{s\leq t}\big(\|D^\alpha(\rho - \rho_{\varepsilon_N,N})(\cdot,s)\|^2 + \int_0^t \|\nabla D^\alpha(\rho - \rho_{\varepsilon_N,N})\|^2 dh\big) \\ &< \|D^\alpha(\rho - \rho_{\varepsilon_N,N})(\cdot,0)\|^2 + c_\alpha(t)N^{-1/6(d+2|\alpha|)}\Big) \geq 1 - c_\alpha(t)N^{-1/6(d+2|\alpha|)}.\end{aligned} \tag{98}$$

where $c_\alpha(t)$, $t > 0$ is positive function dependent only on t, φ and $\|\rho_0\|_{H^k}$ and is defined in (96)

References

1. F. Bolley, A. Guillin and C. Villani, Quantitative concentration inequalities for empirical measures on non-compact spaces, Prob. Theory Rel. Fields, **137** (2007), 541-593.
2. A. J. Chorin, Numerical study of slightly viscous flow, J. Fluid Mech., **57** (1973), 785-796.
3. G. H. Cottet, P. Koumoutsakos, Vortex methods: theory and practice, Cambridge University Press, 2000.
4. R. T. Durrett, Stochastic Calculus: A Practical Introduction, CRCPress, 1996
5. S. N. Ethier, R. G. Kurtz, Markov Processes: Characterization and Convergence John Wiley & Sons, Inc. 2005
6. N. Fournier, M. Hauray and S. Mischler, Propagation of chaos for the 2D viscous vortex model, J. Eur. Math. Soc., **16** (2014), 1425–1466.
7. J. Goodman, Convergence of the random vortex method, Comm. Pure Appl. Math., **40** (1987), 189-220.
8. H. Federer, Geometric measure theory, Springer-Verlag, New York, 1969.
9. P. Koumoutsakos and A. Leonard, High-resolution simulations of the flow around an impulsively started cylinder using vortex methods, J. Fluid Mech. **96**, (1995) 1-38.
10. J.-G. Liu and R. Yang, A random particle blob method for the Keller-Segel equation and convergence analysis, Math. Comp., to appear.
11. D.G. Long, Convergence of the random vortex method in two dimensions, J. Amer. Math. Soc, **1** (1988), 779-804.
12. C. Marchioro and M. Pulvirenti, Hydrodynamics in two dimensions and vortex theory, Commun. Math. Phys., **84** (1982), 483-504.
13. H.P. McKean, Propagation of chaos for a class of non-linear parabolic equation, Lecture Series in Differential Equations, session 7, Catholic Univ., 1967, 177-194.
14. I. Karatzas and S. Shreve Brownian Motion and Stochastic Calculus, 2nd edition Springer, 1991
15. H. Osada, Propagation of chaos for the two dimensional Navier-Stokes equation, Probabilistic methods in mathematical physics (Katata/Kyoto, 1985), 303–334, Academic Press, Boston, MA, 1987.
16. A. Stevens, The derivation of chemotaxis equations as limit dynamics of moderately interacting stochastic many-particle systems, SIAM J. Appl. Math., **61** (2000), 183-212.
17. A.-S. Sznitman, Topics in propagation of chaos, In Ecole d'Eté de Probabilités de Saint-Flour XIX-1989, Lecture Notes in Math. 1464. Springer, Berlin, (1991).