Derivation of macroscopic equations for interacting diffusions with singular kernels

Xianliang Zhao

Derivation of macroscopic equations for interacting diffusions with singular kernels

Xianliang Zhao

A Dissertation Submitted for the Degree of Doctor

at

the Department of Mathematics Bielefeld University

August 8, 2023

Derivation of macroscopic equations for interacting diffusions with singular kernels

Dissertation zur Erlangung des Doktorgrades
der Fakultät für Mathematik
der Universität Bielefeld

vorgelegt von
Xianliang Zhao

August 8, 2023

Gedruckt auf alterungsbeständigem Papier nach DIN–ISO 9706

Preface

This thesis is devoted to derive macroscopic equations for interacting diffusions, showing a flexible type of mean-field convergence and Gaussian fluctuations. We focus on the models with a class of singular kernels, and the guiding example is the vortex model related to 2D Navier-Stokes and Euler equations. We are concerned with the following three related topics:

1. We prove Gaussian fluctuations from the following interacting diffusion system
$$\mathrm{d}X_i = \frac{1}{N}\sum_{j\neq i} K(X_i - X_j)\mathrm{d}t + V(X_i)\mathrm{d}t + \sqrt{2\sigma}\mathrm{d}B_i, \quad i = 1, ..., N,$$
to the nonlinear Fokker-Planck equation
$$\partial_t v_t = \sigma\Delta v_t - \nabla\cdot\Big([K * v_t + V]v_t\Big).$$
By the Donsker-Varadhan variational formula and a uniform relative entropy estimate, we show that the fluctuation measures $\eta^N := \sqrt{N}(\frac{1}{N}\sum_{i=1}^N \delta_{X_i} - v)$ converge to η in suitable negative Sobolev spaces, where η uniquely solves the linear stochastic partial differential equation driven by derivatives of space-time white noise ξ,
$$\partial_t \eta = \sigma\Delta\eta_t - \nabla\cdot\Big([K * v_t + V]\eta_t + K * \eta_t v_t\Big) - \sqrt{2\sigma}\nabla\cdot(\sqrt{v_t}\xi).$$
The Gaussianity and optimal regularity of η are also investigated.

2. We study the mean-field limit of non-exchangeable interacting diffusions, described by
$$\mathrm{d}X_i = \frac{1}{N}\sum_{j\neq i} w_j^N K(X_i - X_j)\mathrm{d}t + \sqrt{2}\mathrm{d}B_i, \quad i = 1, ..., N.$$
Here the deterministic weights $(w_j^N) \subset \mathbb{R}$ are allowed to be non-identical and unbounded. We derive a coupled mean-field PDE from the interacting diffusions, that is
$$\begin{cases} \partial_t v_t = \Delta v_t - \nabla\cdot(K * v_t v_t), \\ \partial_t g_t = \Delta g_t - \nabla\cdot(K * v_t g_t). \end{cases}$$
This result in particular provides mean-field approximations to the dynamics of the passive scalar advected by the 2D Navier-Stokes equation.

3. With more general weights and in a slightly different setting compared to the second topic, the following interacting diffusions with interactions weighted on graphs are investigated,

$$\mathrm{d}X_i = \frac{1}{N}\sum_{j\neq i} w_{ij}^N K(X_i - X_j)\mathrm{d}t + \sqrt{2}\mathrm{d}B_i, \quad i = 1, ..., N.$$

By regarding (w_{ij}^N) as the values of wedges of graphs, we derive the following marcroscopic model

$$\partial_t f(t,x,\xi) = \Delta_x f(t,x,\xi) - \nabla_x \cdot \left(f(t,x,\xi) \int_{\mathbb{R}^d} \int_0^1 G(\xi, eta) K(t, x-y) f(t,y,\zeta)\mathrm{d}\zeta \mathrm{d}y \right),$$

in the sense that $\frac{1}{N}\sum_{i=1}^N \delta_{X_i(t)}$ weakly converges to $\int_0^1 f(t,x,\xi)\mathrm{d}\xi$. Here $G : [0,1] \times [0,1] \to \mathbb{R}$ is the limiting graphon of $\{(w_{ij}^N), N \in \mathbb{N}\}$.

Acknowledgement Before anything else, I want to say how grateful I am to my Ph.D. advisors, Prof. Dr. Zhiming Ma, Prof. Dr. Michael Röckner, and Prof. Dr. Rongchan Zhu, for all the time, energy, and wisdom they put into guiding my research and helping me succeed in my studies. Through the duration of my study and writing, their direction was invaluable.

In addition to my advisors, I would like to express my gratitude to Prof. Dr. Zhao Dong, Prof. Dr. Zhenfu Wang, and Prof. Dr. Xiangchan Zhu, all of whom provided me with insightful suggestions and encouragement during my Ph.D. program and beyond.

As well, I'd want to express my gratitude to the people who have helped me throughout the time of doctoral study, including Bingguang Chen, Mengyu Chen, Zhenxin Ding, Fan Gu, Zimo Hao, Liping Li, Chengcheng Ling, Siyu Liang, Stephen Merkes, Longjie Xie, Huanyu Yang, Li Yang, Chi Zhang, Guohuan Zhao, and many more. Those years were some of the most impressive of my life thanks to them.

I owe a great debt of gratitude to my sister Ting Xie, the medical staff, and many more for all of their support and kindness while I was sick. No one will ever know the depth of my gratitude to them.

Last but not least, I'd like to express my eternal gratitude to my family in China who live far away yet have always been there for me emotionally and spiritually.

Bielefeld, August 8, 2023, Xianliang Zhao

Contents

Chapter 1

Introduction and main results

This thesis is aimed to derive macroscopic continuous models characterizing the asymptotic behavior of interacting particle systems as the number of particles goes to infinity. We shall focus on first order systems describing by singularly interacting diffusions of the following form

$$\mathrm{d}X_i = \frac{1}{N}\sum_{j\neq i} K(X_i - X_j)\mathrm{d}t + \sqrt{2}\mathrm{d}B_i, \quad i = 1, ..., N. \tag{1.1}$$

Here $\{B_i, i \in \mathbb{N}\}$ is a family of independent standard Brownian motions on the Euclidean space $\mathbb{R}^d$ or the torus $\mathbb{T}^d$, and K stands for the interaction kernel.

Many particle systems written in the canonical form (1.1) or variants thereof are now quite ubiquitous. Such systems are usually derived from first-principle agent based models which are conceptually simple. For instance, in physics those particles X_i can represent ions and electrons in plasma physics [Dob79], or molecules in a fluid [JO04] or even large scale galaxies [Jea15] in some cosmological models; in biological sciences, they typically model the collective behavior of animals or micro-organisms (for instance flocking, swarming and chemotaxis and other aggregation phenomena [CCH14]); in economics or social sciences particles are usually individual "agents" or "players" for instance in opinion dynamics [FJ90] or in the study of mean-field games [LL07, HMC$^+$06]. The motivation even extends to the analysis of large biological [BFT15] or artificial [MMN18] neural networks in neuroscience or in machine learning.

Under mild assumptions, it is well-known that the sequence of empirical measures $\mu_N(t) := \frac{1}{N}\sum_{i=1}^N \delta_{X_i(t)}$ of the particle system (1.1) converges to the solution $v(t)$ of the following Fokker-Planck equation (also called mean-field equation)

$$\partial_t v = \Delta v - \mathrm{div}((K * v)v), \tag{1.2}$$

as $N \to \infty$. Such convergence of empirical measures, which is called the mean-field convergence, is equivalent to the *propagation of chaos*, i.e. the $k-$th marginal $F_{N,k}$ of the particle system (1.1) will converge to the product of the limit law $v^{\otimes k}$ as N goes to infinity, for instance for given the i.i.d. initial data. This law of large numbers type result implies that the continuum model (1.2) is a suitable approximation to the particle system (1.1) when N is large, in the sense that $\mu_N \approx v + o(1)$.

The mean field limit and propagation of chaos for the 1st order system given in our canonical form (1.1) have been extensively studied over the last decades. The basic idea of deriving some effective PDE (as in (1.2)) describing the large scale behaviour of interacting particle systems dates back to Maxwell and Boltzmann. But in our setting, the very first mathematical investigation can be traced back to McKean in [McK67]. See also the classical mean field limit from Newton dynamics towards Vlasov Kinetic PDEs in [Dob79, BH77, JH15, Laz16] and the review [Jab14]. Recently much progress has been made on the mean field limit for systems as (2.1) with singular interaction kernels, including those results focusing on the vortex model [Osa86, FHM14], Dyson Brownian motions [BO19, SYY20, LLX20] and very recently quantitative convergence results on general singular kernels for example as in [JW18, BJW20] and [Ser20, Due16, Ros20, NRS21, Lac21, HRZ22]. See also the references therein for a more complete description for the history of the mean field limit.

In this thesis, we generalize the mean-field convergence for interacting diffusions in several directions.

The first part is to study the central limit theorem for interacting diffusions, which provides a better continuum approximation than the mean-field convergence. The study of central limit theory for the system (1.1), in particular for those with singular interactions, is quite limited, due to the lack of proper mathematical tools. The fluctuation problem around a limiting PDE was popularized for the Boltzmann equation in the 1970s-1980s for instance in [McK75, Tan82, Tan83, Uch83]. But those results focus more on jump-type particle systems. We also refer to [BGSRS20, BGSRS22] for the recent breakthrough on the deviation of the hard sphere dynamics from the kinetic Boltzmann equation. For the fluctuations of interacting diffusions, which is the focus of this thesis, to the best of the author's knowledge, one of the earliest results is due to Itô [Itô83] , where he showed that for the system of 1D independent and identically distributed Brownian motions, the limit of the corresponding fluctuations is a Gaussian process. We refer to [TH81, Tan84, Szn84, FM97] for classical results on the fluctuations of interacting diffusions with general coeeficients. In this thesis, we shall show that the sequence of the fluctuation processes $\{\sqrt{N}(\mu_N - v), N \in \mathbb{N}\}$ converges in distribution to a generalized Ornstein-Uhlenbeck process. Our result considerably extends classical results to singular kernels, including the Biot-Savart kernel. The result applies to the point vortex model approximating the 2D incompressible Navier–Stokes equation and the 2D Euler equation. Furthermore, we study Gaussianity and optimal regularity in negative Sobolev spaces of the limiting Ornstein-Uhlenbeck process.

In the second and third parts, we are concerned with the mean-field problem for interacting diffusions with inhomogeneous interactions, i.e. the interactions are weighted. These two parts are studied in slightly different settings.

The study in the second part is motivated by a stochastic vortex model with general intensities:

$$\mathrm{d}X_i = \frac{1}{N}\sum_{j\neq i} w_j^N K(X_i - X_j)\,\mathrm{d}t + \sqrt{2}\mathrm{d}B_i, \quad i = 1,\dots,N,$$

where K is chosen to be the Biot-Savart kernel. The mean-field problem for the stochastic vortex model with general intensities was investigated in [FHM14, Wyn21]. In this thesis, we will study this problem without the usual boundedness conditions and, more importantly, without assuming the symmetry/exchangeability of intensities. We shall demonstrate a flexible type of mean-field convergence for non-exchangeable interacting diffusions with singular kernels, in contrast to the typical convergence of $\frac{1}{N}\sum_{i=1}^N \delta_{X_i}$. More specifically, we show that the sequence of signed empirical measure processes with arbitrary uniform l^r-weights, $r > 1$, weakly converges to a coupled PDE, such as the one for the dynamics describing the passive scalar advected by the 2D Navier-Stokes equation.

In the third part, we study interacting diffusions with weights (w_{ij}^N), which are also called graphon particle systems since one can regard the weights as values of wedges of graphs. Graphon particle systems have attracted increasing attention in recent years, and the mean-field convergence for interacting diffusions weighted on dense graphs has been studied in [BCW20, BCN20, JPS21, Luc20, OR19] in different settings. However, to the best of our knowledge, general mean-field results for systems with singular kernels have not been addressed yet. By the theory of graph limits (see [Lov12, BCCZ]) and the method used in the second part, we prove the mean-field convergence for graphon particle systems with $L^q([0,T],L^p(\mathbb{R}^d))$-type interactions, which is much more singular in comparison to previous results.

1.1 Gaussian fluctuations

In this section, we consider the following interacting diffusions on the torus $\mathbb{T}^d$, $d \geqslant 2$,

$$\mathrm{d}X_i = \frac{1}{N}\sum_{j\neq i} K(X_i - X_j)\mathrm{d}t + V(X_i)\mathrm{d}t + \sqrt{2\sigma_N}\mathrm{d}B_i, \quad i = 1, ..., N, \tag{1.3}$$

with $\mathcal{F}_0$-measurable random initial data $\{X_i(0)\}_{i=1}^N \subset \mathbb{T}^d$. The collection $\{B_i\}_{i=1}^N$ consists of N independent d dimensional Brownian motions on a stochastic basis, i.e. $(\Omega, \mathcal{F}, \mathbb{P})$ with a normal filtration $(\mathcal{F}_t)$, induced by the Laplacian operator on the torus. The coefficient $\sigma_N \geqslant 0$ is a non-negative scalar for simplicity, and the vector fields K and V on the manifold $\mathbb{T}^d$ take values in the tangent spaces. In this model, $X^N(t) := (X_1(t), ..., X_N(t)) \in (\mathbb{T}^d)^N$ represents the positions of particles, which are interacting through the kernel K and confined by the exterior force V. Such stochastic differential equations (SDEs) on manifolds have been studied in different ways (intrinsic, extrinsic, and local). For instance, the extrinsic approach is to embed the manifold into a Euclidean space, based on Whitney's embedding theorem, and then study the equation on the Euclidean space. We refer to monographs [Elw82, Éme12, Str00, Hsu02] for more details about the settings and classical results.

Inspired in particular by quantitative estimates for the propagation of chaos by Jabin and Wang [JW18], which is $\|F_{N,k} - v^{\otimes k}\|_{L^\infty([0,T],L^1)} \leqslant C_T/\sqrt{N}$, we aim to study the fluctuations of (1.3). More precisely, we study the limit of the fluctuation measures around the mean-field law, which are defined by

$$\eta^N := \sqrt{N}(\mu_N - v) = \frac{1}{\sqrt{N}}\sum_{i=1}^N (\delta_{X_i} - v)\,. \tag{1.4}$$

Here v is a probability density-valued solution of the mean-field equation

$$\partial_t v = \sigma\Delta v - \mathrm{div}([V + K * v]v), \tag{1.5}$$

in the distributional sense, the solutions are defined in the following assumptions. In this part Δ denotes the Laplace–Beltrami operator on the torus. We shall prove that the fluctuation measures η^N converge in distribution as $N \to \infty$ to an infinite-dimensional continuous Gaussian process η for a large class of particle systems (1.3). This implies that there exists a continuum model η such that

$$\mu_N \overset{\mathrm{d}}{=} v + \frac{1}{\sqrt{N}}\eta + o(\frac{1}{\sqrt{N}}),$$

where $\stackrel{\mathrm{d}}{=}$ means that the approximation holds in distribution.

1.1.1 Assumptions

To state our main results we first present the framework in this section. Recall that the relative entropy $H(\mu|\nu)$ between probability measures μ and ν on a Polish space E is defined by

$$H(\mu|\nu) := \begin{cases} \int_E \frac{\mathrm{d}\mu}{\mathrm{d}\nu} \log \frac{\mathrm{d}\mu}{\mathrm{d}\nu} \mathrm{d}\nu & \text{if } \mu \ll \nu, \\ \infty & \text{otherwise,} \end{cases}$$

where $\frac{\mathrm{d}\mu}{\mathrm{d}\nu}$ is the Radon-Nikodym derivative of μ with respect to ν. Note that throughout this thesis, the relative entropy is of this classical form. We will not normalize it as what has been done for instance in [JW18].

Our assumptions are listed as follows.

(A1)-CLT for initial values. There exists a random variable η_0, with values in the space of distributions $\mathcal{S}'(\mathbb{T}^d)$, such that the sequence $\{\eta_0^N\}_{N\geqslant 1}$ converges in distribution to η_0. (Here η_0 will be the initial data for our expected limit SPDE (1.7) below.)

(A2)-Regularity of the kernel. The kernel $K : \mathbb{T}^d \to \mathbb{R}^d$, $d \geqslant 2$, satisfies one of the following conditions

1. K is bounded.
2. For each $x \in \mathbb{T}^d$, $K(x) = -K(-x)$ and $|x|K(x) \in L^\infty$.

(A3)-Uniform relative entropy bound. Let $X^N(t) = (X_1(t), ..., X_N(t))$ be a solution to the particle system (1.3), and let F_t^N represent the joint distribution of $X^N(t)$. It holds that

$$\sup_{t\in[0,T]} \sup_N H(F_t^N|v_t^{\otimes N}) < \infty.$$

For simplicity, define $H_t(F^N|v^{\otimes N}) = H(F_t^N|v_t^{\otimes N})$.

The global well-posedness of the limit equation (1.7) will be obtained by two different approaches, depending on whether the diffusion coefficient σ is positive or zero. Hence we distinguish the extra assumptions into the following two cases.

For the case when $\sigma > 0$, in addition to the assumptions **(A1)-(A3)**, we need the following extra assumption:

(A4)-The case with non-vanishing diffusion

1. $\sigma := \lim_{N\to\infty} \sigma_N = 0$ and $|\sigma_N - \sigma| = \mathcal{O}\left(\frac{1}{N}\right)$.
2. $V \in C^\beta(\mathbb{T}^d; \mathbb{R}^d)$.
3. There exists a probability density-valued solution of equation (1.5) v in the sense that $v \in C([0,T], C^\beta(\mathbb{T}^d))$ for some $\beta > d/2$ and v satisfies the equation in the distributional sense.

On the other hand, for the case with vanishing diffusion, besides **(A1)-(A3)**, we require that

(A5)-The case with vanishing diffusion The diffusion coefficients and the mean-field equation (1.5) satisfy,

1. $\sigma := \lim_{N\to\infty} \sigma_N = 0$ and $|\sigma_N - \sigma| = \mathcal{O}\left(\frac{1}{N}\right)$.
2. $\mathrm{div}K \in L^1$.
3. $V \in C^{\beta+1}(\mathbb{T}^d; \mathbb{R}^d)$.
4. There exists a probability density-valued solution of equation (1.5) v in the sense that $v \in C^1([0,T], C^{\beta+2}(\mathbb{T}^d))$ for some $\beta > d/2$ and v satisfies the equation in the distributional sense.

We make several remarks on our assumptions. Firstly, when $\{X_i(0)\}_{i\in\mathbb{N}}$ are i.i.d. with a common probability density function v_0, which is the usual setting to study the fluctuations, one can easily check that **(A1)** holds true. Indeed, for each $\varphi \in C^\infty(\mathbb{T}^d)$, we have

$$\langle \eta_0^N, \varphi\rangle = \frac{1}{\sqrt{N}} \sum_{i=1}^N \Big[\varphi(X_i(0)) - \langle \varphi, v_0\rangle\Big] \xrightarrow{N\to\infty} \mathcal{N}\Big(0, \langle \varphi^2, v_0\rangle - \langle \varphi, v_0\rangle^2\Big),$$

where $\mathcal{N}(0,a)$ denotes the centered Gaussian distribution on $\mathbb{R}$ with variance a. Hereafter we use the bracket $\langle\cdot,\cdot\rangle$ as a shorthand notation for integration, if there is no confusion possible.

Assumption **(A2)** on interaction kernels allows our framework to cover smooth kernels and some singular kernels, in particular, the Biot-Savart kernel related to the vorticity formulation of 2D Navier-Stokes/Euler equation on the torus. See Theorem 1.7 and Section 2.6 for more details.

Assumption **(A3)** seems quite nontrivial and demanding, but fortunately it has been established by Jabin and Wang in [JW18] for a quite large family of interacting kernels. Indeed, once we have that the relative entropy between the joint distribution F^N of the interacting particle system (1.3) and the tensorized law $v^{\otimes N}$ of the mean-field PDE (1.5) is uniformly bounded with respect to N, then easily the particle system (1.3) converges to the mean-field equation (1.5) with a rate $\frac{C_T}{\sqrt{N}}$, in the total variation norm or the Wasserstein metric. More precisely, since all particles in (1.3) are indistinguishable, the joint distribution F^N is thus assumed to be symmetric/exchangeable, so is any k-marginal distribution $F_{N,k}$ of F^N, which is defined as

$$F_{N,k}(t, x_1, ..., x_k) := \int_{\mathbb{T}^{d(N-k)}} F^N(t, x_1, ..., x_N)\mathrm{d}x_{k+1}...\mathrm{d}x_N.$$

Then by the sub-additivity of the relative entropy (in particular $H(F_{N,k}|v^{\otimes k}) \leqslant \frac{k}{N} H(F^N|v^{\otimes N})$) and the classical Csiszár–Kullback–Pinsker inequality [Vil08, (22.25)], it follows that for fixed $k \in \mathbb{N}$,

$$W_1\Big(F_{N,k}(t), v_t^{\otimes k}\Big) \lesssim \|F_{N,k}(t) - v_t^{\otimes k}\|_{TV} \leqslant \sqrt{2H_t(F_{N,k}|v^{\otimes k})} \lesssim \sqrt{\frac{k}{N}} \longrightarrow 0, \qquad (1.6)$$

where $W_1(\cdot,\cdot)$ denotes the 1-Wasserstein distance, $\|\cdot\|_{TV}$ denotes the total variation norm and the first inequality is guaranteed by [Vil08, Theorem 6.15], since $\mathbb{T}^d$ is compact.

1.1.2 The first main results

Under the assumptions **(A1)-(A3)** and either **(A4)** or **(A5)**, depending on $\sigma > 0$ or $\sigma = 0$, we establish that as $N \to \infty$, the sequence of the fluctuation measures $(\eta^N_\cdot)$ converges in distribution to the centered Gaussian process η solving the following stochastic PDE (SPDE)

$$\partial_t \eta = \sigma \Delta \eta - \nabla \cdot (vK * \eta) - \nabla \cdot (\eta K * v) - \nabla \cdot (V\eta) - \sqrt{2\sigma} \nabla \cdot \left(\sqrt{v}\xi\right), \qquad \eta(0) = \eta_0, \quad (1.7)$$

where η_0 is given in Assumption **(A1)** and ξ is a vector-valued space-time white noise on $\mathbb{R}^+ \times \mathbb{T}^d$, i.e. a family of centered Gaussian random variables $\{\xi(h) : h \in L^2(\mathbb{R}^+ \times \mathbb{T}^d; \mathbb{R}^d)\}$ such that $\mathbb{E}[|\xi(h)|^2] = \|h\|^2_{L^2(\mathbb{R}^+ \times \mathbb{T}^d;\mathbb{R}^d)}$, and v is a probability density-valued solution of the mean-field equation (1.5) in the distributional sense. But when $\sigma = 0$, the SPDE (1.7) becomes a deterministic PDE with random initial values.

We now make a proper notion of solutions to the SPDE (1.7). When $\sigma > 0$, it turns out to be the martingale solutions defined as below.

Definition 1.1. *Given that $\sigma > 0$, we call η a martingale solution to the SPDE* (1.7) *on some stochastic basis* $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ *if*

1. *η is a continuous $(\mathcal{F}_t)$-adapted process with values in $H^{-\alpha-2}(\mathbb{T}^d)$ (defined in Section 1.6) and $\eta \in L^2([0,T], H^{-\alpha}(\mathbb{T}^d))$ for every $\alpha > d/2$, $\mathbb{P}$-almost surely.*

2. *For each $\varphi \in C^\infty(\mathbb{T}^d)$ and $t \in [0,T]$,*

$$\langle \eta_t, \varphi \rangle - \langle \eta_0, \varphi \rangle = \int_0^t \langle \sigma \Delta \varphi, \eta \rangle \mathrm{d}s + \int_0^t \langle \nabla \varphi, vK * \eta + \eta K * v + V\eta \rangle \mathrm{d}s + \mathcal{M}_t(\varphi),$$

   where $\mathcal{M}$ is a continuous $(\mathcal{F}_t)$-adapted centered Gaussian process with values in $H^{-\alpha-1}(\mathbb{T}^d)$ for every $\alpha > d/2$ and its covariance is given by

$$\mathbb{E}[\mathcal{M}_t(\varphi_1)\mathcal{M}_s(\varphi_2)] = 2\sigma \int_0^{s\wedge t} \langle \nabla \varphi_1 \cdot \nabla \varphi_2, v_r \rangle \mathrm{d}r,$$

   for each $\varphi_1, \varphi_2 \in C^\infty(\mathbb{T}^d)$ and $s, t \in [0,T]$.

Remark 1.2. *1. The stochastic basis in Definition 1.1 might be different from the stochastic basis for the particle system* (1.3)*.*

2. *By Lemma 2.2 and Lemma 2.4 given in Section 2.1.1, $vK * \eta$, $\eta K * v$, and $V\eta$ are all well-defined under Assumption* **(A4)***.*

3. *The noise $\mathcal{M}$ is equivalently characterized as follows: for each $\varphi \in C^\infty$, $\mathcal{M}(\varphi)$ is a continuous $(\mathcal{F}_t)$-adapted martingale with quadratic variation given by*

$$\mathbb{E}\Big[|\mathcal{M}_t(\varphi)|^2\Big] = 2\sigma \int_0^t \langle |\nabla \varphi|^2, v_r \rangle \mathrm{d}r.$$

When $\sigma = 0$, the equation (1.7) actually becomes a deterministic PDE. We define solutions to this first order PDE as follows.

Definition 1.3. *Given that $\sigma = 0$, we call η a solution to the PDE* (1.7) *with random initial data η_0, if*

1. *$\eta \in L^2([0,T], H^{-\alpha}(\mathbb{T}^d)) \cap C([0,T], H^{-\alpha-2}(\mathbb{T}^d))$ for every $\alpha > d/2$ almost surely.*

2. *For each $\varphi \in C^\infty(\mathbb{T}^d)$ and $t \in [0,T]$,*

$$\langle \eta_t, \varphi \rangle = \langle \eta_0, \varphi \rangle + \int_0^t \langle \nabla \varphi, vK * \eta + \eta K * v + V\eta \rangle \mathrm{d}s.$$

Our first main result gives the convergence of fluctuation measures if the diffusion coefficient σ is strictly positive (which may be generalized to the case with a non-degenerate coefficient matrix though).

Theorem 1.4. *Under the assumptions* **(A1)-(A4)***, the sequence η^N defined in* (1.4) *converges in distribution to η in the space $L^2([0,T], H^{-\alpha}(\mathbb{T}^d)) \cap C([0,T], H^{-\alpha-2}(\mathbb{T}^d))$ for every $\alpha > d/2$, where η is the unique martingale solution to the SPDE* (1.7)*.*

The proof of Theorem 1.4 will be given in Section 2.3.2.

It is worth emphasizing that the condition $\alpha > d/2$ is optimal due to the optimal regularity of η established in Proposition 2.31. Since the driving noise of equation (1.7) is very rough, so are the solutions. In Section 2.4.1, we rewrite the equation (1.7) in the mild form and study the regularity of the stochastic part by Kolmogorov's theorem. Using Schauder estimates, in Proposition 2.31 we obtain the optimal regularity of η to be given by $C([0,T], C^{-\alpha}(\mathbb{T}^d))$ $\mathbb{P}$-a.s. for every $\alpha > d/2$.

Compared to the previous result by Fernandez and Méléard [FM97], Theorem 1.4 requires less regularity of the kernel, but more regularity of the solution to the mean-field equation, which eventually would lead to a more restrictive condition on the initial value $v(0)$. The extra assumption on the mean-field equation is however fairly reasonable. Moreover, by regularity analysis, the condition $\beta > d/2$ in **(A4)** is optimal within the scale of Hölder spaces.

Furthermore, η is a solution to the following linear SPDE in the weak formulation:

$$\langle \eta_t, \varphi \rangle = \langle \eta_0, Q_{0,t}\varphi \rangle + \sqrt{2\sigma} \int_0^t \int_{\mathbb{T}^d} (\nabla Q_{s,t}\varphi) \sqrt{v_s}\, \xi(\mathrm{d}x, \mathrm{d}s), \tag{1.8}$$

for each $\varphi \in C^\infty$, i.e. η is an $\mathcal{S}'(\mathbb{T}^d)$-valued Ornstein-Uhlenbeck process. Here the time evolution operators $\{Q_{s,t}\}_{0 \leqslant s \leqslant t \leqslant T}$ are defined for each $t \in [0,T]$ and $\varphi \in C^\infty$, as

$$Q_{\cdot,t}\varphi := f(\cdot), \tag{1.9}$$

with

1. $f \in L^2([0,t], H^{\beta+2}(\mathbb{T}^d)) \cap C([0,t], H^{\beta+1}(\mathbb{T}^d))$ with $\partial_t f \in L^2([0,t], H^\beta(\mathbb{T}^d))$ for $\beta > d/2$.

2. f is the unique solution with terminal value φ to the following backward equation

$$f_s = \varphi + \sigma \int_s^t \Delta f_r \mathrm{d}r + \int_s^t \Big[K * v_r \cdot \nabla f_r + K(-\cdot) * (\nabla f_r v_r) + V \nabla f_r \Big] \mathrm{d}r, \quad s \in [0,t],$$

where $K(-\cdot) * g(x) := \int K(y-x) g(y) \mathrm{d}y$ and we use this convention throughout the thesis. For the definition of $\{Q_{s,t}\}$ we refer to Section 2.4.2 for more details. The formulation (1.8) gives rise to the Gaussianity of the limit of the fluctuation measures. We state the result as follows and we give the proof in Section 2.4.2.

Proposition 1.5. *Let assumptions* **(A1)-(A4)** *and assume that* $v \in C([0,T], C^{\beta+1}(\mathbb{T}^d))$, $V \in C^{\beta+1}(\mathbb{T}^d)$ *with* $\beta > d/2$*, and that* η_0 *in* **(A1)** *is characterized by*

$$\langle \eta_0, \varphi \rangle \sim \mathcal{N}(0, \langle \varphi^2, v_0 \rangle - \langle \varphi, v_0 \rangle^2), \quad \varphi \in C^\infty(\mathbb{T}^d).$$

Then for η *obtained in Theorem 1.4, it holds for each test function* $\varphi \in C^\infty$ *and* $t \in [0,T]$ *that*

$$\langle \eta_t, \varphi \rangle \sim \mathcal{N}\Big(0, \langle |Q_{0,t}\varphi|^2, v_0 \rangle - \langle Q_{0,t}\varphi, v_0 \rangle^2 + 2\sigma \int_0^t \langle |\nabla Q_{s,t}\varphi|^2, u_s \rangle \,\mathrm{d}s\Big).$$

We now focus on the fluctuation problem for the case with vanishing diffusion. In contrast to the non-degenerate case with $\sigma > 0$, due to the vanishing diffusion, the limit equation becomes a deterministic PDE, but with random initial data. We then analyze the limit equation with the method of characteristics, and obtain the following result in Section 2.5.

Theorem 1.6. *Under the assumptions* **(A1)-(A3)** *and* **(A5)***, assume further that* η_0 *in* **(A1)** *is characterized by*

$$\langle \eta_0, \varphi \rangle \sim \mathcal{N}(0, \langle \varphi^2, v_0 \rangle - \langle \varphi, v_0 \rangle^2), \quad \varphi \in C^\infty(\mathbb{T}^d).$$

Let η *be the unique solution to* (1.7) *with* $\sigma = 0$ *on the same stochastic basis as the particle system* (1.3)*. Then the sequence* η^N *defined in* (1.4) *converges in probability to* η *in the space* $L^2([0,T], H^{-\alpha}(\mathbb{T}^d)) \cap C([0,T], H^{-\alpha-2}(\mathbb{T}^d))$ *for every* $\alpha > d/2$*. Furthermore,* η *satisfies*

$$\langle \eta_t, \varphi \rangle = \langle \eta_0, Q_{0,t}\varphi \rangle \sim \mathcal{N}\Big(0, \langle |Q_{0,t}\varphi|^2, v_0 \rangle - \langle Q_{0,t}\varphi, v_0 \rangle^2\Big),$$

for each test function φ *and* $t \in [0,T]$*. Here the time evolution operator* $\{Q_{0,t}\}_{0 \leqslant t \leqslant T}$ *is given by* (1.9) *with* $\sigma = 0$*.*

Our main results validate that the relative entropy bound $\sup_{t\in[0,T]} \sup_N H(F^N|v^{\otimes N}) \lesssim 1$ which has been established by Jabin and Wang in [JW18] is actually optimal. But the convergence rate for the marginal distributions $\|F_{N,k} - v^{\otimes k}\|_{L^\infty([0,T],L^1)} \lesssim C_T/\sqrt{N}$ is not optimal. Note that very recently Lacker [Lac21] obtains a sharp estimate for marginal distributions, which is $\|F_{N,k} - v^{\otimes k}\|_{L^\infty([0,T],L^1)} \lesssim C_T/N$ by a local relative entropy analysis of the BBGKY hierarchy, but under the stronger assumption $H(F_{N,k}(0)|v_0^{\otimes k}) \lesssim k^2/N^2$. Even though it is well-known that the convergence of empirical measures and the k-marginal distributions are equivalent in the qualitative sense for instance in [Szn], their quantitative behaviors can be quite complicated when it comes to the order of N. See some related discussions in [Lac21, MM13, HM14, MMW15].

As guiding example for our main results in Theorem 1.4 and Theorem 1.6, we consider the famous vortex model for approximating the 2D Navier-Stokes equation in the vorticity formulation when $\sigma > 0$ and also the 2D Euler equation when $\sigma = 0$. More precisely, given a sequence of i.i.d. initial random variables $\{X_i(0)\}_{i\in\mathbb{N}}$ with a common probability density function v_0 on $\mathbb{T}^2$, and consider the particle system

$$\mathrm{d}X_i = \frac{1}{N}\sum_{j\neq i} K(X_i - X_j)\mathrm{d}t + \sqrt{2\sigma}\mathrm{d}B_i, \quad i = 1, 2, \cdots, N, \tag{1.10}$$

with the Biot–Savart law $K : \mathbb{T}^2 \to \mathbb{R}^2$ defined by

$$K = \nabla^\perp G = (-\partial_2 G, \partial_1 G) \tag{1.11}$$

where G is the Green function of the Laplacian on the torus $\mathbb{T}^2$ with mean 0. Note in particular that

$$K(x) = \frac{1}{2\pi}\frac{x^\perp}{|x|^2} + K_0(x),$$

where $x^\perp = (x_1, x_2)^\perp = (-x_2, x_1) \in \mathbb{R}^2$ and K_0 is a smooth correction to periodize K on the torus $\mathbb{T}^2$. Obviously the Biot-Savart kernel K satisfies our assumption **(A2)**.

One major corollary of our main results Theorem 1.4 and Theorem 1.6 is the following result.

Theorem 1.7. *If $v_0 \in C^3(\mathbb{T}^2)$ when $\sigma > 0$ and $v_0 \in C^4(\mathbb{T}^2)$ when $\sigma = 0$, and $\inf v_0 > 0$ for both cases, then the sequence of fluctuation measures $\{\eta^N\}_{N\in\mathbb{N}}$ associated with* (1.10) *converges in distribution to η in the space $L^2([0,T], H^{-\alpha}(\mathbb{T}^d)) \cap C([0,T], H^{-\alpha-2}(\mathbb{T}^d))$ for every $\alpha > 1$. Here η is a generalized Ornstein-Uhlenbeck process solving the equation* (1.7) *with K given by* (1.11) *and $V = 0$. Moreover, $\langle \eta, \varphi\rangle$ is a centered continuous Gaussian process with covariance*

$$\left\langle |Q_{0,t}\varphi|^2, v_0\right\rangle - \left\langle Q_{0,t}\varphi, v_0\right\rangle^2 + 2\sigma\int_0^t \left\langle |\nabla Q_{s,t}\varphi|^2, v_s\right\rangle \mathrm{d}s,$$

where $\{Q_{s,t}\}$ is introduced in (1.9) *with $V = 0$ and u is the solution to the vorticity formulation of the 2D imcompressible Navier-Stokes equation when $\sigma > 0$ and 2D Euler equation when $\sigma = 0$, respectively.*

The point vortex approximation towards 2D Navier-Stokes/Euler equation raised a lot of interest since 1980s. The well-posednesss of the point vortex model (1.10) was established in [Osa85, MP12, Tak85, FM07]. The main part is to show that $X_i(t) \neq X_j(t)$ for all $t \in [0,T]$ and $i \neq j$ almost surely, thus the singularity of the kernel will not be visited almost surely. The routine method for instance in [Tak85] is based on estimating the quantity $\sum_{i\neq j} G(|X_i - X_j|)$, where G is the Green function. Using the fact $\nabla G \cdot \nabla^\perp G = 0$ and by regularization in an intermediate step, it can be shown that $\sum_{i\neq j} G(|X_i - X_j|)$ is finite almost surely for each $t \in [0,T]$. In [MP12] by Marchioro and Pulvirenti and [FM07] by Fontbona and Martinez, the well-posedness of the point vortex model with more general circulations/intensities was established by estimating the displacements of particles. Osada in [Osa85] obtained the same result by an analytic approach, which depends on Gaussian upper and lower bounds for the fundamental solution associated to differential operators of a generalized divergence form and the result from [Kan67].

Osada [Osa86] firstly obtained a propogation of chaos result for (1.10) with bounded initial distribution and large viscosity. More recently, Fournier, Hauray, and Mischler [FHM14] obtained entropic propagation of chaos by a compactness argument and their result applies to arbitrary viscosity, as long as it is strictly positive, and all initial distributions have finite k-moment ($k > 0$) and finite Boltzmann entropy. A quantitative estimate of propagation of chaos has been established by Jabin and Wang in [JW18] by evolving the relative entropy between the joint distribution of (1.10) and the product law at the limit. In particular, note that [JW18] provided the uniform relative entropy bound as in **(A3)** for all the kernels satisfying **(A2)**, including the Biot-Savart kernel. We also mention the result [FL21] by Flandoli and Luo, in which they approximate the 2D Navier-Stokes equations in vorticity form by the point vortex model driven by multiplicative environmental noises. Very recently, a large deviation result on the torus has been obtained by Chen and Ge [CG22].

To the author's best knowledge, Theorem 1.7 is the first result on the fluctuation problem for the 2D Navier-Stokes/Euler equation.

1.2 Mean-field limits for non-exchangeable systems

In this section we consider a generalized version of (1.1), which is non-exchangeable due to inhomogeneous interactions. Note that in this part we work on $\mathbb{R}^d$. Given random initial data $\{X_i(0)\}_{i=1}^N$, the position of each particle X_i is characterized by the following SDEs

$$\mathrm{d}X_i = \frac{1}{N}\sum_{j\neq i} w_j^N K(X_i - X_j)\,\mathrm{d}t + \sqrt{2}\mathrm{d}B_i, \quad i = 1, \dots, N, \tag{1.12}$$

where K denotes the interaction kernel, (B_i) are independent standard Brownian motions on $\mathbb{R}^d$, $d \geqslant 2$, and $\{w_j^N\} \subset \mathbb{R}$ are non-identical deterministic weights that satisfy the following assumption for some $r \in (1, \infty]$

$$\begin{aligned}(\mathbf{W_r}): \quad \frac{1}{N}\sum_{j=1}^N |w_j^N|^r =& O(1), \quad \text{for } r < \infty;\\ \max_{1\leqslant j\leqslant N} |w_j^N| =& O(1), \quad \text{for } r = \infty, \quad \text{as } N \to \infty,\end{aligned} \tag{1.13}$$

where $O(\cdot)$ means "proportional to". Here $r > 1$ ensures that the system is weakly interacting, indeed $|w_j^N|/N \to 0$.

Similar to the last section, the prime example is the famous stochastic vortex model with general intensities, i.e. the kernel K in the system (1.12) is the Biot-Savart kernel on $\mathbb{R}^d$, defined by

$$K = \nabla^\perp G = (-\partial_2 G, \partial_1 G), \tag{1.14}$$

where G is the Green function of the Laplacian on $\mathbb{R}^2$. Note in particular that

$$K(x) = \frac{1}{2\pi}\frac{x^\perp}{|x|^2}$$

where $x^\perp = (x_1, x_2)^\perp = (-x_2, x_1) \in \mathbb{R}^2$. Now the weights w_j^N denotes the intensity/magnitude of the j-th point vortex at position X_j. One may expect that now the (weighted) empirical measures defined as

$$\mu_N(t) = \frac{1}{N}\sum_{j=1}^N w_j^N \delta_{X_j(t)}, \quad N \in \mathbb{N},$$

will converge to the solutions of the 2D Navier-Stokes equations in vorticity form

$$\partial_t v(t,x) + \operatorname{div}(v(t,x) K * v_t(x)) = \Delta v(t,x). \tag{1.15}$$

Our main results validate the above mean-field approximation for 2D Navier-Stokes equation under very general assumptions on the intensities w_j^N, which can be of mixed-sign and unbounded. See in particular Theorem 1.16.

As mentioned before, classical and more recent investigations on the topic of mean-field approximation have mainly focused on the case $w_j^N \equiv 1$ for all $1 \leqslant j \leqslant N$. Classically, such a mean-field limit implies that a continuum model can be found to approximate the associated particle system when N is large. In this section, we do not only establish the mean-field limit for systems with general weights $w^N := (w_1^N, ..., w_N^N)$ as in the system (1.12), but as a byproduct, we demonstrate a more flexible mean-field convergence. Indeed, we can consider the following continuum model given by two coupled PDE's

$$\begin{cases} \partial_t g_t = \Delta g_t - \operatorname{div}\Big(g_t \, K * v_t\Big), \\ \\ \partial_t v_t = \Delta v_t - \operatorname{div}\Big(v_t K * v_t\Big), \end{cases} \tag{1.16}$$

with $g_0, v_0 \in L^1(\mathbb{R}^d)$. The continuum model (1.16) turns out to be a suitable mean-field system for the linear statistics of the interacting diffusions (1.12). Formally, let $\{\tilde{w}^N\}$ be any other sequence of weights that satisfies the assumption $(\mathbf{W}_r)$, then the system (1.16) is a continuous approximation to (1.12), in the sense that as $N \to \infty$,

$$\frac{1}{N} \sum_{i=1}^{N} w_i^N \delta_{X_i} \stackrel{d}{=} v + o(1), \quad \frac{1}{N} \sum_{i=1}^{N} \tilde{w}_i^N \delta_{X_i} \stackrel{d}{=} g + o(1).$$

The novelty of this approximation is that now the choice of weights $\tilde{w}^N$ can be quite flexible, including the classical average type, i.e. $\tilde{w}_i^N = 1$ for all i, or a more general choice based on the relative importance of each particle. When K is the Biot-Savart kernel, our main results do not only provide a viscous vortex model approximation to the vorticity formulation of the 2D Navier-Stokes equation, but now the vorticity is of mixed sign, and also establish a particle approximation to the related passive scalar equation where the flow is given by the Navier-Stokes equation.

Remark 1.8. *If we just consider the deterministic setting, where there is no Brownian motion term in* (1.12)*, and for instance for* $\mu_N = \frac{1}{N}\sum_{i=1}^N w_i^N \delta_{X_i}$ *and* $\tilde{\mu}_N = \frac{1}{N}\sum_{i=1}^N \tilde{w}_i{}^N \delta_{X_i}$*, it is easy to check that* μ_N *and* $\tilde{\mu}_N$ *solves*

$$\partial_t \mu_N + \operatorname{div}_x(\mu_N K * \mu_N) = 0,$$

and

$$\partial_t \tilde{\mu}_N + \operatorname{div}_x(\tilde{\mu}_N K * \mu_N) = 0,$$

respectively. To derive the above PDE or the continuum model (1.16)*, the interacting particle system* (1.12) *is the most general form we can expect. Indeed, if the weights depend both on* i *and* j*, i.e.* $w = (w_{ij}^N)$ *as in [JPS21] or in Section 1.3 below, then there is no simple way to define an empirical measure, let alone to study its corresponding PDEs.*

Now let us state some notations and definitions. We shall use the (non-normalized) Boltzmann entropy functional on $\mathcal{P}_\gamma(\mathbb{R}^{dN})$, which is the subspace of the probability measure space $\mathcal{P}(\mathbb{R}^{dN})$ under the constraint of finite γ-th moment, $\gamma \in (0,1)$. The entropy functional is given by

$$H(f) := \int_{\mathbb{R}^{dN}} f \log f \, \mathrm{d}x^N, \quad f \in \mathcal{P}_\gamma(\mathbb{R}^{dN}) \cap L^1(\mathbb{R}^d),$$

with x^N denoting $(x_1, ..., x_N)$, $x_i \in \mathbb{R}^d$, $1 \leqslant i \leqslant N$. If f has no density, then we set $H(f) = +\infty$. A well-known fact is that the negative part of $H(f)$ is bounded by a universal constant plus the γ-th moment of f. Thus the entropy functional is well-defined on $\mathcal{P}_\gamma(\mathbb{R}^{dN})$.

We assume the following conditions on the initial value and the interaction kernel.

(**H**) Let F_0^N be the joint distribution of $X^N(0) := (X_1(0), .., X_N(0))$. There exists some constant $\gamma \in (0,1)$ such that

$$H(F_0^N) = O(N), \quad \sum_{i=1}^N \mathbb{E}\langle X_i(0)\rangle^\gamma = O(N), \quad \text{as } N \to \infty,$$

where $\langle x\rangle := (1+|x|^2)^{\frac{1}{2}}$.

($\mathbf{K_r}$) Given $r \in (1,\infty]$ from ($\mathbf{W_r}$) and $d \geqslant 2$. The kernel K is of the form $K = K_1 + K_2$, with K_1, K_2 satisfying

1. $\mathrm{div}K_1$, $K_1 \in L^{q_1}([0,T], L^{p_1}(\mathbb{R}^d))$ with $\frac{d}{p_1} + \frac{2}{q_1} + \frac{2}{r} < 2$, where the equality is allowed to be attained when $q_1, r < \infty$;
2. $K_2 \in L^{q_2}([0,T], L^{p_2}(\mathbb{R}^d))$ with $\frac{d}{p_2} + \frac{2}{q_2} + \frac{1}{r} < 1$, where the equality is allowed to be attained when $q_2, r < \infty$.

In the following we give typical examples satisfying condition ($\mathbf{K_r}$).

Examples

1. The Biot-Savart kernel in dimension 2 as in (1.14). It is divergence free and belongs to $L^p + L^\infty$ with $1 < p < 2$, so that it satisfies ($\mathbf{K_r}$) with $r > 2$.
2. $K(x) = \frac{x}{|x|^\alpha}$ or $-\frac{x}{|x|^\alpha}$ with $\alpha \in (1,2)$ and $d \geqslant 2$. Then $K \in L^p + L^\infty$ with $1 < p < \frac{d}{\alpha-1}$ and ($\mathbf{K_r}$) holds with $r > \frac{1}{2-\alpha}$. Since the weights are allowed to be of mixed signs, the force K could be attractive or repulsive.

For the particle system (1.12) on $[0,T]$, $T > 0$, we define the notion of entropy solutions.

Definition 1.9 (Entropy Solutions). *Let $X^N = (X_1, ..., X_N)$ be a $C([0,T], \mathbb{R}^{dN})$-valued random variable satisfying the initial condition* (**H**)*, and let F_t^N denote the law of $X^N(t)$.*

For the system (1.12) *under the condition* ($\mathbf{K_r}$)*, we call X^N an entropy solution if there exists a universal constant $C > 0$ and a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P})$ with a standard dN-dimensional Brownian motions $(B_1, ..., B_N)$ such that X^N solves* (1.12) *$\mathbb{P}$-almost surely and for $t \in [0,T]$,*

$$H(F_t^N) + \sum_{i=1}^N \mathbb{E}\langle X_i(t)\rangle^\gamma + \frac{1}{2}\int_0^t \int_{\mathbb{R}^{dN}} \frac{|\nabla F_t^N|^2}{F_t^N} \mathrm{d}x^N \mathrm{d}t \leqslant H(F_0^N) + \sum_{i=1}^N \mathbb{E}\langle X_i(0)\rangle^\gamma + CN. \tag{1.17}$$

Clearly, each entropy solution to (1.12) is a probabilistically weak solution. The next result gives the existence of entropy solutions.

Proposition 1.10 (Proposition 3.8 below)**.** *Under the conditions* (**H**)*,* ($\mathbf{K_r}$)*, and* ($\mathbf{W_r}$) *for some $r \in (1, \infty]$, for each $N \in \mathbb{N}$, there exists an entropy solution X^N to the particle system* (1.12) *such that the entropy dissipation inequality* (1.17) *holds with some universal constant C that is independent of N.*

The regularity of entropy solutions with a uniform constant will enable us to find the mean-field limits. The entropy solution has proven useful for studying interacting diffusions, and Proposition 1.10 is indeed analogous to [FHM14, Proposition 5.1] and [JW18, Proposition 1]. But we do not need the divergence free or bounded-like conditions on the kernel.

The probabilistically strong well-posedness of the singular interacting system (1.12) with general weights is a fascinating and challenging problem. To the author's best knowledge, existing results are only for specific kernels, such as [FM07] and [FGP11] on the Biot-Savart kernel. The probabilitistically weak well-posedness for a large class of kernels and weights can be found in Section 1.3 below, i.e. Proposition 1.18. There are a lot of results for identical weights, for example [Osa85, Tak85, MP12] on the Biot-Savart kernel and the recent results in [HRZ22] on $L^q([0,T], L^p(\mathbb{R}^d))$-kernels with $d/p + 2/q < 1$.

Now we consider solutions in the space $C([0,T], \mathcal{M}(\mathbb{R}^d))$ for the mean-field PDE system (1.16). Here $\mathcal{M}(\mathbb{R}^d)$ stands for the space of finite signed measures on $\mathbb{R}^d$ with the topology induced by bounded and continuous (test) functions. We use it as the state space for the convergence in our main results below. The solutions to (1.16) are defined as follows.

Definition 1.11. *We call $(v, g) \in C([0,T], \mathcal{M}(\mathbb{R}^d))^{\otimes 2} \cap L^\infty([0,T], L^1(\mathbb{R}^d))^{\otimes 2}$ a solution to the system* (1.16) *if (v, g) satisfies* (1.16) *in the distributional sense and the following estimate holds*

$$\mathbb{E}\|v\|_{L^p_q} + \mathbb{E}\|g\|_{L^p_q} < \infty, \quad \frac{d}{p} + \frac{2(r-1)}{r} \geqslant d, \; \frac{d}{p} + \frac{2}{q} \geqslant d, \; 1 \leqslant p, q < \infty, \tag{1.18}$$

where $\|\cdot\|_{L^p_q} := \|\cdot\|_{L^q([0,T], L^p(\mathbb{R}^d))}$ and $r \in (1, \infty]$. We set $\frac{r-1}{r} := 1$ when $r = \infty$.

Remark 1.12. *The conditions on p, q, r in the above definition ensures that the nonlinear term in the coupled PDEs* (1.16) *is well-defined, and also enables us to deduce uniqueness later.*

The first main result shows that (1.16) characterizes the mean-field limits of the interacting system (1.12).

Theorem 1.13. *Let $\{w^N\}$ and $\{\tilde{w}^N\}$ be two sequences of weights satisfying the condition* ($\mathbf{W_r}$) *with $r \in (1, \infty]$ and suppose that the conditions* (**H**)*,* ($\mathbf{K_r}$) *hold. Let X^N be an entropy solution to* (1.12) *as given by Proposition 1.10. Assume that there exist $v_0, g_0 \in L^1(\mathbb{R}^d)$ such that*

$$\frac{1}{N} \sum_{i=1}^N w_i^N \delta_{X_i(0)} \rightharpoonup v_0, \quad \frac{1}{N} \sum_{i=1}^N \tilde{w}_i^N \delta_{X_i(0)} \rightharpoonup g_0 \tag{1.19}$$

in $\mathcal{M}(\mathbb{R}^d)$ almost surely, where $\rightharpoonup$ means weak convergence in $\mathcal{M}(\mathbb{R}^d)$.

Then the corresponding sequence of laws for the weighted empirical measures $(\mu_N, \tilde{\mu}_N)$ defined by

$$\mu_N(t) := \frac{1}{N} \sum_{i=1}^N w_i^N \delta_{X_i(t)}, \quad \tilde{\mu}_N(t) := \frac{1}{N} \sum_{i=1}^N \tilde{w}_i^N \delta_{X_i(t)}, \tag{1.20}$$

are tight in $C([0,T],\mathcal{M}(\mathbb{R}^d))^{\otimes 2}$ *and every accumulation point is a solution to* (1.16) *with initial value* (v_0, g_0).

Theorem 1.13 gives the mean-field limits of interacting diffusions when the kernel is singular and the system is non-exchangeable at the same time. This result extends classical results on mean-field limits for singular interacting diffusions to non-exchangeable cases, additionally, the weights for the interaction can be unbounded, i.e., $r < \infty$. The system (1.12) with the Biot-Savart kernel and bounded weights has been studied in [FHM14, Wyn21]. However, the result in [Wyn21] only applies to the so-called two pieces interaction, which is basically $w_i^N = a_1 > 0$ when i is odd and $w_i^N = a_2 < 0$ when i is even. Under that restrictive condition, the pairs of particles (X_{2i-1}, X_{2i}) are exchangeable. We also mention that the result in [FHM14] falls into the category of exchangeable $N-$particle systems if we treat $Z_i = (X_i, w_i^N)$ as a single particle in the extended phase space, where X_i and w_i^N denote the position and the magnitude of the $i-$th point vortex, respectively.

As particles/agents in applications are not always identical, more natural assumptions on weights are required. Let us now look at a specific example. Let $K(X_i - X_j)$ be the interaction in N-particles system described by the system of SDEs (1.12). The interaction could be viewed as a function of how X_i is influenced by X_j, with w_j^N representing the intensity associated with X_j. Let the weights be $w^{5N} = (w_1^N, ..., w_N^N, 0, ..., 0)$ with $(w_1^N, ..., w_N^N)$ satisfying

$$|w_i^N| = O(N^{\frac{1}{3}}), \quad \forall 1 \leqslant i \leqslant N^{\frac{1}{3}}; \qquad |w_i^N| = O(1), \quad N^{\frac{1}{3}} < i \leqslant N.$$

Now the model is

$$\begin{cases} \mathrm{d}Y_i = & \frac{1}{5N}\sum_{j\neq i} w_j^N K(Y_i - Y_j)\,\mathrm{d}t + \sqrt{2}\mathrm{d}B_i, \quad i = 1, \dots, N, \\ \\ \mathrm{d}Z_m = & \frac{1}{5N}\sum_{j=1}^{N} w_j^N K(Z_m - Y_j)\,\mathrm{d}t + \sqrt{2}\mathrm{d}B_m, \quad m = 1, \dots, 4N. \end{cases}.$$

Here (Y^N, Z^{4N}) plays the role of X^N in (1.12) and $(\mathbf{W_r})$ holds for $r = 2$. Note that every particle only interacts with the particles of the type Y, in other words, only the particles (Y_i) contribute to the dynamics of the system. Furthermore, there are still differences among the particles of type Y. The particles (Y_i) belongs to two groups, the majority of the number $N - N^{\frac{1}{3}}$ and the minority of the number $N^{\frac{1}{3}}$. Each particle in the majority contributes to the system in a normal way that $|w_j^N| = O(1)$. In contrast, those particles from the minority make significant contributions, at the scale of $N^{\frac{1}{3}}$.

As mentioned earlier, the mean-field convergence of $(\mu_N, \tilde{\mu}_N)$ is much more general than the convergence of (μ_N), since it gives the convergence for any possible weights $(\tilde{w}_j^N)$ in l^r. Formally, one may think that our result fully recovers the linear statistical information of $(X_1, ..., X_N)$ rather than just the average statistics $\frac{1}{N}\sum_{i=1}^N \varphi(X_i)$ or the specific weighted one $\frac{1}{N}\sum_{i=1}^N w_i^N \varphi(X_i)$. In particular, when K is the Biot-Savart kernel, let $u_t = K * v_t$ denote the velocity of the fluid, then Theorem 1.13 gives a mean-field approximation to the dynamics of a passive tracer undergoing advection-diffusion in the fluid described by the Navier-Stokes equation,

$$\begin{cases} \partial_t u_t = \Delta u_t - u_t \cdot \nabla u_t + \nabla P, \quad \mathrm{div} u = 0, \\ \\ \partial_t g_t = \Delta g_t - u_t \cdot \nabla g_t, \end{cases} \tag{1.21}$$

with P being the associated pressure. For more information on the physics behind passive scalars, we refer to [FGV01, SS00, War00] and references therein, and for mathematical studies, see e.g. [ACM19, BBPS22a, BBPS22b, Sei13, ZDE20].

With additional constraints on the kernel and weights, we can show the coincidence of the limiting points of converging subsequences and thus obtain the convergence of the entire sequence.

Theorem 1.14. *If the kernel K belongs to either of the following two cases,*

1. *K is the 2D Biot-Savart kernel and $r \in [3, \infty]$ (since the definition of entropy solutions depends on r).*
2. *Given $r \in (1, \infty]$, and K belongs to $L^{q_2}([0,T], L^{p_2}(\mathbb{R}^d))$ with*

$$\frac{d}{p_2} + \frac{2}{q_2} + \frac{1}{r} \leqslant 1, \quad \frac{d}{p_2} + \frac{2}{q_2} < 1. \tag{1.22}$$

Then there exists a unique solution (v, g) to (1.16) for each given initial value from $L^1(\mathbb{R}^d)^{\otimes 2}$.

Corollary 1.15. *Given two sequences of weights $\{w^N\}$ and $\{\tilde{w}^N\}$ satisfying the condition $(\mathbb{W}_r)$ with $r \in [3, \infty]$. Let X^N be an entropy solution to the stochastic vortex model (i.e. Eq. (1.12) with K the Biot-Savart kernel) as given by Proposition 1.10. Assume that there exist $v_0, g_0 \in L^1(\mathbb{R}^d)$ such that*

$$\frac{1}{N} \sum_{i=1}^{N} w_i^N \delta_{X_i(0)} \rightharpoonup v_0, \quad \frac{1}{N} \sum_{i=1}^{N} \tilde{w}_i^N \delta_{X_i(0)} \rightharpoonup g_0$$

in $\mathcal{M}(\mathbb{R}^2)$ almost surely. Then $(\mu_N, \tilde{\mu}_N)$ defined in Theorem 1.13 converges in law to (v, g) in $C([0,T], \mathcal{M}(\mathbb{R}^2))^{\otimes 2}$. Here (v, g) uniquely solves (1.16) in the sense of Definition 1.11. In particular, $(\nabla^\perp(-\Delta)^{-1} v, g)$ solves the system (1.21) of the passive scalar advected by the 2D Navier-Stokes equation.

Finally, simply by choosing the same weight sequences $\tilde{w}_j^N = w_j^N$ as the intensities of the $j-$th point vortex, one arrives at the following theorem.

Theorem 1.16 (Mean field limit for the stochastic vortex model with general intensities)**.** *Given a sequence of intensities $w^N = (w_j^N)_{1 \leqslant j \leqslant N}$ that satisfies the condition $(\mathbb{W}_r)$ with $r \in [3, \infty]$. Let X^N be an entropy solution to the stochastic vortex model (1.12) with K the Biot-Savart kernel. For the initial data v_0 for the 2D Navier-Stokes equation, assume that (1.15) $v_0 \in L^1(\mathbb{R}^d)$ and*

$$\frac{1}{N} \sum_{i=1}^{N} w_i^N \delta_{X_i(0)} \rightharpoonup v_0,$$

in $\mathcal{M}(\mathbb{R}^2)$ almost surely. Then the empirical measure $\mu_N = \frac{1}{N} \sum_{i=1}^{N} w_i^N \delta_{X_i}$ converges in law to v in $C([0,T], \mathcal{M}(\mathbb{R}^2))$, where v is the unique solution to (1.16) in the sense of Definition 1.11.

1.3 Graphon particle systems

In this section, we consider graphon particle systems on $\mathbb{R}^d$, which are a generalization of (1.12). More precisely, the position of each particle X_i is characterized by the following SDE

$$\mathrm{d}X_i = \frac{1}{N}\sum_{j\neq i} w_{ij}^N K(X_i - X_j)\,\mathrm{d}t + \sqrt{2}\mathrm{d}B_i, \quad i = 1,\ldots,N, \tag{1.23}$$

where $\{w_{ij}^N\}$ are non-identical deterministic weights satisfying the following condition for some $r > 1$

$$\begin{aligned}(\mathbf{W_r'}):\quad & w_{ij}^N = w_{ji}^N, \quad \max_{1\leqslant i\leqslant N}\frac{1}{N}\sum_{j=1}^N |w_{ij}^N|^r = O(1), \quad \text{for } r<\infty,\\ & w_{ij}^N = w_{ji}^N, \quad \max_{1\leqslant i,j\leqslant N}|w_{ij}^N| = O(1), \quad \text{for } r=\infty, \quad \text{as } N\to\infty.\end{aligned} \tag{1.24}$$

The weights $\{w_{ij}^N\}$ can be regarded as the values of the wedges for a graph with N vertices, and this view turns out to be crucial for studying the asymptotic behavior of the particle system as $N \to \infty$. Roughly speaking, in order to compare the interacting particle system (1.23) with its mean-field equation, we also need to study the asymptotic behavior of the variables $\{w_{ij}^N\}$. For the case $w_{ij}^N = w_j^N$ in Section 1.2, we have directly studied weighted empirical measures. Thus deriving the mean-field equation for (1.23) involves the theory of graph limits, c.f. [Lov12, BCCZ]. The objects in graph theory associated to the weights are called graphons: symmetric real-valued measurable functions on $[0,1]\times[0,1]$. Define the graphon G_N as follows

$$G_N(\xi,\zeta) := \sum_{1\leqslant i,j\leqslant N} w_{ij}^N 1_{[\frac{i-1}{N},\frac{i}{N})}(\xi) 1_{[\frac{j-1}{N},\frac{j}{N})}(\zeta), \quad \xi,\zeta\in[0,1],$$

where $w_{ii}^N := 0$ for all i. Then, by Ito's formula, the sequence of empirical measures

$$\nu_N(\mathrm{d}x,\xi)(t) := \sum_{i=1}^N \delta_{X_i(t)}(\mathrm{d}x) 1_{[\frac{i-1}{N},\frac{i}{N})}(\xi)$$

formally converge to f solving

$$\begin{cases}\partial_t f(t,x,\xi) = \Delta_x f(t,x,\xi) - \nabla_x\cdot\Big(\Big[K*(G\circledast f)\Big](t,x,\xi)f(t,x,\xi)\Big),\\ \\ G\circledast f(t,x,\xi) := \int_0^1 G(\xi,\zeta)f(t,x,\zeta)\mathrm{d}\zeta,\end{cases} \tag{1.25}$$

where G is a limiting point of $\{G_N, N\in\mathbb{N}\}$ in a suitable topology. Therefore, (1.25) is the desired macroscopic model.

In this section, we still assume condition (**H**) on the initial value. For the kernels, we assume that

($\mathbf{K_r'}$) Given $r\in(1,\infty]$ from ($\mathbf{W_r'}$) and $d\geqslant 2$, $K\in L^{q_2}([0,T],L^{p_2}(\mathbb{R}^d))$ with $\frac{d}{p_2}+\frac{2}{q_2}+\frac{1}{r}<1$, where the equality is allowed be attained when $q_2, r<\infty$.

Remark 1.17. *This section is devoted to study generalized models with w_j^N in* (1.12) *replaced by w_{ij}^N, and the method and the setting for obtaining mean-field convergences of subsequences actually apply to all the kernels satisfying* ($\mathbf{K_r}$). *We focus on the kernels satisfying* ($\mathbf{K'_r}$) *for the completeness of results, such as the well-posedness of the particle system* (1.23) *and the mean-field equation* (1.25). *To the author's best knowledge, when K is the Biot-Savart kernel, the model in Section 1.2 is the most general version from physics, so in this part we do not study the Graphon system with the Biot-Savart kernel.*

We still use the notation of entropy solutions as in Definition 1.9, but replacing the condition ($\mathbf{K_r}$) by ($\mathbf{K'_r}$). The following result gives existence and uniqueness in law of entropy solutions,

Proposition 1.18. *Under the conditions* ($\mathbf{H}$), ($\mathbf{K'_r}$), *and* ($\mathbf{W'_r}$) *for some $r \in (1, \infty]$, for each $N \in \mathbb{N}$ there exists a unique (in law) entropy solution X^N to the model* (1.23). *Furthermore, the entropy dissipation inequality* (1.17) *holds with some universal constant C that is independent of N.*

For the limiting equation (1.25), we define solutions in the distributional sense.

Definition 1.19. *We call $f \in C([0,T], \mathcal{P}(\mathbb{R}^d \times [0,1]))$ is a solution to* (1.25) *with the graphon G and initial value $f_0 \in \mathcal{P}(\mathbb{R}^d \times [0,1])$ if f satisfies the equation in the sense that, for any $\varphi \in \mathcal{S}(\mathbb{R}^d \times [0,1])$ and $0 \leqslant t \leqslant T$,*

$$\begin{aligned}&\int_{\mathbb{R}^d}\int_0^1 \varphi(x,\xi) f(t,x,\xi)\mathrm{d}x\mathrm{d}\xi - \int_{\mathbb{R}^d}\int_0^1 \varphi(x,\xi) f(0,x,\xi)\mathrm{d}x\mathrm{d}\xi \\ =&\int_0^t\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_0^1\int_0^1 G(\xi,\zeta)\nabla_x\varphi(x,\xi)K(x-y)f(y,\zeta)f(x,\xi)\mathrm{d}\zeta\mathrm{d}\xi\mathrm{d}x\mathrm{d}y\mathrm{d}s\end{aligned}$$

and enjoys finite Fisher information,

$$\int_0^T\int_0^1 I(f(t,\cdot,\xi))\mathrm{d}\xi\mathrm{d}t < \infty.$$

The well-posedness of the equations (1.25) shall be studied in Section 4.4.3 below.

Now we state the first main result, which gives the mean-field convergence of graphon particle systems.

Theorem 1.20. *Let $\{w^N\} := \{(w_{i,j}^N)\}$ be a sequence of weights satisfying the condition* ($\mathbf{W'_r}$) *with $r \in (1,\infty]$ and suppose the conditions* ($\mathbf{H}$) *and* ($\mathbf{K'_r}$) *hold. Let X^N be the unique entropy solution to* (1.23), *which is obtained in Proposition 1.18.*

Then, there exist f solving equations (1.25) *in the sense of Definition 1.19 with a graphon G almost surely, and a subsequence, without relabelling, of (μ_N, G_N), where $\mu_N(\cdot) := \frac{1}{N}\sum_{i=1}^N \delta_{X_i(\cdot)}$, such that μ_N converges to $\int_0^1 f(\cdot,\xi)\mathrm{d}\xi$ in $C([0,T],\mathcal{P}(\mathbb{R}^d))$ almost surely and $\{G_N\}$ converges to G with respect to the cut metric (defined in Section 4.1).*

Theorem 1.20 significantly extends existing results on the mean-field problem of graphon particle systems to cases with general singular kernels. The key point is to use the regularization from the stochastic noise, specifically the Fisher information estimate, as in Section 1.2. The novelty of our work is to show the uniform Fisher information estimate and propagate the Fisher information.

Concerning graphons, we will need the theory of graph limits. To our case, we will use the result from [BCCZ]. However, the condition $(\mathbf{W}'_{\mathbf{r}})$ only ensures the compactness of graphons $\{G_N\}$ with respect to the cut metric, we must impose additional assumptions on $\{G_N\}$ to ensure convergence of the whole sequence. This leads us to the following result.

Theorem 1.21. *Besides the assumptions in Theorem 1.19, assume further that there exist a graphon G such that G_N converges to G with respect to the cut norm, and $f_0 \in \mathcal{P}(\mathbb{R}^d \times [0,1])$ such that*

$$\nu_N(0) \rightharpoonup f_0$$

in $\mathcal{P}(\mathbb{R}^d \times [0,1])$ almost surely. Then the whole sequence $\{\nu_N\}$ weakly converges to f in $C([0,T], \mathcal{P}(\mathbb{R}^d \times [0,1]))$, where f uniquely solves the mean-field equation (1.25) *with G and initial value f_0 in the sense of Definition 1.19.*

This result shows direct convergence from graphon particle systems to the mean-field equation (1.23), which is new to the best of our knowledge and provides more information than the convergence of μ_N. Beyond the setting of the mean-field convergence for systems weighted on dense graphs, the asymptotic behavior of interacting diffusions on sparse graphs, which involves strong interactions, is another related active topic. The local weak convergence for the sparse case has been obtained in [ORS20, LRW19] recently.

1.4 Methodology and difficulties

In this section, we address the challenges encountered in obtaining our results on fluctuations for interacting diffusions, and on the mean-field problem for non-exchangeable inhomogeneous interacting diffusions.

1.4.1 On Gaussian fluctuations

In the literature, there are mainly two type of results for the fluctuations of interacting processes, either in the path space or in the time marginals. This thesis focuses on the latter for bounded kernels and some singular ones. When studying fluctuations in the path space, one treats processes $\{X_i\}$ as random variables with values in some functional space. For instance the fluctuation measures may be defined as $\sqrt{N}\Big(\frac{1}{N}\sum_{i=1}^N \delta_{X_i} - \mathcal{L}(X)\Big)$, with the process $X \in C([0,T], \mathbb{R}^d)$ solving the nonlinear stochastic differential equation

$$X(t) = X(0) + \int_0^t \int_{\mathbb{R}^d} K(X(s) - x)\mathrm{d}\mu_s(x) + \sqrt{2\sigma}B_t, \quad \text{with } \mu_s = \mathcal{L}(X(s)).$$

To the best of our knowledge, a fluctuation in path space type result was firstly obtained by Tanaka and Hitsuda [TH81] and by Tanaka [Tan84] for interacting diffusions. They proved that the fluctuation measures on the path space converges to a Gaussian random field when the interacting kernels are bounded and Lipschitz continuous on $\mathbb{R}^1$, using differentiable test functions on the path space. Later Sznitmann [Szn84] removed the differentiability condition on the test functions and generalized the result to $\mathbb{R}^d$, using Girsanov's formula and the method of U-statistics.

The article by Fernandez and Méléard [FM97] is probably the one closest to our result when it comes to the basic setting, where they studied interacting diffusions with regular enough coefficients, using the so-called Hilbertian approach. Their result cannot cover kernels which are only bounded or even singular. The systems they consider are on the whole space and allow multiplicative independent noises. It is worth emphasizing that the Hilbertian approach introduced in [FM97] has been amplified to study various interacting models, see [JM98, Che17, CF16, LS16] etc. The Hilbertian approach is based on the martingale method (as used in this thesis and many other stochastic problems), the coupling method, and analysis in negative weighted Sobolev spaces. The coupling method, which is based on directly comparing the $N-$ particle system (2.1) and $N-$copies of the limit McKean-Vlasov equation, is also widely used in classical propagation of chaos results [Szn], but usually requires strong assumptions on the interacting kernels and diffusion coefficients.

In contrast to that, our new method enables us to obtain uniform estimates and hence convergence results through directly comparing the Liouville equation and the limit mean-field equation. The main result (Theorem 1.4) follows by the martingale approach, which has also been used to study the fluctuation problem of interacting diffusions with regular kernels as in [Mél96, FM97]. The proof consists of three steps: tightness, identifying the limits of converging subsequences, and well-posedness of the SPDE (1.7). By Itô's formula, we have

$$\begin{aligned} \mathrm{d}\langle\eta_t^N,\varphi\rangle =&\langle\sigma\Delta\varphi,\eta_t^N\rangle\mathrm{d}t+\mathcal{K}_t^N(\varphi)\mathrm{d}t+\langle\nabla\varphi,V\eta_t^N\rangle\mathrm{d}t+\frac{\sqrt{2\sigma_N}}{\sqrt{N}}\sum_{i=1}^N\nabla\varphi(X_i)\mathrm{d}B_t^i\\ &+\sqrt{N}(\sigma_N-\sigma)\langle\Delta\varphi,\mu_N(t)\rangle\mathrm{d}t, \end{aligned}\tag{1.26}$$

$\mathbb{P}$-a.s. for each $\varphi\in C^\infty(\mathbb{T}^d)$. Here the interacting term $\mathcal{K}_t^N: C^\infty(\mathbb{T}^d)\to\mathbb{R}$ is defined by

$$\mathcal{K}_t^N(\varphi)=\sqrt{N}\langle\nabla\varphi,K*\mu_N(t)\mu_N(t)\rangle-\sqrt{N}\langle\nabla\varphi,v_tK*v_t\rangle.\tag{1.27}$$

To show the tightness of η^N, we need to derive some uniform estimates for η^N in (1.26). However, due to the singularity of the kernels K in Assumption **(A2)**, it seems challenging to directly obtain uniform estimates for terms involving η^N in negative Sobolev spaces. In fact, the optimal regularity for the limit η obtained in Section 2.4.1 is in $C_TC^{-\alpha}$ with $\alpha>d/2$. It is natural to consider energy estimates for η^N in $H^{-\alpha}$ using (1.26). For the purpose of illustration, let us assume that $\sigma_N=\sigma\equiv 0$, and $V=0$, so we can rewrite (1.26) in the following form

$$\partial_t\eta^N+\mathrm{div}(\mu_NK*\eta^N)+\mathrm{div}(\eta^NK*\bar\rho)=0.$$

To control nonlinear terms appearing in the time evolution $\frac{\mathrm{d}}{\mathrm{d}t}\langle\eta^N,\eta^N\rangle_{H^{-\alpha}}$, such as $\langle\nabla\eta^N,K*\mu_N\eta^N\rangle_{H^{-\alpha}}$, we need $K\in C^\beta$ with $\beta>d/2$ by the multiplicative inequality in Section 2.1.1, which is much more demanding than the assumptions we made on our kernels K.

We overcome this difficulty caused by the singularity of interaction kernels by using the Donsker-Varadhan variational formula [DE11, Proposition 4.5.1] (see (2.5) below) and two large deviation type estimates, one of which is from [JW18, Theorem 4] and the other is Lemma 2.7 below. More precisely, now the uniform estimate for the fluctuation measures can be controlled by two terms, one of which is the relative entropy $H(F^N|v^{\otimes N})$ and the other is some exponential integral with a product reference measure $v^{\otimes N}$ (see (2.6)). On one hand, the uniform bound on $H(F^N|v^{\otimes N})$, as summarized in Assumption **(A3)**, has already been established by Jabin and Wang in [JW18] for a large family of interaction kernels, in particular including those specified in **(A2)**. On the other hand, exploiting cancellation properties from the interaction terms enables us to obtain a uniform bound for the exponential integrals (see Lemma 2.6 and Lemma 2.7 for details). This large deviation type estimate enables the authors of [JW18] to conclude quantitative estimates for the propagation of chaos.

Recalling the decomposition (1.26), we also need to estimate the martingale part and show its convergence as well. We shall find a pathwise realization $\mathcal{M}^N$ of the martingale part (see Section 2.1.3) and then establish its tightness. The tightness of laws of the fluctuation measures then follows by applying the Arzela-Ascoli theorem.

To identify the limits of converging subsequences, the difficulty still comes from the singularity of kernels. For the illustrating example where $\sigma_N \equiv 0$ and $V = 0$, we note that it has the following representation

$$\partial_t \eta^N + \operatorname{div}(\eta^N K * v) + \operatorname{div}(vK * \bar{\eta}^N) + \frac{1}{\sqrt{N}} \operatorname{div}(\eta^N K * \eta^N) = 0, \tag{1.28}$$

More precisely, the convergence of the interaction term $\mathcal{K}_t^N(\varphi)$ cannot be directly deduced from the convergence of μ_N and η^N. We note that the interaction term can be split into two terms. One term is a continuous function of η^N, namely $vK * \eta^N + \eta^N K * v$, which converges as N goes to infinity. The other term is of the form $\frac{1}{\sqrt{N}}\eta^N K * \eta^N$, which is not easy to handle directly since the formal limit $K * \eta\eta$ is not well-defined (see Lemma 2.2) in the classical sense due to the singularity of K. Instead, we obtain a uniform bound of this singular term by using the variational formula trick again (see Lemma 2.14 below). The remaining part for identifying the limits is classical.

The last step to Theorem 1.4 is to prove the uniqueness of martingale solutions to the SPDE (1.7), which follows by pathwise uniqueness (see Lemma 2.24) and Yamada-Watanabe theorem. Proposition 1.5 is obtained by solving the dual backward equation of (1.7) without noises, which gives the Gaussianity of the limit process of the fluctuation measures.

For the case with vanishing diffusion (which includes the purely deterministic dynamics with $\sigma_N \equiv 0$), the only difference concerns the well-posedness of the limit equation (1.7), which is a first order PDE. The well-posedness follows from the method of characteristics. Since now the limit equation is deterministic, by a useful lemma in [GK96] by Gyöngy and Krylov (see Lemma 2.37 below) we obatin the convergence in probability of the fluctuation measures.

1.4.2 On the mean-field problem for inhomogeneous interacting diffusions

The key feature of interacting diffusions with inhomogeneous interaction (c.f. (1.12), (1.23)) is the lack of exchangeability, which is crucial for scaling limits of interacting diffusions, particularly those with singular interactions. The difficulty is to compare interacting particle systems with mean-field limits in the absence of exchangeability. As mentioned before, the results in [BCW20, JPS21] can be applied to ineracting ODEs/SDEs weighted on graphs, and the symmetry of w_{ij}^N is not required in [JPS21]. In both papers the coupling method and graph limit theory are used. The basic idea in the proof of [BCW20] is to compare the SDEs of the particles and the SDEs of the limiting system, using the convergence of graphons (referring to [Lov12, BCCZ]). In contrast to that, the coupling method was used in [JPS21] to obtain a coupled PDE of McKean-Vlasov type by propagation of independence first. Then a class of new observables was constructed through the combination of weights and laws of independent particles via a family of labeled trees. The authors then transformed the problem into the Vlasov/mean-field hierarchy likewise. Instead of directly using the convergence of graphons, a similar version of Szemerédi's regularity lemma was established in [JPS21, Lemma 4.7] for non-exchangeable systems. However, the coupling method leads to the bounded and Lipschitz continuity restriction on the interactions K. We do not know yet how to extend those methods like the relative entropy method in [JW18] and the modulated energy method in [Ser20], which have been shown powerful for exchangeable systems, to non-exchangeable ones. Indeed, it would be interesting to combine the relative entropy/energy method with the limiting graphon structures of the weights w_{ij}^N.

Our idea of the proof is to apply a compactness/tightness argument, consisting of the classical three steps: the tightness, characterizing the limits, and the uniqueness of the limit equation. For particle systems with singular interactions, the regularity of the joint law is crucial to compensate the singularity. The article by Fournier, Hauray, and Mischler [FHM14] is probably the one closest to our approach regarding tackling the singularity, where they also heavily exploited the Fisher information of the joint laws of the particles. Many features of the Fisher information will also be used in our proof, for instance sub-additivity, the chain rule, and notably the Sobolev regularity estimates following from it. Unfortunately, it is not apparent how to apply the compactness argument, since the non-exchangeability and the singularity cause two difficulties. The first one is to derive uniform estimates (about the Fisher information in our case) for the non-exchangeable system (1.12). For instance, when K is the Biot-Savart kernel, the following frequently used inequality requires exchangeability,

$$\mathbb{E}\,|K(X_i - X_j)| \lesssim 1 + I(\mathcal{L}(X_i, X_j)) \lesssim 1 + \frac{2}{N} I(\mathcal{L}(X^N)),$$

where $I(\mu) : \mathcal{P}(\mathbb{R}^{dk}) \to [0, +\infty]$, $k \in \mathbb{N}$, denotes the Fisher information functional to be defined in Section 3.1. This indicates that the interaction between any two particles can be controlled by $\frac{1}{N}$ of the total Fisher information of the system. However, when the particles are *not* indistinguishable from each other, the joint law is no longer symmetric, and there might be a pair of particles such that $I(\mathcal{L}(X_i, X_j)) > \frac{2}{N} I(\mathcal{L}(X^N))$. To overcome this difficulty, we prove a technical lemma (Lemma 3.4) concerning the average of the interactions, which allows us to derive uniform Fisher information estimates for non-exchangeable systems and to estimate singular interactions. Investigating averaging statistics is a key step to study non-exchangeable systems. Similarly, observables with averaged information of particles also play a crucial role in [JPS21]. Note that the presence of noise is crucial in our analysis since we effectively use the control given by the Fisher information. In the deterministic setting, for instance in the vortex approximation for the 2D incompressible Euler equation, since now the w_j^N in general can be distinct from each other, the symmetrization trick used in [JW18] does not work anymore.

The other major difficulty is to show the regularity of the limiting points of the empirical measures $\{\mu_N, N \in \mathbb{N}\}$. This is closely related to the exchangeability. With the compactness argument, one usually finds that μ_N converges to μ in the distributional sense. To make the mean-field equation with singular coefficients well-defined and show the convergence of the nonlinear interaction term, one must propagate the regularities. Clearly, the empirical measure μ_N enjoys no regularity at all. Again, this difficulty is not problematic for the symmetric case, since there are well-established tools based on the famous DeFinetti–Hewitt–Savage theorem [DF37, HS55] to propagate the Fisher information, c.f. [HM14] and [FHM14]. For the non-exchangeable case, we introduce a sequence of random measures (defined in (3.23) and (4.13)) constructed via disintegration of the joint laws $\{F^N, N \in \mathbb{N}\}$. This sequence of random measures would merge with the sequence of empirical measures as N goes to infinity, thus playing a similar role as the 1-marginal distributions in the symmetric case. An analogous construction can be found in the proof of the Laplace principle via the weak convergence method, see e.g. [DE11, Section 2.5]. The difference is that the proof in [DE11] studies the relative entropy while we focus on the Fisher information functional of probability measures. In the end, uniform Sobolev regularity estimates for the random measures are obtained in Section 3.3, and the uniform bound for the Fisher information is obtained in Section 4.3.2. Consequently, we obtain the required regularity of the limiting points.

The remainder of the compactness argument is almost standard. It is worth mentioning here that the proof of Theorem 1.14 for the Biot-Savart kernel relies on the uniqueness result for the 2D Navier-Stokes equation in [FHM14], which is based on [BA94, Bre94].

1.5 Structure of the thesis

This thesis is organized as follows:

In Chapter 2, we study the Gaussian fluctuations of interacting diffusions. In Section 2.1, we quote some auxiliary results on Besov spaces and concentration estimates from statistics. We also state pathwise realizations of stochastic integrals in negative Sobolev spaces. Section 2.2 is devoted to obtaining three main estimates which are based on the variational formula and concentration estimates, including uniform estimates on terms related to η^N, $\mathcal{K}^N$, and a singular term derived from $\mathcal{K}^N(\varphi)$. The proof of Theorem 1.4 is completed in Section 2.3. In Section 2.4, we show the optimal regularity of solutions to the SPDE (1.7) and prove Proposition 1.5. Section 2.5 is concerned with the case with vanishing diffusion and the proof of Theorem 1.6. Section 2.6 focuses on some examples which fulfill assumptions **(A1)-(A5)**, including the point vortex model approximating the vorticity formulations of the 2D Navier-Stokes/Euler equation on the torus. This part is based on my work [WZZ21] joint with Prof. Zhenfu Wang and Prof. Rongchan Zhu.

In Chapter 3, we investigate the mean-field problem of interacting diffusions with weighted interactions, which are non-exchangeable. We shall state the notations and auxiliary estimates related to the Fisher information in Section 3.1. Section 3.2 is devoted to obtaining the main estimate in this chapter, which gives a uniform estimate on the Fisher information of the joint laws of N-particles. The proof is based on an averaging result for the Fisher information when the probability measure is asymmetric. In Section 3.3, we study a sequence of random measures, which turns out to be close to the sequence of weighted empirical measures and enjoys certain Sobolev regularity estimates uniformly. Lastly, we finish the proofs of Theorem 1.13 and Theorem 1.14 by the compactness argument in Section 3.4. This part is based on my work [WZZ22] joint with Prof. Zhenfu Wang and Prof. Rongchan Zhu.

In Chapter 4, we study models with more general weights than those in Chapter 3, but in a slightly different setting. More precisely, we study the mean-field problem of graphon particle systems. In Section 4.1, we state the notation, topologies, and results about graphons. Section 4.2 is devoted to show existence and uniqueness in law of solutions to graphon particle systems with a class of singular kernels. In Section 4.3, we study the sequence of auxiliary random measures, which are shown to weakly merge with the empirical measures and enjoy uniform Fisher information. Lastly, we prove Theorem 1.20 and Theorem 1.21 in Section 4.4.

1.6 Notations

Throughout the thesis, we use the notation $a \lesssim b$ if there exists a universal constant $C > 0$ such that $a \leqslant Cb$. During the computations, the universal constant may change from line to line. We will point out the dependence on the parameters when it is necessary.

Recall that we have used the notations: $X^N := (X_1, ...X_N)$, $x^N := (x_1, ...x_N)$, $w^N := (w_1^N, ..., w_N^N)$, and $\langle x\rangle := (1+|x|^2)^{\frac{1}{2}}$. The γ-th moment, $\gamma > 0$, of a positive measure μ on $\mathbb{R}^d$ is represented by $\int_{\mathbb{R}^d}\langle x\rangle^\gamma \mu(\mathrm{d}x)$. We shall use $\{e_k\}_{k\in\mathbb{Z}^d}$ to represent the Fourier basis on $\mathbb{T}^d$, i.e. $e_k(x) = e^{\sqrt{-1}k\cdot x}$. For simplicity, we define $\langle k\rangle := \sqrt{1+|k|^2}$.

We will mostly work on Sobolev spaces, Besov spaces, and the space of k-differentiable functions.

Concerning function spaces on the torus, the norm of the Sobolev space $H^\alpha(\mathbb{T}^d)$, $\alpha \in \mathbb{R}$, is defined by

$$\|f\|_{H^\alpha}^2 := \sum_{k\in\mathbb{Z}^d} \langle k\rangle^{2\alpha} |\langle f, e_k\rangle|^2,$$

with the inner product $\langle\cdot,\cdot\rangle_{H^\alpha}$. Moreover, we also use the bracket $\langle\cdot,\cdot\rangle$ to denote integrals when the space and underlying measure are clear from the context. The precise definition and some basic properties of Besov spaces on the torus $B^\alpha_{p,q}(\mathbb{T}^d)$ with $\alpha \in \mathbb{R}$ and $1 \leqslant p, q \leqslant \infty$, will be given in the Section 2.1.1 for completeness. We remark that $B^\alpha_{2,2}(\mathbb{T}^d)$ coincides with Sobolev space $H^\alpha(\mathbb{T}^d)$. We say $f \in C^\alpha(\mathbb{T}^d)$, $\alpha \in \mathbb{N}$, if f is α-times differentiable. For $\alpha \in \mathbb{R}\setminus\mathbb{N}$, $C^\alpha(\mathbb{T}^d)$ is indicated to $B^\alpha_{\infty,\infty}(\mathbb{T}^d)$. We will often write $\|\cdot\|_{C^\alpha}$ instead of $\|\cdot\|_{B^\alpha_{\infty,\infty}}$. In the case $\alpha \in \mathbb{R}^+\setminus\mathbb{N}$, $C^\alpha(\mathbb{T}^d)$ coincides with the usual Hölder space. We use $C^\infty(\mathbb{T}^d)$ to denote the space of infinitely differentiable functions on $\mathbb{T}^d$ and $\mathcal{S}'(\mathbb{T}^d)$ to denote the space of distributions. For simplicity, we may omit the underlying space $\mathbb{T}^d$ in Chapter 2 without causing confusions.

For spaces on $\mathbb{R}^d$, as usual, $C_0(\mathbb{R}^d)$ stands for the space of continuous functions vanishing at infinity, and $C_0^\infty(\mathbb{R}^d)$ stands for the space of smooth functions vanishing at infinity. Let $\mathcal{S}(\mathbb{R}^d)$ be the Schwartz space and $\mathcal{S}'(\mathbb{R}^d)$ be the space of tempered distributions. We also use $C_b^k(\mathbb{R}^d)$ to denote the space of bounded continuous functions with bounded k-th derivative.

Given a Banach space E with a norm $\|\cdot\|_E$ and $T > 0$, we write $C_TE = C([0,T];E)$ to denote for the space of continuous functions from $[0,T]$ to E, equipped with the supremum norm $\|f\|_{C_TE} = \sup_{t\in[0,T]}\|f(t)\|_E$. For $p \in [1,\infty]$ we write $L^p_TE = L^p([0,T];E)$ for the space of L^p-integrable functions from $[0,T]$ to E, equipped with the usual L^p-norm. We use $\|\cdot\|_{L^p_q}$ to denote the $L^q([0,T], L^p(\mathbb{R}^d))$-norm. For notational simplicity we shall not distinguish the space and the norm for the vector valued functions and the scalar valued functions.

For the sequence of weights $w^N = (w_j^N)$, we define

$$\|w^N\|_{l^r} := \left(\frac{1}{N}\sum_{i=1}^N |w_j^N|^r\right)^{\frac{1}{r}},$$

for $r \in [1, \infty)$, and $\|w^N\|_{l^\infty} := \max_{1 \leqslant j \leqslant N} |w_j^N|$. Obviously, $\|w^N\|_{l^{r_1}} \leqslant \|w^N\|_{l^{r_2}}$ when $r_1 \leqslant r_2$. When the class of weights w^N is of the form w_{ij}^N, we define

$$\|w^N\|_{l^r} := \left(\frac{1}{N^2} \sum_{i \neq j}^{N} |w_{ij}^N|^r \right)^{\frac{1}{r}}, \quad \|w^N\|_{l^\infty l^r} := \max_{1 \leqslant i \leqslant N} \left(\frac{1}{N} \sum_{j \neq i}^{N} |w_{ij}^N|^r \right)^{\frac{1}{r}},$$

for $r \in [1, \infty]$. The properties of graphons and related topologies are recalled in Section 4.1.

We use $\mathcal{P}(\mathbb{R}^d)$ to denote the space of probability measures on $\mathbb{R}^d$ and for a given topological space $\mathcal{X}$ we use $\mathscr{B}(\mathcal{X})$ to denote the Borel σ-algebra on $\mathcal{X}$ and use $\mathcal{M}(\mathcal{X})$ to denote the space of finite signed measures on $\mathcal{B}(X)$ endowed with the weak topology induced by all bounded, continuous functions on $\mathcal{X}$. Given $\mu \in \mathcal{M}(\mathcal{X})$, its absolute value is denoted by $|\mu|$, i.e. $|\mu| := \mu^+ + \mu^-$. We use $\| \cdot \|_{TV}$ to denote the total variation norm of elements in $\mathcal{M}(\mathcal{X})$. The notation $L_w^p(\mathcal{X})$, $p \geqslant 1$, denotes the $L^p(\mathcal{X})$ space endowed with the weak topology induced by its dual space.

Chapter 2

Gaussian fluctuations for interacting diffusions

In this chapter, we consider the interacting particles system described by SDEs (1.3). We state the model below for convenience,

$$\mathrm{d}X_i = \frac{1}{N}\sum_{j\neq i} K(X_i - X_j)\mathrm{d}t + V(X_i)\mathrm{d}t + \sqrt{2\sigma_N}\mathrm{d}B_i, \quad i = 1, ..., N. \tag{2.1}$$

with $\mathcal{F}_0$-measurable random initial data $\{X_i(0)\}_{i=1}^N \subset \mathbb{T}^d$. The collection $\{B^i_\cdot\}_{i=1}^N$ consists of N independent d dimensional Brownian motions on a stochastic basis, i.e. $(\Omega, \mathcal{F}, \mathbb{P})$ with a normal filtration $(\mathcal{F}_t)$, induced by the Laplacian operator on the torus. The coefficient $\sigma_N \geqslant 0$ is a non-negative scalar for simplicity, and the vector fields K and V on the torus take values in the tangent spaces. The setting in this chapter is on the torus $\mathbb{T}^d$.

2.1 Preliminary

2.1.1 Besov spaces

We collect useful results related to Besov spaces. Recall that Besov spaces on the torus $B^\alpha_{p,q}(\mathbb{T}^d)$ (c.f [Tri06], [MW17]), with $\alpha \in \mathbb{R}$ and $1 \leqslant p, q \leqslant \infty$, are defined as the completion of C^∞ with respect to the norm

$$\|f\|_{B^\alpha_{p,q}} := \Big(\sum_{n\geqslant -1} \Big(2^{n\alpha q} \|\mathscr{F}^{-1}(\chi_n \mathscr{F}(f))\|^q_{L^p(\mathbb{T}^d)} \Big) \Big)^{\frac{1}{q}},$$

where $\mathscr{F}$ represents Fourier transform on $\mathbb{R}^d$ and $\{\chi_n\}_{n\geqslant -1} : \mathbb{R}^d \to [0, 1]$ are compact supported smooth functions satisfying

$$\mathrm{supp}\chi_{-1} \subseteq B(0, \frac{4}{3}); \quad \mathrm{supp}\chi_0 \subseteq B(0, \frac{8}{3}) \setminus B(0, \frac{4}{3}), \quad \chi_n(\cdot) = \chi_0(2^{-n}\cdot) \text{ for } n \geqslant 0;$$
$$\sum_{n\geqslant -1} \chi_n = 1.$$

Here $B(0, R)$ denotes the ball of center 0 and radius R.

We collect the following results which are frequently used in this chapter.

Lemma 2.1 ([Tri06, Proposition4.6]). *Let $\alpha \in \mathbb{R}$, $\beta \in \mathbb{R}$ and $p_1, p_2, q_1, q_2 \in [1, \infty]$. Then the embedding*

$$B^{\alpha}_{p_1,q_2} \hookrightarrow B^{\beta}_{p_2,q_2}$$

is compact if and only if,

$$\alpha - \beta > d\Big(\frac{1}{p_1} - \frac{1}{p_2}\Big)_+ .$$

Lemma 2.2. *(i) Let $\alpha, \beta \in \mathbb{R}$ and $p, p_1, p_2, q \in [1, \infty]$ be such that $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$. The bilinear map $(u, v) \mapsto uv$ extends to a continuous map from $B^{\alpha}_{p_1,q} \times B^{\beta}_{p_2,q}$ to $B^{\alpha\wedge\beta}_{p,q}$ if $\alpha + \beta > 0$ (cf. [MW17, Corollary 2]).*

(ii) (Duality.) Let $\alpha \in (0, 1)$, $p, q \in [1, \infty]$, p' and q' be their conjugate exponents, respectively. Then the mapping $(u, v) \mapsto \langle u, v\rangle = \int uv\mathrm{d}x$ extends to a continuous bilinear form on $B^{\alpha}_{p,q} \times B^{-\alpha}_{p',q'}$, and one has $|\langle u, v\rangle| \lesssim \|u\|_{B^{\alpha}_{p,q}} \|v\|_{B^{-\alpha}_{p',q'}}$ (cf. [MW17, Proposition 7]).

Lemma 2.3 ([BCD11, Corollary 2.86]). *For any positive real number α and any $p, q \in [1, \infty]$, it holds that*

$$\|fg\|_{B^{\alpha}_{p,q}} \lesssim \|f\|_{L^\infty} \|g\|_{B^{\alpha}_{p,q}} + \|f\|_{B^{\alpha}_{p,q}} \|g\|_{L^\infty},$$

with the proportional constant independent of f, g.

Lemma 2.4 ([KS21, Theorem 2.1 and 2.2]). *Let $\alpha, \beta \in \mathbb{R}, q, q_1, q_2 \in (0, \infty]$ and $p, p_1, p_2 \in [1, \infty]$ be such that*

$$1 + \frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}, \quad \frac{1}{q} \leqslant \frac{1}{q_1} + \frac{1}{q_2}.$$

1. *If $f \in B^{\alpha}_{p_1,q}$ and $g \in L^{p_2}$, then $f * g \in B^{\alpha}_{p,q}$ and*

$$\|f * g\|_{B^{\alpha}_{p,q}} \lesssim \|f\|_{B^{\alpha}_{p_1,q}} \cdot \|g\|_{L^{p_2}} ,$$

with the proportional constant independent of f, g.

2. *If $f \in B^{\alpha}_{p_1,q_1}$ and $g \in B^{\beta}_{p_2,q_2}$, then $f * g \in B^{\alpha+\beta}_{p,q}$ and*

$$\|f * g\|_{B^{\alpha+\beta}_{p,q}} \lesssim \|f\|_{B^{\alpha}_{p_1,q_1}} \cdot \|g\|_{B^{\beta}_{p_2,q_2}} .$$

Recall the result about smoothing effect of the heat kernel Γ.

Lemma 2.5 ([MW17, Propositions 3.11, 3.12]). *Let $u \in B^{\alpha}_{p,q}$ for some $\alpha \in \mathbb{R}$, $1 \leqslant p, q \leqslant \infty$. Then for every $\kappa \geqslant 0$*

$$\|\Gamma_t * u\|_{B^{\alpha+2\kappa}_{p,q}} \lesssim t^{-\kappa} \|u\|_{B^{\alpha}_{p,q}},$$

and

$$\|\Gamma_t * u - u\|_{B^{\alpha}_{p,q}} \lesssim t^{\kappa/2} \|u\|_{B^{\alpha+\kappa}_{p,q}}.$$

2.1.2 Concentration estimates

As we mentioned before, the major difficulty of our main estimates is to bound some exponential integrals, which can be understood as some proper partition functions. To prove Lemma 2.11 below, the following result from Jabin and Wang [JW18, Theorem 4] is crucial.

Lemma 2.6. *For any probability measure $\bar\rho$ on $\mathbb{T}^d$, and any bounded function $\phi(x,y)$ satisfying $\gamma := \|\phi\|_{L^\infty} < \infty$. Assume that ϕ satisfies the following cancellations*

$$\int_{\mathbb{T}^d} \phi(x,y)\bar\rho(x)\mathrm{d}x = 0 \quad \forall y, \qquad \int_{\mathbb{T}^d} \phi(x,y)\bar\rho(y)\mathrm{d}y = 0 \quad \forall x.$$

Then there exists a universal constant η depending on γ only such that

$$\sup_{N\geqslant 2} \int_{\mathbb{T}^{dN}} \bar\rho^{\otimes N} \exp\big(\eta N\langle \phi, \mu_N\otimes\mu_N\rangle\big)\mathrm{d}X^N < \infty,$$

where $\mu_N = \frac{1}{N}\sum_{i=1}^N \delta_{x_i}$, $X^N := (x_1,..,x_N) \in \mathbb{T}^{dN}$.

We shall apply this lemma with the solution to the mean-field equation (1.5) playing the role of $\bar\rho$. In this section we abuse the notations μ_N and X^N, but we shall always point out the dependence on time when we mention the empirical measure and vector associated to the particle system (2.1).

We also need the following novel concentration estimate dating back to [Hoe94], and use it to obtain the uniform estimate of the interaction term.

Lemma 2.7. *For any probability measure $\bar\rho$ on $\mathbb{T}^d$. Assume that the bounded function $\phi(x,y)$ satisfies*

$$\mathbb{E}\phi(X,Y) = 0, \tag{2.2}$$

for independent random variables X, Y with the common law $\bar\rho$.

Then it holds for any constant $\eta > 0$ that

$$\sup_{N\geqslant 2} \int_{\mathbb{T}^{dN}} \bar\rho^{\otimes N} \exp\big(\eta N|\langle \phi, \mu_N\otimes\mu_N\rangle|^2\big)\,\mathrm{d}X^N < \infty.$$

Proof Notice that it suffices to show the following statistics is sub-Gaussian,

$$U := \frac{\sqrt{N}}{N^2}\sum_{i\neq j} \phi(X_i,X_j) = \frac{\sqrt{N}}{N^2}\sum_{i<j}\Big[\phi(X_i,X_j)+\phi(X_j,X_i)\Big], \tag{2.3}$$

where $(X_i, i=1,...,N)$ are i.i.d. random variables with the distribution $\bar\rho$.

Let P be a partition of $\{1,...,N\}$ into $\frac{N}{2}$ pairs of distinct numbers when N is even (or $\frac{N+1}{2}$ groups with exactly one group contaning a single number when N is odd), and let S_N be the set of these partitions. Now we can decompose the summation into

$$U = \frac{1}{|S_N|}\sum_{P\in S_N} \frac{\sqrt{N}|S_N|}{N^2|S_{N-2}|}\sum_{(i,j)\in P} \phi(X_i,X_j)+\phi(X_j,X_i),$$

This decomposition and the relation $|S_N| \leqslant N|S_{N-2}|$ can be found in the proof of Lemma 3.4 below.

By Jensen's inequality, for any $\eta > 0$ we have

$$\begin{aligned}\mathbb{E}e^{\eta|U|^2} \leqslant& \frac{1}{|S_N|}\sum_{P\in S_N}\mathbb{E}\exp^{\eta\left|\frac{\sqrt{N}|S_N|}{N^2|S_{N-2}|}\sum_{(i,j)\in P}\phi(X_i,X_j)+\phi(X_j,X_i)\right|^2}\\ \leqslant& \sup_{P\in S_N}\mathbb{E}\exp^{\frac{\eta}{N}\left|\sum_{(i,j)\in P}\phi(X_i,X_j)+\phi(X_j,X_i)\right|^2}.\end{aligned}$$

Notice taht $\{\phi(X_i, X_j), (i, j) \in P\}$ are i.i.d. bounded random variables with zero expectations. The proof is completed by applying Hoeffding's inequality.

Remark 2.8. *Lemma 2.7 and Lemma 2.6 can be generalized in several aspects. Firstly, the space $\mathbb{T}^d$ could be replaced by any measurable spaces. Also, when ϕ is vector-valued, the result still holds.*

2.1.3 Pathwise realization of stochastic integrals

We introduce pathwise realization of the martingale part in the decomposition (1.26). Recall that the martingale part is given by

$$\frac{\sqrt{2\sigma_N}}{\sqrt{N}}\sum_{i=1}^N\int_0^t\nabla\varphi(X_i)\mathrm{d}B_s^i,$$

for each $\varphi \in C^\infty(\mathbb{T}^d)$. Formally, one could define a random operator $\mathcal{M}_t^N : \Omega \to H^{-\alpha}$, $\alpha > d/2 + 1$, for each $t \in [0, T]$ through

$$\mathcal{M}_t^N(\varphi) = \frac{\sqrt{2\sigma_N}}{\sqrt{N}}\sum_{i=1}^N\int_0^t\nabla\varphi(X_i)\mathrm{d}B_s^i, \quad \mathbb{P}-a.s. \tag{2.4}$$

However, the measurability of $\mathcal{M}_t^N : \Omega \to H^{-\alpha}$ is nontrivial due to the fact that the above stochastic integral is defined as a $\mathbb{P}$-equivalence class for each φ. Finding a measurable map $\mathcal{M}_t^N$ from Ω to $H^{-\alpha}$ requires a pathwise meaning of the map $\mathcal{M}_t^N(\varphi)$, such that $\mathcal{M}_t^N(\varphi)$ is continuous with respect to φ for almost every $\omega \in \Omega$.

Pathwise realization has been studied in different settings, for instance in [Itô83, Theorem 3.1], [fla], [MW17, Lemma 9], etc. Adapting the idea of investigating stochastic currents in [fla] by Flandoli, Gubinelli, Giaquinta, and Tortorelli, a pathwise realization $\mathcal{M}^N$ with values in Hilbert spaces can be obtained with a relatively simple proof. We state the result as follows.

Lemma 2.9. *For each N, there exists a progressively measurable process $\mathcal{M}^N$ with values in $H^{-\alpha}$, for any $\alpha > d/2 + 1$, such that* (2.4) *holds almost surely for all $t \in [0, T]$ and $\varphi \in C^\infty$.*

Now that we give the proof of Lemma 2.9. First recall the following result from [fla].

Lemma 2.10. *Let $\varphi \to S(\varphi)$ be a linear continuous mapping from a separable Banach space E to $L^0(\Omega)$ (random variables with convergence in probability). Assume that there exists a random variable $C(\omega)$ such that for any given $\varphi \in E$ we have*

$$|S(\varphi)(\omega)| \leqslant C(\omega)\|\varphi\|_E \quad \text{for} \quad \mathbb{P}-a.s. \quad \omega \in \Omega.$$

Then there exists a pathwise realization $\mathcal{S}$ of $S(\varphi)$ from $(\Omega, \mathcal{F}, \mathbb{P})$ to the dual space of E in the sense that

$$[S(\varphi)](\omega) =[\mathcal{S}(\omega)](\varphi), \quad \mathbb{P}-a.s,$$

for every $\varphi \in E$.

Proof [Proof of Lemma 2.9] The proof consists of two steps. The first step is to find a pathwise realization for each $t \in [0,T]$. The second step is justifying the pathwise realization forms a progressively measurable process. We denote $\frac{\sqrt{2\sigma_N}}{\sqrt{N}}$ by C_N below for simplicity.

We first apply [fla, Lemma 8] to obtain the following equality, for $\varphi \in C^\infty$,

$$C_N \sum_{i=1}^N \int_0^t \nabla\varphi(X_i)\mathrm{d}B_s^i =C_N \sum_{i=1}^N \sum_{k_i\in\mathbb{Z}^d} \langle \nabla\varphi, e_{-k_i}\rangle \int_0^t e_{k_i}(X_i)\mathrm{d}B_s^i, \quad \mathbb{P}-a.s.$$

Furthermore, using Hölder's inequality we have

$$\begin{aligned}\left|C_N \sum_{i=1}^N \int_0^t \nabla\varphi(X_i)\mathrm{d}B_s^i\right| \leqslant& C_N \sum_{i=1}^N \left(\sum_{k_i\in\mathbb{Z}^d} \langle k_i\rangle^{2\alpha-2} |\langle \nabla\varphi, e_{-k_i}\rangle|^2\right)^{\frac{1}{2}} \\ &\times \left(\sum_{k_i\in\mathbb{Z}^d} \langle k_i\rangle^{-2\alpha+2} \left|\int_0^t e_{k_i}(X_i)\mathrm{d}B_s^i\right|^2\right)^{\frac{1}{2}} \\ \leqslant& \|\varphi\|_{H^\alpha} C_N \sum_{i=1}^N \left(\sum_{k_i\in\mathbb{Z}^d} \langle k_i\rangle^{-2\alpha+2} \left|\int_0^t e_{k_i}(X_i)\mathrm{d}B_s^i\right|^2\right)^{\frac{1}{2}}.\end{aligned}$$

To apply Lemma 2.10 with $E = H^\alpha(\mathbb{T}^d)$ for $\alpha > d/2+1$, it is sufficient to find

$$\begin{aligned}&\mathbb{E}\left(\sum_{i=1}^N \left(\sum_{k_i\in\mathbb{Z}^d} \langle k_i\rangle^{-2\alpha+2} \left|\int_0^t e_{k_i}(X_i)\mathrm{d}B_s^i\right|^2\right)^{\frac{1}{2}}\right)^2 \\ \lesssim_N& \mathbb{E}\left(\sum_{i=1}^N \sum_{k_i\in\mathbb{Z}^d} \langle k_i\rangle^{-2\alpha+2} \left|\int_0^t e_{k_i}(X_i)\mathrm{d}B_s^i\right|^2\right) \\ \lesssim_N& \, t \sum_{i=1}^N \sum_{k_i\in\mathbb{Z}^d} \langle k_i\rangle^{-2\alpha+2} < \infty.\end{aligned}$$

Therefore, we thus obtain a pathwise realization of $C_N \sum_{i=1}^N \int_0^t \nabla\varphi(X_i)\mathrm{d}B_s^i$ for each $t \in [0,T]$, denoted by $\mathcal{M}_t^N$.

Define $\mathcal{M}^N := (\mathcal{M}_t^N)_{t\in[0,T]}$. Since the stochastic integrals are t-continuous, the equality

$$\mathcal{M}_t^N(\varphi) =C_N \sum_{i=1}^N \int_0^t \nabla\varphi(X_i)\mathrm{d}B_s^i$$

holds almost surely for all $t \in [0,T]$ and $\varphi \in C^\infty(\mathbb{T}^d)$. To justify measurability of $\mathcal{M}^N$. Notice that for each $\varphi \in C^\infty$, $\langle \mathcal{M}^N, \varphi\rangle = \mathcal{M}_t^N(\varphi)$ is a continuous adapted process. Hence for each $t \in [0,T]$, $\langle \mathcal{M}^N, \varphi\rangle : \Omega \times [0,t] \to \mathbb{R}$ is $\mathcal{F}_t \times \mathcal{B}([0,t])$-measurable. Since C^∞ is dense in the separable Hilbert space H^α, using Pettis measurability theorem, we thus find $\mathcal{M}^N : \Omega \times [0,T] \to H^{-\alpha}$ is progressively measurable.

2.2 Uniform estimates

This section collects uniform estimates on $\mu_N - v$, the interaction term $\mathcal{K}^N$, and a singular term derived from $\mathcal{K}^N(\varphi)$, where $\mathcal{K}^N$ is defined in (1.27). These estimates shall play crucial roles in obtaining tightness and identifying the limit in Section 2.3. Indeed, proving the uniform estimates is the main difficulty and technical contribution of this chapter. Surprisingly, this type estimate, which has been shown to be very useful for many purposes, can be actually obtained through a simple unified idea. The quantity we want to bound can be put in the integral form $\int \Phi F^N$, where Φ is a nonnegative function on $\mathbb{T}^{dN}$. Applying the famous variational formula from [DE11, Proposition 4.5.1], that is

$$\log \int_{\mathbb{T}^{dN}} \bar\rho_N e^{\Phi} \mathrm{d}X^N = \sup_{\nu \in \mathcal{P}(\mathbb{T}^{dN}), H(\nu|\bar\rho_N)<\infty} \left\{ \int_{\mathbb{T}^{dN}} \Phi \mathrm{d}\nu - H(\nu|\bar\rho_N) \right\}, \quad \forall \Phi \geqslant 0, \tag{2.5}$$

with $X^N := (x_1, .., x_N)$, $\mathcal{P}(\mathbb{T}^{dN})$ the probability measures on $\mathbb{T}^{dN}$, one can easily control $\int \Phi F^N$ as follows

$$\int_{\mathbb{T}^{dN}} \Phi F^N \mathrm{d}X^N \leqslant \frac{1}{\kappa N}\left(H(F^N|v^{\otimes N}) + \log \int_{\mathbb{T}^{dN}} v^{\otimes N} e^{\kappa N \Phi} \mathrm{d}X^N \right), \tag{2.6}$$

for any $\kappa > 0$, simply noticing that F^N plays the role of ν and replacing Φ with $\kappa N \Phi$. See also a direct proof of this inequality (2.6) in [JW18, Lemma 1]. As we will see in Lemma 2.16, the extra factor $\frac{1}{N}$ is essential to obtain uniform estimate for fluctuations $\eta^N = \sqrt{N}(\mu_N - v)$, but it comes with a cost that we have to bound the exponential integral $\int v^{\otimes N} \exp(\kappa N \Phi)$ uniformly in N. Controlling such exponential integrals was achieved in Section 2.1.2, which leads to the following uniform estimates.

The first estimate concerns on the convergence from $\mu_N(t)$ to v_t in $H^{-\alpha}$, for $\alpha > d/2$.

Lemma 2.11. *For each $\alpha > d/2$, there exists a constant C_α such that, for all $t \in [0,T]$,*

$$\mathbb{E}\|\mu_N(t) - v_t\|_{H^{-\alpha}}^2 \leqslant \frac{C_\alpha}{N}(H_t(F^N|v^{\otimes N}) + 1),$$

where we recall $\mu_N(t) = \frac{1}{N}\sum_{i=1}^N \delta_{X_i(t)}$ and the expectation is taken according to the joint distribution $F^N(t,\cdot)$ of the particle system (2.1).

This lemma has a direct consequence. Recall that the fluctuation measure $\eta^N(t) = \sqrt{N}(\mu_N(t) - v_t)$. Under Assumption **(A3)**, i.e. $\sup_{t\in[0,T]} \sup_N H_t(F^N|v^{\otimes N}) \lesssim 1$, one can then immediately obtain

$$\sup_{t\in[0,T]} \sup_N \mathbb{E}\|\eta^N(t)\|_{H^{-\alpha}}^2 \lesssim 1, \quad \text{for } \alpha > d/2.$$

This gives the weakly compactness of the fluctuation measures for every t. But for stochastic processes, some time continuity is required to ensure the measurability of the limiting process.

Proof Since the Dirac measure belongs to $H^{-\alpha}(\mathbb{T}^d)$ for every $\alpha > d/2$, it follows that $\mu_N(t) - v_t \in H^{-\alpha}(\mathbb{T}^d)$. Then by (2.6) , we find for any $\kappa > 0$,

$$\begin{aligned}\mathbb{E}\|\mu_N(t) - v_t\|_{H^{-\alpha}}^2 &= \int_{\mathbb{T}^{dN}} \|\mu_N - v_t\|_{H^{-\alpha}}^2 F^N(t, X^N)\mathrm{d}X^N \\ &\leqslant \frac{1}{\kappa N}\left(H_t(F^N|v^{\otimes N}) + \log\int_{\mathbb{T}^{dN}} \exp\left(\kappa N\|\mu_N - v_t\|_{H^{-\alpha}}^2\right)v_t^{\otimes N}\mathrm{d}X^N\right).\end{aligned} \tag{2.7}$$

Recalling $\{e_k\}_{k\in\mathbb{Z}^d}$ is the Fourier basis, and

$$\|\mu_N - v_t\|_{H^{-\alpha}}^2 = \sum_{k\in\mathbb{Z}^d}\langle k\rangle^{-2\alpha}|\langle e_k, \mu_N - v_t\rangle|^2.$$

Since the exponential function is convex, using Jensen's inequality gives that

$$\begin{aligned}&\int_{\mathbb{T}^{dN}} \exp\left(\kappa N\|\mu_N - v_t\|_{H^{-\alpha}}^2\right)v_t^{\otimes N}\mathrm{d}X^N \\ \leqslant&\frac{1}{C}\sum_{k\in\mathbb{Z}^d}\langle k\rangle^{-2\alpha}\int_{\mathbb{T}^{dN}} \exp\left(\kappa NC|\langle e_k, \mu_N - v_t\rangle|^2\right)v_t^{\otimes N}\mathrm{d}X^N \\ \leqslant&\sup_{k\in\mathbb{Z}^d}\int_{\mathbb{T}^{dN}} \exp\left(\kappa NC|\langle e_k, \mu_N - v_t\rangle|^2\right)v_t^{\otimes N}\mathrm{d}X^N,\end{aligned} \tag{2.8}$$

where the constant $C = \sum_{k\in\mathbb{Z}^d}\langle k\rangle^{-2\alpha}$ depends only on α and is finite since $\alpha > d/2$.

We define

$$\phi_1(t, k, x, y) :=[e_k(x) - \langle e_k, v_t\rangle][e_{-k}(y) - \langle e_{-k}, v_t\rangle],$$

therefore

$$\begin{aligned}&\int_{\mathbb{T}^{dN}} \exp\left(\kappa NC|\langle e_k, \mu_N - v_t\rangle|^2\right)v_t^{\otimes N}\mathrm{d}X^N \\ =&\int_{\mathbb{T}^{dN}} \exp\left(\kappa NC\left\langle \phi_1(t, k, \cdot, \cdot), \mu_N\otimes\mu_N\right\rangle\right)v_t^{\otimes N}\mathrm{d}X^N.\end{aligned}$$

Since v_t is a probability measure, $\|\phi_1\|_{L^\infty}$ is bounded uniformly in t and k. One can also easily check that

$$\int_{\mathbb{T}^d}\phi_1(t, k, x, y)v_t(x)\mathrm{d}x = 0 \quad \forall y, \qquad \int_{\mathbb{T}^d}\phi_1(t, k, x, y)v_t(y)\mathrm{d}y = 0 \quad \forall x.$$

Then by Lemma 2.6 with κ (depending on α) small enough, we deduce that

$$\begin{aligned}&\sup_N\sup_{k\in\mathbb{Z}^d}\int_{\mathbb{T}^{dN}} \exp\left(\kappa NC|\langle e_k, \mu_N - v_t\rangle|^2\right)v_t^{\otimes N}\mathrm{d}X^N \\ =&\sup_N\sup_{k\in\mathbb{Z}^d}\int_{\mathbb{T}^{dN}} v_t^{\otimes N}\exp(\kappa NC\langle\phi_1(t, k, \cdot, \cdot), \mu_N\otimes\mu_N\rangle)\mathrm{d}X^N\end{aligned}$$

$$= \sup_N \sup_{k\in\mathbb{Z}^d} \int_{\mathbb{T}^{dN}} v_t^{\otimes N} \exp(\kappa NC\langle \mathrm{Re}\,\phi_1(t,k,\cdot,\cdot),\mu_N\otimes\mu_N\rangle)\mathrm{d}X^N < \infty, \tag{2.9}$$

where the equalities follows by

$$|\langle e_k, \mu_N - v_t\rangle|^2 = \langle \phi_1(t,k,\cdot,\cdot),\mu_N\otimes\mu_N\rangle \in \mathbb{R}.$$

Combining (2.7)-(2.9) yields

$$\mathbb{E}\|\mu_N(t)-v_t\|^2_{H^{-\alpha}} \leqslant \frac{1}{\kappa N}(H_t(F^N|v^{\otimes N}) + C_\alpha),$$

where C_α is a constant depending only on α. We thus arrive at the result.

In particular, Lemma 2.11 gives the tightness of laws of $\{\eta^N(0)\}$ on $H^{-\alpha}$ under the condition that $H(F^N(0)|v_0^{\otimes N})$ is finite, which together with Assumption **(A1)** yields the convergence of $\{\eta^N(0)\}$ in the negative Sobolev spaces.

Corollary 2.12. *For every $\alpha > d/2$, η_0^N converges in distribution to η_0 given by* **(A1)** *in $H^{-\alpha}$.*

The next lemma concerns on the interaction part in the decomposition (1.26).

Lemma 2.13. *If the kernel K satisfies Assumption* **(A2)**, *then for each $\alpha > d/2+2$, there exists a constant C_α such that, for all $t\in[0,T]$,*

$$\mathbb{E}\|\nabla\cdot[K*\mu_N(t)\mu_N(t) - v_tK*v_t]\|^2_{H^{-\alpha}} \leqslant \frac{C_\alpha}{N}(H_t(F^N|v^{\otimes N})+1),$$

where the expectation is taken according to the joint distribution $F^N(t,\cdot)$ of the particle system (2.1).

Proof The proof is similar to Lemma 2.11. First, by (2.6) we find for any $\kappa>0$,

$$\begin{aligned}&\mathbb{E}\|\nabla\cdot[K*\mu_N(t)\mu_N(t) - v_tK*v_t]\|^2_{H^{-\alpha}}\\ &\leqslant \frac{1}{\kappa N}\left(H_t(F^N|v^{\otimes N}) + \log\int_{\mathbb{T}^{dN}} v_t^{\otimes N}\exp\left(\kappa N\|\nabla\cdot[K*\mu_N\mu_N - v_tK*v_t]\|^2_{H^{-\alpha}}\right)\mathrm{d}X^N\right).\end{aligned} \tag{2.10}$$

Next, we find that

$$\begin{aligned}\|\nabla\cdot[K*\mu_N\mu_N - v_tK*v_t]\|^2_{H^{-\alpha}} &= \sum_{k\in\mathbb{Z}^d}\langle k\rangle^{-2\alpha}|\langle\nabla e_k, K*\mu_N\mu_N - v_tK*v_t\rangle|^2\\ &\leqslant \sum_{k\in\mathbb{Z}^d\setminus\{0\}}\langle k\rangle^{-2\alpha}|k|^2|\langle e_k, K*\mu_N\mu_N - v_tK*v_t\rangle|^2.\end{aligned}$$

For the case $|x|K(x)\in L^\infty$ and $K(x) = -K(-x)$, we do a symmetrization trick. That is, for any $\varphi\in C^\infty(\mathbb{T}^d)$ and a probability measure μ,

$$\begin{aligned}\int_{\mathbb{T}^d}\varphi(x)K*\mu(x)\mu(\mathrm{d}x) &= \int_{\mathbb{T}^{2d}}\varphi(x)K(x-y)\mu^{\otimes 2}(\mathrm{d}x\mathrm{d}y)\\ &= \frac{1}{2}\int_{\mathbb{T}^{2d}}(\varphi(x)-\varphi(y))\cdot K(x-y)\mu^{\otimes 2}(\mathrm{d}x\mathrm{d}y).\end{aligned}$$

We define that

$$\mathbb{K}_\varphi(x,y) := \frac{1}{2}K(x-y)[\varphi(x)-\varphi(y)], \quad \forall \varphi \in C^\infty(\mathbb{T}^d).$$

Thus in this case, $\|\mathbb{K}_\varphi\|_{L^\infty} \lesssim \|\nabla\varphi\|_{L^\infty}\||x|K\|_{L^\infty}$. Consequently, since

$$\langle e_k, K*\mu_N\mu_N - v_tK*v_t\rangle = \langle \mathbb{K}_{e_k}(\cdot,\cdot), \mu_N^{\otimes 2} - \bar{\rho}_t^{\otimes 2}\rangle,$$

and $\|\mathbb{K}_{e_k}\| \lesssim |k|$, one proceeds as

$$\begin{aligned}\|\nabla\cdot[K*\mu_N\mu_N - v_tK*v_t]\|_{H^{-\alpha}}^2 &= \sum_{k\in\mathbb{Z}^d\setminus\{0\}} \langle k\rangle^{-2\alpha}|k|^2|\langle \mathbb{K}_{e_k}, \mu_N\otimes\mu_N - v_t\otimes v_t\rangle|^2 \\ &= \sum_{k\in\mathbb{Z}^d\setminus\{0\}} \langle k\rangle^{-2\alpha}|k|^4|\langle \frac{\mathbb{K}_{e_k}}{|k|}, \mu_N\otimes\mu_N - v_t\otimes v_t\rangle|^2,\end{aligned}$$

where now $\frac{\mathbb{K}_{e_k}}{|k|}$ is bounded.

For the case that $K \in L^\infty$ but K is not necessarily anti-symmetric, we directly write

$$\langle e_k, K*\mu_N\mu_N - v_tK*v_t\rangle = \langle e_k(x)K(x-y), \mu_N^{\otimes 2} - v_t^{\otimes 2}\rangle.$$

To sum it up, we define $\phi_2 : [0,T]\times\{\mathbb{Z}^d\setminus\{0\}\}\times\mathbb{T}^{2d} \to \mathbb{R}^d$ by

$$\phi_2(t,k,x,y) := \begin{cases} K(x-y)e_k(x) - \langle e_k, v_tK*v_t\rangle, & \text{if } K\in L^\infty \\ \frac{\mathbb{K}_{e_k}(x,y)}{|k|} - \langle \frac{\mathbb{K}_{e_k}}{|k|}, v_t\otimes v_t\rangle, & \text{if } |x|K(x)\in L^\infty, K(x) = -K(-x). \end{cases}$$

Using Jensen's inequality, for both cases we have

$$\begin{aligned}&\int_{\mathbb{T}^{dN}} v_t^{\otimes N}\exp\left(\kappa N\|\nabla\cdot[K*\mu_N\mu_N - v_tK*v_t]\|_{H^{-\alpha}}^2\right)\mathrm{d}X^N \\ &= \int_{\mathbb{T}^{dN}} v_t^{\otimes N}\exp\left(\kappa N\sum_{k\in\mathbb{Z}^d\setminus\{0\}}\langle k\rangle^{-2\alpha}|k|^2|\langle e_k, K*\mu_N\mu_N - v_tK*v_t\rangle|^2\right)\mathrm{d}X^N \\ &\leqslant \int_{\mathbb{T}^{dN}} v_t^{\otimes N}\exp\left(\kappa N\sum_{k\in\mathbb{Z}^d\setminus\{0\}}\langle k\rangle^{-2\alpha+4}|\langle \phi_2(t,k,\cdot,\cdot), \mu_N\otimes\mu_N\rangle|^2\right)\mathrm{d}X^N \\ &\leqslant \frac{1}{C}\sum_{k\in\mathbb{Z}^d\setminus\{0\}}\langle k\rangle^{-2\alpha+4}\int_{\mathbb{T}^{dN}} v_t^{\otimes N}\exp\left(\kappa NC|\langle \phi_2(t,k,\cdot,\cdot), \mu_N\otimes\mu_N\rangle|^2\right)\mathrm{d}X^N \\ &\leqslant \sup_{k\in\mathbb{Z}^d\setminus\{0\}}\int_{\mathbb{T}^{dN}} v_t^{\otimes N}\exp\left(\kappa NC|\langle \phi_2(t,k,\cdot,\cdot), \mu_N\otimes\mu_N\rangle|^2\right)\mathrm{d}X^N, \end{aligned} \tag{2.11}$$

where the constant $C := \sum_{k\in\mathbb{Z}^d\setminus 0}\langle k\rangle^{-2\alpha+4}$ depends only on α, and is finite since $\alpha > d/2+2$.

Furthermore, since ϕ_2 is complex-valued, we find

$$\begin{aligned}&\sup_N \sup_{k\in\mathbb{Z}^d\setminus\{0\}}\int_{\mathbb{T}^{dN}} v_t^{\otimes N}\exp\left(\kappa NC|\langle \phi_2(t,k,\cdot,\cdot), \mu_N\otimes\mu_N\rangle|^2\right)\mathrm{d}X^N \\ &\leqslant \frac{1}{2}\sup_N \sup_{k\in\mathbb{Z}^d\setminus\{0\}}\int_{\mathbb{T}^{dN}} v_t^{\otimes N}\exp\left(2\kappa NC|\langle \mathrm{Re}\,\phi_2(t,k,\cdot,\cdot), \mu_N\otimes\mu_N\rangle|^2\right)\mathrm{d}X^N\end{aligned}$$

$$+\frac{1}{2}\sup_N \sup_{k\in\mathbb{Z}^d\setminus\{0\}} \int_{\mathbb{T}^{dN}} v_t^{\otimes N} \exp\left(2\kappa NC|\langle \mathrm{Im}\,\phi_2(t,k,\cdot,\cdot),\mu_N\otimes\mu_N\rangle|^2\right)\mathrm{d}X^N, \qquad (2.12)$$

where the inequality follows by Jensen's inequality and the fact that

$$|\langle \phi_2(t,k,\cdot,\cdot),\mu_N\otimes\mu_N\rangle|^2 = |\langle \mathrm{Re}\,\phi_2(t,k,\cdot,\cdot),\mu_N\otimes\mu_N\rangle|^2 + |\langle \mathrm{Im}\,\phi_2(t,k,\cdot,\cdot),\mu_N\otimes\mu_N\rangle|^2.$$

One can easily find that $\|\phi_2\|_{L^\infty}$ is bounded uniformly in (t,k), and satisfies the cancellation

$$\int_{\mathbb{T}^{2d}} \phi_2(t,k,x,y)v_t(x)v_t(y)\mathrm{d}x\mathrm{d}y = 0,$$

and so do the real and imaginary part of ϕ_2.

Choosing κ (depending on α) sufficiently small, then we are able to apply Lemma 2.7 to obtain that

$$\sup_N \sup_{k\in\mathbb{Z}^d\setminus 0} \int_{\mathbb{T}^{dN}} v_t^{\otimes N} \exp\left(\kappa NC|\langle \phi_2(t,k,\cdot,\cdot),\mu_N\otimes\mu_N\rangle|^2\right)\mathrm{d}X^N \leqslant C_\alpha,$$

where the universal constant C_α only depends on α. The proof is then completed by combining this with (2.10) and (2.11).

The last estimate in this section plays a crucial role in identifying the limit in Section 2.3.2.

Lemma 2.14. *If the kernel K satisfies assumption* **(A2)***, then for each $\varphi \in C^1$, there exists a universal constant C such that, for all $t\in[0,T]$,*

$$\mathbb{E}|\langle \varphi K*(\mu_N(t)-v_t),\mu_N(t)-v_t\rangle| \leqslant \frac{C}{N}(H_t(F^N|v^{\otimes N})+1),$$

where the expectation is taken according to the joint distribution $F^N(t,\cdot)$ of the particle system (2.1).

Proof We first write the quantity in the following form

$$\mathbb{E}|\langle \varphi K*(\mu_N(t)-v_t),\mu_N(t)-v_t\rangle| = \mathbb{E}|\Phi(t,X_t^N)| = \int_{\mathbb{T}^{dN}} |\Phi(t,X^N)|F_t^N\mathrm{d}X^N, \qquad (2.13)$$

where Φ is defined by

$$\Phi(t,X^N) = \langle \varphi K*(\mu_N - v_t),\mu_N - v_t\rangle.$$

For the case $K\in L^\infty$, we find

$$\Phi(t,X^N) = \langle \phi_3(t,\cdot,\cdot),\mu_N\otimes\mu_N\rangle,$$

with ϕ_3 defined by

$$\phi_3(t,x,y) := K(x-y)\varphi(x) - \varphi(x)K*v_t(x) - \langle K(\cdot-y)\varphi, v_t\rangle + \langle \varphi K * v_t, v_t\rangle.$$

For the case $|x|K(x)\in L^\infty$ and $K(x) = -K(-x)$, we do a symmetrization for Φ as in the proof of Lemma 2.13, i.e.

$$\Phi(t,X^N) = \langle \mathbb{K}_\varphi, (\mu_N - v_t)\otimes\mu_N - v_t\rangle = \langle \phi_3(t,\cdot,\cdot),\mu_N\otimes\mu_N\rangle,$$

with ϕ_3 defined by

$$\phi(t,x,y) := \mathbb{K}_\varphi(x,y) - \langle \mathbb{K}_\varphi(x,\cdot), v_t\rangle - \langle \mathbb{K}_\varphi(\cdot,y), v_t\rangle + \langle \mathbb{K}_\varphi, v_t \otimes v_t\rangle.$$

By (2.6), it holds for any $\kappa > 0$ that

$$\begin{aligned}\int_{\mathbb{T}^{dN}} |\Phi(t,X^N)| F_t^N \mathrm{d}X^N &\leqslant \int_{\mathbb{T}^{dN}} |\Phi(t,X^N)|^2 F_t^N \mathrm{d}X^N \\ &\leqslant \frac{1}{\kappa N}\left(H_t(F^N|v^{\otimes N}) + \log \int_{\mathbb{T}^{dN}} v_t^{\otimes N} e^{\kappa N |\Phi|^2} \mathrm{d}X^N\right). \end{aligned} \tag{2.14}$$

On the other hand, one can easily check the following cancellations

$$\int_{\mathbb{T}^d} \phi_3(t,x,y) v_t(x) \mathrm{d}x = 0 \quad \forall y, \qquad \int_{\mathbb{T}^d} \phi_3(t,x,y) v_t(y) \mathrm{d}y = 0 \quad \forall x.$$

Since in both cases, ϕ_3 is bounded uniformly in t, we can choose κ such that $\sqrt{\kappa}\|\phi_3\|_{L^\infty}$ sufficiently small. Letting v_t and $\sqrt{\kappa}\phi_3$ play the roles of $\bar\rho$ and ϕ in Lemma 2.7 respectively, we deduce that

$$\int_{\mathbb{T}^{dN}} v_t^{\otimes N} e^{\kappa N |\Phi|^2} \mathrm{d}X^N \leqslant C,$$

where C is a constant depending only on φ. Combining this with (2.13) and (2.14), we thus arrive at the result.

2.3 The SPDE limit

The aim of this section is to analyze the fluctuation behavior of the empirical measure μ^N for the non-degenerate case, i.e. $\sigma > 0$. It will be shown that $\eta^N = \sqrt{N}(\mu_N - v)$ converges in distribution to the unique solution η to the linear SPDE (1.7). We shall start with proving that the sequence of $(\eta^N)_{N\geqslant 1}$ is tight. Then each tight limit of the subsequence from $(\eta^N)_{N\geqslant 1}$ will be identified as a martingale solution to the equation (1.7). The next step is to show pathwise uniqueness of (1.7), which allows us to conclude the proof of Theorem 1.4.

2.3.1 Tightness

In the following, we are going to prove tightness of $(\eta^N, \mathcal{M}^N)_{N\geqslant 1}$. To start, recall the following tightness criterion given by Arzela-Ascoli theorem [Kel17, Theorem 7.17]. Suppose that $(u^N)_{N\geqslant 1}$ is a class of random variables in $C([0,T],E)$ with a given Polish space E. The sequence of $(u^N)_{N\geqslant 1}$ is tight in $C([0,T],E)$ if and only if the following conditions hold:

1. For each $\epsilon > 0$ and each $t \in [0,T]$, there is a compact set $A \subset E$ (possibly depending on t) such that

$$\sup_N \mathbb{P}(u_t^N \in A) > 1 - \epsilon.$$

2. For each $\epsilon > 0$,

$$\lim_{h\to 0}\sup_N \mathbb{P}(\sup_{s,t\in[0,T]}\sup_{|t-s|\leqslant h}\|u_t^N - u_s^N\|_E > \epsilon) = 0.$$

Since the embedding $H^{-\alpha'} \hookrightarrow H^{-\alpha}$ is compact if $\alpha' < \alpha$ (called the Rellich–Kondrachov theorem, see [Tri06, Proposition 4.6] or [Tay13, Proposition 3.4]), using Chebyshev's inequality, one can get the following sufficient conditions for tightness in $C([0,T], H^{-\alpha})$

1. For each $t \in [0,T]$, there exists some $\alpha' < \alpha$ such that

$$\sup_N \mathbb{E}\|u_t^N\|_{H^{-\alpha'}} < \infty. \tag{2.15}$$

2. There exists $\theta > 0$ such that

$$\sup_N \mathbb{E}\|u_t^N\|_{C^\theta([0,T],H^{-\alpha})} = \sup_N \mathbb{E}\left(\sup_{0\leqslant s<t\leqslant T}\frac{\|u_t^N - u_s^N\|_{H^{-\alpha}}}{(t-s)^\theta}\right) < \infty. \tag{2.16}$$

Therefore to obtain tightness of $\{\eta^N, \mathcal{M}^N\}_{N\in\mathbb{N}}$ it suffices to justify (2.15) and (2.16) with $\mathcal{M}^N$ and η^N playing the role of u^N.

The following lemma gives tightness of the martingale part.

Lemma 2.15. *For every $\alpha > d/2+1$, the sequence of $(\mathcal{M}^N)_{N\geqslant 1}$ is tight in the space $C([0,T], H^{-\alpha})$.*

Proof By the above tightness criterion, it is indeed sufficient to prove that: for each $\alpha > d/2+1$ and $\theta' \in (0, \frac{1}{2})$, it holds that

$$\sup_N \mathbb{E}(\|\mathcal{M}^N\|^2_{C^{\theta'}([0,T],H^{-\alpha})}) < \infty.$$

First, for the Fourier basis $\{e_k\}_{k\in\mathbb{Z}^d}$ and $t \in [0,T]$, we find

$$\mathcal{M}_t^N(e_k) = \frac{\sqrt{2\sigma_N}}{\sqrt{N}}\sum_{i=1}^N\int_0^t \nabla e_k(X_i)\cdot \mathrm{d}B_s^i = \sqrt{-1}\frac{\sqrt{2\sigma_N}}{\sqrt{N}}\sum_{i=1}^N\int_0^t e_k(X_i)k\cdot \mathrm{d}B_s^i. \tag{2.17}$$

For any $\theta > 1$ and $0 \leqslant s < t \leqslant T$, we deduce from Hölder's inequality that

$$\begin{aligned}
&\sup_N \mathbb{E}(\|\mathcal{M}_t^N - \mathcal{M}_s^N\|^{2\theta}_{H^{-\alpha}}) \\
=&\sup_N \mathbb{E}\left[\left(\sum_{k\in\mathbb{Z}^d}\langle k\rangle^{-2\alpha}|\mathcal{M}_t^N(e_k) - \mathcal{M}_s^N(e_k)|^2\right)^\theta\right] \\
\leqslant&\sup_N \mathbb{E}\left(\sum_{k\in\mathbb{Z}^d}\langle k\rangle^{-\alpha_1\theta}|\mathcal{M}_t^N(e_k) - \mathcal{M}_s^N(e_k)|^{2\theta}\right)\left(\sum_{k\in\mathbb{Z}^d}\langle k\rangle^{-\alpha_2\frac{\theta}{\theta-1}}\right)^{\theta-1},
\end{aligned}$$

where $\alpha_1+\alpha_2 = 2\alpha$ and $\alpha_1, \alpha_2 > 0$. Further choosing $\alpha_1 = \alpha + 1 - \frac{d}{2} + \frac{d}{\theta}$ and $\alpha_2 = \alpha - 1 + \frac{d}{2} - \frac{d}{\theta}$, then we have $(\alpha_1 - 2)\theta > d$ and $\alpha_2\frac{\theta}{\theta-1} > d$, due to $\theta > 1$ and the condition $\alpha > d/2+1$. Hence the summation $\sum_{k\in\mathbb{Z}^d}\langle k\rangle^{-\alpha_2\frac{\theta}{\theta-1}}$ is finite. Moreover, using the equality (2.17) gives

$$\sup_N \mathbb{E}(\|\mathcal{M}_t^N - \mathcal{M}_s^N\|^{2\theta}_{H^{-\alpha}}) \lesssim_{\alpha_2,\theta} \sup_N \sum_{k\in\mathbb{Z}^d}\langle k\rangle^{-\alpha_1\theta}\mathbb{E}(|\mathcal{M}_t^N(e_k) - \mathcal{M}_s^N(e_k)|^{2\theta})$$

$$
\begin{aligned}
&\lesssim_{\alpha_2,\theta} \sum_{k\in\mathbb{Z}^d} \langle k\rangle^{-\alpha_1\theta+2\theta} \sup_N \mathbb{E}\left|\frac{\sqrt{2\sigma_N}}{\sqrt{N}}\sum_{i=1}^N \int_s^t e_k(X_i)\mathrm{d}B_r^i\right|^{2\theta}\\
&\lesssim_{\alpha_2,\theta} \sum_{k\in\mathbb{Z}^d} \langle k\rangle^{-\alpha_1\theta+2\theta} \sup_N \mathbb{E}\left(\int_s^t \frac{2\sigma_N}{N}\sum_{i=1}^N |e_k(X_i)|^2\mathrm{d}r\right)^{\theta}\\
&\lesssim_{\alpha_2,\theta} (t-s)^\theta \sum_{k\in\mathbb{Z}^d} \langle k\rangle^{-\alpha_1\theta+2\theta} \lesssim_{\alpha_1,\alpha_2,\theta} (t-s)^\theta,
\end{aligned} \tag{2.18}
$$

where the third inequality follows by the Burkholder-Davis-Gundy's inequality. Therefore, (2.18) allows us to apply the Kolmogorov continuity theorem [BFH18, Theorem 2.3.11], and we find

$$
\sup_N \mathbb{E}(\|\mathcal{M}^N\|^{2\theta}_{C^{\theta'}([0,T],H^{-\alpha})}) < \infty, \tag{2.19}
$$

for any $0<\theta'<\frac{\theta-1}{2\theta}$, $\theta>1$, and $\alpha>d/2+1$. The result follows by arbitrary $\theta>1$.

Next, we need the tightness of the fluctuation measures.

Lemma 2.16. *Under the assumptions* **(A2)-(A4)***, for every $\alpha>d/2+2$, the sequence of $(\eta^N)_{N\geqslant 1}$ is tight in the space $C([0,T],H^{-\alpha})$.*

Proof First, by Assumption **(A3)** and Lemma 2.11, one can easily deduce (2.15) with η^N playing the role of u^N for any $\alpha>d/2+2$. Indeed, taking $\mu_N(\cdot)-v=\frac{1}{\sqrt{N}}\eta^N$ into Lemma 2.11 immediately gives

$$
\sup_{t\in[0,T]}\sup_N \mathbb{E}\|\eta_t^N\|^2_{H^{-\alpha+2}} \lesssim \sup_{t\in[0,T]}\sup_N H(F_t^N|v_t^{\otimes N})(t)+1. \tag{2.20}
$$

Then **(A3)** implies that the right hand side of (2.20) is finite. Thus (2.15) follows by $\alpha-2$ playing the role of α'.

As for (2.16), it suffices to prove the case $\alpha-2\in(d/2,\beta)$, where β is given in Assumption **(A5)**. Recall the decomposition (1.26), $\|\eta_t^N-\eta_s^N\|_{H^{-\alpha}}$, $0\leqslant s<t<T$, is controlled via the following relation

$$
\|\eta_t^N-\eta_s^N\|^2_{H^{-\alpha}} \lesssim \sum_{i=1}^5 J^i_{s,t}, \tag{2.21}
$$

where $J^i_{s,t}$, $i=1,\dots,5$, are defined by

$$
\begin{aligned}
&J^1_{s,t} := \left\|\sigma\int_s^t \Delta\eta_r^N\mathrm{d}r\right\|^2_{H^{-\alpha}}, \qquad J^2_{s,t} := \left\|\int_s^t \mathcal{K}_r^N\mathrm{d}r\right\|^2_{H^{-\alpha}},\\
&J^3_{s,t} := \left\|\int_s^t \nabla\cdot(V\eta_r^N)\mathrm{d}r\right\|^2_{H^{-\alpha}}, \quad J^4_{s,t} := \left\|\int_s^t \sqrt{N}(\sigma_N-\sigma)\Delta\mu_N(r)\mathrm{d}r\right\|^2_{H^{-\alpha}},\\
&J^5_{s,t} := \|\mathcal{M}_t^N-\mathcal{M}_s^N\|^2_{H^{-\alpha}}.
\end{aligned}
$$

For $J^1_{s,t}$, applying Hölder's inequality gives

$$
\sup_N \mathbb{E}\left(\sup_{0\leqslant s<t\leqslant T}\frac{J^1_{s,t}}{t-s}\right) \lesssim \sup_N \mathbb{E}\int_0^T \|\Delta\eta_t^N\|^2_{H^{-\alpha}}\mathrm{d}t \lesssim \sup_N \sup_{t\in[0,T]} \mathbb{E}\|\eta_t^N\|^2_{H^{-\alpha+2}} < \infty, \tag{2.22}
$$

where we used (2.20) in the last step.
For $J_{s,t}^2$, similarly, applying Hölder's inequality gives

$$\begin{aligned}\sup_N \mathbb{E}\left(\sup_{0\leqslant s<t\leqslant T}\frac{J_{s,t}^2}{t-s}\right) &\lesssim \sup_N \mathbb{E}\left(\int_0^T \|\mathcal{K}_t^N\|_{H^{-\alpha}}^2 \mathrm{d}t\right)\\ &\lesssim \sup_N \sup_{t\in[0,T]} \mathbb{E}\|\mathcal{K}_t^N\|_{H^{-\alpha}}^2.\end{aligned}$$

Recall that $\mathcal{K}_t^N = \sqrt{N}\nabla\cdot[K*\mu_N(t)\mu_N(t) - v_t K * v_t]$, and thus Lemma 2.13 and the assumptions **(A2)-(A3)** deduces that

$$\sup_N \mathbb{E}\left(\sup_{0\leqslant s<t\leqslant T}\frac{J_{s,t}^2}{t-s}\right) \lesssim \sup_N \sup_{t\in[0,T]} H(F_t^N|v_t^{\otimes N}) + 1 < \infty. \tag{2.23}$$

Similarly, we obtain

$$\sup_N \mathbb{E}\left(\sup_{0\leqslant s<t\leqslant T}\frac{J_{s,t}^3}{t-s}\right) \lesssim \sup_N \sup_{t\in[0,T]} \mathbb{E}\|V\eta_t^N\|_{H^{-\alpha+2}}^2,$$

and

$$\sup_N \mathbb{E}\left(\sup_{0\leqslant s<t\leqslant T}\frac{J_{s,t}^4}{t-s}\right) \lesssim \sup_N \sup_{t\in[0,T]} N|\sigma_N-\sigma|^2\mathbb{E}\|\mu_N(t)\|_{H^{-\alpha+2}}^2.$$

Furthermore, Lemma 2.1 together with Lemma 2.2 shows that

$$\|V\eta\|_{H^{-\alpha+2}} \lesssim \|V\|_{C^\beta}\|\eta^N\|_{H^{-\alpha+2}}.$$

Hence using (2.20) and Assumption **(A4)** gives

$$\sup_N \mathbb{E}\left(\sup_{0\leqslant s<t\leqslant T}\frac{J_{s,t}^3}{t-s}\right) \lesssim \sup_N \sup_{t\in[0,T]} \|V\|_{C^\beta}^2\mathbb{E}\|\eta_t^N\|_{H^{-\alpha+2}}^2 < \infty. \tag{2.24}$$

On the other hand, Assumption **(A4)**, $\mu_N = N^{-1/2}\eta^N + \bar\rho$ and (2.20) imply that

$$\sup_N \mathbb{E}\left(\sup_{0\leqslant s<t\leqslant T}\frac{J_{s,t}^4}{t-s}\right) \to 0. \tag{2.25}$$

For $J_{s,t}^5$, we deduce from (2.19) that for any $\theta\in\left(0,\frac12\right)$,

$$\sup_N \mathbb{E}\left(\sup_{0\leqslant s<t\leqslant T}\frac{J_{s,t}^5}{(t-s)^{2\theta}}\right) = \sup_N \mathbb{E}(\|\mathcal{M}^N\|_{C^\theta([0,T],H^{-\alpha})}^2) < \infty. \tag{2.26}$$

We are in a position to conclude (2.16) with η^N playing the role of u^N for any $\alpha > d/2+2$, and tightness of the sequence $(\eta^N)_{N\geqslant 1}$ follows. Indeed, combining (2.21)-(2.26) yields that

$$\begin{aligned}\sup_N \mathbb{E}(\|\eta^N\|_{C^\theta([0,T],H^{-\alpha})}^2) &= \sup_N \mathbb{E}\left(\sup_{0\leqslant s<t\leqslant T}\frac{\|\eta_t^N-\eta_s^N\|_{H^{-\alpha}}^2}{(t-s)^{2\theta}}\right)\\ &\lesssim_T \sup_N \sum_{i=1}^5 \mathbb{E}\left(\sup_{0\leqslant s<t\leqslant T}\frac{J_{s,t}^i}{(t-s)^{2\theta}}\right) < \infty,\end{aligned}$$

for any $\theta\in\left(0,\frac12\right)$. The result then follows.

Remark 2.17. *Careful readers may find that it suffices to assume* $|\sigma_N - \sigma| = o(\frac{1}{\sqrt{N}})$ *in order to obtain the tightness of* (η^N). *But we still adopt the assumption that* $|\sigma_N - \sigma| = \mathcal{O}\left(\frac{1}{N}\right)$ *in Assumption* **(A4)** *and Assumption* **(A5)** *since this is one of the assumptions used in [JW18] to obtain the uniform bound for* $H(F^N|\bar{\rho}^{\otimes N})$, *i.e. our Assumption* **(A3)**.

Define the topological space $\mathcal{X}$:

$$\mathcal{X} := \left\{ \bigcap_{k\in\mathbb{N}} \left[C([0,T], H^{-\frac{d}{2}-2-\frac{1}{k}}) \cap L^2([0,T], H^{-\frac{d}{2}-\frac{1}{k}}) \right] \right\} \times \left\{ \bigcap_{k\in\mathbb{N}} C([0,T], H^{-\frac{d}{2}-1-\frac{1}{k}}) \right\}.$$

The space $Y := \cap_{k\in\mathbb{N}} Y_k$ with $C([0,T], H^{-\frac{d}{2}-2-\frac{1}{k}}) \cap L^2([0,T], H^{-\frac{d}{2}-\frac{1}{k}})$ or $C([0,T], H^{-\frac{d}{2}-1-\frac{1}{k}})$ playing the role of Y_k is endowed with the metric $d_Y(f,g) = \sum_{k=1}^{\infty} 2^{-k}(1 \wedge \|f-g\|_{Y_k})$. Thus the convergence in Y is equivalent to the convergence in Y_k for every $k \in \mathbb{N}$. Moreover, $\mathcal{X}$ is a Polish space.

We then deduce the following result by the Skorokhod theorem.

Theorem 2.18. *There exists a subsequence of* $(\eta^N, \mathcal{M}^N)_{N\geqslant 1}$, *still denoted by* $(\eta^N, \mathcal{M}^N)$ *for simplicity, and a probability space* $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ *with* $\mathcal{X}$*-valued random variables* $(\tilde{\eta}, \tilde{M})$ *and* $(\tilde{\eta}^N, \tilde{\mathcal{M}}^N)_{N\geqslant 1}$ *such that*

1. *For each* $N \in \mathbb{N}$, *the law of* $(\tilde{\eta}^N, \tilde{\mathcal{M}}^N)$ *coincides with the law of* $(\eta^N, \mathcal{M}^N)$.
2. *The sequence of* $\mathcal{X}$*-valued random variables* $(\tilde{\eta}^N, \tilde{\mathcal{M}}^N)_{N\geqslant 1}$ *converges to* $(\tilde{\eta}, \tilde{\mathcal{M}})$ *in* $\mathcal{X}$ $\tilde{\mathbb{P}}$*-a.s.*

Proof By the Skorokhod theorem, the result follows by justifying the fact that the joint law of $(\eta^N, \mathcal{M}^N)_{N\geqslant 1}$ is tight on $\mathcal{X}$.

We start with proving the set A defined below is relatively compact in the space $C([0,T], H^{-\alpha-2}) \cap L^2([0,T], H^{-\alpha})$ for each $\alpha > d/2$,

$$A := \left\{ u \in K; \int_0^T \|u(t)\|^2_{H^{-\frac{2\alpha+d}{4}}} \mathrm{d}t \leqslant M, \right\},$$

where K is relatively compact in $C([0,T], H^{-\alpha-2})$. Suppose a sequence $\{u_n\} \subset A$, then there is a subsequence $\{u_{n_m}\}$ converging in $C([0,T], H^{-\alpha-2})$. On the other hand, by the Sobolev interpolation theorem [BCD11, Proposition 1.52], we find for $n, n' \in \mathbb{N}$ and $-\alpha - 2 < -\alpha < -\frac{2\alpha+d}{4}$

$$\begin{aligned}
&\int_0^T \|u_n(t) - u_{n'}(t)\|^2_{H^{-\alpha}} \mathrm{d}t \\
&\leqslant \int_0^T \|u_n(t) - u_{n'}(t)\|^{2\theta}_{H^{-\frac{2\alpha+d}{4}}} \|u_n(t) - u_{n'}(t)\|^{2(1-\theta)}_{H^{-\alpha-2}} \mathrm{d}t \\
&\leqslant \left(\int_0^T \|u_n(t) - u_{n'}(t)\|^2_{H^{-\frac{2\alpha+d}{4}}} \mathrm{d}t \right)^{\theta} \left(\int_0^T \|u_n(t) - u_{n'}(t)\|^2_{H^{-\alpha-2}} \mathrm{d}t \right)^{1-\theta}. \\
&\leqslant \left(\int_0^T \|u_n(t) - u_{n'}(t)\|^2_{H^{-\frac{2\alpha+d}{4}}} \mathrm{d}t \right)^{\theta} \left(T \sup_{t\in[0,T]} \|u_n(t) - u_{n'}(t)\|^2_{H^{-\alpha-2}} \right)^{1-\theta},
\end{aligned}$$

where the interpolation constant $\theta \in (0,1)$ depends on α and d. This implies the convergence of the subsequence $\{u_{n_m}\}$ in $L^2([0,T],H^{-\alpha})$ for each $\alpha > d/2$, and A is thus relatively compact in $C([0,T],H^{-\alpha-2}) \cap L^2([0,T],H^{-\alpha})$. For each $\epsilon > 0$, by (2.20) and Lemma 2.16, one can find M sufficiently large and a compact set K in $C([0,T],H^{-\alpha-2})$ such that

$$\begin{aligned}\mathbb{P}(\eta^N \notin A) \leqslant & P(\eta^N \notin K) + \mathbb{P}\Big(\int_0^T \|\eta^N(t)\|^2_{H^{-\frac{2\alpha+d}{4}}} \mathrm{d}t > M\Big) \\ \leqslant & P(\eta^N \notin K) + \frac{T}{M} \sup_{t\in[0,T]} \sup_N \mathbb{E}\|\eta^N(t)\|^2_{H^{-\frac{2\alpha+d}{4}}} < \epsilon,\end{aligned}$$

where the second line follows by Chebyshev's inequality. Therefore the sequence of laws of $(\eta^N)_{N\geqslant 1}$ is tight on $C([0,T],H^{-\alpha-2}) \cap L^2([0,T],H^{-\alpha})$ for every $\alpha > d/2$.

Furthermore, recall that Lemma 2.15 gives that the sequence of laws of $(\mathcal{M}^N)_{N\geqslant 1}$ is tight on $C([0,T],H^{-\alpha-1})$ for every $\alpha > d/2$. For each $\epsilon > 0$ and $k \in \mathbb{N}$, choose compact sets A_k^ϵ and B_k^ϵ in $C([0,T],H^{-\frac{d}{2}-\frac{1}{k}-2}) \cap L^2([0,T],H^{-\frac{d}{2}-\frac{1}{k}})$ and $C([0,T],H^{-\frac{d}{2}-\frac{1}{k}-1})$, respectively, such that

$$\mathbb{P}(\eta^N \notin A_k^\epsilon) < \epsilon 2^{-k}, \quad \mathbb{P}(\mathcal{M}^N \notin B_k^\epsilon) < \epsilon 2^{-k}, \quad \forall N \in \mathbb{N}.$$

Thus the set $A^\epsilon \times B^\epsilon$ in $\mathcal{X}$ defined by

$$A^\epsilon \times B^\epsilon := \Big(\bigcap_{k\in\mathbb{N}} A_k^\epsilon\Big) \times \Big(\bigcap_{k\in\mathbb{N}} B_k^\epsilon\Big)$$

is relatively compact and satisfies

$$\mathbb{P}\Big((\eta^N,\mathcal{M}^N) \notin A^\epsilon \times B^\epsilon\Big) \leqslant \sum_{k\in\mathbb{N}} \mathbb{P}(\eta^N \notin A_k^\epsilon) + \mathbb{P}(\mathcal{M}^N \notin B_k^\epsilon) < 2\epsilon, \quad \forall N \in \mathbb{N},$$

which shows the tightness of $(\eta^N,\mathcal{M}^N)_{N\geqslant 1}$ in $\mathcal{X}$.

Corollary 2.19. *For every $\alpha > d/2$, it holds that*

$$\tilde{\mathbb{E}} \int_0^T \|\tilde{\eta}_t^N - \tilde{\eta}_t\|_{H^{-\alpha}} \mathrm{d}t \xrightarrow{N\to\infty} 0. \tag{2.27}$$

Proof Notice that

$$\tilde{\mathbb{E}} \int_0^T \|\tilde{\eta}_t\|^2_{H^{-\alpha}} \mathrm{d}t \leqslant \sup_N \tilde{\mathbb{E}} \int_0^T \|\tilde{\eta}_t^N\|^2_{H^{-\alpha}} \mathrm{d}t \leqslant T \sup_{t\in[0,T]} \sup_N \tilde{\mathbb{E}}\|\tilde{\eta}_t^N\|^2_{H^{-\alpha}} < \infty, \quad \forall \alpha > \frac{d}{2},$$

which provides the uniform (in $[0,T]\times\Omega$) integrability of $\|\tilde{\eta}_t^N - \tilde{\eta}_t\|_{H^{-\alpha}}$, thus the convergence of $\|\tilde{\eta}_t^N - \tilde{\eta}_t\|_{H^{-\alpha}}$ $\mathrm{d}t \times \mathrm{d}\tilde{\mathbb{P}}$-a.e. leads to (2.27).

For each N, let $(\tilde{\mathcal{F}}_t^N)_{t\geqslant 0}$ and $(\tilde{\mathcal{F}}_t)_{t\geqslant 0}$ be the normal filtration generated by $(\tilde{\eta}^N, \tilde{\mathcal{M}}^N)$ and $(\tilde{\eta}, \tilde{\mathcal{M}})$, respectively. Then we have

$$\tilde{\mathcal{M}}_t^N = \tilde{\eta}_t^N - \tilde{\eta}_0^N - \sigma \int_0^t \Delta\tilde{\eta}_s^N \mathrm{d}s + \int_0^t \tilde{\mathcal{K}}_s^N \mathrm{d}s + \int_0^t \nabla(V\tilde{\eta}_s^N)\mathrm{d}s - \tilde{R}_t^N, \tag{2.28}$$

where $\tilde{\mathcal{K}}^N$ and $\tilde{R}^N$ are defined with $\tilde{\mu}_N := v + \frac{1}{\sqrt{N}}\tilde{\eta}^N$ and

$$\tilde{\mathcal{K}}_t^N := \sqrt{N}\nabla \cdot \Big(K * \tilde{\mu}_N(t)\tilde{\mu}_N(t) - K * v_t v_t\Big), \quad \tilde{R}_t^N := \sqrt{N}(\sigma_N - \sigma)\int_0^t \Delta\tilde{\mu}_N(s)\mathrm{d}s.$$

Here $\tilde{\mathcal{K}}^N$ is well-defined since $\tilde{\mu}_N$ is linear combination of Dirac measure and $K * \tilde{\mu}_N\tilde{\mu}_N$ is understood as

$$\langle K * \tilde{\mu}_N\tilde{\mu}_N, \varphi\rangle = \int_{\mathbb{T}^d\times\mathbb{T}^d} K(x-y)\varphi(x)\tilde{\mu}_N(\mathrm{d}x)\tilde{\mu}_N(\mathrm{d}y),$$

for $\varphi \in C^1$.

2.3.2 Characterization of the limit

In this section, we conclude that the original sequence $(\eta^N)_{N\geqslant 1}$ converges in distribution to the equation (1.7). Recall that the sequence $(\tilde{\eta}^N)_{N\geqslant 1}$ converges in $C([0,T], H^{-\alpha-2}) \cap L^2([0,T], H^{-\alpha})$ $\tilde{\mathbb{P}}$-a.s. for $\alpha > d/2$ and shares the same distribution with a subsequence of $(\eta^N)_{N\geqslant 1}$. Hence it is sufficient to justify two facts. One is that each limit $\tilde{\eta}$ is a martingale solution to (1.7). The other is that the law of the solution to (1.7) is unique, which would follow by pathwise uniqueness and the Yamada-Watanabe theorem.

Throughout this section, we always assume **(A1)-(A4)**.

Identifying the limit of the interacting term $\tilde{\mathcal{K}}^N$ is one of the main difficulties in this chapter, it derseves to be treated separately from other terms in the decomposition (1.26). The following lemma identifies the limit of the interacting term $\tilde{\mathcal{K}}^N$. The idea of the proof is to split the interacting term into some regular part and a term in the form of a function of $\mu_N - v$, which can be controlled in Lemma 2.14 by the techniques developed in Section 2.2.

Lemma 2.20. *For each $\varphi \in C^\infty(\mathbb{T}^d)$, it holds that*

$$\tilde{\mathbb{E}}\left(\sup_{t\in[0,T]}\left|\int_0^t \tilde{\mathcal{K}}_s^N(\varphi) - \langle v_s K * \tilde{\eta}_s + \tilde{\eta}_s K * v_s, \nabla\varphi\rangle \mathrm{d}s\right|\right) \xrightarrow{N\to\infty} 0.$$

Proof Direct computations give the following identity

$$\sqrt{N}\big(\tilde{\mu}_N K * \tilde{\mu}_N - vK * v\big) = vK * \eta^N + \eta^N K * v + \frac{1}{\sqrt{N}}\eta^N K * \eta^N.$$

Consequently, for each $\varphi \in C^\infty$,

$$\sup_{t\in[0,T]}\left|\int_0^t \tilde{\mathcal{K}}_s^N(\varphi) - \langle v_s K * \tilde{\eta}_s + \tilde{\eta}_s K * v_s, \nabla\varphi\rangle \mathrm{d}s\right| \leqslant J_1^N(\varphi) + J_2^N(\varphi), \tag{2.29}$$

where

$$\begin{aligned} J_1^N(\varphi) :=& \sqrt{N}\int_0^T |\langle \nabla\varphi K * (\tilde{\mu}_N(t) - v_t), \tilde{\mu}_N(t) - v_t\rangle|\mathrm{d}t, \\ J_2^N(\varphi) :=& \int_0^T |\langle v_t K * \tilde{\eta}_t^N + \tilde{\eta}_t^N K * v_t, \nabla\varphi\rangle - \langle v_t K * \tilde{\eta}_t + \tilde{\eta}_t K * v_t, \nabla\varphi\rangle|\mathrm{d}t. \end{aligned}$$

On one hand, we deduce from Lemma 2.14 that

$$\begin{aligned}\tilde{\mathbb{E}}J_1^N(\varphi) &\leqslant T\sqrt{N}\sup_{t\in[0,T]}\tilde{\mathbb{E}}|\langle\nabla\varphi K*(\tilde{\mu}_N(t)-v_t),\tilde{\mu}_N(t)-v_t\rangle|\\ &= T\sqrt{N}\sup_{t\in[0,T]}\mathbb{E}|\langle\nabla\varphi K*(\mu_N(t)-v_t),\mu_N(t)-v_t\rangle|\\ &\lesssim N^{-\frac{1}{2}}\sup_N\sup_{t\in[0,T]}(H(F^N|v_t^{\otimes N})+1)\xrightarrow{N\to\infty}0,\end{aligned}$$

where the limit follows by Assumption **(A3)**. On the other hand, we find

$$\tilde{\mathbb{E}}J_2^N(\varphi)\leqslant\tilde{\mathbb{E}}\int_0^T|\langle v_tK*(\tilde{\eta}_t^N-\tilde{\eta}_t),\nabla\varphi\rangle|+|\langle(\tilde{\eta}_t^N-\tilde{\eta}_t)K*v_t,\nabla\varphi\rangle|\mathrm{d}t. \tag{2.30}$$

For each $t\in[0,T]$, it holds for every $\alpha\in(d/2,\beta)$ that

$$\begin{aligned}|\langle v_tK*(\tilde{\eta}_t^N-\tilde{\eta}_t),\nabla\varphi\rangle| &= |\langle K(-\cdot)*(v_t\nabla\varphi),\tilde{\eta}_t^N-\tilde{\eta}_t\rangle|\\ &\leqslant\|\tilde{\eta}_t^N-\tilde{\eta}_t\|_{H^{-\alpha}}\|K(-\cdot)*(v_t\nabla\varphi)\|_{H^\alpha},\\ |\langle(\tilde{\eta}_t^N-\tilde{\eta}_t)K*v_t,\nabla\varphi\rangle| &\leqslant\|\tilde{\eta}_t^N-\tilde{\eta}_t\|_{H^{-\alpha}}\|\nabla\varphi\cdot K*v_t\|_{H^\alpha},\end{aligned}$$

where

$$K(-\cdot)*g(x):=\int K(y-x)g(y)\mathrm{d}y. \tag{2.31}$$

Applying Lemma 2.4 with $p=p_1=q=2$ and Lemma 2.3 yields that

$$\begin{aligned}|\langle v_tK*(\tilde{\eta}_t^N-\tilde{\eta}_t),\nabla\varphi\rangle| &\lesssim\|\tilde{\eta}_t^N-\tilde{\eta}_t\|_{H^{-\alpha}}\|K\|_{L^1}(\|v_t\|_{H^\alpha}\|\nabla\varphi\|_{L^\infty}+\|v_t\|_{L^\infty}\|\nabla\varphi\|_{H^\alpha}),\\ |\langle(\tilde{\eta}_t^N-\tilde{\eta}_t)K*v_t,\nabla\varphi\rangle| &\lesssim\|\tilde{\eta}_t^N-\tilde{\eta}_t\|_{H^{-\alpha}}\|K\|_{L^1}(\|v_t\|_{H^\alpha}\|\nabla\varphi\|_{L^\infty}+\|v_t\|_{L^\infty}\|\nabla\varphi\|_{H^\alpha}).\end{aligned}$$

Here the fact $K\in L^1$ follows by Assumption **(A2)**. Then taking these two estimates into (2.30), applying Sobolev embedding $H^\alpha\hookrightarrow L^\infty$ with $\alpha>d/2$, we thus arrive at

$$\tilde{\mathbb{E}}J_2^N(\varphi)\lesssim_\varphi\|K\|_{L^1}\sup_{t\in[0,T]}\|v_t\|_{H^\alpha}\mathbb{E}\int_0^T\|\tilde{\eta}_t^N-\tilde{\eta}_t\|_{H^{-\alpha}}\mathrm{d}t\xrightarrow{N\to\infty}0,$$

where the limit follows by (2.27). Using inequality (2.29) and $\mathbb{E}J_1^N(\varphi)\to 0$, the proof is completed.

Remark 2.21. *One may easily find that in the "identifying the limit" part, we only need to assume that the relative entropy grow slower than the order $\sqrt{N}$, i.e. $H(F^N|v^{\otimes N})=o(\sqrt{N})$ as $N\to\infty$. However, in the tightness part we need a stronger assumption, namely our Assumption* **(A3)***. As a separate question, it would be interesting to show whether or not there exists some symmetric probability measure $F^N\in\mathcal{P}_{Sym}(S^N)$ such that $H(F^N|\bar{\rho}^{\otimes N})=N^\theta$ with $\theta\in(0,1)$, where $\bar{\rho}$ is a given probability measure on the Polish space S.*

Now we are in the position to conclude that $\tilde{\eta}$ solves (1.7).

Theorem 2.22. *The limit $\tilde{\eta}$ is a martingale solution to* (1.7) *in the sense of Definition 1.1.*

Proof We deduce from (2.28) that

$$\begin{aligned}\tilde{\mathcal{M}}_t^N(\varphi) =&\langle \tilde{\eta}_t^N, \varphi\rangle - \langle \tilde{\eta}_0^N, \varphi\rangle - \sigma \int_0^t \langle \Delta\varphi, \tilde{\eta}_s^N\rangle \mathrm{d}s - \int_0^t \tilde{\mathcal{K}}_s^N(\varphi)\mathrm{d}s - \int_0^t \langle \nabla\varphi, V\tilde{\eta}_s^N\rangle \mathrm{d}s\\ &- \sqrt{N}(\sigma_N - \sigma)\int_0^t \langle \Delta\varphi, \tilde{\mu}_N(s)\rangle \mathrm{d}s,\end{aligned}$$

for each $\varphi \in C^\infty(\mathbb{T}^d)$ and $t \in [0,T]$. By Lemma 2.20, $\sigma_N - \sigma = \mathcal{O}\left(\frac{1}{N}\right)$, and the fact that $\tilde{\eta}^N$ converges to $\tilde{\eta}$ in $C([0,T], H^{-\alpha-2}) \cap L^2([0,T], H^{-\alpha})$ for every $\alpha > d/2$ $\tilde{\mathbb{P}}$-a.s., one can take limit of every term above on both sides and have

$$\tilde{\mathcal{M}}_t(\varphi) = \langle \tilde{\eta}_t, \varphi\rangle - \langle \tilde{\eta}_0, \varphi\rangle - \sigma\int_0^t \langle \Delta\varphi, \tilde{\eta}_s\rangle \,\mathrm{d}s - \int_0^t \langle \nabla\varphi, v_s K * \tilde{\eta}_s + \tilde{\eta}_s K * v_s + V\tilde{\eta}_s\rangle \,\mathrm{d}s,$$

$\tilde{\mathbb{P}}-$ almost surely. To indentify $\tilde{\eta}$ is a martingale solution, we need to justify properties of $\tilde{\mathcal{M}}$. Since $\tilde{\mathcal{M}}^N$ are centered Gaussian process and by Theorem 2.18, the limit $\tilde{\mathcal{M}}$ is a centered Gaussian process with values in $H^{-\alpha-1}$ for every $\alpha > d/2$.

As for the covariance functions, on one hand, applying Burkholder-Davis-Gundy's inequality, we have for each $1 < \theta \leqslant 2$

$$\begin{aligned}\sup_N \tilde{\mathbb{E}}[\sup_{t\in[0,T]} |\tilde{\mathcal{M}}_t^N(\varphi)|^{2\theta}] = \sup_N \mathbb{E}[\sup_{t\in[0,T]} |\mathcal{M}_t^N(\varphi)|^{2\theta}]&\\ \lesssim \sup_N \mathbb{E}\left(\int_0^T \sum_{i=1}^N \frac{\sigma_N}{N}|\nabla\varphi(X_i)|^2 \,\mathrm{d}t\right)^\theta&\\ = \sup_N \mathbb{E}\left(\int_0^T \sigma_N\langle |\nabla\varphi|^2, \mu_N(t)\rangle \mathrm{d}t\right)^\theta&\\ \lesssim_\varphi \sup_N \sigma_N^\theta T^\theta \|\nabla\varphi\|_{L^\infty}^{2\theta} < \infty.&\end{aligned}$$

This implies uniform integrability of $|\tilde{\mathcal{M}}_t^N(\varphi)|^2$ for each $t \in [0,T]$.

On the other hand, we have that $|\tilde{\mathcal{M}}_t^N(\varphi_1)\tilde{\mathcal{M}}_s^N(\varphi_2)|$ converges to $|\tilde{\mathcal{M}}_t(\varphi_1)\tilde{\mathcal{M}}_s(\varphi_2)|$ $\tilde{\mathbb{P}}$-a.s. for $s, t \in [0,T]$ and $\varphi_1, \varphi_2 \in C^\infty$. Thus by the uniform integrability of $|\tilde{\mathcal{M}}_t^N(\varphi)|^2$ and

$$|\tilde{\mathcal{M}}_t^N(\varphi_1)\tilde{\mathcal{M}}_s^N(\varphi_2)| \leqslant |\tilde{\mathcal{M}}_t^N(\varphi_1)|^2 + |\tilde{\mathcal{M}}_s^N(\varphi_2)|^2,$$

we arrive at

$$\tilde{\mathbb{E}}[\tilde{\mathcal{M}}_t(\varphi_1)\tilde{\mathcal{M}}_s(\varphi_2)] = \lim_{N\to\infty} \tilde{\mathbb{E}}[\tilde{\mathcal{M}}_t^N(\varphi_1)\tilde{\mathcal{M}}_s^N(\varphi_2)] = \lim_{N\to\infty} \mathbb{E}[\mathcal{M}_t^N(\varphi_1)\mathcal{M}_s^N(\varphi_2)].$$

Furthermore, using (2.4) and Ito's isometry we obtain that

$$\begin{aligned}\mathbb{E}[\mathcal{M}_t^N(\varphi_1)\mathcal{M}_s^N(\varphi_2)] =&\frac{2\sigma_N}{N}\mathbb{E}\left[\left(\sum_{i=1}^N \int_0^t \nabla\varphi_1(X_i)\mathrm{d}B_r^i\right)\left(\sum_{i=1}^N \int_0^s \nabla\varphi_2(X_i)\mathrm{d}B_r^i\right)\right]\\ =&\frac{2\sigma_N}{N}\mathbb{E}\left[\sum_{i=1}^N \int_0^{s\wedge t} \nabla\varphi_1(X_i)\nabla\varphi_2(X_i)\mathrm{d}r\right]\\ =&2\sigma_N \int_0^{s\wedge t} \mathbb{E}\langle \nabla\varphi_1 \cdot \nabla\varphi_2, \mu_N(r)\rangle \mathrm{d}r.\end{aligned}$$

The proof is thus completed since

$$2\sigma_N \int_0^{s\wedge t} \mathbb{E}\langle \nabla\varphi_1 \cdot \nabla\varphi_2, \mu_N(r)\rangle \mathrm{d}r \xrightarrow{N\to\infty} 2\sigma \int_0^{s\wedge t} \langle \nabla\varphi_1 \cdot \nabla\varphi_2, v_r\rangle \mathrm{d}r.$$

The rest of this subsection is devoted to obtain the well-posedness of the SPDE (1.7), and finish the proof of Theorem 1.4.

Let us first introduce an equivalent definition of martingale solutions to (1.7), which is used in the proof of Theorem 1.4. For notations' simplicity, we omit the tildes in the following.

Definition 2.23. *We call $(\eta, \mathcal{M})$ a probabilistically weak solution to* (1.7) *on stochastic basis $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ with initial data η_0 if*

1. *η is a continuous $(\mathcal{F}_t)$-adapted process with values in $H^{-\alpha-2}$ and $\eta \in L^2([0,T], H^{-\alpha})$ for every $\alpha > d/2$, $\mathbb{P}$-a.s.*

2. *$\mathcal{M}$ is a continuous $(\mathcal{F}_t)$-adapted centered Gaussian process with values in $H^{-\alpha-1}$ for every $\alpha > d/2$, with covariance given by*

$$\mathbb{E}[\mathcal{M}_t(\varphi_1)\mathcal{M}_s(\varphi_2)] = 2\sigma \int_0^{s\wedge t} \langle \nabla\varphi_1 \cdot \nabla\varphi_2, v_r\rangle \mathrm{d}r, \tag{2.32}$$

for each $\varphi_1, \varphi_2 \in C^\infty$ and $s, t \in [0,T]$.

3. *For each $\varphi \in C^\infty(\mathbb{T}^d)$ and $t \in [0,T]$, it holds that*

$$\begin{aligned}\mathcal{M}_t(\varphi) =& \langle \eta_t, \varphi\rangle - \langle \eta_0, \varphi\rangle - \int_0^t \langle \sigma\Delta\varphi, \eta\rangle \mathrm{d}s - \int_0^t \langle \nabla\varphi, vK * \eta\rangle \mathrm{d}s \\ &- \int_0^t \langle \nabla\varphi, \eta K * v\rangle \mathrm{d}s - \int_0^t \langle \nabla\varphi, V\eta\rangle \mathrm{d}s.\end{aligned}$$

Furthermore, given a centered Gaussian process $\mathcal{M}$ on stochastic basis $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ with covariance characterized by (2.32)*, we call η is a probabilistically strong solution to* (1.5) *if $(\eta, \mathcal{M})$ is a probabilisttically weak solution and η is adapted to the normal filtration generated by $\mathcal{M}$.*

Uniqueness in law of the solutions to (1.7) usually follows by the Yamada-Watanabe theorem, which requires existence of probabilistically weak solutions and pathwise uniqueness. Since the martingale solutions and the probabilistically weak solutions are equivalent, Theorem 2.22 means that there exists a stochastic basis $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ such that $(\eta, \mathcal{M})$ is a probabilistically weak solution to (1.7), it thus suffices to prove the pathwise uniqueness.

We now briefly explain the concept of pathwise uniqueness of probabilistically weak solutions introduced before. Equation (1.7) can be viewed as a system, for which the information of the initial data and the nosie is given (i.e. the distribution of $(\mathcal{M}, \eta_0)$ is fixed), and $(\mathcal{M}, \eta_0)$ can be seen as the input and η is the output. Pathwise uniqueness means that if on some fixed stochastic basis there exist two outputs η and $\tilde{\eta}$ with given η_0 and $\mathcal{M}$, then η coincides with $\tilde{\eta}$ $\mathbb{P}$-a.s..

Notice that the covariance function of $\mathcal{M}$ and Assumption **(A1)** have determined the distribution of $(\mathcal{M}, \eta_0)$. Since equation (1.7) is linear and is driven by additive noise, pathwise uniqueness of solutions to the equation (1.7) follows from uniqueness of solutions to the following PDE

$$\partial_t u = \sigma\Delta u - \nabla\cdot(vK * u) - \nabla\cdot(uK * v) - \nabla\cdot(Vu), \quad u_0 = 0. \tag{2.33}$$

Lemma 2.24. *Under the assumptions* **(A2)** *and* **(A4)** *with parameter β, for each $\alpha \in (d/2, \beta)$, $u \equiv 0$ is the only solution with zero initial value to* (2.33) *in the sense that*

1. *$u \in L^2([0,T], H^{-\alpha}) \cap C([0,T], H^{-\alpha-2})$.*
2. *For each $\varphi \in C^\infty$ and $t \in [0,T]$,*

$$\langle u_t, \varphi\rangle = \int_0^t \langle \sigma u_s, \Delta\varphi\rangle \mathrm{d}s + \int_0^t \langle v_s K * u_s + u_s K * v_s + V u_s, \nabla\varphi\rangle \mathrm{d}s.$$

Proof Testing u with the Fourier basis $\{e_k\}_{k\in\mathbb{Z}^d}$, then we find for every $t \in [0,T]$ and $k \in \mathbb{Z}^d$,

$$\begin{aligned}\partial_t |\langle u_t, e_k\rangle|^2 =& -2\sigma|k|^2\langle u_t, e_k\rangle\langle u_t, e_{-k}\rangle + \langle u_t, e_{-k}\rangle[J_t^1(k) + J_t^2(k)] \\ &+ \langle u_t, e_k\rangle[J_t^1(-k) + J_t^2(-k)] + \sqrt{-1}k\langle u_t, e_{-k}\rangle\langle V u_t, e_k\rangle \\ &- \sqrt{-1}k\langle u_t, e_k\rangle\langle V u_t, e_{-k}\rangle,\end{aligned} \tag{2.34}$$

where $J_t^1(k)$ and $J_t^2(k)$, for each $k \in \mathbb{Z}^d$, are defined by

$$\begin{aligned}J_t^1(k) :=& \sqrt{-1}k\int_{\mathbb{T}^d} K * u_t(x) e_k(x) v_t(x)\mathrm{d}x, \\ J_t^2(k) :=& \sqrt{-1}k\int_{\mathbb{T}^d} K * v_t(x) e_k(x) u_t(x)\mathrm{d}x.\end{aligned}$$

Integrating (2.34) over time, summing up over k with weight $\langle k\rangle^{-2\alpha-2}$, and applying Young's inequality yields that there exists a constant C_ϵ for each $\epsilon > 0$ such that

$$\begin{aligned}&\sum_{k\in\mathbb{Z}^d}\langle k\rangle^{-2\alpha-2}|\langle u_t, e_k\rangle|^2 + 2\sigma\sum_{k\in\mathbb{Z}^d}\langle k\rangle^{-2\alpha}\int_0^t |\langle u_t, e_k\rangle|^2\mathrm{d}s \\ \leqslant& C_\epsilon\sum_{k\in\mathbb{Z}^d}\langle k\rangle^{-2\alpha-2}\int_0^t |\langle u_s, e_k\rangle|^2\mathrm{d}s + \epsilon\sum_{k\in\mathbb{Z}^d}\langle k\rangle^{-2\alpha-2}\int_0^t |J_s^1(-k) + J_s^2(-k)|^2\mathrm{d}s \\ &+ \epsilon\sum_{k\in\mathbb{Z}^d}\langle k\rangle^{-2\alpha-2}|k|^2\int_0^t |\langle V u_s, e_k\rangle|^2\mathrm{d}s.\end{aligned} \tag{2.35}$$

To make (2.35) suitable for applying Gronwall's lemma, we first find estimates related to $J_t^1(k)$ and $J_t^2(k)$,

$$\sum_{k\in\mathbb{Z}^d}\langle k\rangle^{-2\alpha-2}|J^1(k)|^2 = \sum_{k\in\mathbb{Z}^d}\langle k\rangle^{-2\alpha-2}|k|^2\langle K * uv, e_k\rangle\langle K * uv, e_{-k}\rangle \leqslant \|K * uv\|_{H^{-\alpha}}^2,$$

and

$$\sum_{k\in\mathbb{Z}^d}\langle k\rangle^{-2\alpha-2}|J^2(k)|^2 = \sum_{k\in\mathbb{Z}^d}\langle k\rangle^{-2\alpha-2}|k|^2\langle K * vu, e_k\rangle\langle K * vu, e_{-k}\rangle \leqslant \|K * vu\|_{H^{-\alpha}}^2.$$

Then applying Lemmas 2.1 and 2.2 gives that

$$\|K * uv\|_{H^{-\alpha}} \leqslant C_\alpha\|K * u\|_{H^{-\alpha}}\|v\|_{C^\beta}, \quad \|K * vu\|_{H^{-\alpha}} \leqslant C_\alpha\|u\|_{H^{-\alpha}}\|K * v\|_{C^\beta}.$$

Furthermore, by Lemma 2.4, we deduce

$$\sum_{k\in\mathbb{Z}^d}\langle k\rangle^{-2\alpha-2}|J^1(k)|^2+\sum_{k\in\mathbb{Z}^d}\langle k\rangle^{-2\alpha-2}|J^2(k)|^2\leqslant C_\alpha\|u\|^2_{H^{-\alpha}}\|K\|^2_{L^1}\|v\|^2_{C^\beta}. \tag{2.36}$$

Similarly, we obtain

$$\sum_{k\in\mathbb{Z}^d}\langle k\rangle^{-2\alpha-2}|k|^2|\langle Vu,e_k\rangle|^2\leqslant\|Vu\|^2_{H^{-\alpha}}\leqslant C_\alpha\|V\|^2_{C^\beta}\|u\|^2_{H^{-\alpha}}. \tag{2.37}$$

Since $u\in L^2([0,T],H^{-\alpha})$, we obtain $\partial_t u\in L^2([0,T],H^{-\alpha-2})$, which by Lions-Magenes Lemma implies $u\in C([0,T],H^{-\alpha-1})$. Combining (2.35)-(2.37) leads to

$$\begin{aligned}&\|u_t\|^2_{H^{-\alpha-1}}+2\sigma\int_0^t\|u_s\|^2_{H^{-\alpha}}\mathrm{d}s\\ \leqslant&C_\epsilon\int_0^t\|u_s\|^2_{H^{-\alpha-1}}\mathrm{d}s+\epsilon C_\alpha\int_0^t\left(\|K\|^2_{L^1}\|v_s\|^2_{C^\beta}+\|V\|^2_{C^\beta}\right)\|u_s\|^2_{H^{-\alpha}}\mathrm{d}s. \end{aligned}\tag{2.38}$$

Choosing ϵ such that

$$\epsilon C_\alpha\|K\|^2_{L^1}\Big(\sup_{s\in[0,t]}\|v_s\|^2_{C^\beta}+\|V\|^2_{C^\beta}\Big)<2\sigma,$$

then using Gronwall's inequality gives

$$\|u_t\|^2_{H^{-\alpha-1}}+\int_0^t\|u_s\|^2_{H^{-\alpha}}\mathrm{d}s=0.$$

This completes the proof.

Proof [Proof of Theorem 1.4] We have proved that the sequence of laws of $\{\eta^N\}_{N\in\mathbb{N}}$ is tight and every tight limit is a martingale solution to (1.7) (Theorem 2.22). As a result, existence of martingale solutions (equivalently probabilistically weak solutions) follows. On the other hand, Lemma 2.24 together with Corollary 2.12 implies pathwise uniqueness of probabilistically weak solutions. Then applying the general Yamada-Watanabe theorem [Kur14, Theorem 1.5] gives that the law of martingale solutions starting from the same initial distribution is unique, and every probabilistically weak solution is a probabilistically strong solution. Therefore η^N converges in distribution to the unique (in distribution) martingale solution η.

Remark 2.25. *From the proof of Theorem 1.4, we also obtain the well-posedness of probabilistically strong solutions to the SPDE* (1.7).

2.4 Gaussianity and optimal regularity

In Section 2.4.1, we prove the optimal regularity of solutions to the limit SPDE (1.7). Then, the proof of Proposition 1.5, which gives the Gaussianity of the unique limit of fluctuation measures, is given in Section 2.4.2.

2.4.1 Optimal regularity

In this section we improve the regularity of η by using the mild formulation and the smooth effect of the heat kernel.

Recall that $\mathcal{M}$ is a centered Gaussian process with covariance given by

$$\mathbb{E}[\mathcal{M}_t(\varphi_1)\mathcal{M}_s(\varphi_2)] = 2\sigma \int_0^{s\wedge t} \langle \nabla\varphi_1 \cdot \nabla\varphi_2, v_r\rangle \mathrm{d}r,$$

for $\varphi_1, \varphi_2 \in C^\infty(\mathbb{T}^d)$. Therefore, the distribution of $\mathcal{M}$ is uniquely determined, and one can regard $\mathcal{M}$ as $\nabla \cdot \int_0^\cdot \sqrt{v}\xi(\mathrm{d}s, \mathrm{d}x)$ with $\xi = (\xi^i)_{i=1}^d$ being vector valued space-time white noise on $\mathbb{R}^+ \times \mathbb{T}^d$. In fact, for every $\varphi \in C^\infty$,

$$\mathcal{M}_t(\varphi) \stackrel{d}{=} -\sqrt{2\sigma} \int_0^t \int_{\mathbb{T}^d} \nabla\varphi(x)\sqrt{v_s(x)}\xi(\mathrm{d}s, \mathrm{d}x), \tag{2.39}$$

where $\stackrel{d}{=}$ means equal in distribution and we omit the inner product in $\mathbb{R}^d$ between ξ and $\nabla\varphi$. We start with investigating the regularity of a stochastic integral, which will be the stochastic term in the mild form of equation (1.7). Define a stochastic process Z as

$$Z_t := \int_0^t \int_{\mathbb{T}^d} \nabla\Gamma_{t-s}(\cdot - y)\sqrt{v_s(y)}\xi(\mathrm{d}s, \mathrm{d}y), \tag{2.40}$$

where Γ is the heat kernel of $\sigma\Delta$ on $\mathbb{T}^d$.

Recall that $\{\chi_n\}_{n\geqslant -1}$ is the Littlewood-Paley partition functions and $\chi_n(\cdot) = \chi_0(2^{-n}\cdot)$ for $n \geqslant 0$ (see Section 2.1.1). Denote ψ_n be the inverse Fourier transform of χ_n for every n, we then have the following result.

Lemma 2.26. *For each $\kappa > 0$ and every $n \geqslant -1$, it holds for all $s, t \in [0, T]$ and $x \in \mathbb{T}^d$ that*

$$\mathbb{E}\,|\langle Z_t, \psi_n(\cdot - x)\rangle|^2 \lesssim 2^{dn},$$
$$\mathbb{E}\,|\langle Z_t, \psi_n(\cdot - x)\rangle - \langle Z_s, \psi_n(\cdot - x)\rangle|^2 \lesssim 2^{dn+2\kappa n}(t-s)^\kappa,$$

where the proportional constants depend on $\|v\|_{C_T L^\infty}$.

Proof For simplicity we set $\sigma = 1$ in the proof. First, we use Fourier transform to represent $\langle Z_t, \psi_n(\cdot - x)\rangle$ as follows,

$$\begin{aligned}
\langle Z_t, \psi_n(\cdot - x)\rangle &= \int_0^t \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \nabla\Gamma_{t-r}(y - z)\sqrt{v_r(z)}\psi_n(y - x)\mathrm{d}y\xi(\mathrm{d}r, \mathrm{d}z) \\
&= \int_0^t \int_{\mathbb{T}^d} \langle \nabla\Gamma_{t-r}(\cdot - z), \psi_n(\cdot - x)\rangle \sqrt{v_r(z)}\xi(\mathrm{d}r, \mathrm{d}z) \\
&= \int_0^t \int_{\mathbb{T}^d} \sum_{k\in\mathbb{Z}^d} G_{t-r}(k)e_{-k}(z)\sqrt{v_r(z)}\xi(\mathrm{d}r, \mathrm{d}z),
\end{aligned} \tag{2.41}$$

where $G_t(k)$ is defined by

$$G_t(k) := \int_{\mathbb{T}^d} \langle \nabla\Gamma_t(\cdot - z'), \psi_n(\cdot - x)\rangle\, e_k(z')\mathrm{d}z'$$

and we used $\langle \nabla\Gamma_{t-r}(\cdot - z), \psi_n(\cdot - x)\rangle \in L^2(\mathbb{T}^d)$ and the sum in (2.41) converges in $L^2(\mathbb{T}^d)$. Furthermore, noticing that $G_t(-k)$ is the complex conjugate of $G_t(k)$, we thus have

$$\begin{aligned}
\mathbb{E}\,|\langle Z_t, \psi_n(\cdot - x)\rangle|^2 &= \int_0^t \int_{\mathbb{T}^d} \Big| \sum_{k\in\mathbb{Z}^d} G_{t-r}(k) e_{-k}(z) \Big|^2 v_r(z)\mathrm{d}z\mathrm{d}r \\
&\lesssim \|v\|_{C_T L^\infty} \sum_{k_1\in\mathbb{Z}^d} \sum_{k_2\in\mathbb{Z}^d} \int_0^t \int_{\mathbb{T}^d} G_{t-r}(k_1) G_{t-r}(-k_2) e_{-k_1}(z) e_{k_2}(z)\mathrm{d}z\mathrm{d}r \\
&\lesssim \|v\|_{C_T L^\infty} \sum_{k\in\mathbb{Z}^d} \int_0^t G_{t-r}(k) G_{t-r}(-k)\mathrm{d}r, \qquad (2.42)
\end{aligned}$$

where the last inequality follows by $\int_{\mathbb{T}^d} e_{k_2-k_1}(z)\mathrm{d}z = C_d \delta_{k_2=k_1}$, C_d is the volume of $\mathbb{T}^d$.

For each $k \in \mathbb{Z}^d$ and $t \in [0,T]$, we find that

$$\begin{aligned}
G_t(k) &= \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \nabla\Gamma_t(y - z')\psi_n(y - x) e_k(z')\mathrm{d}y\mathrm{d}z' \\
&= \langle \nabla\Gamma_t, e_{-k}\rangle \langle \psi_n, e_k\rangle e_k(x) \\
&= -\sqrt{-1} k e^{-t|k|^2} \chi_n(-k) e_k(x). \qquad (2.43)
\end{aligned}$$

Here we used the facts that Γ is the heat kernel on $\mathbb{T}^d$ and $\langle \psi_n, e_{-k}\rangle = \chi_n(k)$.

Combining (2.42) and (2.43) yields that

$$\mathbb{E}\,|\langle Z_t, \psi_n(\cdot - x)\rangle|^2 \lesssim \|v\|_{C_T L^\infty} \sum_{k\in\mathbb{Z}^d} \int_0^t |k|^2 e^{-2|k|^2(t-r)} \chi_n(-k)\chi_n(k)\mathrm{d}r.$$

Notice that

$$\int_0^t |k|^2 e^{-2|k|^2(t-r)}\mathrm{d}r = \frac{1}{2}\Big(1 - e^{-2|k|^2 t}\Big),$$

which implies that for $n \geqslant 0$

$$\begin{aligned}
\mathbb{E}\,|\langle Z_t, \psi_n(\cdot - x)\rangle|^2 &\lesssim_v \sum_{k\in\mathbb{Z}^d} \chi_n(-k)\chi_n(k)\Big(1 - e^{-2|k|^2 t}\Big) \\
&\lesssim_v 2^{dn} \int \chi_0(-k')\chi_0(k')\Big(1 - e^{-2^{2n+1}|k'|^2 t}\Big)\mathrm{d}k' \lesssim_v 2^{dn}. \qquad (2.44)
\end{aligned}$$

Here we used $\chi_n(\cdot) = \chi_0(2^{-n}\cdot)$ and the fact that χ_0 is of compact support. The case $n = -1$ is similar.

Next, we deduce by (2.41) that

$$\begin{aligned}
&\langle Z_t, \psi_n(\cdot - x)\rangle - \langle Z_s, \psi_n(\cdot - x)\rangle \\
&= \int_0^t \int_{\mathbb{T}^d} \sum_{k\in\mathbb{Z}^d} G_{t-r}(k) e_{-k}(z)\sqrt{v_r(z)}\xi(\mathrm{d}r, \mathrm{d}z) - \int_0^s \int_{\mathbb{T}^d} \sum_{k\in\mathbb{Z}^d} G_{s-r}(k) e_{-k}(z)\sqrt{v_r(z)}\xi(\mathrm{d}r, \mathrm{d}z) \\
&= \int_s^t \int_{\mathbb{T}^d} \sum_{k\in\mathbb{Z}^d} G_{t-r}(k) e_{-k}(z)\sqrt{v_r(z)}\xi(\mathrm{d}r, \mathrm{d}z) \\
&\quad + \int_0^s \int_{\mathbb{T}^d} \sum_{k\in\mathbb{Z}^d} \Big[G_{t-r}(k) - G_{s-r}(k)\Big] e_{-k}(z)\sqrt{v_r(z)}\xi(\mathrm{d}r, \mathrm{d}z)
\end{aligned}$$

$:= J_1^n + J_2^n.$

Moreover, we have

$$\mathbb{E}\,|\langle Z_t, \psi_n(\cdot - x)\rangle - \langle Z_s, \psi_n(\cdot - x)\rangle|^2 \leqslant 2\mathbb{E}|J_1^n|^2 + 2\mathbb{E}|J_2^n|^2. \tag{2.45}$$

Again, it suffices to check the cases with $n \geqslant 0$. Similar as in (2.44), we have

$$\begin{aligned}
\mathbb{E}|J_1^n|^2 \lesssim_v & \sum_{k\in\mathbb{Z}^d} \int_s^t |k|^2 e^{-2|k|^2(t-r)} \chi_n(k)\chi_n(-k)\mathrm{d}r \\
\lesssim_v & \sum_{k\in\mathbb{Z}^d} \Big(1 - e^{-2|k|^2(t-s)}\Big)\chi_n(k)\chi_n(-k) \\
\lesssim_v & 2^{dn} \int \Big(1 - e^{-2^{2n+1}|k'|^2(t-s)}\Big)\chi_0(k')\chi_0(-k')\mathrm{d}k' \;\lesssim_v 2^{dn}(1 - e^{-C2^{2n}(t-s)}),
\end{aligned}$$

where $C > 0$ is a universal constant determinded by the support of χ_0. Notice that for each $\kappa > 0$, it holds that $1 - e^{-a} \lesssim a^\kappa$ for $a \geqslant 0$. Therefore, for each $\kappa > 0$, let $C2^{2n}(t-s)$ in the above inequality play the role of a, we arrive at

$$\mathbb{E}|J_1^n|^2 \lesssim_v 2^{dn+2\kappa n}(t-s)^\kappa. \tag{2.46}$$

Similarly, one can study J_2^n and find

$$\begin{aligned}
\mathbb{E}|J_2^n|^2 \lesssim_v & \sum_{k\in\mathbb{Z}^d} \int_0^s |k|^2 \chi_n(k)\chi_n(-k)\Big(e^{-|k|^2(s-r)} - e^{-|k|^2(t-r)}\Big)^2 \mathrm{d}r \\
\lesssim_v & \sum_{k\in\mathbb{Z}^d} \Big(1 - e^{-2|k|^2 s}\Big)\chi_n(k)\chi_n(-k)\Big(1 - e^{-|k|^2(t-s)}\Big)^2 \\
\lesssim_v & 2^{dn} \int_{k'\in\mathbb{R}^d} \chi_0(k')\chi_0(-k')\Big(1 - e^{-2^{2n}|k'|^2(t-s)}\Big)^2 \mathrm{d}k' \lesssim_v 2^{dn+2\kappa n}(t-s)^\kappa,
\end{aligned}$$

for each $\kappa > 0$. This together with (2.45) and (2.46) leads to

$$\mathbb{E}\,|\langle Z_t, \psi_n(\cdot - x)\rangle - \langle Z_s, \psi_n(\cdot - x)\rangle|^2 \lesssim_v 2^{dn+2\kappa n}(t-s)^\kappa.$$

The proof is thus completed.

We now apply the above result to study regularity of the process Z.

Lemma 2.27. *Suppose that $v \in C([0,T], L^\infty)$. It holds that $Z \in C([0,T], C^{-\alpha})$ $\mathbb{P}$-a.s. for every $\alpha > d/2$. Moreover, for all $p > 2$,*

$$\mathbb{E} \sup_{t\in[0,T]} \|Z_t\|_{C^{-\alpha}}^p < \infty.$$

Proof Since Z is a centered Gaussian process, Lemma 2.26 together with the hypercontractivity property [Nua06, Theorem 1.4.1] implies that

$$\mathbb{E}\,|\langle Z_t, \psi_n(\cdot - x)\rangle|^p \lesssim \Big(\mathbb{E}\,|\langle Z_t, \psi_n(\cdot - x)\rangle|^2\Big)^{\frac{p}{2}} \lesssim 2^{\frac{dnp}{2}},$$

$$\mathbb{E}\,|\langle Z_t, \psi_n(\cdot - x)\rangle - \langle Z_s, \psi_n(\cdot - x)\rangle|^p \lesssim \Big(\mathbb{E}\,|\langle Z_t, \psi_n(\cdot - x)\rangle - \langle Z_s, \psi_n(\cdot - x)\rangle|^2\Big)^{\frac{p}{2}}$$

$$\lesssim 2^{(\frac{d}{2}+\kappa)np}(t-s)^{\frac{\kappa p}{2}},$$

for each $\kappa > 0$, $p > 2$, and every $n \geqslant -1$. This allows us to apply the Kolmogorov criterion [MW17, Lemma 10] to conclude that $Z \in C([0,T], B^{-\alpha}_{p,p})$ $\mathbb{P}$-a.s., for each $p > 2$ and every $\alpha > d/2 + 2/p$. Moreover,

$$\mathbb{E} \sup_{t\in[0,T]} \|Z_t\|^p_{B^{-\alpha}_{p,p}} < \infty.$$

The result follows by the embedding $B^{-\alpha}_{p,p} \hookrightarrow B^{-\beta}_{\infty,\infty}$ for $\beta > \alpha + d/p$ (see Lemma 2.1).

Next we rewrite the unique solution η to (1.7), which has been obtained in Section 2.3.2, in the mild form.

Proposition 2.28. *Under the assumptions* **(A1)-(A4)**, *the unique solution η to* (1.7) *satisfies*

$$\eta_t = \Gamma_t * \eta_0 - \int_0^t \nabla\Gamma_{t-s} * (vK*\eta + \eta K * v + V\eta)\mathrm{d}s - \sqrt{2\sigma}\tilde{Z}_t, \quad \mathbb{P} - a.s,$$

where $\tilde{Z}$ has the same distribution as Z.

Proof We start with the following statement: for every function φ of class $C^1([0,t], C^\infty(\mathbb{T}^d))$ and $t \in [0,T]$, it holds that,

$$\begin{aligned}\langle \eta_t, \varphi(t)\rangle - \langle \eta_0, \varphi(0)\rangle = &\int_0^t \langle \eta_s, \partial_s\varphi + \sigma\Delta\varphi\rangle\,\mathrm{d}s + \int_0^t \langle v_s K * \eta_s + \eta_s K * v_s + V\eta_s, \nabla\varphi\rangle \mathrm{d}s \\ &+ \sqrt{2\sigma}\int_0^t\int_{\mathbb{T}^d} \nabla\varphi(x)\sqrt{v_s(x)}\xi(\mathrm{d}s, \mathrm{d}x). \end{aligned} \tag{2.47}$$

It is straightforward to check the statement for finite linear combinations of functions φ of the form $\varphi(s,x) = \varphi_1(s)\varphi_2(x)$, where $\varphi_1 \in C^\infty([0,t])$ and $\varphi_2 \in C^\infty(\mathbb{T}^d)$. Then one can uniformly approximate functions in $C^1([0,t], C^\infty(\mathbb{T}^d))$ with such combinations and find (2.47).

For every $\varphi_0 \in C^\infty(\mathbb{T}^d)$ and $0 \leqslant s \leqslant t$, define $\varphi(s) := \Gamma_{t-s} * \varphi_0$, then $\partial_s\varphi(s) = -\sigma\Delta\varphi(s)$ and $\varphi(t) = \varphi_0$. By (2.47), we find

$$\begin{aligned}\langle \eta_t, \varphi_0\rangle - \langle \eta_0, \Gamma_t * \varphi_0\rangle &= \int_0^t \langle v_s K * \eta_s + \eta_s K * v_s + F\eta_s, \nabla\Gamma_{t-s} * \varphi_0\rangle \mathrm{d}s \\ &\quad + \sqrt{2\sigma}\int_0^t\int_{\mathbb{T}^d} (\nabla\Gamma_{t-s} * \varphi_0)(x)\sqrt{v_s(z)}\xi(\mathrm{d}s, \mathrm{d}x) \\ &= -\int_0^t \langle \nabla\Gamma_{t-s} * (vK * \eta + \eta K * v + V\eta), \varphi_0\rangle\,\mathrm{d}s \\ &\quad - \sqrt{2\sigma}\int_{\mathbb{T}^d}\varphi_0(x)\left(\int_0^t\int_{\mathbb{T}^d}\nabla\Gamma_{t-s}(x-y)\sqrt{v_s(y)}\xi(\mathrm{d}s,\mathrm{d}y)\right)\mathrm{d}x,\end{aligned}$$

where we used symmetry of Γ at the last inequality. The result then follows by arbitrary $\varphi_0 \in C^\infty(\mathbb{T}^d)$ and the definition of Z.

This result gives rise to the definition of mild solutions to (1.7) on a stochastic basis $(\Omega, \mathcal{F}_t, \mathcal{F}, \mathbb{P})$. We set Z given by (2.40) with ξ being vector-valued space-time white noise on $(\Omega, \mathcal{F}_t, \mathcal{F}, \mathbb{P})$.

Definition 2.29. *Assume that* $K \in L^1$, $v \in C([0,T], C^\beta)$, *and* $F \in C^\beta$ *for some* $\beta > d/2$. *We call* $\eta \in C([0,T], \mathcal{S}'(\mathbb{T}^d)) \cap L^2([0,T], B^{-\alpha}_{p,q})$ *with* $\alpha < \beta$, $p, q \in [1,\infty]$ *a mild solution to* (1.7) *with initial condition* η_0 *if for every* $\varphi \in C^\infty$

$$\langle \eta_t, \varphi\rangle = \langle \Gamma_t * \eta_0, \varphi\rangle - \int_0^t \langle \nabla \Gamma_{t-s} * (vK * \eta + \eta K * v + V\eta), \varphi\rangle \, \mathrm{d}s - \sqrt{2\sigma}\, \langle Z_t, \varphi\rangle \, .$$

Remark 2.30. *By Proposition 2.28, we know, under the assumptions* **(A1)-(A4)**, *the solutions to* (1.7) *obtained from Theorem 2.22 have the same law as the mild solutions.*

To make sense of $vK * \eta$, $\eta K * v$, *and* $V\eta$ *in the definition of mild solutions, we need the condition* $\eta_t \in B^{-\alpha}_{p,q}$ *for a.e.* $t \in [0,T]$ *with* $\alpha < \beta$ *and* $p, q \in [1,\infty]$, *where* β *is from* **(A4)**.

The following result based on the smoothing effect of heat kernel (see Lemma 2.5) gives the optimal regularity of η.

Proposition 2.31. *Suppose that Assumption* **(A4)** *holds with parameter* $\beta > d/2$ *and* η *is a mild solution to* (1.7), *and assume* $\eta_0 \in L^1(\Omega, B^{-\alpha}_{p,q})$, $\alpha \in (d/2, \beta)$, *and* $p, q \in [1,\infty]$. *Then* $\eta \in C([0,T], B^{-\alpha}_{p,q})$ *almost surely. Moreover,*

$$\mathbb{E} \sup_{t\in[0,T]} \|\eta_t\|_{B^{-\alpha}_{p,q}} < \infty.$$

Proof Firstly, applying Lemma 2.5, we have

$$\begin{aligned}\|\eta_t\|_{B^{-\alpha}_{p,q}} \lesssim & \|\eta_0\|_{B^{-\alpha}_{p,q}} + \int_0^t (t-s)^{-\frac{1}{2}} \Big[\|vK * \eta\|_{B^{-\alpha}_{p,q}} + \|\eta K * v\|_{B^{-\alpha}_{p,q}} + \|V\eta\|_{B^{-\alpha}_{p,q}} \Big] \mathrm{d}s \\ & + \|Z_t\|_{B^{-\alpha}_{p,q}}.\end{aligned}$$

To further estimate the right hand side of the above inequality, by $\alpha < \beta$, applying Lemmas 2.1-2.4 gives that

$$\|vK * \eta\|_{B^{-\alpha}_{p,q}} + \|\eta K * v\|_{B^{-\alpha}_{p,q}} \lesssim \|v\|_{C^\beta} \|K\|_{L^1} \|\eta\|_{B^{-\alpha}_{p,q}}, \quad \|V\eta\|_{B^{-\alpha}_{p,q}} \lesssim \|V\|_{C^\beta} \|\eta\|_{B^{-\alpha}_{p,q}}.$$

Hence,

$$\|\eta_t\|_{B^{-\alpha}_{p,q}} \lesssim \|\eta_0\|_{B^{-\alpha}_{p,q}} + \int_0^t (t-s)^{-\frac{1}{2}} \|\eta_s\|_{B^{-\alpha}_{p,q}} \mathrm{d}s + \|Z_t\|_{B^{-\alpha}_{p,q}}. \tag{2.48}$$

Applying Gronwall's inequality of Volterra type, see for instance [Zha10, Example 2.4], yields that

$$\mathbb{E}\mathrm{ess} \sup_{t\in[0,T]} \|\eta_t\|_{B^{-\alpha}_{p,q}} \lesssim \mathbb{E}\|\eta_0\|_{B^{-\alpha}_{p,q}} + \mathbb{E}\|Z\|_{C_T B^{-\alpha}_{p,q}} + 1.$$

By the assumption on η_0 and Lemma 2.27, the right hand side of the above inequality is thus finite.

Using (2.48), the continuity of η on $[0,T]$ follows by the continuity of Z and continuity of Γ_t from Lemma 2.5. The proof is thus completed.

Remark 2.32. *By [Hai14] we know the space-time white noise* $\xi \in C^{-\frac{d}{2}-1-\varepsilon}_{t,x}$ $\mathbb{P}$*-a.s. for every* $\varepsilon > 0$, *where* $C^{-\frac{d}{2}-1-\varepsilon}_{t,x}$ *is endowed with suitable parabolic time space scaling. Hence by Schauder estimates* $\eta \in C([0,T], C^{-\alpha})$ *for* $\alpha > d/2$ *gives the best regularity by taking* $p, q = \infty$ *in Proposition 2.31.*

Remark 2.33. *By the optimal regularity of η and Lemma 2.2, $K * v$ needs to stay in C^β with $\beta > d/2$ so that the term $K * v\eta$ appearing in SPDE* (1.7) *is well-defined. Applying Lemma 2.4 and noticing $K \in L^1$ for K satisfying* **(A2)***, the assumption about v in* **(A4)** *is thus a sufficient condition for $K * v \in C^\beta$. Moreover, $\beta > d/2$ is optimal in general. With appropriate modifications of the proof in Section 2.3.2 and by the convolution inequality in Lemma 2.4, we could also weaken the condition of v to $v \in C^{\beta-\beta_1}$ with $\beta > d/2$ and $\beta_1 \in (0, \frac{d}{2})$, at the cost of stronger condition on the interacting kernel: $K \in C^{\beta_1}$.*

2.4.2 Gaussianity

This section is devoted to the proof of Proposition 1.5. As mentioned in the introduction, we need a class of time evolution operators $\{Q_{s,t}\}$ in order to rewrite η as the generalized Ornstein-Uhlenbeck process (1.8), which would be given by the following result.

Lemma 2.34. *Assume that $v \in C([0,T], C^{\beta+1}(\mathbb{T}^d))$ and $F \in C^{\beta+1}(\mathbb{T}^d)$ with $\beta > d/2$, for each $\varphi \in C^\infty(\mathbb{T}^d)$ and $t \in [0,T]$, there exists a unique solution $f \in L^2([0,t], H^{\beta+2}) \cap C([0,t], H^{\beta+1})$ with $\partial_s f \in L^2([0,t], H^\beta)$ to the following backward equation:*

$$f_s = \varphi + \sigma \int_s^t \Delta f_r \mathrm{d}r + \int_s^t \Big[K * v_r \cdot \nabla f_r + K(-\cdot) * (\nabla f_r v_r) + V \cdot \nabla f_r\Big] \mathrm{d}r, \quad s \in [0,t], \tag{2.49}$$

where $K(-\cdot) * g$ is given in (2.31)*.*

Proof Similar to (2.38), we obtain the following a priori energy estimate for any $\epsilon > 0$

$$\begin{aligned}
&\|f_s\|_{H^{\beta+1}}^2 + 2\sigma \int_s^t \|f_r\|_{H^{\beta+2}}^2 \mathrm{d}r \\
\leqslant &\|\varphi\|_{H^{\beta+1}}^2 + C_\epsilon \int_s^t \|f_r\|_{H^{\beta+1}}^2 \mathrm{d}r + \epsilon(\|K\|_{L^1}^2 \|v\|_{C_T C^{\beta+1}}^2 + \|V\|_{C^{\beta+1}}^2) \int_s^t \|f_r\|_{H^{\beta+2}}^2 \mathrm{d}r.
\end{aligned}$$

Choosing $\epsilon > 0$ sufficiently small, $f \in L^2([0,t], H^{\beta+2})$ follows from the Gronwall's inequality. Furthermore, by Lemma 2.2 and Lemma 2.4, we find

$$\begin{aligned}
&\|K * v \cdot \nabla f\|_{H^\beta} \lesssim \|K * v\|_{C^\beta} \|f\|_{H^{\beta+1}} \lesssim \|K\|_{L^1} \|v\|_{C^\beta} \|f\|_{H^{\beta+1}}, \\
&\|V \cdot \nabla f\|_{H^\beta} \lesssim \|V\|_{C^\beta} \|f\|_{H^{\beta+1}}, \\
&\|K(-\cdot) * (v \nabla f)\|_{H^\beta} \lesssim \|K\|_{L^1} \|v \nabla f\|_{H^\beta} \lesssim \|K\|_{L^1} \|v\|_{C^\beta} \|f\|_{H^{\beta+1}}.
\end{aligned}$$

Hence we deduce from equation (2.49) that $\partial_t f \in L^2([0,t], H^\beta)$, which combined with $f \in L^2([0,t], H^{\beta+2})$ implies that $f \in C([0,t], H^{\beta+1})$ by Lions-Magenes Lemma. When $\varphi = 0$, the above energy estimate implies that $f = 0$. This fact together with linearity of equation implies the uniqueness of solutions. On the other hand, one can obtain the existence of solutions to (2.49) by classical Galerkin method (cf. [Eva98, Chapter 7]).

Define the space $\mathcal{X}_t^\beta$ and time evolution operators $\{Q_{\cdot,t}\}_{0 \leqslant t \leqslant T} : C^\infty(\mathbb{T}^d) \to \mathcal{X}_t^\beta$ as

$$\begin{gathered}
\mathcal{X}_t^\beta := \left\{ f \in L^2([0,t], H^{\beta+2}) \cap C([0,t], H^{\beta+1}); \partial_s f \in L^2([0,t], H^\beta) \right\}, \\
Q_{s,t}\varphi := f(s).
\end{gathered}$$

where f is the unique solution to (2.49) with terminal value φ at time t and is given by Lemma 2.34.

Now we are in the position to justify the Gaussianity of the unique (in distribution) limit of fluctuation measures.

Proof [Proof of Proposition 1.5] Recall that $\eta \in L^2([0,T], H^{-\alpha})$ for any $\alpha > d/2$. Then for each test function $f \in \mathcal{X}_t^\beta$ with $\beta > d/2$, by Lemma 2.2 and Lemma 2.4, we have

$$
\begin{aligned}
&\|\langle \eta_s, \partial_s f\rangle\|_{L^1_T} \lesssim \|\eta\|_{L^2_T H^{-\beta}}\|\partial_s f\|_{L^2_T H^\beta}, \quad \|\langle \eta_s, \Delta f_s\rangle\|_{L^1_T} \lesssim \|\eta\|_{L^2_T H^{-\beta}}\|f\|_{L^2_T H^{\beta+2}}, \\
&\|\langle v_s K * \eta_s + \eta_s K * v_s + V\eta_s, \nabla f_s\rangle\|_{L^1_T} \\
\lesssim &\Big(\|K\|_{L^1}\|v\|_{C_T C^\beta} + \|V\|_{C^\beta}\Big)\|\eta\|_{L^2_T H^{-\beta+\epsilon}}\|f\|_{L^2_T H^{\beta+1}} \lesssim 1,
\end{aligned}
$$

where $\epsilon > 0$ is sufficiently small, so that the weak formulation (2.47) extends to all the test fucntions $f \in \mathcal{X}_t^\beta$ with $\beta > d/2$. For each $\varphi \in C^\infty$, choosing $Q_{\cdot,t}\varphi$ as the test function in (2.47), we find

$$
\begin{aligned}
&\langle \eta_t, \varphi\rangle - \langle \eta_0, Q_{0,t}\varphi\rangle \\
=&\int_0^t \langle \eta_s, \partial_s Q_{s,t}\varphi + \sigma\Delta Q_{s,t}\varphi\rangle\,\mathrm{d}s + \int_0^t \langle v_s K * \eta_s + \eta_s K * v_s + V\eta_s, \nabla Q_{s,t}\varphi\rangle \mathrm{d}s \\
&+ \sqrt{2\sigma}\int_0^t\int_{\mathbb{T}^d} \nabla Q_{s,t}\varphi(x)\sqrt{v_s(x)}\xi(\mathrm{d}s, \mathrm{d}x) \\
=&\sqrt{2\sigma}\int_0^t\int_{\mathbb{T}^d} \nabla Q_{s,t}\varphi(x)\sqrt{v_s(x)}\xi(\mathrm{d}s, \mathrm{d}x),
\end{aligned}
$$

where we used Lemma 2.34 with $Q_{s,t}\varphi = f(s)$. The result follows by the assumption on η_0 and the fact that the stochastic integral is a centered Gaussian process with quadratic variation

$$
2\sigma\int_0^t \left\langle |\nabla Q_{s,t}\varphi|^2, v_s\right\rangle \mathrm{d}s.
$$

2.5 The Vanishing Diffusion Case

In this section, we study particle systems with vanishing diffusion, i.e. $\sigma = 0$. We denote the fluctuation measures by $\eta^N := \sqrt{N}(\mu^N - v)$ as well. Instead of the SPDE limit (1.7) in the case when $\sigma > 0$, now the fluctuation measures converge to a deterministic first order nonlocal PDE with random initial value η_0, which reads

$$
\partial_t \eta = -\nabla\cdot(vK * \eta) - \nabla\cdot(\eta K * v) - \nabla\cdot(V\eta). \tag{2.50}
$$

With the same proof as in Section 2.3.1 and Section 2.3.2, we can deduce that under the assumptions **(A1)-(A3), (A5)**, we have

1. The sequence of laws of $(\eta^N)_{N\geqslant 1}$ is tight in the space $L^2([0,T], H^{-\alpha}) \cap C([0,T], H^{-\alpha-2})$, for every $\alpha > d/2$.

2. Any limit η of converging (in distribution) subsequence of $(\eta^N)_{N\geqslant 1}$ is an analytic weak solution to (2.50) in the sense that

$$
\langle \eta_t, \varphi(t)\rangle = \langle \eta_0, \varphi(0)\rangle + \int_0^t\int_{\mathbb{T}^d} \eta_s[\partial_s\varphi + K(-\cdot) * (v\nabla\varphi) + K * v\cdot\nabla\varphi + V\nabla\varphi]\mathrm{d}x\mathrm{d}s, \tag{2.51}
$$

for every $\varphi \in C^1([0,t], C^{\beta+1})$ with $\beta > d/2$, $\mathbb{P}$-a.s. and $K(-\cdot) * g$ is given in (2.31).

However, for the case with vanishing diffusion, we cannot deduce uniqueness of the solutions to (2.50) by the proof of Lemma 2.24, due to the lack of the energy inequality (2.38). The following uniqueness result follows by the method of characteristics. We also recall the definition of flow from [Kun97, Chapter 4][1], which is used in the following proof. $\phi_{s,t}$ be a continuous map from $\mathbb{T}^d$ into itself for any $s,t \in [0,T]$ is called a flow if it satisfies the following property

1. $\phi_{s,u} = \phi_{t,u} \circ \phi_{s,t}$ holds for all s,t,u, where $\circ$ denotes the composition of maps;

2. $\phi_{s,s} = \mathrm{Id}$;

3. $\phi_{s,t} : \mathbb{T}^d \to \mathbb{T}^d$ is an onto homeomorphism for all s,t.

We refer to [Kun97, Chapter 4] for the relation between flows and ODEs. In general the solutions to ODEs with regular coefficients could generate a flow.

Proposition 2.35. *Under the assumptions* **(A2)** *and* **(A5)**, *$\eta \equiv 0$ is the only solution with zero initial value to* (2.50) *in the space $L^2([0,T], H^{-\alpha}) \cap C([0,T], H^{-\alpha-2})$ for $\alpha \in (d/2, \beta)$, the parameter β is from* **(A5)**.

Proof We first claim that a similar result to Lemma 2.34 holds for $\sigma = 0$. That is, there exists a unique solution $\varphi \in C^1([0,t], C^{\beta'})$ with $\beta' \in (\alpha+1, \beta+1)$ for any $\varphi(t,x) = \psi(x) \in C^\infty$ to the following backward equation

$$\partial_s \varphi + K(-\cdot) * (v \nabla \varphi) + K * v \cdot \nabla \varphi + V \cdot \nabla \varphi = 0, \quad \forall s \in [0,t]. \tag{2.52}$$

Suppose that the claim holds. Then for every $\psi \in C^\infty$ and $t \in [0,T]$, we use (2.51) with the test function given by the solution φ to (2.52) and $\eta_0 = 0$. Then we conclude that $\int_{\mathbb{T}^d} \eta(t,x)\psi(x)\mathrm{d}x = 0$, which implies the result. It thus suffices to justify the claim.

In the following we verify the claim by considering the backward flow $(\phi_{t,s})_{0 \leqslant s \leqslant t \leqslant T}$ generated by

$$\phi_{t,s} = x + \int_s^t (K * v_r(\phi_{t,r}) + V(\phi_{t,r}))\mathrm{d}r, \quad x \in \mathbb{T}^d, s \in [0,t]. \tag{2.53}$$

Define the forward flow $\phi_{s,t} := \phi_{t,s}^{-1}$, $0 \leqslant s \leqslant t \leqslant T$. Since $V \in C^{\beta+1}$ and $K * v \in C^1([0,T], C^{\beta+2})$ by Assumption **(A5)** and Lemma 2.4, the existence of the flow $(\phi_{t,s})_{t,s\in[0,T]}$ in $C^{\beta'}$ for any $\beta' < \beta + 1$ follows from [Kun97, Theorem 4.6.5]. For fixed $t \in [0,T]$, denote $\phi_s := \phi_{t,s}$ and $\phi_{-s} := \phi_{s,t}$ for $s \in [0,t]$, then $\phi_{-s} \circ \phi_s = \mathrm{Id}$.

The next step is the one-to-one correspondence between the solutions to (2.52) and the solutions to the following equation

$$g_s(x) = \psi(x) - \int_s^t \Big[K(-\cdot) * (g_r \circ \phi_{-r} \nabla v) \Big] \circ \phi_r(x) \mathrm{d}r$$

[1] Although the framework in [Kun97] is for the flow on $\mathbb{R}^d$, it also holds for the periodic case since the functions on $\mathbb{T}^d$ could be viewed as periodic functions on $\mathbb{R}^d$ and the framework has been extended to Riemannian manifolds in [Kun97, Chapter 4].

$$
\begin{aligned}
&-\int_s^t \Big[\mathrm{div}K(-\cdot) * (g_r \circ \phi_{-r} v)\Big] \circ \phi_r(x) \mathrm{d}r \\
:=&\psi(x) + \Phi_s(g), \quad s \in [0,t].
\end{aligned}
\tag{2.54}
$$

Here the notation $K(-\cdot) * f$ is given in (2.31). We also write $g(s,x) := g_s(x)$. We will prove that $g(s,x) := \varphi(s,\phi_s)$ satisfies (2.54). Indeed, suppose that $\varphi \in C^1([0,t], C^{\beta'})$ with $\beta' > \alpha + 1$ solves (2.52), then by the chain rule, we have

$$
\begin{aligned}
\partial_s\{\varphi(s,\phi_s(x))\} =&\partial_s \varphi(s,\phi_s(x)) + \nabla\varphi(s,\phi_s(x)) \cdot \partial_s \phi_s(x) = -K(-\cdot) * (v\nabla\varphi)(\phi_s(x)) \\
=&K(-\cdot) * (\varphi \nabla v)(\phi_s(x)) + \mathrm{div}\, K(-\cdot) * (\varphi v)(\phi_s(x)) \\
=&\Big[K(-\cdot) * \{(\varphi \circ \phi_s \circ \phi_{-s})\nabla v\}\Big] \circ \phi_s(x) + \\
&\Big[\mathrm{div}K(-\cdot) * \{(\varphi \circ \phi_s \circ \phi_{-s}) v\}\Big] \circ \phi_s(x),
\end{aligned}
$$

where we used integration by parts formula in the third step. Therefore $\varphi(s,\phi_s)$ satisfies (2.54).

Conversely, if $g \in C^1([0,t], C^{\beta'})$ is a solution to equation (2.54), let $\varphi(s,x) := g(s,\phi_{-s}(x))$, then $g(s,x) = \varphi(s,\phi_s(x))$. Similarly, we have

$$
\begin{aligned}
\partial_s g(s,x) =&\partial_s\{\varphi(s,\phi_s(x))\} = \partial_s\varphi(s,\phi_s(x)) + \nabla\varphi(s,\phi_s(x)) \cdot \partial_s\phi_s(x), \\
\partial_s g(s,x) =&K(-\cdot) * (\varphi\nabla v)(\phi_s(x)) + \mathrm{div}\, K(-\cdot) * (\varphi v)(\phi_s(x)) = -K(-\cdot) * (v\nabla\varphi)(\phi_s(x)),
\end{aligned}
$$

where the first line follows by the chain rule, while the second line follows by integration by parts. Substituting (2.53) into the first line, we obtain that φ is a solution to (2.52). Hence justifying the claim is turned into obtaining the well-posedness of backward equation (2.54) in $C^1([0,t], C^{\beta'})$.

Notice that $\Phi_\cdot : C([s,t], C^{\beta'}) \to C([s,t], C^{\beta'})$ satisfies

$$
\begin{aligned}
\sup_{r\in[s,t]} \|\Phi_r(g)\|_{C^{\beta'}} \leqslant& \int_s^t \Big\| \Big[K(-\cdot) * (g_r \circ \phi_{-r} \nabla v)\Big] \circ \phi_r \Big\|_{C^{\beta'}} \\
&+ \Big\| \Big[\mathrm{div}K(-\cdot) * (g_r \circ \phi_{-r} v)\Big] \circ \phi_r \Big\|_{C^{\beta'}} \mathrm{d}r \\
\lesssim& \int_s^t \|K\|_{L^1} \|g_r\|_{C^{\beta'}} \Big(1 + \|\phi_{-r}\|_{C^{\beta'}}^{\beta'}\Big) \|v\|_{C^{\beta'+1}} (1 + \|\phi_r\|_{C^{\beta'}}^{\beta'}) \mathrm{d}r \\
&+ \int_s^t \|\mathrm{div}K\|_{L^1} \|g_r\|_{C^{\beta'}} \Big(1 + \|\phi_{-r}\|_{C^{\beta'}}^{\beta'}\Big) \|v\|_{C^{\beta'}} \Big(1 + \|\phi_r\|_{C^{\beta'}}^{\beta'}\Big) \mathrm{d}r,
\end{aligned}
\tag{2.55}
$$

where we used Lemma 2.2, Lemma 2.4, and the fact that $\|f_1 \circ f_2\|_{C^\alpha} \lesssim \|f_1\|_{C^\alpha}(1 + \|f_2\|_{C^\alpha}^\alpha)$ when $\alpha \geqslant 1$. Recalling Assumption **(A5)** and $\phi \in C([0,t], C^{\beta'})$, we find

$$
\sup_{r\in[s,t]} \|\Phi_r(g)\|_{C^{\beta'}} \lesssim (t-s) \sup_{r\in[s,t]} \|g_r\|_{C^{\beta'}}.
$$

Choosing s close to t enough, by the linearity of Φ, we find $g \mapsto \psi + \Phi_\cdot(g)$ is a contraction mapping on $C([s,t], C^{\beta'})$, hence it has a unique fixed point solving (2.54) on $[s,t]$. Applying this argument a finite number of times, we remove the constraint on s. Therefore, there exists a unique solution $g \in C([0,t], C^{\beta'})$ to the ODE (2.54). Finally, we deduce $g \in C^1([0,t], \beta')$ from $\partial_s\Phi_s \in C([0,t], C^{\beta'})$, which follows by the calculation in (2.55). Now we obtain the global well-posedness of (2.54) in $C^1([0,t], \beta')$, which by the one-to-one correspondence between φ and g concludes the result.

Now we are able to conclude the result for vanishing diffusion cases similar to Theorem 1.4.

Theorem 2.36. *Under the assumptions* **(A1)-(A3), (A5)**, *let η be the unique solution to* (2.50) *on the same stochastic basis with the particle system* (2.1), *the sequence $(\eta^N)_{N\geqslant 1}$ converges in probability to η in $L^2([0,T],H^{-\alpha})\cap C([0,T],H^{-\alpha-2})$, for every $\alpha>d/2$.*

Proof By the facts that $(\eta^N)_{N\geqslant 1}$ are tight, the tight limits of converging subsequences solve (2.50), and there exists a unique analytic weak solution to the equation (2.50), which is ensured by Proposition 2.35, it follows immediately that the sequence $(\eta^N)_{N\geqslant 1}$ converges in distribution to the unique solution η .

Similar to $(\eta^N)_{N\geqslant 1}$, one can first obtain tightness of laws of $(\eta^l,\eta^m)_{l,m\in\mathbb{N}}$. Without loss of generality, we assume $(\eta^l,\eta^m)_{l,m\in\mathbb{N}}$ to be two converging subsequences. Then using Skorohod theorem and identifying the limit we deduce that (η^l) and (η^m) converge in distribution to η and η', which both solve random PDE (2.50) with the same initial value η_0. Furthermore, Proposition 2.35 leads to $\eta=\eta'$ $\mathbb{P}$-a.s.. Therefore, we can deduce by Lemma 2.37 below that $(\eta^N)_{N\geqslant 1}$ converges in probability to the unique solution η.

Lemma 2.37 (Gyöngy and Krylov [GK96])**.** *Let $(Z^N)_{N\geqslant 1}$ be a sequence of random elements in a Polish space E equipped with the Borel σ-algebra. Then $(Z^N)_{N\geqslant 1}$ converges in probability to an E-valued random element if and only if for every pair of subsequences (Z^l) and (Z^m) there exists a subsequence $u^k:=(Z^{l(k)},Z^{m(k)})$ converging in distribution to a random element u supported on the diagonal $\{(x,y)\in E\times E: x=y\}$.*

Proof [Proof of Theorem 1.6] The proof is similar to Proposition 1.5. In addition to the convergence obtained in Theorem 2.36, we need to check the Gaussianity of the unique solution η to (2.50) with Gaussian initial value η_0.

Define the time evolution operators $\{Q_{s,t}\}_{0\leqslant s\leqslant t\leqslant T}$ by $Q_{s,t}\varphi:=f(s)$, where $f\in C^1([0,t],C^{\beta'})$, $\beta'\in(d/2+1,\beta+1)$, is the unique solution to (2.52) with terminal value φ at time t. Then for each $\varphi\in C^\infty$ and $t\in[0,T]$, let $Q_{\cdot,t}\varphi$ play the role of test function in (2.51). We find

$$\langle\eta_t,\varphi\rangle=\langle\eta_0,Q_{0,t}\varphi\rangle\,.$$

Finally, the result follows by the assumption on η_0.

2.6 Applications

In this section we finish the proof of Theorem 1.7 and then give a similar result for the particle system (2.1) with C^1 kernels.

Let us start with proving Theorem 1.7, which concerns on the most important example of this chapter: the Biot-Savart kernel.

Proof [Proof of Theorem 1.7] By our main results Theorem 1.4, Proposition 1.5, and Theorem 1.6, it suffices to check the assumptions **(A1)-(A5)**.

(A1) is automatically satisfied since the point vortex model (1.10) is of i.i.d initial data, and one can easily check that the Biot-Savart kernel satisfies the second case of

Assumption **(A2)**. Moreover, by [JW18, Theorem 2], the following condition (2.56) yields **(A3)**,

$$v \in C([0,T], C^3) \quad \text{and} \quad \inf v > 0. \tag{2.56}$$

Now we check (2.56). On one hand, when $\sigma > 0$ we deduce $v \in C([0,T], C^3)$ under the assumption that $v_0 \in C^3$ by [BA94, Theorem A]. The fact that $v \in C([0,T], C^3)$ for the case $\sigma = 0$ follows by [MP12, Theorem 2.4.1]. On the other hand, $v \in C([0,T], C^3)$ and Lemma 2.4 implies that $K * v$ is bounded and Lipschitz continuous, which yields the global well-posedness for the Cauchy problem to the following SDE:

$$\varphi_t = x + \int_0^t K * v_s(\varphi_s)\mathrm{d}s + \sqrt{2\sigma}B_t. \tag{2.57}$$

where B is a d-dimensional Brownian motion. Let $\{\varphi_t\}_{t\in[0,T]}$ be the unique solution to (2.57), and notice that v is the density of the time marginal law of solution to (2.57) with initial value v_0-distributed.

Since K is divergence free and σ is a constant, the flow $\{\varphi_t\}_{t\in[-T,T]}$ (recall that $\varphi_{-t} := \varphi_t^{-1}$) is measure preserving (see [Kun97, Lemma 4.3.1]). Then for any bounded measurable function f, we have

$$\langle f, v_t\rangle = \mathbb{E}\int_{\mathbb{T}^d} f(\varphi_t(x))v_0(x)\mathrm{d}x = \mathbb{E}\int_{\mathbb{T}^d} f(x)v_0(\varphi_{-t}(x))\mathrm{d}x,$$

which implies $\inf v > 0$ since $\inf v_0 > 0$. Thus we obtain (2.56) and thus obviously Assumption **(A4)** for the 2D Navier-Stokes equations. Lastly, for the 2D Euler equations, we deduce Assumption **(A5)** from [MP12, Theorem 2.4.1].

For general cases, our main result Theorem 1.4 only requires bounded kernels. However, in order to check Assumption **(A3)** using [JW18] , the extra condtion $\mathrm{div}K \in L^\infty$ is necessary. Thus we consider the system (2.1) with C^1 kernels below, and give a complete result with the only assumption on the initial data and the confined potential V. Nevertheless, the following result considerably relaxes the condition on kernels in the classical work by Fernandez and Méléard [FM97], where the kernel K they considered should be regular enough, for instance in $C^{2+d/2}$.

[C^1 kernels] Consider the particle system (2.1) on $\mathbb{T}^d$ and a sequence of independent initial random variables $\{X_i(0)\}_{i\in\mathbb{N}}$ with identical probability density v_0. Assume that $\sigma > 0$, $K \in C^1$, $V, v_0 \in C^\beta$ for some $\beta > 2 \vee d/2$, and $\inf v_0 > 0$. Then the assumptions **(A1)-(A4)** hold. In particular, Theorem 1.4 and Proposition 1.5 hold in this case.

Proof The assumptions **(A1)-(A2)** follow immediately. For simplicity, we set $\sigma = 1$ and prove the required regularity for v. Consider the McKean-Vlasov equation:

$$\mathrm{d}X_t = K * v_t(X_t)\mathrm{d}t + V(X_t)\mathrm{d}t + \sqrt{2}\mathrm{d}B_t, \quad v_t = \mathcal{L}(X_t), \tag{2.58}$$

where B is a d-dimensional Brownian motion. Since K and F are bounded and Lipschitz continuous, one can obtain the well-posedness of (2.58), c.f. [CG19, Theorem 3.3]. Furthermore, applying Itô's formula and superposition principle (see [BRS21, Tre16]) we have the one-to-one correspondance between the solutions to the Mckean-Vlasov equation (2.58) and the solutions to the mean field equation (1.5), which implies the global well-posedness of the mean field equation in the space $C([0,T],\mathcal{P}(\mathbb{T}^d))$. Recall that the mild form of the mean-field equation (1.5) is stated as

$$v_t = \Gamma_t * v_0 + \int_0^t \nabla\Gamma_{t-s} * (K * v_s v_s + V v_s)\mathrm{d}s,$$

where Γ is the heat kernel of Δ. Since K is bounded and Lipschitz continuous, $K * \rho_t$ is bounded and Lipschitz continuous as well, hence belongs to the Besov space $B^1_{\infty,\infty}$. Next we consider the following linearized equation

$$\rho_t = \Gamma_t * v_0 + \int_0^t \nabla\Gamma_{t-s} * (K * v_s \rho_s + V\rho_s)\mathrm{d}s, \tag{2.59}$$

where v is the unique solution to the mean field equation in $C([0,T],\mathcal{P}(\mathbb{T}^d))$. We are going to exploit the regularity of the kernel to improve the regularity of v. Notice first that $\Gamma_\cdot * v_0 \in C([0,T],C^\beta)$ and we use Lemma 2.5 to have

$$\left\|\int_0^t \nabla\Gamma_{t-s} * (K * v_s\rho_s + V\rho_s)\mathrm{d}s\right\|^p_{B^1_{\infty,\infty}} \lesssim \left(\int_0^t (t-s)^{-\frac{1}{2}} \|K * v_s\rho + V\rho\|_{B^1_{\infty,\infty}} \mathrm{d}s\right)^p.$$

Furthermore, let $p>2$, then Lemma 2.2 and Hölder's inequality yield that

$$\begin{aligned}\left\|\int_0^t \nabla\Gamma_{t-s} * (K * v_s\rho_s + V\rho_s)\mathrm{d}s\right\|^p_{B^1_{\infty,\infty}} &\lesssim \left(\int_0^t (t-s)^{-\frac{1}{2}} \|K * v_s + V\|_{B^1_{\infty,\infty}} \|\rho\|_{B^1_{\infty,\infty}}\mathrm{d}s\right)^p \\ &\lesssim_{K,v,V} \left(\int_0^t (t-s)^{-\frac{p}{2(p-1)}}\mathrm{d}s\right)^{\frac{p-1}{p}} \int_0^t \|\rho_s\|^p_{B^1_{\infty,\infty}}\mathrm{d}s \\ &\lesssim_{K,v,V} t^{\frac{p-2}{2p}} \int_0^t \|\rho_s\|^p_{B^1_{\infty,\infty}}\mathrm{d}s.\end{aligned}$$

The constant omitted here depends on $\|V\|_{C^\beta}$ and $\sup_{t\in[0,T]}\|K * v_s\|_{B^1_{\infty,\infty}}$. Thus we have

$$\|\rho_t\|^p_{B^1_{\infty,\infty}} \lesssim \|v_0\|^p_{B^1_{\infty,\infty}} + \int_0^t \|\rho_s\|^p_{B^1_{\infty,\infty}}\mathrm{d}s,$$

for any ρ satisfies the linearized equation (2.59). Since the solution v to the mean field equation also satisfies (2.59), we find $v \in C([0,T],B^1_{\infty,\infty})$ by Gronwall's inequality.

Recall that we first obtained probability measure-valued solution to (1.5). As $v \in C([0,T],B^1_{\infty,\infty})$, the coefficient $K * v$ has better regularity, which provides the possibility to improve the regularity of v. In fact, by $K \in B^1_{\infty,\infty}$, Lemma 2.2 and Lemma 2.4, we deduce $K * v_t \in B^{\alpha+1-\epsilon}_{\infty,\infty}$ whenever $v \in B^\alpha_{\infty,\infty}$, for sufficiently small $\epsilon>0$. Therefore, v helps improving the regularity of the coefficient to the linearized equation (2.59). As a result, we could repeat the above estimates with $B^1_{\infty,\infty}$-norm replaced by $B^{2-\epsilon}_{\infty,\infty}$-norm for some $\epsilon>0$ and conclude $v \in C([0,T],B^{2-\epsilon}_{\infty,\infty})$. We iterate this procedure again and we get $v \in C([0,T],B^\beta_{\infty,\infty})$ for $\beta > 2\vee d/2$, which implies Assumption **(A4)**.

As to Assumption **(A3)**, by [JW18, Theorem 2] and $v \in C([0,T],B^\beta_{\infty,\infty})$, it is sufficient to check $\inf v > 0$. Similar to the proof of Theorem 1.7, we need the auxiliary SDE

$$\varphi_t = x + \int_0^t K * v_s(\varphi_s)\mathrm{d}s + \int_0^t V(\varphi_s)\mathrm{d}s + \sqrt{2}B_t.$$

Then for any nonnegative measurable function f on $\mathbb{T}^d$, we have

$$\begin{aligned}\langle f, v_t\rangle &= \mathbb{E}\int_{\mathbb{T}^d} f(\varphi_t(x))v_0(x)\mathrm{d}x = \mathbb{E}\int_{\mathbb{T}^d} f(x)v_0(\varphi_{-t}(x))|\det\partial\varphi_{-s}(x)|\mathrm{d}x,\\ &\geqslant \inf v_0\int_{\mathbb{T}^d} f(x)e^{-\int_0^s \|\mathrm{div}(K*v_r+F)\|_{L^\infty}dr}\mathrm{d}x\\ &\geqslant \inf v_0 e^{-T(\|K\|_{C^1}+\|F\|_{C^1})}\int_{\mathbb{T}^d} f(x)\mathrm{d}x,\end{aligned}$$

where the first inequality follows by the representation of the determinant for the Jacobian matrix $\det\partial\varphi_{-s}$, see [Kun97, Lemma 4.3.1]. This implies **(A3)**.

Chapter 3

Mean-field limits of non-exchangeable interacting diffusions

In this chapter, we show a flexible type of mean-field convergence for weighted interacting diffusions (1.12), the particles are allowed to be non-exchangeable. We recall the model here,

$$\mathrm{d}X_i = \frac{1}{N}\sum_{j\neq i} w_j^N K(X_i - X_j)\,\mathrm{d}t + \sqrt{2}\mathrm{d}B_i, \quad i = 1,\ldots,N, \tag{3.1}$$

See in Section 1.2 the condition $(\mathbf{W_r})$ on the weights (w_j^N). The setting in this Chapter is on the whole space $\mathbb{R}^d$.

3.1 Preliminaries

In this section, we define the Fisher information functional for N-particle distribution functions, and collect some related auxiliary estimates.

For $F \in \mathcal{P}(\mathbb{R}^{dN})$, the (non-normalized) Fisher information functional is defined as follows

$$I(F) := \int_{\mathbb{R}^{dN}} \frac{|\nabla F(x^N)|^2}{F(x^N)} \mathrm{d}x^N.$$

When F has no density, we set $I(F) = +\infty$.

The following lemma (a modification of [HM14, Lemma 3.7]) shows the Fisher information of probability measures is sub-additive. The general study of this topic can be found in [HM14]. See also [Car91, Theorem 3] for an analytic proof.

Lemma 3.1. *Let $X^N := (X_1, \ldots X_N)$ be a random variable on $\mathbb{R}^{dN}$ with joint law F^N, and denote by F_i the law of X_i, namely,*

$$F_i(\mathrm{d}x_i) = \int_{\mathbb{R}^{d(N-1)}} F^N(\mathrm{d}x_1 \cdots \mathrm{d}x_{i-1}\mathrm{d}x_{i+1} \cdots \mathrm{d}x_N).$$

Then it holds that

$$\sum_{i=1}^N I(F_i) \leqslant I(F^N)$$

where $I(F_i)$ is the Fisher information for distributions in $\mathcal{P}(\mathbb{R}^d)$, while $I(F^N)$ is the one for the joint law $F^N \in \mathcal{P}(\mathbb{R}^{dN})$.

Proof One could assume $I(F^N)$ is finite, and use the variational formulation of Fisher information (see [HM14, Lemma 3.5]) to obtain

$$\begin{aligned}
I(F^N) &= \sup_{\varphi \in C_b^1(\mathbb{R}^{dN};\mathbb{R}^{dN})} \left\langle F^N, -\frac{|\varphi|^2}{4} - \operatorname{div}\varphi \right\rangle \\
&\geqslant \sup_{\varphi_i \in C_b^1(\mathbb{R}^d;\mathbb{R}^d),\, 1\leqslant i\leqslant N} \left\langle F^N, -\sum_{i=1}^N \left(\frac{|\varphi_i|^2}{4} + \operatorname{div}_i \varphi_i\right) \right\rangle \\
&= \sum_{i=1}^N \sup_{\varphi_i \in C_b^1(\mathbb{R}^d;\mathbb{R}^d)} \left\langle F_i, -\left(\frac{|\varphi_i|^2}{4} + \operatorname{div}_i \varphi_i\right) \right\rangle \\
&= \sum_{i=1}^N I(F_i),
\end{aligned}$$

where φ_i depends only on the i-th variable. This completes the proof.

One can control L^p norms and $W^{1,p}$ norms of a probability density function by its Fisher information. More precisely

Lemma 3.2. *For $d \geqslant 3$, a probability measure F on $\mathbb{R}^d$ with finite Fisher information, one has*

1. *For all $p \in [1, \frac{d}{d-2}]$, it holds that $\|F\|_{L^p(\mathbb{R}^d)} \leqslant C_{p,d} I(F)^{\frac{d}{2}(1-\frac{1}{p})}$.*
2. *For all $q \in [1, \frac{d}{d-1}]$, it holds that $\|\nabla F\|_{L^q(\mathbb{R}^d)} \leqslant C_{q,d} I(F)^{\frac{d+1}{2}-\frac{d}{2q}}$.*

Note that if $d = 2$, then the 1st control holds for all $p \in [1, +\infty)$, while the 2nd estimate holds for $q \in [1, 2)$.

Proof These estimates are quite standard. We refer the interested readers to Lemma 3.2 in [FHM14] for the 2-dimensional case and also Lemma 2.4 in [LLY19] for the general case $d \geqslant 3$. The proof is essentially based on the interpolation inequality, Sobolev inequality, and also the fact that $\|F\|_{L^1} = 1$ since it is a probability density.

Lemma 3.3. *For $d \geqslant 2$, consider an $\mathbb{R}^d-$valued function $K \in L^q([0,T], L^p(\mathbb{R}^d))$ with*

$$\frac{d}{p} + \frac{2}{q} + \frac{2}{r} \leqslant 2, \quad \frac{d}{p} + \frac{2}{r} < 2, \quad r \in (1, \infty],$$

and probability measures $F(t,\cdot)$ on $\mathbb{R}^d$ with finite Fisher information for a.e. $t \in [0,T]$. Then for any $\varepsilon > 0$, we have

$$\int_0^T \int_{\mathbb{R}^d} |K(t,x)|^{\frac{r}{r-1}} F(t,x) \mathrm{d}x\mathrm{d}t \leqslant \|K\|_{L_q^p}^{\frac{r}{r-1}} \left(C_{\varepsilon,p,q,r,d} + \varepsilon \int_0^T I(F(t,\cdot))\mathrm{d}t \right). \tag{3.2}$$

Proof When $p = +\infty$, the result is trivial. So we prove only for the case when $p < \infty$, and we then have $1/q + 1/r < 1$. Notice that when $d \geqslant 2$, the condition $d/p + 2/r < 2$ implies that $p > r/(r-1)$.

Repeatedly applying Hölder's inequality gives

$$\begin{aligned}\int_0^T \int_{\mathbb{R}^d} |K(t,x)|^{\frac{r}{r-1}} F(t,x)\mathrm{d}x\mathrm{d}t \leqslant & \|K\|_{L^p_q}^{\frac{r}{r-1}} \Big(\int_0^T \Big(\|F(t,\cdot)\|_{L^{\frac{p}{p-\frac{r}{r-1}}}} \Big)^{\frac{q}{q-\frac{r}{r-1}}} \mathrm{d}t \Big)^{\frac{q-\frac{r}{r-1}}{q}} \\ \leqslant & C_{p,r,d} \|K\|_{L^p_q}^{\frac{r}{r-1}} \Big(\int_0^T I(F(t,\cdot))^{\frac{dq\frac{r}{r-1}}{2p(q-\frac{r}{r-1})}} \mathrm{d}t \Big)^{\frac{q-\frac{r}{r-1}}{q}},\end{aligned}$$

where the constant $C_{p,r,d}$ is from applying Lemma 3.2.

The conditions $d/p + 2/q + 2/r \leqslant 2$ and $1/q + 1/r < 1$ imply

$$\frac{dq\frac{r}{r-1}}{2p(q-\frac{r}{r-1})} = \frac{\frac{d}{p}}{2\Big(1-\frac{1}{r}-\frac{1}{q}\Big)} \leqslant 1, \quad \frac{r}{r-1} < q.$$

Therefore,

$$\int_0^T \int_{\mathbb{R}^d} |K(t,x)|^{\frac{r}{r-1}} F(t,x)\mathrm{d}x\mathrm{d}t \leqslant C_{p,r,d} \|K\|_{L^p_q}^{\frac{r}{r-1}} \Big(\int_0^T I(F(t,\cdot))^{\alpha_1} \mathrm{d}t \Big)^{\alpha_2},$$

with $0 < \alpha_1 \leqslant 1$ and $0 < \alpha_2 < 1$. The result is then concluded by Young's inequality.

3.2 Uniform Fisher information

Uniform Fisher information for N-particles system is quite useful when the interaction is singular, see for instance applications in [FHM14] on the Biot-Savart kernel and [FH16] on the homogenous Landau equation with moderate soft potential. The key observation is that Fisher information provides Sobolev regularities, see Lemma 3.2, and controls the singularity of interaction, see Lemma 3.3. In this section, we derive uniform Fisher information of the joint laws $\{F^N, N \in \mathbb{N}\}$. As mentioned in the introduction, the difficulty is the lack of symmetry. However, we are fortunate enough to establish the following estimate for the average of singular interactions, which will be applied to derive the main estimate Proposition 3.8 and to identify the limits in the subsequent section.

Lemma 3.4. *Assume that the function $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}^+$ satisfies the following property*

$$\int_0^T \int_{\mathbb{R}^d} f(t,x)F(t,\mathrm{d}x)\mathrm{d}t \leqslant \alpha + \beta \int_0^T I(F(t,\cdot))\mathrm{d}t,$$

for some constants $\alpha, \beta > 0$ and any probability measures $F(t,\cdot)$ on $\mathbb{R}^d$ with finite Fisher information for a.e. $t \in [0,T]$.

Then given $F^N(t,\cdot)$ the joint distribution of $(X_i(t))$, one has the estimate

$$\frac{1}{N^2}\sum_{i\neq j} \int_0^T \mathbb{E}\Big[f\Big(t, \frac{1}{\sqrt{2}}(X_i(t)-X_j(t))\Big)\Big]\mathrm{d}t \leqslant \alpha + \frac{2\beta}{N}\int_0^T I(F^N(t,\cdot))\mathrm{d}t. \tag{3.3}$$

Note that the typical choices of f are of the forms $f(x) = |K(x)|^\theta$ as in Lemma 3.3.

Remark 3.5. *If the joint distribution F^N is symmetric/exchangeable, then the conclusion in Lemma 3.4 is almost trivial. The novelty of this lemma is that we do not impose the symmetry constraint on F_N. It would be an interesting topic to study the tensorized property of entropy and Fisher information without the usual symmetry assumption.*

Remark 3.6. *The static version of this lemma holds as well. More precisely, if the function/kernel above doesn't depend on t, i.e. $f : \mathbb{R}^d \to \mathbb{R}^+$, and $\int_{\mathbb{R}^d} f(x)F(\mathrm{d}x) \leqslant \alpha + \beta I(F)$ for all $F \in \mathcal{P}(\mathbb{R}^d)$, then it holds*

$$\frac{1}{N^2}\sum_{i\neq j}\mathbb{E}\Big[f\Big(\frac{1}{\sqrt{2}}(X_i - X_j)\Big)\Big] \leqslant \alpha + \frac{2\beta}{N} I(F^N).$$

Proof Since we study the 2-body interactions, the case $N = 2$ has nothing different compared to the exchangeable case (for any N) and its proof has been essentially verified in the proof of Lemma 3.3 in [FHM14].The only difference here is that we now write an abstract function f, instead of a particular form $f(x) = 1/|x|^\theta$ as in [FHM14].

Now we consider the general case $N \geqslant 3$, where the proof is quite non-trivial and the key point is the following novel decomposition strategy for some average statistics (dating back to [Hoe94]).

We start with rewriting the right-hand side of Eq. (3.3). Let σ be a partition which divides the set $\{1, \dots, N\}$ into $\frac{N}{2}$ groups of pairs of distinct numbers when N is even, or $\frac{N+1}{2}$ groups with one group containing a single number and the other $\frac{N-1}{2}$ groups consisting pairs of distinct numbers when N is odd. Denote the collection of such partitions by S_N. Indeed, when N is even, then up to a permutation, any partition in S_N can be reduced to the following canonical form

$$\Big\{(1,2),(3,4),\cdots,(N-1,N)\Big\}.$$

When N is odd, then again any partition in S_N can be reduce to the canonical form

$$\Big\{(1,2),(3,4),\cdots,(N-2,N-1),\{N\}\Big\}.$$

Note that we keep the order in those pairs in $\sigma \in S_N$, i.e. when we write that $(i, j) \in \sigma$, by default we mean $i < j$.

Given non-negative variables $(x_{i,j})_{i\neq j}$, one has

$$\sum_{i\neq j} x_{i,j} = \sum_{i>j} x_{i,j} + \sum_{i<j} x_{i,j}. \tag{3.4}$$

Below we focus on the summation of $i < j$, the case $i > j$ can be dealt in the same manner. When $x_{i,j} = x_{j,i}$, the two summations are identical. We further find

$$\sum_{i<j} x_{i,j} = \frac{1}{|S_{N-2}|}\sum_{i<j}|S_{N-2}|x_{i,j} = \frac{1}{|S_{N-2}|}\sum_{\sigma\in S_N}\sum_{(i,j)\in\sigma} x_{i,j}, \tag{3.5}$$

where the last equality follows by the fact that for each pair (i, j) with $i < j$, it appears exactly at $|S_{N-2}|$ times in the summation $\sum_{\sigma\in S_N}$. This is more evident by regarding $\{x_{i,j}\}$ as variables and comparing the coefficient of each $x_{i,j}$.

Furthermore, when counting $|S_N|$, the cardinality of S_N, one can proceed by first arranging a number j for the number 1 to get a pair $(1, j)$, then there are $|S_{N-2}|$ possible ways to do the partition of the remaining numbers $\{1, 2, \cdots, N\} \setminus \{1, j\}$ when N is even. This reasoning gives the conclusion that

$$|S_N| = (N-1)|S_{N-2}|, \quad N = 4, 6, \cdots, \tag{3.6}$$

When N is odd, we have $|S_N| = (N-1)|S_{N-2}| + |S_{N-1}|$, since now we can first arrange a pair like $(1, 2)$ or simply the single one $\{1\}$. By induction, one has for even N, $|S_N| = |S_{N-1}|$. Consequently,

$$|S_N| = N|S_{N-2}|, \quad N = 3, 5, \cdots. \tag{3.7}$$

Combining Eq. (3.5), (3.6) and (3.7), one obtains that

$$\begin{aligned} \frac{1}{N^2}\sum_{i<j} x_{i,j} &= \frac{1}{|S_N|}\sum_{\sigma\in S_N}\frac{|S_N|}{N^2|S_{N-2}|}\sum_{(i,j)\in\sigma} x_{i,j} \\ &\leqslant \frac{1}{|S_N|}\sum_{\sigma\in S_N}\frac{1}{N}\sum_{(i,j)\in\sigma} x_{i,j}. \end{aligned} \tag{3.8}$$

This holds as well for the summation of $i > j$ pairs. Now letting $\int_0^T \mathbb{E}[f(\frac{1}{\sqrt{2}}(X_i - X_j))]$ play the role of $x_{i,j}$, it thus suffices to show for every partition σ,

$$\frac{1}{N}\sum_{(i,j)\in\sigma}\int_0^T \mathbb{E}\Big[f\Big(\frac{1}{\sqrt{2}}(X_i(t) - X_j(t))\Big)\Big]\mathrm{d}t \leqslant \frac{\alpha}{2} + \frac{\beta}{N}\int_0^T I(F^N(t,\cdot))\mathrm{d}t.$$

To this end, for each partition $\sigma \in S_N$, we define $(Y_i)_{1\leqslant i\leqslant N}$ as

$$\begin{aligned} &Y_i := \frac{1}{\sqrt{2}}(X_i - X_j), \quad Y_j := \frac{1}{\sqrt{2}}(X_i + X_j), \quad \text{for } (i,j)\in\sigma \,(\text{with } i<j); \\ &Y_i := X_i, \quad \text{for } \{i\}\in\sigma. \end{aligned} \tag{3.9}$$

Indeed, without loss of generality, one can always reduce all σ to the canonical form by a permutation of N indices, for instance in the following let us assume that N is even and $\sigma = \{(1, 2), (3, 4), \ldots, (X_{N-1}, X_N)\}$. We denote $Y^N = (Y_1, \cdots, Y_N)$ as a function of $X^N = (X_1, \cdots, X_N)$, or simply $Y^N = \Phi(X^N)$, according to the definition in Eq. (3.9).

Consequently, by change of variables,

$$\frac{1}{N}\sum_{(i,j)\in\sigma}\int_0^T \mathbb{E}\Big[f\Big(t, \frac{1}{\sqrt{2}}(X_i(t) - X_j(t))\Big)\Big]\mathrm{d}t = \frac{1}{N}\sum_{k=1}^{N/2}\int_0^T \mathbb{E}\Big[f(t, Y_{2k-1}(t))\Big]\mathrm{d}t. \tag{3.10}$$

Denote that $\bar{F}^N = F^N \circ \Phi^{-1}$. Then $\bar{F}^N$ is nothing but the law of the random variable Y^N, and in particular $I(\bar{F}^N) = I(F^N)$ since the determinant of the Jacobian matrix of Φ is 1. Furthermore, let $\bar{F}_i$ be the distribution of Y_i. Then recalling our assumption on the function f and applying Lemma 3.1, the right-hand side of Eq. (3.10) can be further bounded by

$$\begin{aligned} \sum_{k=1}^{N/2}\int_0^T \mathbb{E}\Big[f(t, Y_{2k-1}(t))\Big]\mathrm{d}t &= \sum_{k=1}^{N/2}\int_0^T\int_{\mathbb{R}^d} f(t,y)\bar{F}_{2k-1}(t,\mathrm{d}y)\mathrm{d}t \\ &\leqslant \sum_{k=1}^{N/2}\Big(\alpha + \beta\int_0^T I(\bar{F}_{2k-1}(t,\cdot))\mathrm{d}t\Big) \\ &\leqslant \frac{N\alpha}{2} + \beta\int_0^T I(\bar{F}^N(t,\cdot))\mathrm{d}t = \frac{N\alpha}{2} + \beta\int_0^t I(F^N(t,\cdot))\mathrm{d}t. \end{aligned}$$

When N is odd, the bound is simply replacing $N\alpha/2$ by $(N-1)\alpha/2$.

This completes the proof.

Applying Lemma 3.4 with $|\tilde{K}|^{\frac{r}{r-1}}$ in Lemma 3.3 playing the role of f, we arrive at the following result.

Corollary 3.7. *For $d \geqslant 2$, consider an $\mathbb{R}^d-$valued function $\tilde{K} \in L^q([0,T], L^p(\mathbb{R}^d))$ with*

$$\frac{d}{p} + \frac{2}{q} + \frac{2}{r} \leqslant 2, \quad \frac{d}{p} + \frac{2}{r} < 2, \quad r \in (1, \infty].$$

Then for any $\varepsilon > 0$ and any $F^N \in C([0,T], \mathcal{P}(\mathbb{R}^{dN}))$, we have

$$\begin{aligned}&\frac{1}{N^2}\sum_{i\neq j}\int_0^T\int_{\mathbb{R}^d}|\tilde{K}(t,x_i-x_j)|^{\frac{r}{r-1}}F^N(t,x)\mathrm{d}x^N\mathrm{d}t\\ \leqslant&\|\tilde{K}\|_{L_q^p}^{\frac{r}{r-1}}\Big(C_{\varepsilon,p,q,r,d} + \frac{\varepsilon}{N}\int_0^T I(F^N(t,\cdot))\mathrm{d}t\Big).\end{aligned}$$

Now we are in the position to show the main estimate of this chapter. Due to the previously established technical lemmas 3.3 and 3.4, the proof is quite neat.

Proposition 3.8. *Suppose that* $(\mathbf{K_r})$ *and* $(\mathbf{H})$ *hold for some $r \in (1, \infty]$. For each $N \in \mathbb{N}$ and $T \geqslant 0$, there exists an entropy solution to* (3.1)*. Furthermore, let $\{w^N\}$ be a bounded sequence in l^r, there exists a positive constant C_T such that for all $t \in [0,T]$, $N \in \mathbb{N}$ and $\gamma \in (0,1)$,*

$$\begin{aligned}&H(F_t^N) + \sum_{i=1}^N\int_{\mathbb{R}^{dN}}\langle x_i\rangle^\gamma F_t^N\mathrm{d}x^N + \frac{1}{2}\int_0^t I(F_s^N)\mathrm{d}s\\ \leqslant&H(F_0^N) + \sum_{i=1}^N\int_{\mathbb{R}^{dN}}\langle x_i\rangle^\gamma F_0^N\mathrm{d}x^N + C_T N. \end{aligned} \tag{3.11}$$

Proof We start with showing the a priori estimate uniformly in N.

For any $\varphi \in C^2(\mathbb{R}^{dN})$ vanishing at infinity, applying Itô's formula to $\varphi(X^N)$ and taking expectation, we arrive at the Liouville equation of F^N as

$$\partial_t F^N = \Delta F^N - \sum_{i=1}^N \mathrm{div}_{x_i}\Big(F^N\frac{1}{N}\sum_{j\neq i}w_j^N K(x_i-x_j)\Big) \tag{3.12}$$

in the distributional sense. We then do some formal computations that can be made rigorous by approximating the singular kernel K by smooth functions. Writing $K = K_1 + K_2$ as in $(\mathbf{K_r})$, we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}H(F^N) =& -I(F^N) + \frac{1}{N}\sum_{i\neq j}\int_{\mathbb{R}^{dN}}\nabla_i F^N\cdot w_j^N K(x_i-x_j)\mathrm{d}x^N\\ =& -I(F^N) - \frac{1}{N}\sum_{i\neq j}\int_{\mathbb{R}^{dN}}F^N w_j^N \mathrm{div}K_1(x_i-x_j)\mathrm{d}x^N\\ &+\frac{1}{N}\sum_{i\neq j}\int_{\mathbb{R}^{dN}}\nabla_i F^N\cdot w_j^N K_2(x_i-x_j)\mathrm{d}x^N\end{aligned}$$

$$:= -\, I(F^N) + J_1 + J_2. \tag{3.13}$$

In the following we apply Corollary 3.7 to handle the interaction terms J_1 and J_2. Applying Hölder's inequality, we obtain

$$\begin{aligned}
|J_1| \leqslant & N \int_{\mathbb{R}^{dN}} F^N \|w^N\|_{l^r} \left(\frac{1}{N^2} \sum_{i\neq j} |\mathrm{div} K_1(x_i - x_j)|^{\frac{r}{r-1}} \right)^{\frac{r-1}{r}} \mathrm{d}x^N \\
\lesssim & N \|w^N\|_{l^r} \int_{\mathbb{R}^{dN}} F^N \left(1 + \frac{1}{N^2} \sum_{i\neq j} |\mathrm{div} K_1(x_i - x_j)|^{\frac{r}{r-1}} \right) \mathrm{d}x^N \\
\lesssim & N \|w^N\|_{l^r} + \|w^N\|_{l^r} \frac{1}{N} \sum_{i\neq j} \int_{\mathbb{R}^{dN}} F^N |\mathrm{div} K_1(x_i - x_j)|^{\frac{r}{r-1}} \mathrm{d}x^N.
\end{aligned}$$

By the condition ($\mathbf{K_r}$), we are allowed to apply Corollary 3.7 with $\mathrm{div} K_1$ playing the role of $\tilde{K}$. That means, there exists a positive constant independent of N such that

$$\frac{1}{N} \sum_{i\neq j} \int_0^t \int_{\mathbb{R}^{dN}} F^N |\mathrm{div} K_1(x_i - x_j)|^{\frac{r}{r-1}} \mathrm{d}x^N \mathrm{d}s \leqslant CN + \frac{1}{8\|w^N\|_{l^r}} \int_0^t I(F^N) \mathrm{d}s.$$

Therefore, integrating $|J_1|$ w.r.t. time then gives

$$\int_0^t |J_1| \mathrm{d}s \leqslant CN + \frac{1}{8} \int_0^t I(F^N) \mathrm{d}s. \tag{3.14}$$

The procedure for K_2 is similar. We first apply the Young's inequality to find

$$\begin{aligned}
|J_2| \leqslant & \frac{1}{N} \sum_{i\neq j} |w_j^N| \left(\epsilon \int_{\mathbb{R}^{dN}} \frac{|\nabla_i F^N|^2}{F^N} \mathrm{d}x^N + C_\epsilon \int_{\mathbb{R}^{dN}} F^N \left| K_2(x_i - x_j) \right|^2 \mathrm{d}x^N \right) \\
\leqslant & \epsilon \|w^N\|_{l^1} I(F^N) + \frac{C_\epsilon}{N} \sum_{i\neq j} |w_j^N| \int_{\mathbb{R}^{dN}} F^N \left| K_2(x_i - x_j) \right|^2 \mathrm{d}x^N.
\end{aligned}$$

Similarly, let $|K_2|^2$ play the role of $\tilde{K}$ in Corollary 3.7, there exists a constant $C'_\epsilon > 0$, depending on ϵ only, such that

$$\frac{C_\epsilon}{N} \sum_{i\neq j} |w_j^N| \int_0^t \int_{\mathbb{R}^{dN}} F^N \left| K_2(x_i - x_j) \right|^2 \mathrm{d}x^N \mathrm{d}s \leqslant C'_\epsilon N + \frac{1}{8} \int_0^t I(F^N) \mathrm{d}s.$$

Choosing ϵ such that $\varepsilon \sup_N \|w^N\|_{l^1}$ less than $1/8$, we have

$$\int_0^t |J_2| \mathrm{d}s \leqslant CN + \frac{1}{8} \int_0^t I(F^N) \mathrm{d}s. \tag{3.15}$$

Combining (3.13), (3.14), and (3.15) then yields that

$$\begin{aligned}
H(F_t^N) - H(F_0^N) \leqslant & - \int_0^t I(F_s^N) \mathrm{d}s + CN + \frac{1}{4} \int_0^t I(F_s^N) \mathrm{d}s \\
\leqslant & - \frac{3}{4} \int_0^t I(F_s^N) \mathrm{d}s + C_\Theta N.
\end{aligned} \tag{3.16}$$

Here the constant C_Θ depends on $\Theta = \{w^N, K_2, \mathrm{div}K_1, p_1, q_1, p_2, q_2, r, d\}$.

On the other hand, testing $\partial_t F_t^N$ with $\sum_i |x_i|^\gamma$ gives

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}\sum_{i=1}^N \int_{\mathbb{R}^{dN}} \langle x_i\rangle^\gamma F^N \mathrm{d}x^N \\ =&\sum_{i=1}^N \int_{\mathbb{R}^{dN}} F^N \Delta_i \langle x_i\rangle^\gamma \mathrm{d}x^N + \frac{1}{N}\sum_{i\neq j}\int_{\mathbb{R}^{dN}} F^N \nabla_i \langle x_i\rangle^\gamma w_j^N K(x_i-x_j)\mathrm{d}x^N.\end{aligned}$$

Since $\gamma \in (0,1)$, the functions $\Delta\langle\cdot\rangle^\gamma$ and $\nabla\langle\cdot\rangle^\gamma$ are bounded. This implies

$$\sum_{i=1}^N \left|\int_{\mathbb{R}^{dN}} F^N \Delta_i \langle x_i\rangle^\gamma \mathrm{d}x^N\right| \leqslant CN,$$

and

$$\frac{1}{N}\sum_{i\neq j}\left|\int_{\mathbb{R}^{dN}} F^N \nabla_i \langle x_i\rangle^\gamma w_j^N K(x_i-x_j)\mathrm{d}x^N\right| \leqslant \frac{C}{N}\sum_{i\neq j}|w_j^N|\int_{\mathbb{R}^{dN}} F^N |K(x_i-x_j)|\,\mathrm{d}x^N, \tag{3.17}$$

with a constant C depending on γ only. One may find the right hand side of (3.17) familiar, which enjoys the same formulation as J_1 and J_2. Similarly, by the condition $(\mathbf{K_r})$, we can apply Corollary 3.7 with $|K|$ playing the role of $\tilde{K}$, and obtain

$$\frac{1}{N}\sum_{i\neq j}\int_0^t \left|\int_{\mathbb{R}^{dN}} F^N \nabla_i \langle x_i\rangle^\gamma w_j^N K(x_i-x_j)\mathrm{d}x^N\right|\mathrm{d}s \leqslant CN + \frac{1}{4}\int_0^t I(F^N)\mathrm{d}s. \tag{3.18}$$

Therefore, we have

$$\sum_{i=1}^N \int_{\mathbb{R}^{dN}} \langle x_i\rangle^\gamma F_t^N \mathrm{d}x^N \leqslant \sum_{i=1}^N \int_{\mathbb{R}^{dN}} \langle x_i\rangle^\gamma F_0^N \mathrm{d}x^N + C_T N + \frac{1}{4}\int_0^t I(F_s^N)\mathrm{d}s. \tag{3.19}$$

Now that we conclude the uniform estimate (3.11) by summing up (3.16) and (3.19).

In the following we prove the existence of entropy solutions to the particle systems (3.1). We consider the approximating systems to (3.1) with K in (3.1) replaced by regularized kernels $\{K_\varepsilon\}$. When $p, q < \infty$, we can construct $K_\varepsilon := K * \rho_\varepsilon \chi_{1/\varepsilon}$ with $\rho_\varepsilon = \varepsilon^{-d}\rho(\varepsilon^{-1}x)$ being the mollifiers and $\chi_R \in C_c^\infty(\mathbb{R}^d)$ with $\chi_R = 1$ for $|x| \leqslant R$ and $\chi_R = 0$ for $|x| > 2R$. We then have

$$\|K_{1,\varepsilon} - K_1\|_{L^{p_1}_{q_1}} + \|\mathrm{div}K_{1,\varepsilon} - \mathrm{div}K_1\|_{L^{p_1}_{q_1}} + \|K_{2,\varepsilon} - K_2\|_{L^{p_2}_{q_2}} \xrightarrow{\varepsilon\to 0} 0. \tag{3.20}$$

When $p_1 = \infty$ or $p_2 = \infty$, i.e. the bounded case, it is intuitively less singular in our setting, but requires additionally truncations in the approximating procedure. For instance $p_2 = \infty$, we decompose K_2 into $K_2 1_{|x|\leqslant R}$ plus the reminder $K_2 1_{|x|>R}$. Thus $K_2 1_{|x|\leqslant R}$ is L^p-integrable for any $p > 1$, we then proceed the approximations $\{K_{2,\varepsilon}^R\}$ for $K_2 1_{|x|<R}$ as the case when $p < \infty$. The reminder part is controlled by the γ-moment estimate.

Since for the approximating system the coefficients K^ε are smooth and have compact support, there exist unique solutions $X^{\varepsilon,N}$ to the approximating system. Moreover, the related infinitesimal generator for the approximating system is uniform elliptic and has smooth coefficients, which implies that the law of $X^{\varepsilon,N}$ has a smooth density $F^{\varepsilon,N} \in C((0,T], C^\infty(\mathbb{R}^d))$. Furthermore, the above computation for (3.11) holds for $F^{\varepsilon,N}$ with C_T independent of ε and N.

To pass the limit $\varepsilon \to 0$ and construct a entropy solutions to (3.1), we could use a standard tightness argument, which is similar as the tightness argument in Section 3.4 below. Here we only give a sketch of the proof and refer the readers to Section 3.4 for more details. We could obtain uniform in ε estimates for $X^{\varepsilon,N}$ as in (3.37) below, which gives that the sequence of $\{X^{\varepsilon,N}, \varepsilon > 0\}$ is tight in $C([0,T], \mathbb{R}^{dN})$. Extracting a subsequence of $\{X^{\varepsilon,N}\}$, using the Skorohod theorem to modify the stochastic basis, we obtain a limiting point X^N. We then show that the limiting point solves (3.1). As usual, due to the singularity of the kernel, one needs to regularize the kernel when identifying the limits, the error term produced by regularizing eventually vanishes. More precisely, notice that

$$\begin{aligned}&K_\varepsilon(X^{\varepsilon,N}) - K(X^N)\\ =&(K_\delta(X^{\varepsilon,N}) - K_\delta(X^N)) + (K_\varepsilon(X^{\varepsilon,N}) - K_\delta(X^{\varepsilon,N})) + (K_\delta(X^N) - K(X^N)),\end{aligned} \tag{3.21}$$

where $K_\varepsilon(X^{\varepsilon,N})$ is short for $\sum_j w_j^N K_\varepsilon(X_i^\varepsilon - X_j^\varepsilon)$, and other abbreviations are analogous. By (3.20), the uniform in ε estimate (3.11), and similar calculation as in the proof of (3.14), (3.15), we then have that the time integrals of the second term and the third term on the right hand side of (3.21) converge to zero as $\varepsilon, \delta \to 0$. For fixed small δ, we also have the first term on the right hand side of (3.21) go to zero as $\varepsilon \to 0$. Hence, we get the convergence of the interacting term. By Lévy's characterization theorem, X^N satisfies (3.1). Finally, since the Boltzmann entropy, the γ-th moment, and the Fisher information functionals are lower semicontinuous with respect to the weak convergence, the estimate (3.11) holds uniformly for F^N. Therefore, X^N is an entropy solution.

3.3 Random measures with Sobolev regularity

In this section, we investigate a sequence of random measures g^N in order to propagate the regularities.

When the systems are exchangeable, every accumulation point of $\{\frac{1}{N}\sum_{i=1}^N \delta_{X_i(t)} := \nu_N\}$ enjoys finite Fisher information once the normalized Fisher information of the joint laws F^N, i.e. $\frac{1}{N} I(F_N)$, is uniformly bounded, see [HM14, Theorem 5.7]. However, the exchangeability plays a crucial role in the above argument, so it cannot be applied in our setting. In order to propagate the regularity of empirical measures for non-exchangeable systems, we introduce a sequence of auxiliary random measures $\{g^N\}$ as follows.

We use the disintegrate theorem from [AGS08, Theorem 5.3.1] to write the product measure $\mathrm{d}t \times F_t^N(\mathrm{d}x_1, \ldots, \mathrm{d}x_N)$ as

$$\begin{aligned}F_t^N(\mathrm{d}x_1, ..., \mathrm{d}x_N)\mathrm{d}t =&\mathrm{d}t \times f_t^1(\mathrm{d}x_1) f_t^2(x_1, \mathrm{d}x_2) ... f_t^N(x_1, ..., x_{N-1}, \mathrm{d}x_N)\\ :=&\mathrm{d}t \times \Pi_{i=1}^N f_t^i(x^{i-1,N}, \mathrm{d}x_i),\end{aligned}$$

where $f_t^i(x^{i-1,N}, \mathrm{d}x_i)$ is a transition probability kernel from $[0,T] \times \mathbb{R}^{d(i-1)}$ to $\mathscr{B}(\mathbb{R}^d)$, i.e. for every $A \in \mathscr{B}(\mathbb{R}^d)$, $(t, x^{i-1,N}) \to f_t^i(x^{i-1,N}, A)$ is $\mathscr{B}([0,T] \times \mathbb{R}^{(i-1)d})$-measurable and for every $t \in [0,T], x^{i-1,N} \in \mathbb{R}^{d(i-1)}$, $f_t^i(x^{i-1,N}, \mathrm{d}x_i)$ is a probability on $\mathbb{R}^d$. Furthermore, there exists a zero measure set $\mathcal{N} \subset [0,T]$ such that for $t \in \mathcal{N}^c$

$$f_t^i(X^{i-1,N}(t), \mathrm{d}x_i) = \mathcal{L}(X_i(t) | X^{i-1,N}(t)) \quad \mathbb{P} - a.s., \tag{3.22}$$

where $\mathcal{L}(X_i(t)|X^{i-1,N}(t))$ is the conditional probability of $X_i(t)$ w.r.t. the σ-algebra generated by $X^{i-1,N}(t)$.

Given a set of deterministic weights $\{\tilde{w}_i^N, 1 \leqslant i \leqslant N\}$, we define the random measures $\{g^N\}$ as

$$g^N(t, \mathrm{d}x) := \frac{1}{N}\sum_{i=1}^N \tilde{w}_i^N f_t^i(X_t^{i-1,N}, \mathrm{d}x), \tag{3.23}$$

where $X_t^{i-1,N} = (X_1(t), \ldots, X_{i-1}(t))$. Since $f_t^i(x^{i-1,N}, \mathrm{d}x_i)$ is a transition probability kernel and $t \to X_t^{i-1,N}$ is continuous a.s., $g^N(t, \mathrm{d}x)$ is also a transition kernel from $[0,T]$ to $\mathscr{B}(\mathbb{R}^d)$ a.s..

The main results in this section are Lemma 3.10 and Lemma 3.13. Lemma 3.10 tells us that the sequence $\{g^N\}$ merges with the sequence of empirical measures $\{\tilde{\mu}_N\}$ as $N \to \infty$. We use Lemma 3.13 to obtain the unform regularities of g^N.

3.3.1 Weakly merging sequences

We use the concept *weakly merging* to describe how "close" is g_t^N to $\tilde{\mu}_N(t)$.

Definition 3.9. *Two sequences of finite measure valued stochastic processes $\{\mu_N(t)\}_{t\in[0,T]}$ and $\{\nu_N(t)\}_{t\in[0,T]}$ on $\mathbb{R}^d$ are called weakly merging if for each $\varphi \in C_b(\mathbb{R}^d)$, the sequence of random variables $\{\langle \varphi, \mu_N(t) - \nu_N(t)\rangle\}$ converges to zero for almost all $(t,\omega) \in [0,T]\times\Omega$.*

When the sequences are deterministic and static, Definition 3.9 agrees with classical version as in [DDF88, Bog07, Dud18]. The first result below shows that $\{g_t^N\}_{t\in[0,T]}$ and $\{\tilde{\mu}_N(t)\}_{t\in[0,T]}$ defined in Theorem 1.13 are weakly merging.

Lemma 3.10. *Given a family $\{\tilde{w}^N, N \in \mathbb{N}\}$ bounded in l^r for some $r \in (1,\infty]$. Then the sequences of finite measure valued stochastic processes $\{\tilde{\mu}_N, N \in \mathbb{N}\}$ and $\{g^N, N \in \mathbb{N}\}$ are weakly merging.*

Proof We start with representing $\tilde{\mu}_N(t) - g_t^N$ via a martingale difference sequence for a.e. $t \in [0,T]$. For $t \in [0,T]$ we use $\mathcal{F}_i$, $i = 0, ..., N-1$ to denote the $\sigma-$fields generated by $(X_1(t), ...X_i(t))$, where we omit the dependence of $\mathcal{F}_i$ on t for simplicity. Observe that for each bounded Borel measurable function φ on $\mathbb{R}^d$ and $t \in \mathcal{N}^c$

$$\mathbb{E}\Big(\varphi(X_i(t))|\mathcal{F}_{i-1}\Big) = \int_{\mathbb{R}^d} \varphi(x) f_t^i(X_1(t), ..., X_{i-1}(t)|\mathrm{d}x) = \langle \varphi, f_t^i(X_t^{i-1,N}, \cdot)\rangle, \quad \mathbb{P}-\text{a.s.},$$

which leads to for $t \in \mathcal{N}^c$

$$\langle \varphi, \tilde{\mu}_N(t)\rangle - \langle \varphi, g_t^N\rangle = \frac{1}{N}\sum_{i=1}^N \tilde{w}_i^N \Big(\varphi(X_i(t)) - \langle \varphi, f_t^i(X_t^{i-1,N}, \cdot)\rangle\Big) := \frac{1}{N}\sum_{i=1}^N M_i, \tag{3.24}$$

where $\{M_i, i = 1, ..., N\}$ is a martingale difference sequence with respect to $(\mathcal{F}_i)$. Applying the Azuma–Hoeffding inequality [AS16, Theorem 7.2.1] thus gives, for all $N \in \mathbb{N}$ and $\varepsilon > 0$

$$\mathbb{P}\left(\left|\langle \varphi, \tilde{\mu}_N(t)\rangle - \langle \varphi, g_t^N\rangle\right| > \varepsilon\right) \leqslant 2\exp\left(-\frac{N^2\varepsilon^2}{8\|\varphi\|_{L^\infty}^2 \sum_{i=1}^N |\tilde{w}_i^N|^2}\right). \tag{3.25}$$

When $r \geqslant 2$, the fact that $\|\tilde{w}^N\|_{l^2} \leqslant \|\tilde{w}^N\|_{l^r}$ gives for $t \in \mathcal{N}^c$

$$\mathbb{P}\left(\left|\langle\varphi, \tilde{\mu}_N(t)\rangle - \langle\varphi, g_t^N\rangle\right| > \varepsilon\right) \leqslant 2\exp\Big(-CN\varepsilon^2\Big).$$

When $r \in (1,2)$, we use $\|\tilde{w}^N\|_{l^2}^2 \leqslant \|\tilde{w}^N\|_{l^r}^r \|\tilde{w}^N\|_{l^\infty}^{2-r}$ and $\|\tilde{w}^N\|_{l^\infty} \lesssim N^{\frac{1}{r}}$ to obtain for $t \in \mathcal{N}^c$

$$\mathbb{P}\left(\left|\langle\varphi, \tilde{\mu}_N(t)\rangle - \langle\varphi, g_t^N\rangle\right| > \varepsilon\right) \leqslant 2\exp\Big(-CN\|\tilde{w}^N\|_{l^\infty}^{r-2}\varepsilon^2\Big) \leqslant 2\exp\Big(-CN^{2-\frac{2}{r}}\varepsilon^2\Big).$$

We note that the universal constant C is independent of t and N, which yields that

$$\begin{aligned}\int_{[0,T]\times\Omega} 1_{\left\{\left|\langle\varphi,\tilde{\mu}_N(t)\rangle - \langle\varphi, g_t^N\rangle\right|>\varepsilon\right\}} \mathrm{d}t \times \mathrm{d}\mathbb{P} \leqslant& \sup_{t\in\mathcal{N}^c} T\mathbb{P}\left(\left|\langle\varphi, \tilde{\mu}_N(t)\rangle - \langle\varphi, g_t^N\rangle\right| > \varepsilon\right)\\ \leqslant& 2T\exp\Big(-CN^{\theta_r}\varepsilon^2\Big),\end{aligned}$$

where $\theta_r > 0$. Therefore, for each φ, the sequence

$$\big\{\langle\varphi, \tilde{\mu}_N(\cdot)\rangle - \langle\varphi, g_\cdot^N\rangle, N \in \mathbb{N}\big\}$$

converges to zero in measure. Furthermore, by the Borel-Cantelli lemma and

$$\sum_{N\geqslant 1}\int_0^T \mathbb{P}\left(\left|\langle\varphi, \tilde{\mu}_N(t)\rangle - \langle\varphi, g_t^N\rangle\right| > \varepsilon\right)\mathrm{d}t < \infty,$$

we conclude that $\langle\varphi, \tilde{\mu}_N(t)\rangle - \langle\varphi, g_t^N\rangle$ converges to zero $\mathrm{d}t \times \mathrm{d}\mathbb{P}$-almost everywhere for each $\varphi \in C_b(\mathbb{R}^d)$.

3.3.2 Regularity of g^N

In the subsequent lemmas, we shall study the regularity of g^N. Classically, we first justify the absolutely continuity of the random measures.

Lemma 3.11. *Assume that* $(\mathbf{H})$, $(\mathbf{K_r})$ *and* $(\mathbf{W_r})$ *hold for some* $r \in (1,\infty]$. *For all* $1 \leqslant i \leqslant N$, $f_t^i(X_t^{i-1,N}, \mathrm{d}x)$ *is absolutely continuous with respect to the Lebesgue measure for a.s.* $(t,\omega) \in [0,T]\times\Omega$. *Furthermore, we have for* $\gamma \in (0,1)$,

$$\frac{1}{N}\sum_{i=1}^N \mathbb{E}\left(\int_{\mathbb{R}^d}|x|^\gamma f_t^i(X_t^{i-1,N},x)\mathrm{d}x + H(f_t^i(X_t^{i-1,N}))\right) \leqslant \frac{1}{N}\sum_{i=1}^N \mathbb{E}|X_i(t)|^\gamma + \frac{1}{N}H(F_t^N).$$

Proof For each $\theta \in \mathcal{P}(\mathbb{R}^d)$, the chain rule for relative entropy [DE11, Theorem C.3.1] gives that

$$\begin{aligned}H(F_t^N|\theta^{\otimes N}) =& \sum_{i=1}^N \int_{\mathbb{R}^{d(i-1)}} H\Big(f_t^i(x^{i-1,N},\cdot)|\theta\Big)\Pi_{k=1}^{i-1} f_t^k(x^{k-1,N},\mathrm{d}x_k)\\ =& \sum_{i=1}^N \int_{\mathbb{R}^{dN}} H\Big(f_t^i(x^{i-1,N},\cdot)|\theta\Big)F_t^N(\mathrm{d}x^N)\end{aligned}$$

$$= \sum_{i=1}^{N} \mathbb{E}\Big[H\Big(f_t^i(X_t^{i-1,N},\cdot)|\theta\Big)\Big].$$

Then we choose θ to be $Ce^{-|x|^\gamma}$, where C is the normalizing constant such that $\|\theta\|_{L^1} = 1$. We thus find

$$\sum_{i=1}^{N} \mathbb{E}\Big[H\Big(f_t^i(X_t^{i-1,N},\cdot)|\theta\Big)\Big] = H(F_t^N) - N\log C + \sum_{i=1}^{N} \mathbb{E}|X_i(t)|^\gamma < \infty, \tag{3.26}$$

which implies the absolutely continuity of $f_t^i(X_t^{i-1,N},\cdot)$ for each i.

On the other hand, we find

$$H(f_t^i(X_t^{i-1,N},\cdot)|\theta) = H(f_t^i(X_t^{i-1,N},\cdot)) - \log C + \int_{\mathbb{R}^d} |x|^\gamma f_t^i(X_t^{i-1,N},x)\mathrm{d}x. \tag{3.27}$$

The proof is thus completed by combining (3.26) and (3.27).

From Proposition 3.8 we know F_t^N is absolutely continuous w.r.t. the Lebesgue measure, which combined with Lemma 3.11 implies the absolute continuity of $f_t^i(x^{i-1,N},\mathrm{d}x_i)$, denoted as $f_t^i(x_1,\dots,x_i)\mathrm{d}x_i$.

Lemma 3.11 also implies the absolute continuity of g^N.

Corollary 3.12. *For each N, g^N has a density, still denoted by g^N, with respect to the to the Lebesgue measure for a.s. $(t,\omega) \in [0,T]\times\Omega$,*

The following lemma ensures that the Fisher information does not increase under the construction of random measures.

Lemma 3.13. *For the conditional distributions $\{f_t^i, 1 \leqslant i \leqslant N\}$ constructed by disintegration as in (3.22), it holds that*

$$\sum_{i=1}^{N} \mathbb{E}\int_0^T I\Big(f_t^i(X_t^{i-1,N},\cdot)\Big)\mathrm{d}t \leqslant \int_0^T I(F_t^N)\mathrm{d}t. \tag{3.28}$$

In particular, if $\tilde{w}_i^N = 1$ in (3.23) for all $1 \leqslant i \leqslant N$, then

$$\mathbb{E}\int_0^T I(g_t^N)\mathrm{d}t \leqslant \frac{1}{N}\int_0^T I(F_t^N)\mathrm{d}t.$$

Proof We first rewrite the left side of (4.15) using the definition of the Fisher information, and find

$$\begin{aligned}\sum_{i=1}^{N} \mathbb{E}\int_0^T I\Big(f_t^i(X_t^{i-1,N},\cdot)\Big)\mathrm{d}t &= \sum_{i=1}^{N} \mathbb{E}\int_0^T \int_{\mathbb{R}^d} \frac{|\nabla_x f_t^i(X_t^{i-1,N},x)|^2}{f_t^i(X_t^{i-1,N},x)}\mathrm{d}x\mathrm{d}t \\ &= \sum_{i=1}^{N} \int_0^T \int_{\mathbb{R}^{dN}} \int_{\mathbb{R}^d} \frac{|\nabla_x f_t^i(x_1,...x_{i-1,},x)|^2}{f_t^i(x_1,...,x_{i-1,},x)}\mathrm{d}xF_t^N(x^N)\mathrm{d}x^N\mathrm{d}t.\end{aligned}$$

Using the disintegration of F_t^N, we have

$$\sum_{i=1}^{N} \mathbb{E}\int_0^T I\Big(f_t^i(X_t^{i-1,N})\Big)\mathrm{d}t$$

$$
\begin{aligned}
&= \sum_{i=1}^{N} \int_0^T \int_{\mathbb{R}^{dN}} \int_{\mathbb{R}^d} \frac{|\nabla_x f_t^i(x_1,...,x_{i-1},x)|^2}{f_t^i(x_1,...,x_{i-1},x)} \Pi_{j=1}^N f_t^j(x_1,...,x_j) \mathrm{d}x \mathrm{d}x^N \mathrm{d}t \\
&= \sum_{i=1}^{N} \int_0^T \int_{\mathbb{R}^{di}} \frac{|\nabla_{x_i} f_t^i(x_1,...,x_{i-1},x_i)|^2}{f_t^i(x_1,...,x_{i-1},x_i)} \Pi_{j=1}^{i-1} f_t^j(x_1,...,x_j) \mathrm{d}x_1 ... \mathrm{d}x_i \mathrm{d}t \\
&= \sum_{i=1}^{N} \int_0^T \int_{\mathbb{R}^{dN}} \left| \nabla_{x_i} \log f_t^i(x_1,...,x_{i-1},x_i) \right|^2 F_t^N(x^N) \mathrm{d}x^N \mathrm{d}t,
\end{aligned} \tag{3.29}
$$

where we used $\int_{\mathbb{R}^d} f_t^i(x_1,...,x_{i-1},x_i)\mathrm{d}x_i = 1$. On the other hand, we find

$$
\begin{aligned}
I(F_t^N) &= \int_{\mathbb{R}^{dN}} \left| \frac{\nabla_{x^N} F_t^N(x^N)}{F_t^N(x^N)} \right|^2 F_t^N(x^N) \mathrm{d}x^N \\
&= \int_{\mathbb{R}^{dN}} \left| \sum_{i=1}^{N} \nabla_{x^N} \log f_t^i(x_1,...,x_i) \right|^2 F_t^N(x^N) \mathrm{d}x^N \\
&= \sum_{i=1}^{N} \int_{\mathbb{R}^{dN}} \left| \nabla_{x^N} \log f_t^i(x_1,...,x_i) \right|^2 F_t^N(x^N) \mathrm{d}x^N.
\end{aligned}
$$

At the last equality we used the chain rule of Fisher information [Zam98], equivalent to the fact that the summation of cross terms equals to zero. More precisely, the summation of all the cross terms consists of the following summations with $k \leqslant i < j$ (we use abbreviated notations for simplicity) ,

$$
\begin{aligned}
&\sum_{j>i} \int_{\mathbb{R}^{dN}} \nabla_{x_k} \log f^i(x_1,...,x_i) \nabla_{x_k} \log f^j(x_1,...,x_i,...,x_j) F^N(x^N) \\
=& \sum_{j>i} \int_{\mathbb{R}^{di}} \nabla_{x_k} \log f^i(x_1,...,x_i) \Pi_{m\leqslant i} f^m \int_{\mathbb{R}^{d(N-i)}} \frac{\nabla_{x_k} f^j(x_1,...,x_i,...,x_j)}{f^j} \Pi_{l>i} f^l \\
=& \int_{\mathbb{R}^{di}} \nabla_{x_k} \log f^i(x_1,...,x_i) \Pi_{m\leqslant i} f^m \left(\sum_{j>i} \int_{\mathbb{R}^{d(N-i)}} \nabla_{x_k} f^j \Pi_{l>i, l\neq j} f^l \right) \\
=& \int_{\mathbb{R}^{di}} \nabla_{x_k} \log f^i(x_1,...,x_i) \Pi_{m\leqslant i} f^m \left(\nabla_{x_k} \int_{\mathbb{R}^{d(N-i)}} \Pi_{l>i} f^l \right) \\
=& \int_{\mathbb{R}^{di}} \nabla_{x_k} \log f^i(x_1,...,x_i) \Pi_{m\leqslant i} f^m \left(\nabla_{x_k} 1 \right) = 0.
\end{aligned}
$$

Therefore, we arrive at

$$
\begin{aligned}
\sum_{i=1}^{N} \mathbb{E} \int_0^T I\Big(f_t^i(X_t^{i-1,N})\Big) \mathrm{d}t &= \sum_{i=1}^{N} \int_0^T \int_{\mathbb{R}^{dN}} \left| \nabla_{x_i} \log f_t^i(x_1,...,x_i) \right|^2 F_t^N(x^N) \mathrm{d}x^N \mathrm{d}t \\
&\leqslant \sum_{i=1}^{N} \int_0^T \int_{\mathbb{R}^{dN}} \left| \nabla_{x^N} \log f_t^i(x_1,...,x_i) \right|^2 F_t^N(x^N) \mathrm{d}x^N \mathrm{d}t \\
&= \int_0^T I(F_t^N) \mathrm{d}t,
\end{aligned}
$$

which is exactly (4.15).

When $\tilde{w}_i^N = 1$ for all $1 \leqslant i \leqslant N$, the convexity of the Fisher information yields that

$$I(g_t^N) \leqslant \frac{1}{N} \sum_{i=1}^{N} I\Big(f_t^i(X_t^{i-1,N}, \cdot)\Big),$$

the result thus follows.

Since we are interested in non-identical and even unbounded weights, there would be loss of regularity for g^N defined in (3.23), due to the appearance of the weights and the corresponding condition $(\mathbf{W_r})$. Fortunately, we can still obtain certain Sobolev regularity.

Lemma 3.14. *Given a family $\{\tilde{w}^N, N \in \mathbb{N}\}$ satisfying the condition $(\mathbf{W_r})$ (i.e. uniformly bounded in l^r) for some $r \in (1, \infty]$ and assume $(\mathbf{H})$, $(\mathbf{W_r})$ and $(\mathbf{K_r})$, then one has the following results:*

1. *It holds that*

$$\mathbb{E}\|g_t^N\|_{L^\infty([0,T],L^1(\mathbb{R}^d))} \leqslant \|w^N\|_{l^1} \leqslant \|w^N\|_{l^r}. \tag{3.30}$$

2. *For any $1 \leqslant p, q < \infty$ satisfying*

$$\frac{d}{p} + \frac{2(r-1)}{r} \geqslant d, \quad \frac{d}{p} + \frac{2}{q} \geqslant d, \tag{3.31}$$

it holds that

$$\mathbb{E}\int_0^T \|g_t^N\|_{L^p}^q \mathrm{d}t \leqslant C\|\tilde{w}^N\|_{l^r}^q T + C\frac{\|\tilde{w}^N\|_{l^r}^q}{N}\mathbb{E}\int_0^T I(F_t^N)\mathrm{d}t, \tag{3.32}$$

where $p \in [1, \frac{d}{d-2}]$ when $d \geqslant 3$ and $p \in [1, \infty)$ when $d = 2$.

3. *When $d \geqslant 3$, for any $1 \leqslant p, q < \infty$ satisfying*

$$\frac{d}{p} + \frac{2(r-1)}{r} \geqslant d+1, \quad \frac{d}{p} + \frac{2}{q} \geqslant d+1,, \tag{3.33}$$

it holds that

$$\mathbb{E}\int_0^T \|\nabla g_t^N\|_{L^p}^q \mathrm{d}t \leqslant C\|\tilde{w}^N\|_{l^r}^q T + C\frac{\|\tilde{w}^N\|_{l^r}^q}{N}\mathbb{E}\int_0^T I(F_t^N)\mathrm{d}t, \tag{3.34}$$

When $d = 2$, the result holds for $p \in [1, 2)$.

Here the proportional constants are independent of N.

Proof The part (1) follows by the fact that f_t^i is a probability density for each i. More precisely,

$$\|g_t^N\|_{L^1(\mathbb{R}^d)} \leqslant \frac{1}{N}\sum_{i=1}^{N} |\tilde{w}_i^N| \int_{\mathbb{R}^d} f_t^i(X_t^{i-1,N}, x)\mathrm{d}x = \|\tilde{w}^N\|_{l^1}.$$

The proof of other parts is based on Lemma 3.2, Lemma 3.13, and repeatedly applying Hölder's inequality.

For the part (2), by Jensen's inequality and Hölder's inequality , we find

$$\begin{aligned}\mathbb{E}\int_0^T \|g_t^N\|_{L^p}^q \mathrm{d}t \leqslant & \mathbb{E}\int_0^T \Big(\frac{1}{N}\sum_{i=1}^N |\tilde{w}_i^N| \|f_t^i(X_t^{i-1,N})\|_{L^p}\Big)^q \mathrm{d}t \\ \leqslant & \mathbb{E}\int_0^T \Big[\|\tilde{w}^N\|_{l^r}\Big(\frac{1}{N}\sum_{i=1}^N \|f_t^i(X_t^{i-1,N})\|_{L^p}^{\frac{r}{r-1}}\Big)^{\frac{(r-1)}{r}}\Big]^q \mathrm{d}t.\end{aligned}$$

Then applying the first point of Lemma 3.2 gives that

$$\begin{aligned}\mathbb{E}\int_0^T \|g_t^N\|_{L^p}^q \mathrm{d}t \leqslant & C\|\tilde{w}^N\|_{l^r}^q \mathbb{E}\int_0^T \Big(\frac{1}{N}\sum_{i=1}^N I\Big(f_t^i(X_t^{i-1,N})\Big)^{\frac{d}{2}(1-\frac{1}{p})\frac{r}{r-1}}\Big)^{\frac{q(r-1)}{r}} \mathrm{d}t \\ \leqslant & C\|\tilde{w}^N\|_{l^r}^q \mathbb{E}\int_0^T \Big(1+\frac{1}{N}\sum_{i=1}^N I\Big(f_t^i(X_t^{i-1,N})\Big)^{\frac{d}{2}(1-\frac{1}{p})\max\{q,\frac{r}{r-1}\}}\Big)\mathrm{d}t,\end{aligned}$$

where the constant term "1" comes from applying Young's inequality and appears only when $q < r/(r-1)$. Observe that the condition (3.31) is equivalent to

$$\frac{d}{2}\Big(1-\frac{1}{p}\Big)\max\{q, \frac{r}{r-1}\} \leqslant 1,$$

thus the estimate (3.32) is concluded by Lemma 3.13.

The proof of Part (3) is almost the same, except that we use the 2nd part of Lemma 3.2 to control ∇g_t^N and also $\nabla f_t^i(X_t^{i-1,N},\cdot)$, instead of the first one. In this case, the condition (3.33) is equivalent to

$$\Big(\frac{d+1}{2}-\frac{d}{2p}\Big)\max\{q, \frac{r}{r-1}\} \leqslant 1.$$

We omit the rest of the proof to avoid repeating.

Lemma 3.15. *Suppose the same setting as in Lemma 3.14 and that g_t^N converges to g_t in the space of distributions $\mathcal{S}'(\mathbb{R}^d)$, i.e. the dual space of Schwartz functions,* $\mathrm{d}t \times \mathrm{d}\mathbb{P}$ *almost everywhere. Then the Sobolev regularity estimates* (3.30), (3.32) *and* (3.34) *hold for g. In particular, $g_t(\omega)$ has a density w.r.t. the Lebesgue measure for a.e. (t,ω).*

Proof Let A be the set $\{\varphi \in \mathcal{S}(\mathbb{R}^d), \|\varphi\|_{L^{\frac{p}{p-1}}} \leqslant 1\}$. By the convergence of g_t^N to g_t in $\mathcal{S}'(\mathbb{R}^d)$ for almost every (t,ω), we find

$$\begin{aligned}\mathbb{E}\int_0^T \|g_t\|_{L^p}^q \mathrm{d}t = & \mathbb{E}\int_0^T \Big(\sup_{\varphi\in A}\langle\varphi, g_t\rangle\Big)^q \mathrm{d}t \leqslant \mathbb{E}\int_0^T \Big(\liminf_{N\to\infty}\sup_{\varphi\in A}\langle\varphi, g_t^N\rangle\Big)^q \mathrm{d}t \\ = & \mathbb{E}\int_0^T \Big(\liminf_{N\to\infty}\|g_t^N\|_{L^p}\Big)^q \mathrm{d}t \leqslant \liminf_{N\to\infty}\mathbb{E}\int_0^T \|g_t^N\|_{L^p}^q \mathrm{d}t,\end{aligned}$$

where we used the lower semi-continuity of supremum and Fatou's lemma. By (3.32), Proposition 3.8, and the condition $(\mathbf{W_r})$, we conclude that

$$\mathbb{E}\int_0^T \|g_t\|_{L^p}^q \mathrm{d}t \leqslant C\|\tilde{w}^N\|_{l^r}^q T + C\liminf_{N\to\infty}\frac{\|\tilde{w}^N\|_{l^r}^q}{N}\mathbb{E}\int_0^T I(F_t^N)\mathrm{d}t < \infty.$$

Since ∇g_t^N converges to ∇g_t in $\mathcal{S}'(\mathbb{R}^d;\mathbb{R}^d)$ for a.e. (t,ω). Again, by lower semi-continuity we find the estimate (3.34) holds for the limit ∇g_t.

3.4 Mean-fields limits

This section is devoted to show the mean-field limits of the interacting system (3.1), or more precisely the convergence of empirical measures $\mu_N(t)$ and $\tilde{\mu}_N(t)$ as in (1.20), and completes the proof of Theorem 1.13 and Theorem 1.14.

3.4.1 Tightness

To apply the classical tightness argument (also known as stochastic compactness method), it requires to find a suitable topology, which is weak enough to show the tightness of laws while sufficiently strong such that the equation (primarily the nonlinear singular part) as a functional of solutions is continuous. However, the topology for the convergence and the tightness of the empirical measures on $\mathbb{R}^d$ are too weak to ensure the convergence of the nonlinear term. To solve this problem, we consider the joint laws of the random measures $\{g^N\}$ introduced in Section 3.3, together with the associated weighted empirical measures, to handle singular interacting kernels. More precisely, we show the tightness of joint laws of $\{(\tilde{Q}^N, g^N)\}$, where $\tilde{Q}^N$ defined below is the path version of $\tilde{\mu}_N$.

The definition of tightness and the Prokhorov theorem for probability measures is well-known. We recall the generalizations for signed measures for convenience, which could be easily found in the textbook [Bog07].

Definition 3.16. *A family $\mathcal{V}$ of Radon measures on a topological space $\mathcal{Y}$ is called tight if for every $\varepsilon > 0$, there exists a compact set A_ε such that $|\nu|(\mathcal{Y}\backslash A_\varepsilon) < \varepsilon$ for all $\nu \in \mathcal{V}$.*

The following theorem due to Prokhorov connects tightness, weak convergence sequences, and compactness, c.f. [Bog07, Theorems 8.6.2 and 8.6.7].

Lemma 3.17. *Let $\mathcal{Y}$ be a complete separable metric space and let $\mathcal{V}$ be a family of Radon measures on $\mathcal{Y}$. Then the following two statements are equivalent:*

1. *every sequence $\{\nu_N\} \subset \mathcal{V}$ contains a weakly convergent subsequence;*

2. *the family $\mathcal{V}$ is tight and uniformly bounded in total variation norm.*

Let $\mathcal{V} \subset \mathcal{M}(\mathcal{Y})$ be a uniformly bounded in total variation norm and tight family of Radon measures on $\mathcal{Y}$. Then $\mathcal{V}$ has compact closure in the weak topology.

We first show the tightness of laws of the empirical measures on the path space $C([0,T],\mathbb{R}^d)$, which are defined by

$$\tilde{Q}^N(\cdot) = \frac{1}{N}\sum_i \tilde{w}_i^N \delta_{X_i} \in \mathcal{M}(C([0,T];\mathbb{R}^d)).$$

Notice that $\tilde{\mu}_N(t) = \tilde{Q}^N \circ \pi_t^{-1}$, where π_t for $t \in [0,T]$ is the canonical projection from $C([0,T],\mathbb{R}^d)$ to $\mathbb{R}^d$ defined by $\pi_t(X) = X(t)$ for $X \in C([0,T],\mathbb{R}^d)$. Let $\phi : C([0,T],\mathbb{R}^d) \to [0,\infty]$ be the function

$$\phi(X) := \sup_{0\leqslant s<t\leqslant T} \frac{|X(t)-X(s)|}{(t-s)^{1-\alpha}} + |X(0)|^{\frac{(r-1)\gamma}{r}}, \tag{3.35}$$

where $\gamma \in (0,1)$ and $\alpha \in (\max\{\frac{1}{2}, \frac{d}{2p_1} + \frac{1}{q_1}\}, 1)$. The choice for such α ensures that for $\alpha^* := \frac{1}{1-\alpha}$,

$$2(1-\alpha) < 1, \quad 1 < \alpha^*, \quad \frac{d}{p_1} + \frac{2}{q_1} + \frac{2}{\alpha^*} < 2, \quad \frac{d}{p_2} + \frac{2}{q_2} + \frac{2}{\alpha^*} < 2, \quad \frac{1}{\alpha} = \frac{\alpha^*}{\alpha^* - 1}.$$

We will apply Corollary 3.7 with (α^*, p_1, q_1) and (α^*, p_2, q_2) playing the role of (r, p, q), under the condition $(\mathbf{K_r})$ to obtain the following key uniform estimate.

Lemma 3.18. *Suppose that* $(\mathbf{H})$, $(\mathbf{W_r})$ *and* $(\mathbf{K_r})$ *hold for some* $r \in (1, \infty]$. *Given a family* $\{\tilde{w}^N, N \in \mathbb{N}\}$ *satisfying the condition* $(\mathbf{W_r})$, *it holds that*

$$\sup_N \mathbb{E}\langle \phi, |\tilde{Q}^N| \rangle = \sup_N \mathbb{E} \int_{C([0,T],\mathbb{R}^d)} \phi(X) |\tilde{Q}^N|(\mathrm{d}X) < \infty. \tag{3.36}$$

Proof By the definition of ϕ, we indeed need to show

$$\sup_N \left(\frac{1}{N} \sum_{i=1}^N |\tilde{w}_i^N| \, \mathbb{E} \, |X_i(0)|^{\frac{(r-1)\gamma}{r}} + \frac{1}{N} \sum_{i=1}^N |\tilde{w}_i^N| \, \mathbb{E} \sup_{0 \leqslant s < t \leqslant T} \frac{|X_i(t) - X_i(s)|}{(t-s)^{1-\alpha}} \right) < \infty. \tag{3.37}$$

The first summation in the bracket concerns on the γ-th moments of the initial values. Using Hölder's inequality, we find

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N |\tilde{w}_i^N| \mathbb{E}|X_i(0)|^{\frac{(r-1)\gamma}{r}} \leqslant& \|\tilde{w}^N\|_{l^r} \left(\frac{1}{N} \sum_{i=1}^N \left[\mathbb{E}|X_i(0)|^{\frac{(r-1)\gamma}{r}} \right]^{\frac{r}{r-1}} \right)^{\frac{r-1}{r}} \\ \leqslant& \|\tilde{w}^N\|_{l^r} \left(\frac{1}{N} \sum_{i=1}^N \mathbb{E}|X_i(0)|^{\gamma} \right)^{\frac{r-1}{r}}, \end{aligned}$$

which is uniformly bounded under the condition $(\mathbf{H})$.

We then investigate the second summation. Observe that

$$\begin{aligned} &\frac{1}{N} \sum_{i=1}^N |\tilde{w}_i^N| \mathbb{E} \sup_{s<t} \frac{|X_i(t) - X_i(s)|}{(t-s)^{1-\alpha}} \\ =& \frac{1}{N} \sum_{i=1}^N |\tilde{w}_i^N| \mathbb{E} \sup_{s<t} \frac{\left| \int_s^t \frac{1}{N} \sum_{j \neq i} w_j^N K(X_i(\tau) - X_j(\tau)) \mathrm{d}\tau + \sqrt{2}(B_i(t) - B_i(s)) \right|}{(t-s)^{1-\alpha}} \\ \leqslant& J_1^N + J_2^N \end{aligned}$$

where J_i^N, $i = 1, 2$, are defined by

$$\begin{aligned} J_1^N :=& \frac{1}{N} \sum_{i=1}^N |\tilde{w}_i^N| \mathbb{E} \sup_{s<t} \frac{\left| \int_s^t \frac{1}{N} \sum_{j \neq i} w_j^N K(X_i(\tau) - X_j(\tau)) \mathrm{d}\tau \right|}{(t-s)^{1-\alpha}}, \\ J_2^N :=& \frac{\sqrt{2}}{N} \sum_{i=1}^N |\tilde{w}_i^N| \mathbb{E} \sup_{s<t} \frac{|(B_i(t) - B_i(s))|}{(t-s)^{1-\alpha}}. \end{aligned}$$

For J_1^N involving the interactions, we have

$$J_1^N \leqslant \frac{1}{N^2} \sum_{i \neq j} |\tilde{w}_i^N w_j^N| \mathbb{E} \left(\sup_{s<t} \frac{\int_s^t |K(X_i - X_j)| \, \mathrm{d}\tau}{(t-s)^{1-\alpha}} \right)$$

$$\leqslant \|w^N\|_{l^r}\|\tilde{w}^N\|_{l^r}\left(\frac{1}{N^2}\sum_{i\neq j}\left[\mathbb{E}\left(\sup_{s<t}\frac{\int_s^t |K(X_i-X_j)|\,\mathrm{d}\tau}{(t-s)^{1-\alpha}}\right)\right]^{\frac{r}{r-1}}\right)^{\frac{r-1}{r}}$$

$$\leqslant \|w^N\|_{l^r}\|\tilde{w}^N\|_{l^r}\left(\frac{1}{N^2}\sum_{i\neq j}\left[\mathbb{E}\int_0^T |K(X_i-X_j)|^{\frac{1}{\alpha}}\,\mathrm{d}t\right]^{\frac{\alpha r}{r-1}}\right)^{\frac{r-1}{r}}$$

$$\leqslant C\|w^N\|_{l^r}\|\tilde{w}^N\|_{l^r}\left(T+\frac{1}{N^2}\sum_{i\neq j}\mathbb{E}\int_0^T |K(X_i-X_j)|^{\max\{\frac{r}{r-1},\frac{1}{\alpha}\}}\,\mathrm{d}t\right)^{\frac{r-1}{r}},$$

where the constant T is given by Young's inequality $|x|^{\alpha r/(r-1)} \leqslant |x|+1$ when $\alpha r/(r-1) \leqslant 1$. We thus obtain

$$J_1^N \leqslant C\|w^N\|_{l^r}\|\tilde{w}^N\|_{l^r}\left(T+\frac{1}{N^2}\sum_{i\neq j}\mathbb{E}\int_0^T \left(|K(X_i-X_j)|^{\frac{r}{r-1}}+|K(X_i-X_j)|^{\frac{1}{\alpha}}\right)\mathrm{d}t\right)^{\frac{r-1}{r}}.$$

Using Corollary 3.7, we find J_1^N is bounded by the Fisher information. That is

$$J_1^N \leqslant C\|w^N\|_{l^r}\|\tilde{w}^N\|_{l^r}\left(C+\frac{1}{N}\int_0^T I(F_t^N)\mathrm{d}t\right)^{\frac{r-1}{r}},$$

for all $N\in\mathbb{N}$. Proposition 3.8 thus implies that J_1^N is uniformly bounded for all $N\in\mathbb{N}$.

Next, $\sup_N J_2^N<\infty$ follows by the modulus continuity of Brownian motions,

$$J_2^N \leqslant \sqrt{2}\|\tilde{w}^N\|_{l^1}\left[\mathbb{E}\sup_{s<t}\frac{|B_1(t)-B_1(s)|}{(t-s)^{1-\alpha}}\right].$$

The proof of (3.36) is thus completed, and the result follows.

Lemma 3.19. *Suppose that* $(\mathbf{H})$, $(\mathbf{W_r})$ *and* $(\mathbf{K_r})$ *hold for some* $r\in(1,\infty]$. *Given a family* $\{\tilde{w}^N, N\in\mathbb{N}\}$ *satisfying the condition* $(\mathbf{W_r})$, *the laws of the sequence* $\{\tilde{Q}^N, N\in\mathbb{N}\}$ *are tight on* $\mathcal{M}(C([0,T],\mathbb{R}^d))$.

Proof By the Arzelà–Ascoli theorem, the measurable function ϕ on $C([0,T],\mathbb{R}^d)$ is lower bounded and has precompact level sets, i.e. the set $\{X|\phi(X)\leqslant c\}$ is precompact for any positive number c. Such function with precompact level sets is called a tightness function in the literature, c.f. [DE11].

We claim that $\Phi:\mathcal{M}(C([0,T]),\mathbb{R}^d)\to[0,\infty]$ defined by

$$\Phi(\mu):=\langle\phi,|\mu|\rangle+\|\mu\|_{TV}$$

has precompact level sets. One may deduce by the Chebyshev's inequality that

$$|\mu|(\{\phi>c\})\leqslant\frac{1}{c}\langle\phi,|\mu|\rangle.$$

Thus for any given level set $A_R:=\{\mu|\Phi(\mu)\leqslant R\}$, $R>0$, the family A_R is tight in $\mathcal{M}(C([0,T],\mathbb{R}^d))$ and uniformly bounded in total variation norm. By the generalized Prokhorov's theorem (Lemma 3.17), the closure of A_R is compact in the weak topology.

Furthermore, by $(\mathbf{W_r})$ and Lemma 3.18, we obtain

$$\mathbb{P}(\tilde{Q}^N\notin A_R)=\mathbb{P}(\Phi(\tilde{Q}^N)>R)\leqslant\frac{1}{R}\mathbb{E}[\Phi(\tilde{Q}^N)]$$

$$\leqslant \frac{1}{R}\Big(\|\tilde{w}^N\|_{l^1} + \mathbb{E}\langle \phi, |\tilde{Q}^N|\rangle\Big) \xrightarrow{R\to\infty} 0,$$

which is uniformly in N. The tightness of the laws of $\{\tilde{Q}^N, N \in \mathbb{N}\}$ thus follows.

The next result concerns on tightness of laws of $\{g^N\}$.

Lemma 3.20. *Suppose that* $(\mathbf{H})$, $(\mathbf{W_r})$ *and* $(\mathbf{K_r})$ *hold for some* $r \in (1, \infty]$. *There exists* $p^* > 1$ *such that the laws of* $\{g^N, N \in \mathbb{N}\}$ *are tight on* $L^{p^*}_w([0,T] \times \mathbb{R}^d)$.

Proof By the part (1) of Lemma 3.14, one may choose $p^* > 1$ such that $p^* = p = q$ such that

$$\frac{d}{p^*} + \frac{2(r-1)}{r} \geqslant d, \quad \frac{d+2}{p^*} \geqslant d,$$

which equals to

$$1 < p^* \leqslant \min\left\{\frac{d}{d - \frac{2(r-1)}{r}}, \frac{d+2}{d}\right\}, \quad p* < \infty.$$

Thus Lemma 3.14 shows there exists such $p^* > 1$ such that

$$\sup_N \mathbb{E}\|g^N\|^{p^*}_{L^{p^*}([0,T]\times\mathbb{R}^d)} \leqslant C + C\left(\sup_N \|\tilde{w}^N\|^{p^*}_{l^r}\right)\left(\sup_N \frac{1}{N}\mathbb{E}\int_0^T I(F_t^N)\mathrm{d}t\right) < \infty.$$

This uniform bound of Fisher information is ensured by Proposition 3.8. Furthermore, applying Chebyshev's inequality yields that

$$\sup_N \mathbb{P}\Big(g^N(t,x) \in L^{p^*}([0, T \times \mathbb{R}^d]), \|g^N\|_{L^{p^*}} > R\Big) \leqslant \frac{1}{R^{p^*}} \sup_N \mathbb{E}\|g^N\|^{p^*}_{L^{p^*}([0,T]\times\mathbb{R}^d)} \xrightarrow{R\to\infty} 0.$$

The proof is thus completed.

3.4.2 Identify the limits

Now we extract a subsequence of $\{(Q^N, v^N, \tilde{Q}^N, g^N)\}$, where (Q^N, v^N) is defined by replacing $\tilde{w}^N$ in the definition of $(\tilde{Q}^N, g^N)$ by w^N, and identify the limiting point as a solution to (1.16).

Observe that $\{w^N\}$ is just a specific example of $\{\tilde{w}^N\}$, by Lemma 3.19 and Lemma 3.20, one may deduce that the sequence of laws of $\{(Q^N, v^N, \tilde{Q}^N, g^N), N \in \mathbb{N}\}$ is tight on the space $\mathcal{X}$ defined by

$$\mathcal{X} := \mathcal{M}\Big(C([0,T], \mathbb{R}^d)\Big) \times L^{p^*}_w([0,T] \times \mathbb{R}^d) \times \mathcal{M}\Big(C([0,T], \mathbb{R}^d)\Big) \times L^{p^*}_w([0,T] \times \mathbb{R}^d)$$

By the generalized Skorokhod Representation Theorem/ Jakubowski Theorem (c.f. [BFH18, Theorem 2.7.1]), we deduce the following result.

Proposition 3.21. *There exists a subsequence of* $\{(Q^N, v^N, \tilde{Q}^N, g^N), N \in \mathbb{N}\}$, *without relabeling for simplicity, and a probability space* $(\Omega^*, \mathcal{F}^*, \mathbb{P}^*)$ *with* $\mathcal{X}$*-valued random variables* $\{(Q^{*N}, v^{*N}, \tilde{Q}^{*N}, g^{*N}),\ N \in \mathbb{N}\}$ *and* $(Q, v, \tilde{Q}, g)$ *such that*

1. *For each N, the law of $(Q^N, v^N, \tilde{Q}^N, g^N)$ coincides with the law of $(Q^{*N}, v^{*N}, \tilde{Q}^{*N}, g^{*N})$.*

2. *The sequence of random variables $(Q^{*N}, v^{*N}, \tilde{Q}^{*N}, g^{*N})$ converges to $(Q, v, \tilde{Q}, g)$ in $\mathcal{X}$ $\mathbb{P}^*$-almost surely.*

Remark 3.22. *To apply the Jakubowski theorem, one needs to check a topological property, that is the space $\mathcal{X}$ is countably separated. This property is closely related to submetrizability (or metrizability for compact spaces). For our case, the weak topology of the Polish space L^{p*} is clearly countably separated since its dual space is separable. As to $\mathcal{M}(C([0,T],\mathbb{R}^d))$, the required topological property follows by Koumoullis and Sapounakis [KS84, Theorem 4.1].*

For simplicity, we omit the superscript $*$ in the following text.

Now we are able to deduce the convergence of $\{\tilde{\mu}_N\}$ (and the specific sequence $\{\mu_N\}$).

Corollary 3.23. *The sequence of the empirical measure processes $\tilde{\mu}_N$ converges to $\tilde{\mu}$ in $C([0,T],\mathcal{M}(\mathbb{R}^d))$ almost surely, where $\tilde{\mu} := (\tilde{Q} \circ \pi_t^{-1})_{t\in[0,T]}$.*

Proof Recall that the canonical projection π_t from $C([0,T],\mathbb{R}^d)$ to $\mathbb{R}^d$. Clearly, $\tilde{\mu}$ belongs to the space $C([0,T],\mathcal{M}(\mathbb{R}^d))$. Given any function $\varphi \in C_b(\mathbb{R}^d)$, it is straightforward to check that the family of functions $\{\Phi_t, t \in [0,T]\}$ on $C([0,T],\mathbb{R}^d)$, defined by $\Phi_t(X) := \varphi(X_t)$ for $X \in C([0,T],\mathbb{R}^d)$, is uniformly bounded and pointwise equicontinuous. Therefore, we have

$$\begin{aligned}\sup_{t\in[0,T]}\left|\int_{\mathbb{R}^d}\varphi(x)\Big(\tilde{\mu}_N(t)(\mathrm{d}x)-\tilde{\mu}_t(\mathrm{d}x)\Big)\right| &= \sup_{t\in[0,T]}\left|\int_{\mathbb{R}^d}\varphi(x)\Big(\tilde{Q}^N\circ\pi_t^{-1}(\mathrm{d}x)-\tilde{Q}\circ\pi_t^{-1}(\mathrm{d}x)\Big)\right| \\ &= \sup_{t\in[0,T]}\left|\int_{C([0,T],\mathbb{R}^d)}\Phi_t(X)\Big(\tilde{Q}^N(\mathrm{d}X)-\tilde{Q}(\mathrm{d}X)\Big)\right| \\ &\xrightarrow{N\to\infty} 0,\end{aligned}$$

where the convergence follows by the convergence of $\tilde{Q}^N$ and applying [Bog07, Exercise 8.10.134].

The next lemma connects the limiting points of weakly merging sequences $\{\tilde{\mu}_N\}$ and g^N.

Lemma 3.24. *The subsequence $\{g_t^N\}$ converges to $\tilde{\mu}_t$ in $\mathcal{S}'(\mathbb{R}^d)$ for almost every (t,ω). Furthermore, g_t is a density of $\tilde{\mu}_t$ for almost every (t,ω).*

Proof For a.e. $(t,\omega) \in [0,T]\times\Omega$ and any $\varphi \in C_0(\mathbb{R}^d)$, we have

$$\begin{aligned}\langle\varphi,\tilde{\mu}_t\rangle &= \lim_{N\to\infty}\langle\varphi,\tilde{\mu}_N(t)\rangle \\ &= \lim_{N\to\infty}\left(\left\langle\varphi,\tilde{\mu}_N(t)-g_t^N\right\rangle+\left\langle\varphi,g_t^N\right\rangle\right) \\ &= \lim_{N\to\infty}\left\langle\varphi,g_t^N\right\rangle.\end{aligned}$$

The last equality follows by the fact that $\{\tilde{\mu}_N(t)\}$ and $\{g_t^N\}$ are weakly merging, proved in Lemma 3.10.

On the other hand, since g^N converges to g in $L^{p^*}_w([0,T]\times\mathbb{R}^d)$, we have

$$\int_{[0,T]\times\mathbb{R}^d}\varphi(t,x)g_t(x)\mathrm{d}x\mathrm{d}t=\int_{[0,T]\times\mathbb{R}^d}\varphi(t,x)\tilde{\mu}_t(\mathrm{d}x)\mathrm{d}t,\quad \forall\varphi\in C_0([0,T]\times\mathbb{R}^d),$$

almost surely. Choosing a countable dense subset of $C_0(\mathbb{R}^d)$ leads to $g_t(x)\mathrm{d}x\times\mathrm{d}t=\tilde{\mu}_t(\mathrm{d}x)\times\mathrm{d}t$ in $\mathcal{M}([0,T]\times\mathbb{R}^d)$ almost surely. Here we used the Riesz-Markov-Kakutani representation theorem (see [Bog07, Section 7.10]), which characterizes $\mathcal{M}(\mathbb{R}^d)$ as the space of bounded linear functionals on $C_0(\mathbb{R}^d)$. Furthermore, by the uniqueness of disintegration (we refer to [Bog07, Lemma 10.4.3] for the result on signed measures), we conclude that $g_t(x)\mathrm{d}x=\tilde{\mu}_t(\mathrm{d}x)$ for almost every (t,ω).

In the following, we shall not distinguish g and $\tilde{\mu}$. The previous lemma together with Lemma 3.15 gives the following.

Corollary 3.25. *The Sobolev regularity estimates* (3.30), (3.32) *and* (3.34) *hold for g.*

Now we are in the position to identify the limiting point (v,g).

Proposition 3.26. *Suppose that* (**H**) *and* ($\mathbf{K_r}$) *hold for some $r\in(1,\infty]$. Given two sequences $\{\tilde{w}^N,N\in\mathbb{N}\}$ and $\{w^N,N\in\mathbb{N}\}$ satisfying the condition* ($\mathbf{W_r}$)*, each limiting point (v,g) obtained from Proposition 3.21 is a solution to* (1.16) *in the sense of Definition 1.11.*

Proof Clearly

$$M_t^N(\varphi)=\langle\varphi,\tilde{\mu}_N(t)\rangle-\langle\varphi,\tilde{\mu}_N(0)\rangle-\int_0^t\langle\Delta\varphi,\tilde{\mu}_N(s)\rangle\,\mathrm{d}s-\int_0^t\langle\nabla\varphi\cdot K*\mu_N(s),\tilde{\mu}_N(s)\rangle\,\mathrm{d}s \tag{3.38}$$

is a martingale w.r.t. the filtration generated by $\tilde{\mu}_N$ and μ_N for $\varphi\in C_0^\infty(\mathbb{R}^d)$. Observe that the covariance of martingale $M_t^N(\varphi)$ is of $O(\frac{1}{N})$ by the independence of Brownian motions, we thus have

$$M_t^N(\varphi)\to 0,\quad \text{as } N\to\infty,$$

in probability. Up to a subsequence, the martingale converges to zero almost surely.

Since $\tilde{\mu}_N$ converges to g (equivalently to $\tilde{\mu}$) in $C([0,T],\mathcal{M}(\mathbb{R}^d))$, letting $N\to\infty$ on the both sides of the equality (3.38) leads to

$$\begin{aligned}\langle\varphi,g_t\rangle&=\lim_{N\to\infty}\langle\varphi,\tilde{\mu}_N(t)\rangle\\&=\langle\varphi,g_0\rangle+\int_0^t\langle\Delta\varphi,g_s\rangle\,\mathrm{d}s+\lim_{N\to\infty}\int_0^t\langle\nabla\varphi\cdot K*\mu_N(s),\tilde{\mu}_N(s)\rangle\,\mathrm{d}s.\end{aligned}\tag{3.39}$$

It suffices to identify the limits of the interacting term. The main difficulty is the lack of continuity of the singular interacting term with respect to the weak topology in $\mathcal{M}(\mathbb{R}^d)$. Fortunately, with the estimates Proposition 3.8 and Corollary 3.25, later we shall show

$$\lim_{N\to\infty}\int_0^t\langle\nabla\varphi\cdot K*\mu_N(s),\tilde{\mu}_N(s)\rangle\,\mathrm{d}s=\int_0^t\langle\nabla\varphi\cdot K*v_s,g_s\rangle\mathrm{d}s. \tag{3.40}$$

To obtain (3.40), we approximate $K = K_1 + K_2$ by $K_\varepsilon = K_{1,\varepsilon} + K_{2,\varepsilon}$ as in the proof of Proposition 3.8, where $K_{1,\varepsilon}$ and $K_{2,\varepsilon}$, $\varepsilon \in (0,1)$, are smooth and compactly supported functions satisfying

$$\|K_{1,\varepsilon} - K_1\|_{L^{p_1}_{q_1}} \to 0; \quad \|K_{2,\varepsilon} - K_2\|_{L^{p_2}_{q_2}} \to 0,$$

for $p_1, p_2 < \infty$. When $p_1 = \infty$ or $p_2 = \infty$, we first truncate K by letting $K = K1_{|\cdot|\leqslant R} + K1_{|\cdot|>R}$, then proceed the regularization on the local term $K1_{|\cdot|\leqslant R}$. The term $K1_{|\cdot|>R}$ is controlled by finite moments of particles and causes no difficulty in singularity, we ingore this term in the following. Therefore, one may divide the singular interacting term into a continuous functional on $C([0,T], \mathcal{M}(\mathbb{R}^d))$ and a correction. More precisely,

$$\int_0^t \langle \nabla\varphi \cdot K * \mu_N(s), \tilde{\mu}_N(s)\rangle \,\mathrm{d}s = \int_0^t \langle \nabla\varphi \cdot K_\varepsilon * \mu_N(s), \tilde{\mu}_N(s)\rangle \,\mathrm{d}s + R^N_{\varepsilon,\varphi},$$

where the correction $R^N_{\varepsilon,\varphi}$ is taken as

$$R^N_{\varepsilon,\varphi} = \int_0^t \langle \nabla\varphi \cdot [K_1 - K_{1,\varepsilon}] * \mu_N(s), \tilde{\mu}_N(s)\rangle \,\mathrm{d}s + \int_0^t \langle \nabla\varphi \cdot [K_2 - K_{2,\varepsilon}] * \mu_N(s), \tilde{\mu}_N(s)\rangle \,\mathrm{d}s.$$

Similarly, the notation $R_{\varepsilon,\varphi}$ stands for

$$\begin{aligned} R_{\varepsilon,\varphi} &:= \int_0^t \langle \nabla\varphi \cdot K * v_s, g_s\rangle \mathrm{d}s - \int_0^t \langle \nabla\varphi \cdot K_\varepsilon * v_s, g_s\rangle \mathrm{d}s \\ &= \int_0^t \langle \nabla\varphi \cdot [K_1 - K_{1,\varepsilon}] * v_s, g_s\rangle \mathrm{d}s + \int_0^t \langle \nabla\varphi \cdot [K_2 - K_{2,\varepsilon}] * v_s, g_s\rangle \mathrm{d}s. \end{aligned}$$

We now claim that for each $\varphi \in C^2_b(\mathbb{R}^d)$,

$$\mathbb{E}\left(\sup_{t\in[0,T]} |R_{\varepsilon,\varphi}(t)|\right) \xrightarrow{\varepsilon\to 0} 0; \quad \sup_N \mathbb{E}\left(\sup_{t\in[0,T]} \left|R^N_{\varepsilon,\varphi}(t)\right|\right) \xrightarrow{\varepsilon\to 0} 0. \tag{3.41}$$

This uniform convergence of the corrections is the key ingredient to deduce (3.40). Indeed, the approximations for the kernel K implies that

$$\begin{aligned} &\left|\int_0^t \langle \nabla\varphi \cdot K * \mu_N(s), \tilde{\mu}_N(s)\rangle \,\mathrm{d}s - \int_0^t \langle \nabla\varphi \cdot K * v_s, g_s\rangle \mathrm{d}s\right| \\ \leqslant &\left|\int_0^t \langle \nabla\varphi \cdot K_\varepsilon * \mu_N(s), \tilde{\mu}_N(s)\rangle \,\mathrm{d}s - \int_0^t \langle \nabla\varphi \cdot K_\varepsilon * v_s, g_s\rangle \mathrm{d}s\right| + \left|R^N_{\varepsilon,\varphi}(t)\right| + |R_{\varepsilon,\varphi}(t)|. \end{aligned}$$

By the convergence of $(\mu_N, \tilde{\mu}_N)$ to (v, g) in $C([0,T]; \mathcal{M}(\mathbb{R}^d))^{\otimes 2}$, the first absolute value at the second line vanishes almost surely as N goes to infinity. Thus for $\varepsilon > 0$

$$\begin{aligned} &\lim_{N\to\infty} \mathbb{E}\left|\int_0^t \langle \nabla\varphi \cdot K * \mu_N(s), \tilde{\mu}_N(s)\rangle \,\mathrm{d}s - \int_0^t \langle \nabla\varphi \cdot K * v_s, g_s\rangle \mathrm{d}s\right| \\ \leqslant &\lim_{N\to\infty} \left|\int_0^t \langle \nabla\varphi \cdot K_\varepsilon * \mu_N(s), \tilde{\mu}_N(s)\rangle \,\mathrm{d}s - \int_0^t \langle \nabla\varphi \cdot K_\varepsilon * v_s, g_s\rangle \mathrm{d}s\right| \\ &+ \sup_N \mathbb{E}\left|R^N_{\varepsilon,\varphi}(t)\right| + \mathbb{E}\,|R_{\varepsilon,\varphi}(t)| \\ \leqslant &\sup_N \mathbb{E}\left|R^N_{\varepsilon,\varphi}(t)\right| + \mathbb{E}\,|R_{\varepsilon,\varphi}(t)|\,. \end{aligned}$$

Choosing ε sufficient small and applying (3.41), we arrive at (3.40).

Now that it remains to prove the claim (3.41).

Recall the definition of $R_{\varepsilon,\varphi}$, we have

$$\begin{aligned}\mathbb{E}\Big(\sup_{t\in[0,T]}|R_{\varepsilon,\varphi}(t)|\Big)\leqslant&\|\nabla\varphi\|_{L^\infty}\mathbb{E}\Big(\int_0^T\int_{\mathbb{R}^d}|[K_1-K_{1,\varepsilon}]*v_t(x)g_t(x)|\,\mathrm{d}x\mathrm{d}t\Big)\\&+\|\nabla\varphi\|_{L^\infty}\mathbb{E}\Big(\int_0^T\int_{\mathbb{R}^d}|[K_2-K_{2,\varepsilon}]*v_t(x)g_t(x)|\,\mathrm{d}x\mathrm{d}t\Big)\\:=&J_1^\varepsilon+J_2^\varepsilon.\end{aligned}$$

For J_1^ε, applying Young's inequality for the convolution of two functions and Hölder's inequality gives

$$\begin{aligned}J_1^\varepsilon\leqslant&C\mathbb{E}\Big(\int_0^T\|K_1-K_{1,\varepsilon}\|_{L^{p_1}}\|v_t\|_{L^1}\|g_t\|_{L^p}\mathrm{d}t\Big)\\\leqslant&C\|K_1-K_{1,\varepsilon}\|_{L^{p_1}_{q_1}}\mathbb{E}\|g\|_{L^p_q},\qquad\Big(\frac{1}{p_1}+\frac{1}{p}=\frac{1}{q_1}+\frac{1}{q}=1\Big).\end{aligned}$$

The condition ($\mathbf{K_r}$) together with the relationship between (p,q) and (p_1,q_1) exactly leads to the condition (3.31), so that we can apply Corollary 3.25 to find $\mathbb{E}\|g\|_{L^p_q}<\infty$. We thus have

$$J_1^\varepsilon\leqslant C\|K_1-K_{1,\varepsilon}\|_{L^{p_1}_{q_1}}\xrightarrow{\varepsilon\to0}0.$$

Since it does not involve the divergence of K_1, the computation for K_1 applies to the less singular part K_2 as well, with p_1,q_1 replaced by (p_2,q_2). We obtain the convergence of J_2^ε and arrive at

$$\mathbb{E}\Big(\sup_{t\in[0,T]}|R_{\varepsilon,\varphi}(t)|\Big)\xrightarrow{\varepsilon\to0}0.$$

The second uniform convergence in (3.41) is similar. Since it concerns on N-particles, the regularity result we shall apply is Proposition 3.8 instead of Corollary 3.25. Again, the technical result Corollary 3.7 will be used to handle non-exchangeability. We start with a simple bound for $|R^N_{\varepsilon,\varphi}|$,

$$\begin{aligned}\mathbb{E}\Big(\sup_{t\in[0,T]}|R^N_{\varepsilon,\varphi}|\Big)\leqslant&\|\nabla\varphi\|_{L^\infty}\mathbb{E}\Big(\int_0^T\frac{1}{N^2}\sum_{i\neq j}|\tilde{w}_i^N||w_j^N|\Big|\Big[K_1-K_{1,\varepsilon}\Big](X_i-X_j)\Big|\,\mathrm{d}t\Big)\\&+\|\nabla\varphi\|_{L^\infty}\mathbb{E}\Big(\int_0^T\frac{1}{N^2}\sum_{i\neq j}|\tilde{w}_i^N||w_j^N|\Big|\Big[K_2-K_{2,\varepsilon}\Big](X_i-X_j)\Big|\,\mathrm{d}t\Big)\\:=&J_1^{\varepsilon,N}+J_2^{\varepsilon,N}.\end{aligned}$$

We only give the details for the bound of $J_1^{\varepsilon,N}$ explicitly and the required bound for $J_2^{\varepsilon,N}$ follows similarly. First, applying Hölder's inequality w.r.t. the sum over i,j leads to

$$J_1^{\varepsilon,N}=\|\nabla\varphi\|_{L^\infty}\int_0^T\frac{1}{N^2}\sum_{i\neq j}\mathbb{E}\left(|\tilde{w}_i^N||w_j^N|\Big|\Big[K_1-K_{1,\varepsilon}\Big](X_i-X_j)\Big|\right)\mathrm{d}t$$

$$\leqslant C_\varphi \int_0^T \|w^N\|_{l^r}\|\tilde{w}^N\|_{l^r}\bigg(\frac{1}{N^2}\sum_{i\neq j}\mathbb{E}\Big|\Big[K_1-K_{1,\varepsilon}\Big](X_i-X_j)\Big|^{\frac{r}{r-1}}\bigg)^{\frac{r-1}{r}}\mathrm{d}t.$$

Since the two sequences $\{w^N\}$ and $\{\tilde{w}^N\}$ are uniformly bounded in l^r, there exists a universal constant $C>0$ such that

$$J_1^{\varepsilon,N}\leqslant C\int_0^T \bigg(\frac{1}{N^2}\sum_{i\neq j}\mathbb{E}\Big|\Big[K_1-K_{1,\varepsilon}\Big](X_i-X_j)\Big|^{\frac{r}{r-1}}\bigg)^{\frac{r-1}{r}}\mathrm{d}t.$$

Applying Corollary 3.7 with $|K_1-K_{1,\varepsilon}|$ playing the role of $\tilde{K}$, we get

$$J_1^{\varepsilon,N}\leqslant C\|K_1-K_{1,\varepsilon}\|_{L^{p_1}_{q_1}}\int_0^T\Big(1+\frac{1}{N}I(F_t^N)\Big)\mathrm{d}t.$$

By Proposition 3.8, Assumptions (**H**), ($\mathbf{K_r}$) and the convergence of $K_{1,\varepsilon}$ to K_1, we conclude that

$$J_1^{\varepsilon,N}\leqslant C\|K_1-K_{1,\varepsilon}\|_{L^{p_1}_{q_1}}\xrightarrow{\varepsilon\to 0}0.$$

The claim (3.41) is thus proved.

Proof [Proof of Theorem 1.13] The result follows by combining Corollary 3.23 and Proposition 3.26. The regularity estimates in Definition 1.11 are obtained by Corollary 3.25.

3.4.3 Uniqueness

In this section, we prove the uniqueness of the mean-field system (1.16). We divide Theorem 1.14 into the following Theorem 3.27 and Theorem 3.28.

Theorem 3.27. *There exists a unique solution $(v,g)\in C([0,T],\mathcal{M}(\mathbb{R}^d))$ to* (1.16) *in the sense of Definition 1.11, if the kernel K belongs to $L^{q_2}([0,T],L^{p_2}(\mathbb{R}^d))$ with*

$$\frac{d}{p_2}+\frac{2}{q_2}+\frac{1}{r}\leqslant 1,\quad \frac{d}{p_2}+\frac{2}{q_2}<1. \tag{3.42}$$

Proof The proof consists of two parts: the uniqueness of the solution v to the first equation in (1.16) and the uniqueness of the solution g to the second equation (1.16) in the sense of Definition 1.11. Observe that g solves a linear equation depending on v, then it is natural to study the equation of v first.

Uniqueness of v:

For general $L^{p_2}_{q_2}$-type kernel, the proof is through the mild formulation of (1.16). We consider the equation for v,

$$v_t=\Gamma_t * v_0-\int_0^t\nabla\Gamma_{t-s}*\Big(K*v_sv_s\Big)\mathrm{d}s.$$

Let $\kappa>0$ be a positive number satisfying

$$0<\kappa<\min\bigg\{\frac{1}{p_2},\frac{2}{d}(1-\frac{1}{r}),\frac{1}{2d}\Big(\frac{1}{r}+1-[\frac{d}{p_2}+\frac{2}{q_2}+\frac{1}{r}]\Big)\bigg\}<\frac{1}{d}. \tag{3.43}$$

The constraint (3.42) implies that it happens either $\frac{d}{p_2} + \frac{2}{q_2} + \frac{1}{r} < 1$ or $r < \infty$, which together with $r > 1$ ensures the existence of κ.

Suppose that there exist two solutions v^1 and v^2 starting from the same initial data. Since $\kappa < \frac{2}{d}(1 - \frac{1}{r})$, we deduce from Definition 1.11 that v^1 and v^2 belong to $L^{\frac{2}{d\kappa}}([0, T, L^{\frac{1}{1-\kappa}}(\mathbb{R}^d))$. Computing the $L^{\frac{1}{1-\kappa}}$-norm of $v^1 - v^2$ then leads to

$$\begin{aligned}
\|v_t^1 - v_t^2\|_{L^{\frac{1}{1-\kappa}}} &\leqslant \int_0^t \left\| \nabla\Gamma_{t-s} * \left(K * v_s^1 v_s^1 - K * v_s^2 v_s^2 \right) \right\|_{L^{\frac{1}{1-\kappa}}} \mathrm{d}s \\
&\leqslant \int_0^t \left\| \nabla\Gamma_{t-s} * \left(K * v_s^1 \left[v_s^1 - v_s^2 \right] \right) \right\|_{L^{\frac{1}{1-\kappa}}} \mathrm{d}s \\
&\quad + \int_0^t \left\| \nabla\Gamma_{t-s} * \left(K * \left[v_s^1 - v_s^2 \right] v_s^2 \right) \right\|_{L^{\frac{1}{1-\kappa}}} \mathrm{d}s,
\end{aligned} \tag{3.44}$$

where Γ_t is the the heat kernel of Δ. Using Young's convolution inequality, we have

$$\begin{aligned}
&\int_0^t \left\| \nabla\Gamma_{t-s} * \left(K * v_s^1 \left[v_s^1 - v_s^2 \right] \right) \right\|_{L^{\frac{1}{1-\kappa}}} \mathrm{d}s \\
\leqslant &\int_0^t \|\nabla\Gamma_{t-s}\|_{L^{\frac{1}{1-\kappa^2}}} \left\| K * v_s^1 \left[v_s^1 - v_s^2 \right] \right\|_{L^{\frac{1}{1-\kappa(1-\kappa)}}} \mathrm{d}s.
\end{aligned}$$

Furthermore, by the property of heat kernel $\|\nabla\Gamma_t\|_{L^q} \lesssim t^{\frac{d}{2q} - \frac{d+1}{2}}$ for $q \geqslant 1$, taking $q = \frac{1}{1-\kappa^2}$ gives

$$\begin{aligned}
&\int_0^t \left\| \nabla\Gamma_{t-s} * \left(K * v_s^1 \left[v_s^1 - v_s^2 \right] \right) \right\|_{L^{\frac{1}{1-\kappa}}} \mathrm{d}s \\
\lesssim &\int_0^t (t-s)^{\frac{d}{2}(1-\kappa^2) - \frac{d+1}{2}} \left\| K * v_s^1 \right\|_{L^{\frac{1}{\kappa^2}}} \|v_s^1 - v_s^2\|_{L^{\frac{1}{1-\kappa}}} \mathrm{d}s \\
\lesssim &\int_0^t (t-s)^{-\frac{1+d\kappa^2}{2}} \|K\|_{L^{p_2}} \left\| v_s^1 \right\|_{L^{p_3}} \|v_s^1 - v_s^2\|_{L^{\frac{1}{1-\kappa}}} \mathrm{d}s, \qquad \frac{1}{p_3} = 1 + \kappa^2 - \frac{1}{p_2}.
\end{aligned} \tag{3.45}$$

In the Definition 1.11, the maximal spatial integrability we obtained for v_t^1 and v_t^2 is $L^p(\mathbb{R}^d)$ with $p = \frac{d}{d-2+\frac{2}{r}} < \infty$; this excludes the case $d = 2$ and $r = \infty$, where the range of p is $[1, \infty)$. Now we check that $p_3 \in (1, p)$. On one hand, $p_3 > 1$ follows by

$$1 + \kappa^2 - \frac{1}{p_2} < 1 + \kappa - \frac{1}{p_2} < 1,$$

where we used (3.43). On the other hand, we find the upper bound $p_3 < p$ by noticing

$$1 + \kappa^2 - \frac{1}{p_2} \geqslant 1 + \kappa^2 - \frac{1}{d}(1 - \frac{1}{r}) = 1 - \frac{2}{d}(1 - \frac{1}{r}) + \kappa^2 + \frac{1}{d}(1 - \frac{1}{r}) > \frac{1}{p},$$

where the first inequality follows by (3.42), while the last inequality is given by $\frac{1}{p} = 1 - \frac{2}{d}(1 - \frac{1}{r})$.

By Corollary 3.25, we can take $q_3 > 1$ such that $\frac{1}{q_3} = \frac{d}{2}(1 - \frac{1}{p_3})$. Let $m \geqslant 1$ such that $\frac{1}{q_3} + \frac{1}{q_2} + \frac{1}{m} + \frac{d\kappa}{2} = 1$, we find

$$\frac{1}{m} = 1 - \frac{1}{q_3} - \frac{1}{q_2} - \frac{d\kappa}{2}$$

$$
\begin{aligned}
&=1+\frac{d}{2}(\frac{1}{p_3}-1)-\frac{1}{2}\Big(\frac{d}{p_2}+\frac{2}{q_2}+\frac{1}{r}-1\Big)+\frac{d}{2p_2}+\frac{1}{2r}-\frac{1}{2}-\frac{d\kappa}{2}\\
&=\frac{1}{2}+\frac{d}{2}\Big(\frac{1}{p_3}+\frac{1}{p_2}-1\Big)+\frac{1}{2r}-\frac{1}{2}\Big(\frac{d}{p_2}+\frac{2}{q_2}+\frac{1}{r}-1\Big)-\frac{d\kappa}{2}\\
&=\frac{1}{2}+\frac{d\kappa^2}{2}+\frac{1}{2r}-\frac{1}{2}\Big(\frac{d}{p_2}+\frac{2}{q_2}+\frac{1}{r}-1\Big)-\frac{d\kappa}{2},
\end{aligned}
$$

where we used the condition on (κ, p_3, p_2) in (3.45) to find the last equality. Recall the condition on κ, we have

$$
\frac{1}{m}>\frac{1+d\kappa^2}{2}. \tag{3.46}
$$

Applying Hölder's inequality to (3.45), we find

$$
\begin{aligned}
&\int_0^t \Big\|\nabla\Gamma_{t-s}*\Big(K*v_s^1\Big[v_s^1-v_s^2\Big]\Big)\Big\|_{L^{\frac{1}{1-\kappa}}}\,\mathrm{d}s\\
\lesssim&\|K\|_{L^{p_2}_{q_2}}\|v^1\|_{L^{p_3}_{q_3}}\|v^1-v^2\|_{L^{\frac{1}{1-\kappa}}_{\frac{2}{d\kappa}}}\Big(\int_0^t (t-s)^{-\frac{1+d\kappa^2}{2}m}\mathrm{d}s\Big)^{\frac{1}{m}}.
\end{aligned}
$$

By (3.46) and the regularity estimate in the Definition 1.11, we conclude that

$$
\int_0^t \Big\|\nabla\Gamma_{t-s}*\Big(K*v_s^1\Big[v_s^1-v_s^2\Big]\Big)\Big\|_{L^{\frac{1}{1-\kappa}}}\,\mathrm{d}s\lesssim\|v^1-v^2\|_{L^{\frac{1}{1-\kappa}}_{\frac{2}{d\kappa}}}. \tag{3.47}
$$

Similarly, we have

$$
\begin{aligned}
&\int_0^t \Big\|\nabla\Gamma_{t-s}*\Big(K*\Big[v_s^1-v_s^2\Big]v_s^2\Big)\Big\|_{L^{\frac{1}{1-\kappa}}}\,\mathrm{d}s\\
\lesssim&\int_0^t (t-s)^{-\frac{1+d\kappa^2}{2}}\Big\|K*\Big[v_s^1-v_s^2\Big]v_s^2\Big\|_{L^{\frac{1}{1-\kappa(1-\kappa)}}}\,\mathrm{d}s\\
\lesssim&\int_0^t (t-s)^{-\frac{1+d\kappa^2}{2}}\|K\|_{L^{p_2}}\big\|v_s^2\big\|_{L^{p_3}}\|v_s^1-v_s^2\|_{L^{\frac{1}{1-\kappa}}}\mathrm{d}s\\
\lesssim&\|v^1-v^2\|_{L^{\frac{1}{1-\kappa}}_{\frac{2}{d\kappa}}}.
\end{aligned} \tag{3.48}
$$

Combining (3.44)-(3.48), we arrive at

$$
\|v_t^1-v_t^2\|^{\frac{2}{d\kappa}}_{L^{\frac{1}{1-\kappa}}}\lesssim\int_0^t\|v_s^1-v_s^2\|^{\frac{2}{d\kappa}}_{L^{\frac{1}{1-\kappa}}}\mathrm{d}s. \tag{3.49}
$$

Therefore we obtain the $v^1=v^2$ in $L^{\frac{1}{1-\kappa}}(\mathbb{R}^d)$ for almost all $t\in(0,T]$ by applying Gronwall's inequality. We then conclude the uniqueness.

Uniqueness of g:

Now we consider the mild formulation of g,

$$
g_t=\Gamma_t*g_0-\int_0^t\nabla\Gamma_{t-s}*\Big(K*v_sg_s\Big)\mathrm{d}s.
$$

Observe that this is a linearized version of the equation for g. Similarly, suppose that there exist two solutions g^1 and g^2, then studying the $L^{\frac{1}{1-\kappa}}$-norm of $g^1 - g^2$ leads to

$$\|g_t^1 - g_t^2\|_{L^{\frac{1}{1-\kappa}}} \leqslant \int_0^t \left\|\nabla\Gamma_{t-s} * \left(K * v_s\left[g_s^1 - g_s^2\right]\right)\right\|_{L^{\frac{1}{1-\kappa}}} \mathrm{d}s.$$

Similar to (3.47), we find

$$\|g_t^1 - g_t^2\|_{L^{\frac{1}{1-\kappa}}} \lesssim \left(\int_0^t \|g_s^1 - g_s^2\|_{L^{\frac{1}{1-\kappa}}}^{\frac{2}{d\kappa}} \mathrm{d}s\right)^{\frac{d\kappa}{2}},$$

the proof is thus completed by applying Gronwall's inequality.

When K is the Biot-Savart kernel on dimension two, Theorem 1.14 is indeed the uniqueness of solutions to the passive scalar advented by the 2D Navier-Stokes equation.

Theorem 3.28. *There exists a unique solution* $(v, g) \in C([0,T], \mathcal{M}(\mathbb{R}^2))$ *to* (1.16) *in the sense of Definition 1.11 if the kernel K is the 2D Biot-Savart kernel*(1.11) *and* $r \in [3, \infty]$.

Proof When K is the Biot-Savart law, v solves the vorticity formulation of 2D Navier-Stokes equation. For this case, the uniqueness of solutions with the regularity properties [FHM14, (2.6)] (i.e. Corollary 3.25 with $r = \infty$) is already obtained in [FHM14], using the well-posedness result in the space $C([0,T), L^1(\mathbb{R}^2) \cap C((0,T), L^\infty(\mathbb{R}^d))$ from [BA94] and the remark [Bre94]. The strategy in [FHM14] is to improve the regularity of solutions by the DiPerna-Lions' renormalized solution and the maximal regularity of the heat equation so that the solution v meets the conditions in [BA94] and [Bre94].

The regularity result Corollary 3.25 is in fact a generalization of [FHM14, (2.6)]. In particular, Corollary 3.25 implies

$$v \in L^\infty([0,T], L^1(\mathbb{R}^2)) \cap L^{\frac{p}{p-1}}([0,T], L^p(\mathbb{R}^2)), \quad p \leqslant r \text{ and } 1 < p < \infty;$$

and

$$\nabla v \in L^{\frac{2q}{3q-2}}([0,T], L^q(\mathbb{R}^2), \quad q \leqslant \frac{2r}{r+2},\ q < 2, \text{ and } 1 \leqslant q < 2.$$

Although here we have the extra restrictions $p \leqslant r$ and $q \leqslant 2r/(r+2)$, by letting $r \geqslant 3$, one can track the proof [FHM14, Theorem 2.5] to get the uniqueness of the solutions to the vorticity form of the 2D Navier-Stokes equation. Therefore, v is the unique solution in the sense of Definition 1.11. Furthermore, the remark [Bre94] by Brezis shows that for L^1-valued initial data,

$$\lim_{t \to 0} t\|v_t\|_{L^\infty} = 0. \tag{3.50}$$

We then use $|K(y)| \lesssim |y|^{-1}$ to obtain

$$\begin{aligned}
|K * v_t(x)| \leqslant & \left|\int_{|y| \leqslant \sqrt{\frac{t}{c(t)}}} K(y) v_t(x-y) \mathrm{d}y\right| + \left|\int_{|y| > \sqrt{\frac{t}{c(t)}}} K(y) v_t(x-y) \mathrm{d}y\right| \\
\leqslant & \frac{1}{2\pi}\|v_t\|_{L^\infty} \int_{|y| \leqslant \sqrt{\frac{t}{c(t)}}} \frac{1}{|y|} \mathrm{d}y + \frac{1}{2\pi}\sqrt{\frac{c(t)}{t}}\|v_t\|_{L^1}
\end{aligned}$$

$$\lesssim \sqrt{\frac{t}{c(t)}} \|v_t\|_{L^\infty} + \sqrt{\frac{c(t)}{t}},$$

for $c(t) > 0$ and all $x \in \mathbb{R}^2$. Letting $c(t) = t\|v_t\|_{L^\infty}$ and applying (3.50), we arrive at

$$\lim_{t\to 0} t^{\frac{1}{2}} \|K * v_t\|_{L^\infty} \lesssim \lim_{t\to 0} \sqrt{t\|v_t\|_{L^\infty}} = 0. \tag{3.51}$$

Suppose there exist two solutions g^1 and g^2 to the second equation in (1.16). By the mild formulations of solutions, we have

$$\begin{aligned} \|g_t^1 - g_t^2\|_{L^1} &\leqslant \int_0^t \left\| \nabla \Gamma_{t-s} * \left(K * v_s \left[g_s^1 - g_s^2 \right] \right) \right\|_{L^1} \mathrm{d}s \\ &\lesssim \int_0^t (t-s)^{-\frac{1}{2}} \|K * v_s\|_{L^\infty} \left\| g_s^1 - g_s^2 \right\|_{L^1} \mathrm{d}s. \end{aligned}$$

Using the time regularity result (3.51), we find

$$\|g_t^1 - g_t^2\|_{L^1} \lesssim c_0(t) \int_0^t (t-s)^{-\frac{1}{2}} s^{-\frac{1}{2}} \left\| g_s^1 - g_s^2 \right\|_{L^1} \mathrm{d}s,$$

where $c_0(t) = \sup_{s\in[0,t]} s^{\frac{1}{2}} \|K * v_s\|_{L^\infty} \to 0$ as $t \to 0$. We then deduce $g_t^1 = g_t^2$ up to a short time $t_0 > 0$ by Gronwall's inequality of Volterra type, see for instance [Zha10, Example 2.4]. Applying this argument for finite times, we conclude the the uniqueness for all $t \in [0, T]$.

Chapter 4

Graphon particle systems

In this chapter, we consider interacting diffusions with weights on graphs (1.23), and we recall the model for convenience,

$$\mathrm{d}X_i = \frac{1}{N}\sum_{j\neq i} w_{ij}^N K(X_i - X_j)\,\mathrm{d}t + \sqrt{2}\mathrm{d}B_i, \quad i = 1,\dots,N, \tag{4.1}$$

where $\{w_{ij}^N\}$ are non-identical deterministic weights satisfying the condition $(\mathbf{W}'_{\mathbf{r}})$ for some $r > 1$, see Section 1.3. We shall show the mean-field convergence for the graphon particle system (4.1). The setting in this chapter is on $\mathbb{R}^d$.

4.1 Preliminary

In the graph limit theory [Lov12, BCCZ], a graphon G is a symmetric integrable function mapping from $[0,1]\times[0,1]$ into $\mathbb{R}$. The most important metric on the space of graphons is the cut metric, defined through the cut norm. Given a graphon G, define the cut norm by

$$\|G\|_\square := \sup_{S,T\subset\mathcal{B}([0,1])} \left| \int_{S\times T} G(\xi,\zeta)\mathrm{d}\xi\mathrm{d}\zeta \right|.$$

And given two graphons G and G', we define the cut metric (or the cut distance) by

$$\delta_\square(G,G') := \inf_\sigma \|G^\sigma - G'\|_\square,$$

where σ ranges over all measure-preserving bijections $[0,1]\to[0,1]$ and $G^\sigma(\xi,\zeta) := G(\sigma(\xi),\sigma(\zeta))$.

It's worth noting that the cut norm is equivalent to the operator norm from $L^\infty([0,1])$ to $L^1([0,1])$, where

$$\|G\|_{\infty\to 1} := \sup_{\|\varphi\|_{L^\infty},\|\phi\|_{L^\infty}\leqslant 1} \left| \int_0^1\int_0^1 G(\xi,\zeta)\varphi(\xi)\phi(\zeta)\mathrm{d}\xi\mathrm{d}\zeta \right|.$$

In particular, $\|G\|_\square \leqslant \|G\|_{\infty\to 1} \leqslant 4\|G\|_\square$.

The assumption $(\mathbf{W}'_{\mathbf{r}})$ implies that the sequence of graphons $\{G_N\}$ is uniformly $L^\infty L^r$-bounded, since

$$\sup_N \|G_N\|_{L^\infty L^r}^r = \sup_N \operatorname{ess\,sup}_{\xi\in[0,1]} \int_0^1 |G_N(\xi,\zeta)|^r \,\mathrm{d}\zeta$$

$$
\begin{aligned}
&= \sup_N \text{ess sup}_{\xi\in[0,1]} \int_0^1 \sum_{i,j} \left|w_{ij}^N\right|^r 1_{[\frac{i-1}{N},\frac{i}{N})}(\xi) 1_{[\frac{j-1}{N},\frac{j}{N})}(\zeta)\mathrm{d}\zeta \\
&= \sup_N \text{ess sup}_{\xi\in[0,1]} \sum_{i=1}^N \left(\frac{1}{N}\sum_{j\neq i}\left|w_{ij}^N\right|^r\right) 1_{[\frac{i-1}{N},\frac{i}{N})}(\xi) \\
&= \sup_N \max_{1\leqslant i\leqslant N}\left(\frac{1}{N}\sum_{j\neq i}\left|w_{ij}^N\right|^r\right) < \infty.
\end{aligned}
$$

Therefore, we find $\{G_N\}_{N\geqslant 2}$ is precompact with respect to the weak-*topology of $L^r([0,1]\times[0,1])$. Indeed, $\{G_N\}_{N\geqslant 2}$ is precompact with respect to the cut metric as well, which is deduced by the result about compactness of graphons [BCCZ, Theorem 2.13]:

Lemma 4.1. *Let $r\in(1,\infty]$ and $C>0$, and let $(G_n)_{n\geqslant 0}$ be a sequence of graphons with $\|G_n\|_{L^r}\leqslant C$ for all n. Then there exists a graphon G with $\|G\|_{L^r}\leqslant C$ such that*

$$\liminf_{n\to\infty}\delta_\square(G_n,G)=0.$$

In other words, the set $\{$graphons G; $\|G\|_{L^r}\leqslant C\}$ is compact with respect to the cut metric.

Furthermore, by [Lov12, Theorem 11.59], we have

Lemma 4.2. *Let $\{G_N\}$ be a converging subsequence (without relabeling) with repect to the cut metric and let G be the limiting point. There exist permutations $\{\sigma_N\}$ such that $\|G_N^{\sigma_N}-G\|_\square\to 0$.*

4.2 Well-posedness of graphon particle systems

In this section, we show the existence and uniqueness in law of probabilistically weak solutions to (4.1).

We start with the regularized system:

$$\mathrm{d}X_i^\varepsilon = \frac{1}{N}\sum_{j\neq i} w_{ij}^N K_\varepsilon(X_i^\varepsilon - X_j^\varepsilon) + \mathrm{d}B_i, \quad i=1,...,N., \tag{4.2}$$

where the regularized kernel $K_\varepsilon\in C_0^\infty([0,T]\times\mathbb{R}^d;\mathbb{R}^d)$ is constructed by convolution with mollifiers and trucations such that

$$\|K_\varepsilon - K\|_{L_q^p}\xrightarrow{\varepsilon\to 0} 0.$$

Although it was mentioned in Chapter 3, we note again that such regularizations are unavailable when $p=\infty$, because of the simple fact that L^∞ is not separable. However, we can still regularized the kernel locally. Since bounded kernels are less singular in our setting, we omit the details of regularizations for the case $p=\infty$.

Classically, the regularized SDEs (4.2) with the initial value X_0^N has a unique probabilistically strong solution $X^{\varepsilon,N}$ with smooth time-marginal densities $F_t^{\varepsilon,N}$, $t>0$, for each ε and N. The next lemma gives uniform Fisher information of the regularized systems.

Lemma 4.3. *Suppose that* $(\mathbf{W}_\mathbf{r}')$, $(\mathbf{K}_\mathbf{r}')$ *and* $(\mathbf{H})$ *hold for some* $r \in (1,\infty]$, *there exists a positive constant* C_T *such that for all* $t \in [0,T]$, $\varepsilon > 0$, $N \in \mathbb{N}$ *and* $\gamma \in (0,1)$, *such that*

$$\begin{aligned} &H(F_t^{\varepsilon,N}) + \sum_{i=1}^N \int_{\mathbb{R}^{dN}} \langle x_i\rangle^\gamma F_t^{\varepsilon,N} \mathrm{d}x^N + \frac{1}{2}\int_0^t I(F_s^{\varepsilon,N})\mathrm{d}s \\ \leqslant &H(F_0^N) + \sum_{i=1}^N \int_{\mathbb{R}^{dN}} \langle x_i\rangle^\gamma F_0^N \mathrm{d}x^N + C_T N. \end{aligned}$$

Proof The proof is similar to the proof of Proposition 3.8. For any $\varphi \in C^2(\mathbb{R}^{dN})$ vanishing at infinity, applying Itô's formula to $\varphi(X^{\varepsilon,N})$ and taking expectation, we arrive at the Liouville equation of $F^{\varepsilon,N}$ as

$$\partial_t F^{\varepsilon,N} = \Delta F^{\varepsilon,N} - \sum_{i=1}^N \mathrm{div}_{x_i}\Big(F^{\varepsilon,N}\frac{1}{N}\sum_{j\neq i} w_{ij}^N K_\varepsilon(x_i - x_j)\Big). \tag{4.3}$$

Evolving the entropy of $F^{\varepsilon,N}$, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}H(F^{\varepsilon,N}) = -\, I(F^{\varepsilon,N}) + \frac{1}{N}\sum_{i\neq j}\int_{\mathbb{R}^{dN}} \nabla_i F^{\varepsilon,N}\cdot w_{ij}^N K_\varepsilon(x_i-x_j)\mathrm{d}x^N. \tag{4.4}$$

By Young's inequality, we obtain

$$\begin{aligned} &\left|\frac{1}{N}\sum_{i\neq j}\int_{\mathbb{R}^{dN}} \nabla_i F^{\varepsilon,N}\cdot w_{ij}^N K_\varepsilon(x_i-x_j)\mathrm{d}x^N\right| \\ \leqslant &\frac{1}{N}\sum_{i\neq j}|w_{ij}^N|\left(\epsilon\int_{\mathbb{R}^{dN}}\frac{|\nabla_i F^{\varepsilon,N}|^2}{F^{\varepsilon,N}}\mathrm{d}x^N + C_\epsilon\int_{\mathbb{R}^{dN}} F^{\varepsilon,N}\left|K_\varepsilon(x_i-x_j)\right|^2\mathrm{d}x^N\right) \\ \leqslant &\epsilon\|w^N\|_{l^\infty l^1} I(F^{\varepsilon,N}) + \frac{C_\epsilon}{N}\sum_{i\neq j}|w_{ij}^N|\int_{\mathbb{R}^{dN}} F^{\varepsilon,N}\left|K_\varepsilon(x_i-x_j)\right|^2\mathrm{d}x^N. \end{aligned}$$

Furthermore, applying Hölder's inequality, we have

$$\begin{aligned} &\frac{1}{N}\sum_{i\neq j}|w_{ij}^N|\int_{\mathbb{R}^{dN}} F^{\varepsilon,N}\left|K_\varepsilon(x_i-x_j)\right|^2\mathrm{d}x^N. \\ \leqslant &N\int_{\mathbb{R}^{dN}} F^{\varepsilon,N}\|w^N\|_{l^r}\left(\frac{1}{N^2}\sum_{i\neq j}|K_\varepsilon(x_i-x_j)|^{\frac{2r}{r-1}}\right)^{\frac{r-1}{r}}\mathrm{d}x^N \\ \lesssim &N\|w^N\|_{l^r}\int_{\mathbb{R}^{dN}} F^{\varepsilon,N}\left(1+\frac{1}{N^2}\sum_{i\neq j}|K_\varepsilon(x_i-x_j)|^{\frac{2r}{r-1}}\right)\mathrm{d}x^N \\ \lesssim &N\|w^N\|_{l^r} + \|w^N\|_{l^r}\frac{1}{N}\sum_{i\neq j}\int_{\mathbb{R}^{dN}} F^{\varepsilon,N}|K_\varepsilon(x_i-x_j)|^{\frac{2r}{r-1}}\mathrm{d}x^N. \end{aligned}$$

By the condition $(\mathbf{K}_\mathbf{r}')$, we are allowed to apply Corollary 3.7 with $|K_\varepsilon|^2$ playing the role of $\tilde{K}$. There exists a constant $C_\epsilon' > 0$, depending on ϵ only, such that

$$\frac{C_\epsilon}{N}\sum_{i\neq j}|w_{ij}^N|\int_{\mathbb{R}^{dN}} F^{\varepsilon,N}\left|K_\varepsilon(x_i-x_j)\right|^2\mathrm{d}x^N \leqslant C_\epsilon' N + \frac{1}{8}\int_0^t I(F^{\varepsilon,N})\mathrm{d}s. \tag{4.5}$$

Choosing ϵ less than 1/8, we have

$$\left|\frac{1}{N}\sum_{i\neq j}\int_{\mathbb{R}^{dN}}\nabla_i F^{\varepsilon,N}\cdot w_{ij}^N K_\varepsilon(x_i-x_j)\mathrm{d}x^N\right|\leqslant CN+\frac{1}{4}\int_0^t I(F^{\varepsilon,N})\mathrm{d}s. \tag{4.6}$$

Combining (4.4) and (4.6) then yields that

$$\begin{aligned}H(F_t^{\varepsilon,N})-H(F_0^{\varepsilon,N})&\leqslant-\int_0^t I(F_s^{\varepsilon,N})\mathrm{d}s+CN+\frac{1}{4}\int_0^t I(F_s^{\varepsilon,N})\mathrm{d}s\\&\leqslant-\frac{3}{4}\int_0^t I(F_s^{\varepsilon,N})\mathrm{d}s+C_\Theta N.\end{aligned} \tag{4.7}$$

Here the constant C_Θ depends on $\Theta=\{\|w^N\|_{l^\infty l^r},\|K_\varepsilon\|_{L_q^p},p,q,r,d\}$.

On the other hand, testing $\partial_t F_t^{\varepsilon,N}$ with $\sum_i|x_i|^\gamma$ gives

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}\sum_{i=1}^N\int_{\mathbb{R}^{dN}}\langle x_i\rangle^\gamma F^{\varepsilon,N}\mathrm{d}x^N\\=&\sum_{i=1}^N\int_{\mathbb{R}^{dN}}F^{\varepsilon,N}\Delta_i\langle x_i\rangle^\gamma\mathrm{d}x^N+\frac{1}{N}\sum_{i\neq j}\int_{\mathbb{R}^{dN}}F^{\varepsilon,N}\nabla_i\langle x_i\rangle^\gamma w_{ij}^N K_\varepsilon(x_i-x_j)\mathrm{d}x^N.\end{aligned}$$

Since $\gamma\in(0,1)$, the functions $\Delta\langle\cdot\rangle^\gamma$ and $\nabla\langle\cdot\rangle^\gamma$ are bounded. This implies

$$\sum_{i=1}^N\left|\int_{\mathbb{R}^{dN}}F^{\varepsilon,N}\Delta_i\langle x_i\rangle^\gamma\mathrm{d}x^N\right|\leqslant CN,$$

and

$$\begin{aligned}&\frac{1}{N}\sum_{i\neq j}\left|\int_{\mathbb{R}^{dN}}F^{\varepsilon,N}\nabla_i\langle x_i\rangle^\gamma w_{ij}^N K_\varepsilon(x_i-x_j)\mathrm{d}x^N\right|\\\leqslant&\frac{C}{N}\sum_{i\neq j}|w_{ij}^N|\int_{\mathbb{R}^{dN}}F^{\varepsilon,N}\left|K_\varepsilon(x_i-x_j)\right|\mathrm{d}x^N,\end{aligned} \tag{4.8}$$

with a constant C depending on γ only. Handling the right hand side of (4.8) is similar to deriving the estimate (4.5), the difference is that here we apply Corollary 3.7 with $|K|$ playing the role of $\tilde{K}$. We thus conclude

$$\frac{1}{N}\sum_{i\neq j}\int_0^t\left|\int_{\mathbb{R}^{dN}}F^{\varepsilon,N}\nabla_i\langle x_i\rangle^\gamma w_{ij}^N K_\varepsilon(x_i-x_j)\mathrm{d}x^N\right|\mathrm{d}s\leqslant CN+\frac{1}{4}\int_0^t I(F^{\varepsilon,N})\mathrm{d}s. \tag{4.9}$$

Therefore, we have

$$\sum_{i=1}^N\int_{\mathbb{R}^{dN}}\langle x_i\rangle^\gamma F_t^{\varepsilon,N}\mathrm{d}x^N\leqslant\sum_{i=1}^N\int_{\mathbb{R}^{dN}}\langle x_i\rangle^\gamma F_0^{\varepsilon,N}\mathrm{d}x^N+CN+\frac{1}{4}\int_0^t I(F_s^{\varepsilon,N})\mathrm{d}s. \tag{4.10}$$

Now we conclude the uniform estimate by summing up (4.7) and (4.10).

Next we show the existence of probabilistically weak solution to the original system (4.1) by a tightness argument.

Lemma 4.4. *For each $N \in \mathbb{N}$, there exists a probabilistically weak solution X^N to (4.1). Furthermore, it holds uniformly for N that*

$$\begin{aligned} &H(F_t^N) + \sum_{i=1}^N \int_{\mathbb{R}^{dN}} \langle x_i \rangle^\gamma F_t^N \mathrm{d}x^N + \frac{1}{2}\int_0^t I(F_s^N)\mathrm{d}s \\ \lesssim &H(F_0^N) + \sum_{i=1}^N \int_{\mathbb{R}^{dN}} \langle x_i \rangle^\gamma F_0^N \mathrm{d}x^N + C_T N. \end{aligned} \tag{4.11}$$

Proof The proof follows by the tightness argument. We firstly show that the sequence of laws of $X^{\varepsilon,N}$ is tight on $C([0,T],\mathbb{R}^{dN})$. To see this, it suffices to find the uniform estimate

$$\sup_{\varepsilon>0} \mathbb{E}\left(\sup_{0\leqslant s<t\leqslant T} \frac{|X^{\varepsilon,N}(t) - X^{\varepsilon,N}(s)|}{(t-s)^{1-\alpha}}\right) + \mathbb{E}|X^N(0)|^\gamma < \infty, \tag{4.12}$$

where $\alpha \in (\frac{1}{2}, 1)$. The dependence on N is omitted since the tightness holds for fixed N.

By the equation of $X^{\varepsilon,N}$, we find

$$\begin{aligned} &\mathbb{E}\left(\sup_{0\leqslant s<t\leqslant T} \frac{|X^{\varepsilon,N}(t) - X^{\varepsilon,N}(s)|}{(t-s)^{1-\alpha}}\right) \\ =&\sum_{i=1}^N \mathbb{E}\sup_{s<t} \frac{\left|\int_s^t \frac{1}{N}\sum_{j\neq i} w_{ij}^N K(X_i^\varepsilon(\tau) - X_j^\varepsilon(\tau))\mathrm{d}\tau + \sqrt{2}(B_i(t) - B_i(s))\right|}{(t-s)^{1-\alpha}} \\ \leqslant &\sum_{i=1}^N \mathbb{E}\sup_{s<t} \frac{\left|\int_s^t \frac{1}{N}\sum_{j\neq i} w_{ij}^N K(X_i^\varepsilon(\tau) - X_j^\varepsilon(\tau))\mathrm{d}\tau\right|}{(t-s)^{1-\alpha}} + \sum_{i=1}^N \mathbb{E}\sup_{s<t} \frac{|(B_i(t) - B_i(s))|}{(t-s)^{1-\alpha}}. \end{aligned}$$

The second summation is independent of ε, and is finite due to the modulus continuity of Brownian motions. As to the first summation involving the interactions, we have

$$\begin{aligned} &\sum_{i=1}^N \mathbb{E}\sup_{s<t} \frac{\left|\int_s^t \frac{1}{N}\sum_{j\neq i} w_{ij}^N K(X_i^\varepsilon(\tau) - X_j^\varepsilon(\tau))\mathrm{d}\tau\right|}{(t-s)^{1-\alpha}} \\ \lesssim &\sum_{i\neq j} |w_{ij}^N| \mathbb{E}\left(\sup_{s<t} \frac{\int_s^t \left|K(X_i^\varepsilon - X_j^\varepsilon)\right|\mathrm{d}\tau}{(t-s)^{1-\alpha}}\right) \\ \lesssim &\|w^N\|_{l^r} \left(\sum_{i\neq j}\left[\mathbb{E}\left(\sup_{s<t} \frac{\int_s^t \left|K(X_i^\varepsilon - X_j^\varepsilon)\right|\mathrm{d}\tau}{(t-s)^{1-\alpha}}\right)\right]^{\frac{r}{r-1}}\right)^{\frac{r-1}{r}} \\ \lesssim &\|w^N\|_{l^r} \left(\sum_{i\neq j}\left[\mathbb{E}\int_0^T \left|K(X_i^\varepsilon - X_j^\varepsilon)\right|^{\frac{1}{\alpha}}\mathrm{d}t\right]^{\frac{\alpha r}{r-1}}\right)^{\frac{r-1}{r}} \\ \lesssim &\|w^N\|_{l^r} \left(T + \sum_{i\neq j}\mathbb{E}\int_0^T \left|K(X_i^\varepsilon - X_j^\varepsilon)\right|^{\max\{\frac{r}{r-1},\frac{1}{\alpha}\}}\mathrm{d}t\right)^{\frac{r-1}{r}}, \end{aligned}$$

where the constant T is given by Young's inequality $|x|^{\alpha r/(r-1)} \leqslant |x| + 1$ when $\alpha r/(r-1) \leqslant 1$. Using Corollary 3.7, we find the above summation is bounded by the Fisher information. That is

$$\sum_{i=1}^N \mathbb{E}\sup_{s<t} \frac{\left|\int_s^t \frac{1}{N}\sum_{j\neq i} w_{ij}^N K(X_i^\varepsilon(\tau) - X_j^\varepsilon(\tau))\mathrm{d}\tau\right|}{(t-s)^{1-\alpha}} \lesssim \|w^N\|_{l^r}\left(C + \int_0^T I(F_t^{\varepsilon,N})\mathrm{d}t\right)^{\frac{r-1}{r}},$$

for all $\varepsilon > 0$. We then deduce the uniform estimate (4.12) by Lemma 4.3. As a consequence of the Arzelà–Ascoli theorem, the sequence of laws of $X^{\varepsilon,N}$ is tight on $C([0,T],\mathbb{R}^{dN})$.

By the Skrohod theorem, up to a subsequence and on a new stochastic basis, there exists a continuous $\mathbb{R}^{dN}$-valued process X^N such that $X^{\varepsilon,N}$ converges to X^N in $C([0,T],\mathbb{R}^{dN})$ almosts surely. By the lower-semicontinuities of the Fisher information, Boltzmann entropy, finite moments functionals, we deduce that the time marginal law of X^N, denoted by F_t^N, satisfies

$$\begin{aligned}&H(F_t^N) + \sum_{i=1}^N \int_{\mathbb{R}^{dN}} \langle x_i \rangle^\gamma F_t^N \mathrm{d}x^N + \frac{1}{2}\int_0^t I(F_s^N)\mathrm{d}s \\ \leqslant &H(F_0^N) + \sum_{i=1}^N \int_{\mathbb{R}^{dN}} \langle x_i \rangle^\gamma F_0^N \mathrm{d}x^N + C_T N.\end{aligned}$$

To justify that X^N solves (4.1), we need to take the limit of the regularized system (4.2) through associated martingale problem, and the only non-trival part is the term of singular interactions. We find that

$$\begin{aligned}&\sum_{j\neq i} w_{ij}^N K_\varepsilon(X_i^\varepsilon - X_j^\varepsilon) - \sum_{j\neq i} w_{ij}^N K(X_i - X_j) \\ =&\left(\sum_{j\neq i} w_{ij}^N K_\sigma(X_i^\varepsilon - X_j^\varepsilon) - \sum_{j\neq i} w_{ij}^N K_\sigma(X_i - X_j)\right) \\ &+\left(\sum_{j\neq i} w_{ij}^N K_\varepsilon(X_i^\varepsilon - X_j^\varepsilon) - \sum_{j\neq i} w_{ij}^N K_\sigma(X_i^\varepsilon - X_j^\varepsilon)\right) \\ &+\left(\sum_{j\neq i} w_{ij}^N K_\sigma(X_i - X_j) - \sum_{j\neq i} w_{ij}^N K(X_i - X_j)\right) \\ :=&J_1(\sigma,\varepsilon) + J_2(\sigma,\varepsilon) + J_3(\sigma).\end{aligned}$$

Clearly, for fixed $\sigma > 0$, $J_1(\sigma,\varepsilon)$ converges to zero almost surely as $\varepsilon \to 0$. For $J_2(\sigma,\varepsilon)$, we have

$$\mathbb{E}\sup_{t\in[0,T]}\left|\int_0^t J_2(\sigma,\varepsilon)\right| \leqslant \left(CN^2 + N\int_0^T I(F^{\varepsilon,N})\mathrm{d}t\right)\|K_\varepsilon - K_\sigma\|_{L_q^p}^\alpha,$$

where $\alpha > 0$ depends on p,q,r and the constant C is universal. Deriving the above estimate is indeed in the spirit with obtaining (4.5), here we point out the constant (from Lemma 3.3) on the kernel explicitly. Therefore, we have

$$\mathbb{E}\sup_{t\in[0,T]}\left|\int_0^t J_2(\sigma,\varepsilon)\right| \lesssim \left(\|K_\varepsilon - K\|_{L_q^p} + \|K_\sigma - K\|_{L_q^p}\right)^\alpha \to 0.$$

Analogously, we have

$$\mathbb{E}\sup_{t\in[0,T]}\left|\int_0^t J_3(\sigma)\right| \to 0.$$

In conclusion, up to a subsequence, the term of singular interactions between (X_i^ε) converges to the interactions between (X_i) almost surely. Thus X^N solves (4.1).

Before we show the uniqueness in law of solutions to (4.1), we quote a relative entropy estimate given by Girsanov theory, see [Lac21, Lemma 4.4].

Lemma 4.5. *Given progressively measurable drifts $b^1, b^2 := [0,T]\times\mathbb{R}^k \to \mathbb{R}^k$, $k \in \mathbb{N}$. Suppose that for each $i = 1, 2$, the SDE*

$$\mathrm{d}Y_t^i = b^i(t, Y_t^i)\mathrm{d}t + \mathrm{d}B_t^i$$

admits a weak solution. Suppose further that the SDE with the drift b^2 is well-posed in the sense that unique in law for any initial value $x \in \mathbb{R}^k$. Denote the law of Y^i by $\mathcal{L}(Y^i) \in \mathcal{P}(C([0,T],\mathbb{R}^k))$. Then we have

$$H(\mathcal{L}(Y^1)|\mathcal{L}(Y^2)) = H(\mathcal{L}(Y_0^1)|\mathcal{L}(Y_0^2)) + \frac{1}{2}\mathbb{E}\int_0^T |b^1(t,Y_t^1) - b^2(t,Y_t^1)|^2\mathrm{d}t,$$

if the right hand side is finite.

Theorem 4.6. *Under the conditions $(\mathbf{W}_\mathbf{r}')$, $(\mathbf{K}_\mathbf{r}')$ and $(\mathbf{H})$ for some $r \in (1,\infty]$, there exists a unique (in law) probabilistically weak solution X^N to* (4.1) *which satisfying* (4.11).

Proof Suppose that there exist two solutions X^N and Y^N with the same initial distribution F_0^N satisfying $(\mathbf{H})$. Denote the law of X^N and Y^N by $\mathbb{F}$ and $\mathbb{F}'$ in $\mathcal{P}(C([0,T],\mathbb{R}^{dN}))$, respectively. Denote the law of regularized system by $\mathbb{F}^\varepsilon$.

Applying Lemma 4.5 gives that

$$\begin{aligned}
&H(\mathbb{F}|\mathbb{F}^\varepsilon) - H(F_0^N|F_0^N)\\
=&\sum_{i=1}^N \mathbb{E}\int_0^T \left|\frac{1}{N}\sum_{j\neq i} w_{ij}^N\Big(K(X_i - X_j) - K_\varepsilon(X_i - X_j)\Big)\right|^2 \mathrm{d}t\\
=&\sum_{i=1}^N \int_0^T\int_{\mathbb{R}^{dN}} \left|\frac{1}{N}\sum_{j\neq i} w_{ij}^N\Big(K(x_i - x_j) - K_\varepsilon(x_i - x_j)\Big)\right|^2 F_t^N\mathrm{d}x^N\mathrm{d}t\\
\leqslant&\sum_{i=1}^N \int_0^T\int_{\mathbb{R}^{dN}} \|w_{i\cdot}^N\|_{l^r}^2 \left(\frac{1}{N}\sum_{j\neq i}|K(x_i - x_j) - K_\varepsilon(x_i - x_j)|^{\frac{r}{r-1}}\right)^{\frac{2(r-1)}{r}} F_t^N\mathrm{d}x^N\mathrm{d}t\\
\leqslant&\|w^N\|_{l^\infty l^r}^2\sum_{i=1}^N \int_0^T\int_{\mathbb{R}^{dN}} \left(\frac{1}{N}\sum_{j\neq i}|K(x_i - x_j) - K_\varepsilon(x_i - x_j)|^{\frac{r}{r-1}}\right)^{\frac{2(r-1)}{r}} F_t^N\mathrm{d}x^N\mathrm{d}t
\end{aligned}$$

By Hölder's inequality, we have that

$$\begin{aligned}
H(\mathbb{F}|\mathbb{F}^\varepsilon) \lesssim&\sum_{i=1}^N \int_0^T\int_{\mathbb{R}^{dN}} \left(\frac{1}{N}\sum_{j\neq i}|K(x_i - x_j) - K_\varepsilon(x_i - x_j)|^{\frac{r}{r-1}}\right)^{\frac{2(r-1)}{r}} F_t^N\mathrm{d}x^N\mathrm{d}t\\
\lesssim&1 + \frac{1}{N^2}\sum_{i\neq j}\int_0^T\int_{\mathbb{R}^{dN}} |K(x_i - x_j) - K_\varepsilon(x_i - x_j)|^{\frac{2r}{r-1}} F_t^N\mathrm{d}x^N\mathrm{d}t.
\end{aligned}$$

Here the omiited constant is independent of ε. Now we apply Corollary 3.7 to find that

$$H(\mathbb{F}|\mathbb{F}^\varepsilon) \lesssim\|K - K_\varepsilon\|_{L_q^p}^{\frac{2r}{r-1}}\Big(1 + \int_0^T I(F_t^N)\mathrm{d}t\Big)$$

$$\lesssim \|K - K_\varepsilon\|_{L_q^p}^{\frac{2r}{r-1}},$$

where the last line follows by (4.11). Analogously, we have

$$H(\mathbb{F}'|\mathbb{F}^\varepsilon) \lesssim \|K - K_\varepsilon\|_{L_q^p}^{\frac{2r}{r-1}}.$$

Therefore, we deduce by the Csiszár-Kullback-Pinsker inequality that

$$\begin{aligned}\|\mathbb{F} - \mathbb{F}'\|_{TV} \leqslant& \|\mathbb{F} - \mathbb{F}^\varepsilon\|_{TV} + \|\mathbb{F}' - \mathbb{F}^\varepsilon\|_{TV} \\ \leqslant& \sqrt{\frac{1}{2}H(\mathbb{F}|\mathbb{F}^\varepsilon)} + \sqrt{\frac{1}{2}H(\mathbb{F}'|\mathbb{F}^\varepsilon)} \\ \lesssim& \|K - K_\varepsilon\|_{L_q^p}^{\frac{r}{r-1}}.\end{aligned}$$

Let $\varepsilon \to 0$, we conclude that $\mathbb{F} = \mathbb{F}'$.

4.3 Random measures with uniform Fisher information

This section is similar to Section 3.3. In this section, we study a sequence of random measures which is close to the sequence of empirical measures $\{\nu_N\}$, recall that $\nu_N(\mathrm{d}x, \xi)(t) = \sum_{i=1}^N \delta_{X_i(t)}(\mathrm{d}x) 1_{[\frac{i-1}{N}, \frac{i}{N})}(\xi)$.

4.3.1 Weakly merging sequences

We start with disintegrating $\mathrm{d}t \times F_t^N$ as

$$\begin{aligned}\mathrm{d}t \times F_t^N(\mathrm{d}x_1, ..., \mathrm{d}x_N) =& \mathrm{d}t \times f_t^1(\mathrm{d}x_1) f_t^2(x_1, \mathrm{d}x_2) ... f_t^N(x_1, ..., x_{N-1}, \mathrm{d}x_N) \\ :=& \mathrm{d}t \times \Pi_{i=1}^N f_t^i(x^{i-1,N}, \mathrm{d}x_i),\end{aligned}$$

The random measures $\{g^N\}$ on $\mathbb{R}^d \times [0, 1]$ are defined as

$$g_t^N(\mathrm{d}x, \xi) := \sum_{i=1}^N f_t^i(X_t^{i-1,N}, \mathrm{d}x) 1_{[\frac{i-1}{N}, \frac{i}{N})}(\xi), \tag{4.13}$$

for $X_t^{i-1,N} = (X_1(t), \dots, X_{i-1}(t))$. Notice that $g^N \in \mathcal{P}(\mathbb{R}^d \times [0, 1])$ for almost all $(t, \omega) \in [0, T] \times \Omega$.

Next we define *weakly merging* for sequence of measures on $\mathbb{R}^d \times [0, 1]$.

Definition 4.7. *Two sequences of finite Borel random measures $\{\mu_N(t, \omega)\}$ and $\{\nu_N(t, \omega)\}$ on $\mathbb{R}^d \times [0, 1]$ are called weakly merging if for each $\varphi \in C_b(\mathbb{R}^d \times [0, 1])$, the sequence of random variables $\{\langle \varphi, \mu_N - \nu_N \rangle\}$ converges to zero for all most all (t, ω).*

Lemma 4.8. *The sequences of random measures $\{\nu_N, N \in \mathbb{N}\}$ and $\{g^N, N \in \mathbb{N}\}$ are weakly merging.*

Proof For any test function $\varphi \in C_b(\mathbb{R}^d \times [0,1])$, we have

$$\begin{aligned}\left\langle \nu_N(t) - g^N(t), \varphi(x,\xi)\right\rangle =& \sum_{i=1}^N \int_{[\frac{i-1}{N},\frac{i}{N})} \varphi(X_i(t),\xi)\mathrm{d}\xi - \int_{[\frac{i-1}{N},\frac{i}{N})} \mathbb{E}(\varphi(X_i(t),\xi)|X^{i-1,N})\mathrm{d}\xi \\ :=& \sum_{i=1}^N M_i,\end{aligned}$$

Define the $\sigma-$fields generated by $(X_1(t), ...X_i(t))$ as $\mathcal{F}_i$, $i = 0, ..., N-1$. Observe that for each bounded continuous function φ on $\mathbb{R}^d \times [0,1]$,

$$\begin{aligned}&\mathbb{E}\left(\int_{[\frac{i-1}{N},\frac{i}{N})} \varphi(X_i(t),\xi)\mathrm{d}\xi \mid X^{i-1,N}\right) \\ =& \int_{\mathbb{R}^d}\left(\int_{[\frac{i-1}{N},\frac{i}{N})} \varphi(x,\xi)\mathrm{d}\xi\right) f_t^i(X_1(t),...,X_{i-1}(t)|\mathrm{d}x) \\ =& \int_{[\frac{i-1}{N},\frac{i}{N})}\int_{\mathbb{R}^d} \varphi(x,\xi)\mathrm{d}f_t^i(X_1(t),...,X_{i-1}(t)|\mathrm{d}x)\mathrm{d}\xi \\ =& \int_{[\frac{i-1}{N},\frac{i}{N})} \mathbb{E}\left(\varphi(X_i(t),\xi)|X^{i-1,N}\right)\mathrm{d}\xi, \quad a.s.,\end{aligned}$$

which implies that $\{M_i, i = 1, ..., N\}$ is a martingale difference sequence with respect to $(\mathcal{F}_i)$. The rest of proof is completed by the same argument as in Lemma 3.10.

4.3.2 Fisher information of the random measures

Now we study the regularity of $\{g^N\}$ as in Section 3.3.2. We first show the absolutely continuities of the random measures.

Lemma 4.9. *For each $N \in \mathbb{N}$, the random measure $g^N(t, \mathrm{d}x, \xi, \omega)$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^d$ for almost all (t, ξ, ω). Furthermore, we have for $\gamma \in (0,1)$,*

$$\mathbb{E}\int_0^1 H(g^N(t)|\theta_\gamma)\mathrm{d}\xi \leqslant \frac{1}{N} H(F_t^N|\theta_\gamma^{\otimes N}).$$

Proof Let θ be $\theta_\gamma := Ce^{-|x|^\gamma}$, where C is the normalizing constant such that $\|\theta\|_{L^1} = 1$. As in Lemma 3.11, we have

$$\mathbb{E}\sum_{i=1}^N H(f_t^i(X_t^{i-1,N},\cdot)|\theta_\gamma) = H(F_t^N|\theta_\gamma^{\otimes N}). \tag{4.14}$$

Then applying Jensen's inequality gives

$$\begin{aligned}\mathbb{E}\int_0^1 H(g^N(t)|\theta_\gamma)\mathrm{d}\xi \leqslant& \mathbb{E}\int_0^1 \sum_{i=1}^N 1_{[\frac{i-1}{N},\frac{i}{N})}(\xi) H(f_t^i(X_t^{i-1,N},\mathrm{d}x)|\theta_\gamma)\mathrm{d}\xi \\ =& \frac{1}{N}\mathbb{E}\sum_{i=1}^N H(f_t^i(X_t^{i-1,N},\mathrm{d}x)|\theta_\gamma).\end{aligned}$$

The result follows by (4.14).

The following lemma provides the Fisher information estimate for the auxiliary random measures.

Lemma 4.10. *It holds that*

$$\mathbb{E}\int_0^T\int_0^1 I(g^N(t))\mathrm{d}t\mathrm{d}\xi \leqslant \frac{1}{N}\int_0^T I(F_t^N)\mathrm{d}t. \tag{4.15}$$

Proof According to Lemma 3.13, it holds that

$$\sum_{i=1}^N \mathbb{E}\int_0^T I\Big(f_t^i(X_t^{i-1,N},\cdot)\Big)\mathrm{d}t \leqslant \int_0^T I(F_t^N)\mathrm{d}t.$$

On the other hand, since the Fisher information functional is convex, applying Jensen's inequality gives

$$\begin{aligned}\mathbb{E}\int_0^T\int_0^1 I(g^N(t))\mathrm{d}t\mathrm{d}\xi \leqslant& \mathbb{E}\int_0^T\int_0^1\sum_{i=1}^N 1_{[\frac{i-1}{N},\frac{i}{N})}(\xi) I\Big(f_t^i(X_t^{i-1,N},\cdot)\Big)\mathrm{d}t\mathrm{d}\xi\\ =&\frac{1}{N}\sum_{i=1}^N\mathbb{E}\int_0^T I\Big(f_t^i(X_t^{i-1,N},\cdot)\Big)\mathrm{d}t\\ \leqslant&\frac{1}{N}\int_0^T I(F_t^N)\mathrm{d}t.\end{aligned}$$

The result thus follows.

4.4 Mean-field convergence

In this section, we show the mean-field convergence of the graphon particle system (1.23). The proof is based on the compactness argument, and studying auxiliary random measures $\{g^N\}$ enables us to deal with models with non-exchangeability and singular interactions.

4.4.1 Tightness

We define the empirical measure Q^N on the path space $C([0,T],\mathbb{R}^d)\times[0,1]$,

$$Q^N := \sum_{i=1}^N \delta_{X_i} 1_{[\frac{i-1}{N},\frac{i}{N})}(\xi).$$

Note that $Q^N \in \mathcal{P}(C([0,T],\mathbb{R}^d)\times[0,1])$ and $Q^N\circ\pi_t^{-1} = \mu_N(t)$, where $\pi_t : C([0,T],\mathbb{R}^d)\times[0,1]\to\mathbb{R}^d\times[0,1]$ is the canonical mapping.

Let $\phi : C([0,T],\mathbb{R}^d)\to[0,\infty]$ be the function

$$\phi(X) := \sup_{0\leqslant s<t\leqslant T}\frac{|X(t)-X(s)|}{(t-s)^{1-\alpha}} + |X(0)|^{\frac{(r-1)\gamma}{r}}, \tag{4.16}$$

where $\gamma \in (0,1)$ and $\alpha \in (\frac{1}{2}, 1)$. The choice for such α ensures that for $\alpha^* := \frac{1}{1-\alpha}$,

$$2(1-\alpha) < 1, \quad 1 < \alpha^*, \quad \frac{d}{p_2} + \frac{2}{q_2} + \frac{2}{\alpha^*} < 2, \quad \frac{1}{\alpha} = \frac{\alpha^*}{\alpha^* - 1}.$$

Now that we show the following key uniform estimate.

Lemma 4.11. *Suppose that* (**H**), ($\mathbf{W}'_\mathbf{r}$) *and* ($\mathbf{K}'_\mathbf{r}$) *hold for some* $r \in (1, \infty)$. *It holds that*

$$\sup_N \mathbb{E} \int_0^1 \langle \phi, Q^N \rangle \mathrm{d}\xi = \sup_N \mathbb{E} \left\langle \phi, \frac{1}{N} \sum_{i=1}^N \delta_{X_i} \right\rangle < \infty. \tag{4.17}$$

Proof By the definition of ϕ, we indeed need to show

$$\sup_N \left(\frac{1}{N} \sum_{i=1}^N \mathbb{E} \, |X_i(0)|^{\frac{(r-1)\gamma}{r}} + \frac{1}{N} \sum_{i=1}^N \mathbb{E} \sup_{0 \leqslant s < t \leqslant T} \frac{|X_i(t) - X_i(s)|}{(t-s)^{1-\alpha}} \right) < \infty. \tag{4.18}$$

This is indeed the estimate (3.37) in Chapter 3. The only difference in the proof is that the weights $\{w_{ij}^N\}$ now depends on both i and j, we need to estimate

$$\frac{1}{N} \sum_{i=1}^N \mathbb{E} \sup_{s<t} \frac{\left| \int_s^t \frac{1}{N} \sum_{j \neq i} w_{ij}^N K(X_i(\tau) - X_j(\tau)) \mathrm{d}\tau \right|}{(t-s)^{1-\alpha}}.$$

By Hölder ineqeuality, we find

$$\begin{aligned}
&\frac{1}{N} \sum_{i=1}^N \mathbb{E} \sup_{s<t} \frac{\left| \int_s^t \frac{1}{N} \sum_{j \neq i} w_{ij}^N K(X_i(\tau) - X_j(\tau)) \mathrm{d}\tau \right|}{(t-s)^{1-\alpha}} \\
\leqslant & \frac{1}{N^2} \sum_{i \neq j} |\tilde{w}_{ij}^N| \mathbb{E} \Bigg(\sup_{s<t} \frac{\int_s^t |K(X_i - X_j)| \, \mathrm{d}\tau}{(t-s)^{1-\alpha}} \Bigg) \\
\leqslant & \|w^N\|_{l^r} \left(\frac{1}{N^2} \sum_{i \neq j} \left[\mathbb{E} \Bigg(\sup_{s<t} \frac{\int_s^t |K(X_i - X_j)| \, \mathrm{d}\tau}{(t-s)^{1-\alpha}} \Bigg) \right]^{\frac{r}{r-1}} \right)^{\frac{r-1}{r}} \\
\leqslant & \|w^N\|_{l^r} \left(\frac{1}{N^2} \sum_{i \neq j} \left[\mathbb{E} \int_0^T |K(X_i - X_j)|^{\frac{1}{\alpha}} \, \mathrm{d}t \right]^{\frac{\alpha r}{r-1}} \right)^{\frac{r-1}{r}} \\
\leqslant & \|w^N\|_{l^r} \left(T + \frac{1}{N^2} \sum_{i \neq j} \mathbb{E} \int_0^T |K(X_i - X_j)|^{\max\{\frac{r}{r-1}, \frac{1}{\alpha}\}} \, \mathrm{d}t \right)^{\frac{r-1}{r}},
\end{aligned}$$

where the constant T is given by Young's inequality $|x|^{\alpha r/(r-1)} \leqslant |x| + 1$ when $\alpha r/(r-1) \leqslant 1$. Now the proof is reduced to the case in the proof of Lemma 3.18. The result then follows.

Lemma 4.12. *Suppose that* (**H**), ($\mathbf{W}'_\mathbf{r}$) *and* ($\mathbf{K}'_\mathbf{r}$) *hold for some* $r \in (1, \infty)$. *The laws of the sequence* $\{Q^N, N \in \mathbb{N}\}$ *are tight on* $\mathcal{P}(C([0,T], \mathbb{R}^d) \times [0,1])$.

Proof It suffices to show the sequences of marginal laws of $\{Q^N\}$ are tight respectively.

Firstly, the marginal distributions on $[0,1]$ is independent of N by noticing

$$\int_{C([0,T],\mathbb{R}^d)} Q^N(\mathrm{d}X\mathrm{d}\xi) = \sum_{i=1}^{N} 1_{[\frac{i-1}{N},\frac{i}{N})}(\xi) = 1_{[0,1)}(\xi).$$

As to the marginal distributions on $C([0,T],\mathbb{R}^d)$, we need to show the tightness of the laws of $\{\frac{1}{N}\sum_{i=1}^N \delta_{X_i}\}$. This is the case studied in Lemma 3.19, and the estimate (4.17) plays a similar role of (3.36). We omitted the rest of proof for simplicity.

The next result concerns on tightness of laws of $\{g^N\}$.

Lemma 4.13. *Suppose that* $(\mathbf{H})$, $(\mathbf{W}_\mathbf{r}')$ *and* $(\mathbf{K}_\mathbf{r}')$ *hold for some* $r \in (1,\infty]$. *There exists* $p^* > 1$ *such that the laws of* $\{g^N, N \in \mathbb{N}\}$ *are tight on* $L_w^{p^*}([0,T]\times\mathbb{R}^d\times[0,1])$.

Proof Choosing $p^* > 1$ such that $\frac{d}{2}(1-\frac{1}{p*}) < 1$, then we have

$$\begin{aligned}
&\sup_N \mathbb{E}\int_0^T\int_0^1\int_{\mathbb{R}^d} g^N(t,x,\xi)|^{p^*}\mathrm{d}x\mathrm{d}\xi\mathrm{d}t \\
\leqslant &\sup_N \mathbb{E}\int_0^T\int_0^1 I(g^N(t,\cdot,\xi))^{\frac{d}{2}(1-\frac{1}{p*})}\mathrm{d}\xi\mathrm{d}t < \infty.
\end{aligned}$$

where we used Lemma 3.3 and Lemma 4.10. The result thus follows.

Similar to Section 3.4, we extract a subsequence of $\{(Q^N, g^N)\}$ and identify the limiting point as a solution to (1.16). Define the space $\mathcal{X}$ by

$$\mathcal{X} := \mathcal{P}\Big(C([0,T],\mathbb{R}^d)\times[0,1]\Big) \times L_w^{p^*}\left([0,T]\times\mathbb{R}^d\times[0,1]\right)$$

By the generalized Skorokhod Representation Theorem/ Jakubowski Theorem (c.f. [BFH18, Theorem 2.7.1]), we find

Proposition 4.14. *There exists a subsequence of* $\{(Q^N, g^N), N \in \mathbb{N}\}$, *without relabeling for simplicity, and a probability space* $(\Omega^*, \mathcal{F}^*, \mathbb{P}^*)$ *with* $\mathcal{X}$*-valued random variables* $\{(Q^{*N}, g^{*N}),\ N \in \mathbb{N}\}$ *and* (Q, f) *such that*

1. *For each* N, *the law of* (Q^N, g^N) *coincides with the law of* (Q^{*N}, g^{*N}).
2. *The sequence of random variables* (Q^{*N}, g^{*N}) *converges to* (Q, f) *in* $\mathcal{X}$ $\mathbb{P}^*$*-almost surely.*

For simplicity, we omit the superscript $*$ in the following text. And we derive the convergence of $\{\nu_N\}$ from the convergence of $\{Q^N\}$ as deriving Corollary 3.23.

Corollary 4.15. *The sequence of the empirical measure processes* ν_N *converges to* ν *in* $C([0,T],\mathcal{P}(\mathbb{R}^d\times[0,1]))$ *almost surely, where* $\nu := (Q\circ\pi_t^{-1})_{t\in[0,T]}$.

By a simialr argument with Lemma 3.24, using the weakly merging result Lemma 4.8, we identifies f with ν.

Corollary 4.16. *The subsequence* $\{g_t^N\}$ *converges to* ν_t *in* $\mathcal{P}(\mathbb{R}^d\times[0,1])$ *for a.e.* (t,ω). *Furthermore,* $f(t,\omega)$ *is a density of* $\nu_t(\omega)$ *for a.e.* (t,ω).

Again, we shall not distinguish f and ν. The next result shows that f has finite partial Fisher information.

Lemma 4.17. *For each limiting point f, it holds that*

$$\mathbb{E}\int_0^T\int_0^1 I(f(t,\cdot,\xi))\mathrm{d}\xi\mathrm{d}t<\infty.$$

Proof We use an alternative presentation of the partial Fisher information,

$$\int_0^1 I(f(t,\cdot,\xi))\mathrm{d}\xi=\sup_{\varphi\in C_b^1([0,1]\times\mathbb{R}^d;\mathbb{R}^d)}\left\langle f(t,\cdot,\cdot),-\frac{|\varphi|^2}{4}-\operatorname{div}_x\varphi\right\rangle.$$

Therefore, applying Corollary 4.16 gives that

$$\begin{aligned}\int_0^1 I(f(t,\cdot,\xi))\mathrm{d}\xi=&\sup_{\varphi\in C_b^1([0,1]\times\mathbb{R}^d;\mathbb{R}^d)}\lim_{N\to\infty}\left\langle g_t^N(\cdot,\cdot),-\frac{|\varphi|^2}{4}-\operatorname{div}_x\varphi\right\rangle\\ \leqslant&\liminf_{N\to\infty}\sup_{\varphi\in C_b^1([0,1]\times\mathbb{R}^d;\mathbb{R}^d)}\left\langle g_t^N(\cdot,\cdot),-\frac{|\varphi|^2}{4}-\operatorname{div}_x\varphi\right\rangle\\ =&\liminf_{N\to\infty}\int_0^1 I(g_t^N(\cdot,\xi))\mathrm{d}\xi,\end{aligned}$$

for a.e. (t,ω). The result then follows by Lemma 4.10.

As a byprofuct, we have the following estimate.

Corollary 4.18. *For any $p,q,\beta\geqslant 1$ satisfying*

$$\frac{d}{p}+\frac{2}{q}\geqslant d,\quad \frac{d}{p}+\frac{2}{\beta}\geqslant d,\quad p<\infty,$$

it holds that

$$\mathbb{E}\int_0^T\left(\int_0^1\left(\int_{\mathbb{R}^d}|f(t,x,\xi)|^p\mathrm{d}x\right)^{\frac{\beta}{p}}\mathrm{d}\xi\right)^{\frac{q}{\beta}}\mathrm{d}t<\infty$$

Proof Applying the Lemma 3.3 yields

$$\begin{aligned}&\mathbb{E}\int_0^T\left(\int_0^1\left(\int_{\mathbb{R}^d}|f(t,x,\xi)|^p\mathrm{d}x\right)^{\frac{\beta}{p}}\mathrm{d}\xi\right)^{\frac{q}{\beta}}\mathrm{d}t\\ \leqslant&\mathbb{E}\int_0^T\left(\int_0^1 I(f(t,\cdot,\xi))^{\frac{d}{2}(1-\frac{1}{p})\beta}\mathrm{d}\xi\right)^{\frac{q}{\beta}}\mathrm{d}t\\ \lesssim&1+\mathbb{E}\int_0^T\int_0^1 I\left(f(t,\cdot,\xi)\right)^{\frac{d}{2}(1-\frac{1}{p})\max\{\beta,q\}}\mathrm{d}\xi\mathrm{d}t.\end{aligned}$$

By the constraint on (p,q,β), we find $\frac{d}{2}(1-\frac{1}{p})\max\{\beta,q\}\leqslant 1$, the result then follows by Lemma 4.17.

Lastly, we find f belongs to $L^\infty([0,1],L^1(\mathbb{R}^d))$.

Lemma 4.19. *For each limiting point f, it holds that*

$$\|f(t,\cdot,\cdot)\|_{L^\infty_\xi L^1_x} = 1$$

for a.e. (t,ω).

Proof Since f is nonnegative and $C^\infty([0,1])$ is dense in $L^1([0,1])$, we conclude the result by

$$\begin{aligned}
\|f(t,\cdot,\cdot)\|_{L^\infty_\xi L^1_x} &= \sup_{\varphi\in C^\infty([0,1]),\|\varphi\|_{L^1}\leqslant 1} \left\langle \varphi, \int_{\mathbb{R}^d} f(t,x,\cdot)\mathrm{d}x \right\rangle \\
&= \sup_{\varphi\in C^\infty([0,1]),\|\varphi\|_{L^1}\leqslant 1} \int_0^1 \int_{\mathbb{R}^d} \varphi(\xi) f(t,x,\xi)\mathrm{d}x\mathrm{d}\xi \\
&= \sup_{\varphi\in C^\infty([0,1]),\|\varphi\|_{L^1}\leqslant 1} \lim_{N\to\infty} \int_0^1 \int_{\mathbb{R}^d} \varphi(\xi)\nu_N(t,\mathrm{d}x,\mathrm{d}\xi) \\
&= \sup_{\varphi\in C^\infty([0,1]),\|\varphi\|_{L^1}\leqslant 1} \int_0^1 \varphi(\xi)\mathrm{d}\xi = 1.
\end{aligned}$$

for a.e. (t,ω).

4.4.2 Characterizing the limits

Next we identify those limiting points of $\{\nu_N\}$ as solutions to (1.25).

Proposition 4.20. *Suppose that there exists a graphon G with finite $L^\infty L^r$-norm such that $\|G_N - G\|_\square \to 0$. Then every limiting point of ν_N is a solution to* (1.25) *in the sense of Definition 1.19.*

Proof We already know that (v_N, g^N) converges in $\Big(C([0,T],\mathcal{P}(\mathbb{R}^d\times[0,1]))\Big)\times L^{p*}_w([0,T]\times\mathbb{R}^d\times[0,1])$ to the same limit f.

For each $\varphi\in C_0^\infty(\mathbb{R}^d\times[0,1])$, applying Ito's formula yields that

$$\begin{aligned}
\mathrm{d}\varphi(X_i,\xi) =& \nabla_x\varphi(X_i,\xi)\mathrm{d}X_i + \Delta_x\varphi(X_i,\xi)\mathrm{d}t \\
=& \Delta_x\varphi(X_i,\xi)\mathrm{d}t + \nabla_x\varphi(X_i,\xi)\cdot\left[\frac{1}{N}\sum_{j\neq i} w_{ij}K(X_i-X_j)\right]\mathrm{d}t + \sqrt{2}\nabla_x\varphi(X_i,\xi)\mathrm{d}B_i.
\end{aligned}$$

Therefore, we have

$$\langle\varphi,\nu_N(t)\rangle = \langle\varphi,\nu_N(0)\rangle + \int_0^t \langle\Delta_x\varphi,\nu_N(s)\rangle\,\mathrm{d}s + \mathcal{K}_N + M_N, \tag{4.19}$$

where

$$\begin{aligned}
M_N :=& \sqrt{2}\sum_{i=1}^N \int_0^t \int_{\frac{i-1}{N}}^{\frac{i}{N}} \nabla_x\varphi(X_i,\xi)\mathrm{d}\xi\mathrm{d}B_i, \\
\mathcal{K}_N :=& \frac{1}{N}\sum_{i\neq j} w_{ij}^N \int_0^t \int_{\frac{i-1}{N}}^{\frac{i}{N}} \nabla_x\varphi(X_i,\xi)\mathrm{d}\xi K(X_i-X_j)\mathrm{d}s.
\end{aligned}$$

By the following computation,

$$
\begin{aligned}
\Phi(G_N, K, \nu_N) :=& \int_{\mathbb{R}^d \times \mathbb{R}^d \setminus \{x=y\}} \int_0^1 \int_0^1 G_N(\xi,\zeta) \nabla_x \varphi(x,\xi) K(x-y) \nu_N(\mathrm{d}y,\zeta)\nu_N(\mathrm{d}x,\xi)\mathrm{d}\zeta\mathrm{d}\xi \\
=& \sum_{i\neq j} \int_{\frac{i-1}{N}}^{\frac{j}{N}} \int_{\frac{j-1}{N}}^{\frac{i}{N}} G_N(\xi,\zeta)\nabla_x\varphi(X_i,\xi)K(X_i-X_j)\mathrm{d}\zeta\mathrm{d}\xi \\
=& \sum_{i\neq j} \int_{\frac{i-1}{N}}^{\frac{j}{N}} \int_{\frac{j-1}{N}}^{\frac{i}{N}} \left(\sum_{k,l} w_{kl}^N 1_{[\frac{k-1}{N},\frac{k}{N})}(\xi) 1_{[\frac{l-1}{N},\frac{l}{N})}(\zeta) \right) \nabla_x\varphi(X_i,\xi)K(X_i-X_j)\mathrm{d}\zeta\mathrm{d}\xi \\
=& \sum_{i\neq j} w_{ij}^N \int_{\frac{i-1}{N}}^{\frac{j}{N}} \int_{\frac{j-1}{N}}^{\frac{i}{N}} \nabla_x\varphi(X_i,\xi)K(X_i-X_j)\mathrm{d}\zeta\mathrm{d}\xi \\
=& \frac{1}{N}\sum_{i\neq j} w_{ij}^N \int_{\frac{i-1}{N}}^{\frac{i}{N}} \nabla_x\varphi(X_i,\xi)\mathrm{d}\xi K(X_i-X_j),
\end{aligned}
$$

we find that

$$
\mathcal{K}_N = \int_0^t \Phi(G_N, K, \nu_N(s))\mathrm{d}s. \tag{4.20}
$$

Now we split $\Phi(G_N, K, \nu_N)$ into different parts. We firstly approximate K by smooth kernel $K_\varepsilon \in C_0^\infty([0,T]\times\mathbb{R}^d)$ (again, we igonore differences about bounded kernels), $\varepsilon > 0$, such that

$$
\|K_\varepsilon - K\|_{L^{q_2}([0,T],L^{p_2}(\mathbb{R}^d))} \to 0.
$$

Notice that the map Φ is linear with respect to G, K, and $\nu\otimes\nu$, then we have

$$
\begin{aligned}
&\Phi(G_N, K, \nu_N) - \Phi(G, K, f) \\
=& \Big[\Phi(G_{N'}, K_\varepsilon, \nu_N) - \Phi(G_{N'}, K_\varepsilon, f)\Big] + \Phi(G_N - G_{N'}, K_\varepsilon, \nu_N) \\
&+ \Big[\Phi(G_N, K-K_\varepsilon, \nu_N) - \Phi(G, K-K_\varepsilon, f)\Big] + \Phi(G_{N'} - G, K_\varepsilon, f) \\
:=& J_1 + J_2 + J_3 + J_4,
\end{aligned}
$$

where $N, N' \in \mathbb{N}$.

The first term vanishes as N goes to infinity for arbitary fixed ε and N' because of $\nu_N \rightharpoonup f$. The second term is the technical part involving the convergence of graphons. We need to rewrite it into the following form,

$$
\begin{aligned}
J_2 =& \sum_{i\neq j} \int_{\frac{i-1}{N}}^{\frac{j}{N}} \int_{\frac{j-1}{N}}^{\frac{i}{N}} [G_N(\xi,\zeta) - G_{N'}(\xi,\zeta)]\nabla_x\varphi(X_i,\xi)K_\varepsilon(X_i-X_j)\mathrm{d}\zeta\mathrm{d}\xi \\
=& \sum_{i\neq j} \int_{\frac{i-1}{N}}^{\frac{j}{N}} \int_{\frac{j-1}{N}}^{\frac{i}{N}} [G_N(\xi,\zeta) - G_{N'}(\xi,\zeta)]\nabla_x\varphi(X_i,\xi)K_\varepsilon(X_i-X_j)\Big[1-\chi_R(X_i,X_j)\Big]\mathrm{d}\zeta\mathrm{d}\xi \\
&+ \sum_{i\neq j} \int_{\frac{i-1}{N}}^{\frac{j}{N}} \int_{\frac{j-1}{N}}^{\frac{i}{N}} [G_N(\xi,\zeta) - G_{N'}(\xi,\zeta)]\nabla_x\varphi(X_i,\xi)K_\varepsilon(X_i-X_j)\chi_R(X_i,X_j)\mathrm{d}\zeta\mathrm{d}\xi,
\end{aligned}
$$

where χ_R is a non-negative smooth function bounded by 1, and equals to 1 on $\{(x,y), |x| \leqslant \frac{R}{2}, |y| \leqslant \frac{R}{2}\}$, supporting on $\{(x,y), |x| \leqslant R, |y| \leqslant R\}$. Furthermore, we can uniformly approximate the function $K_\varepsilon(x-y)\chi_R(x,y) : [0,T]\times B_R \times B_R \to \mathbb{R}^d$ by linear conbinations of smooth functions of the form $\phi(t,x)\tilde{\phi}(t,y)$. That is, for all (t,x,y) and any $R>0$, there exists a finite of set smooth functions (for instance polynomials) $\{\phi_k^m, \tilde{\phi}_k^m, 1 \leqslant k \leqslant m\}$ on $[0,T]\times B_R$ such that

$$\left| K_\varepsilon(x-y)\chi_R(x,y) - \sum_{k=1}^m \phi_k^m(t,x)\tilde{\phi}_k^m(t,y) \right| < \frac{1}{R}. \tag{4.21}$$

Here the sequence of product functions depends on both ε and R. Then we have

$$\begin{aligned}
J_2 =& \sum_{i\neq j} \int_{\frac{i-1}{N}}^{\frac{j}{N}} \int_{\frac{j-1}{N}}^{\frac{i}{N}} [G_N(\xi,\zeta) - G_{N'}(\xi,\zeta)] \nabla_x \varphi(X_i,\xi) K_\varepsilon(X_i - X_j) \Big[1 - \chi_R(X_i,X_j)\Big] \mathrm{d}\zeta \mathrm{d}\xi \\
&+ \Bigg[\sum_{i\neq j} \int_{\frac{i-1}{N}}^{\frac{j}{N}} \int_{\frac{j-1}{N}}^{\frac{i}{N}} [G_N(\xi,\zeta) - G_{N'}(\xi,\zeta)] \nabla_x \varphi(X_i,\xi) K_\varepsilon(X_i - X_j) \chi_R(X_i,X_j) \mathrm{d}\zeta \mathrm{d}\xi \\
&- \sum_{i\neq j} \int_{\frac{i-1}{N}}^{\frac{j}{N}} \int_{\frac{j-1}{N}}^{\frac{i}{N}} [G_N(\xi,\zeta) - G_{N'}(\xi,\zeta)] \nabla_x \varphi(X_i,\xi) \sum_{k=1}^m \phi_k^m(t,X_i)\tilde{\phi}_k^m(t,X_j) \mathrm{d}\zeta \mathrm{d}\xi \Bigg] \\
&+ \sum_{k=1}^m \sum_{i\neq j} \int_{\frac{i-1}{N}}^{\frac{j}{N}} \int_{\frac{j-1}{N}}^{\frac{i}{N}} [G_N(\xi,\zeta) - G_{N'}(\xi,\zeta)] \nabla_x \varphi(X_i,\xi) \phi_k^m(t,X_i)\tilde{\phi}_k^m(t,X_j) \mathrm{d}\zeta \mathrm{d}\xi \\
:=& J_{21} + J_{22} + J_{23}.
\end{aligned}$$

The large area part J_{21} is controlled as following,

$$\begin{aligned}
\mathbb{E}|J_{21}| \leqslant& \|\varphi\|_{C^1} \|K_\varepsilon\|_{L^\infty} \mathbb{E} \sum_{i\neq j} \int_{\frac{i-1}{N}}^{\frac{j}{N}} \int_{\frac{j-1}{N}}^{\frac{i}{N}} |G_N(\xi,\zeta) - G_{N'}(\xi,\zeta)| \Big[1 - \chi_R(X_i,X_j)\Big] \mathrm{d}\zeta \mathrm{d}\xi \\
\lesssim& \frac{1}{R^{\frac{\gamma(r-1)}{r}}} \mathbb{E} \sum_{i\neq j} \int_{\frac{i-1}{N}}^{\frac{j}{N}} \int_{\frac{j-1}{N}}^{\frac{i}{N}} |G_N(\xi,\zeta) - G_{N'}(\xi,\zeta)| \Big(|X_i|^{\frac{\gamma(r-1)}{r}} + |X_j|^{\frac{\gamma(r-1)}{r}} \Big) \mathrm{d}\zeta \mathrm{d}\xi \\
\lesssim& \frac{1}{R^{\frac{\gamma(r-1)}{r}}} \mathbb{E} \sum_{i\neq j} \int_{\frac{i-1}{N}}^{\frac{j}{N}} \int_{\frac{j-1}{N}}^{\frac{i}{N}} |G_N(\xi,\zeta) - G_{N'}(\xi,\zeta)| \, |X_i|^{\frac{\gamma(r-1)}{r}} \mathrm{d}\zeta \mathrm{d}\xi \\
\lesssim& \frac{1}{R^{\frac{\gamma(r-1)}{r}}} \left(\mathbb{E} \sum_{i\neq j} \int_{\frac{i-1}{N}}^{\frac{j}{N}} \int_{\frac{j-1}{N}}^{\frac{i}{N}} |G_N(\xi,\zeta) - G_{N'}(\xi,\zeta)|^r \mathrm{d}\zeta \mathrm{d}\xi \right)^{\frac{1}{r}} \cdot \left(\mathbb{E} \frac{1}{N^2} \sum_{i\neq j} |X_i|^\gamma \right)^{\frac{r-1}{r}} \\
\lesssim& \frac{\Big(\sup_N \|G_N\|_{L^r}\Big)(1 + \sup_N \frac{1}{N}\sum_i \mathbb{E}|X_i|^\gamma)}{R^{\frac{\gamma(r-1)}{r}}} \\
\lesssim& \frac{\Big(\sup_N \|w^N\|_{l^\infty l^r}\Big)(1 + \sup_N \frac{1}{N}\sum_i \mathbb{E}|X_i|^\gamma)}{R^{\frac{\gamma(r-1)}{r}}}.
\end{aligned}$$

By (4.11) and ($\mathbf{W}'_\mathbf{r}$), this implies that for fixed ε, $\mathbb{E}|J_{21}|$ vanishes as R goes to infinity, which is unifom for N, N'. Then up to subsequences, the convegence of $|J_{21}|$ holds almost surely.

As to J_{22}, using (4.21) gives

$$
\begin{aligned}
|J_{22}| \lesssim & \frac{1}{R} \sum_{i \neq j} \int_{\frac{i-1}{N}}^{\frac{j}{N}} \int_{\frac{j-1}{N}}^{\frac{i}{N}} |G_N(\xi,\zeta) - G_{N'}(\xi,\zeta)| \, \mathrm{d}\zeta \mathrm{d}\xi \\
\lesssim & \frac{1}{R} \sup_N \|G_N\|_{l^1} \lesssim \frac{\sup_N \|w^N\|_{l^\infty l^r}}{R}.
\end{aligned}
$$

Lastly, the term J_{23} is handled by the convergence of graphons. More precisely, observe that

$$
\begin{aligned}
J_{23} = & \sum_{k=1}^{m} \int_0^1 \int_0^1 [G_N(\xi,\zeta) - G_{N'}(\xi,\zeta)] \Big[\sum_{i=1}^{N} \nabla_x \varphi(X_i,\xi) \phi_k^m(t,X_i) 1_{[\frac{i-1}{N},\frac{i}{N})}(\xi) \Big] \cdot \\
& \Big[\sum_{j=1}^{N} \tilde{\phi}_k^m(t,X_j) 1_{[\frac{j-1}{N},\frac{j}{N})}(\zeta) \Big] \mathrm{d}\zeta \mathrm{d}\xi.
\end{aligned}
$$

Since the convergence with respect to the cut norm is equivalent to the convergence with respect to the operator norm $L^\infty \to L^1$, we have

$$
\begin{aligned}
|J_{23}| \leqslant & \sum_{k=1}^{m} \|G_N - G_{N'}\|_{\square} \|\nabla_x \varphi \phi_k^m\|_{L^\infty} \|\tilde{\phi}_k^m\|_{L^\infty} \\
\lesssim & \left(\sum_{k=1}^{m} \|\phi_k^m\|_{L^\infty} \|\tilde{\phi}_k^m\|_{L^\infty} \right) (\|G_N - G\|_{\square} + \|G_{N'} - G\|_{\square}).
\end{aligned}
$$

We then obtain that $|J_{23}|$ converges to 0 as N and N' go to infinity, for fixed ε, R, since the functions $\{\phi_k^m, \tilde{\phi}_k^m\}$ depend on ε and R only. In conclusion, for fixed $\varepsilon > 0$, we have

$$
\begin{aligned}
\lim_{N'\to\infty} \lim_{N\to\infty} |J_2| \leqslant & \lim_{N'\to\infty} \lim_{N\to\infty} \Big(|J_{21}(R)| + |J_{22}(R) + |J_{23}(R)| \Big) \\
\lesssim & \frac{1}{R^{\frac{\gamma(r-1)}{r}}} + \frac{1}{R} + \lim_{N'\to\infty} \lim_{N\to\infty} |J_{23}(R)| \\
\lesssim & \frac{1}{R^{\frac{\gamma(r-1)}{r}}} + \frac{1}{R}.
\end{aligned}
\tag{4.22}
$$

By choosing R sufficiently large, we arrive at $\lim_{N'\to\infty} \lim_{N\to\infty} |J_2| = 0$ in probability.

The term J_3 is a typical error term that we shall control with uniform estimates. Notice that J_3 consists of two terms, the term depending on N is written as follows

$$
\Phi(G_N, K - K_\varepsilon, \nu_N) = \frac{1}{N} \sum_{i \neq j} w_{ij}^N \int_{\frac{i-1}{N}}^{\frac{i}{N}} \nabla_x \varphi(X_i,\xi) \mathrm{d}\xi \Big(K(X_i - X_j) - K_\varepsilon(X_i - X_j) \Big).
$$

Then we have

$$
\begin{aligned}
& \mathbb{E} \int_0^T |\Phi(G_N, K - K_\varepsilon, \nu_N(t))| \, \mathrm{d}t \\
\leqslant & \mathbb{E} \int_0^T \frac{1}{N^2} \sum_{i \neq j} |w_{ij}^N| \, |K(X_i - X_j) - K_\varepsilon(X_i - X_j)| \, \mathrm{d}t.
\end{aligned}
$$

The right hand side likewise has appeared multiple times in Section 4.2, handled by using Corollary 3.7. Similarly, we obtain

$$\sup_N \mathbb{E}\int_0^T |\Phi(G_N, K-K_\varepsilon, \nu_N(t))|\,\mathrm{d}t \xrightarrow{\varepsilon\to 0} 0.$$

On the other hand, the other term in J_3 is handled by the regularity of f. That is,

$$\begin{aligned}
&\mathbb{E}\int_0^T |\Phi(G, K-K_\varepsilon, f)|\,\mathrm{d}t\\
=&\mathbb{E}\int_0^T \left|\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_0^1\int_0^1 G(\xi,\zeta)\nabla_x\varphi(x,\xi)\Big[K(x-y)-K_\varepsilon(x-y)\Big]f(x,\xi)f(y,\zeta)\mathrm{d}\zeta\mathrm{d}\xi\mathrm{d}x\mathrm{d}y\right|\mathrm{d}t\\
\leqslant&\|\varphi\|_{C^1}\mathbb{E}\int_0^T\int_0^1\int_0^1 |G(\xi,\zeta)|\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\left|\Big[K(x-y)-K_\varepsilon(x-y)\Big]f(x,\xi)f(y,\zeta)\right|\mathrm{d}x\mathrm{d}y\mathrm{d}\zeta\mathrm{d}\xi\mathrm{d}t\\
\lesssim&\mathbb{E}\int_0^T\int_0^1\int_0^1 |G(\xi,\zeta)|\|K-K_\varepsilon\|_{L^p}\|f(\cdot,\xi)\|_{L^1}\|f(\cdot,\zeta)\|_{L^{\frac{p}{p-1}}}\mathrm{d}\zeta\mathrm{d}\xi\mathrm{d}t.
\end{aligned}$$

Furthermore, using Hölder's inequality and Young's inequality gives that

$$\begin{aligned}
&\mathbb{E}\int_0^T |\Phi(G, K-K_\varepsilon, f)|\,\mathrm{d}t\\
\lesssim&\|G\|_{L^\infty L^r}\|K-K_\varepsilon\|_{L^p_q}\left(1+\mathbb{E}\int_0^T \|f\|_{L^1(\mathbb{R}^d\times[0,1])}\int_0^1 \|f(\cdot,\zeta)\|_{L^{\frac{p}{p-1}}}^{\frac{q}{q-1}\cdot\frac{r}{r-1}}\mathrm{d}\zeta\mathrm{d}t\right).\\
\lesssim&\|G\|_{L^\infty L^r}\|K-K_\varepsilon\|_{L^p_q}\left(1+\mathbb{E}\int_0^T\int_0^1 I(f(t,\cdot,\zeta))^{\frac{d}{2p}\cdot\frac{q}{q-1}\cdot\frac{r}{r-1}}\mathrm{d}\zeta\mathrm{d}t\right),
\end{aligned}$$

where the last inequality follows by the fact that $f(t,\cdot,\cdot)$ is a probability density and applying Lemma 3.2.

Notice that

$$\begin{aligned}
(\mathbf{K_r'}) \Longrightarrow&\frac{d}{p} < 2(1-\frac{1}{q}) - 1 - \frac{1}{r} \Longrightarrow \frac{d}{p} < 2(1-\frac{1}{q}) - \frac{2}{r}\\
\Longrightarrow&\frac{d}{p} < 2(1-\frac{1}{q}) - \frac{2}{r}(1-\frac{1}{q})\\
\Longrightarrow&\frac{d}{p} < 2(1-\frac{1}{q})(1-\frac{1}{r})\\
\Longrightarrow&\frac{d}{2p}\cdot\frac{q}{q-1}\cdot\frac{r}{r-1} < 1.
\end{aligned}$$

And since the $L^\infty L^r$-norm of $\{G_N\}$ is bounded, we have

$$\begin{aligned}
\|G\|_{L^\infty L^r} =&\operatorname{ess\,sup}_{\xi\in[0,1]}\sup_{\|\varphi\|_{L^{\frac{r}{r-1}}}\leqslant 1} |\langle G(\xi,\cdot),\varphi\rangle|\\
=&\sup_{\|\varphi\|_{L^{\frac{r}{r-1}}}\leqslant 1}\operatorname{ess\,sup}_{\xi\in[0,1]} |\langle G(\xi,\cdot),\varphi\rangle|\\
=&\sup_{\|\varphi\|_{L^{\frac{r}{r-1}}}\leqslant 1}\sup_{\|\phi\|_{L^1}\leqslant 1}\int_0^1\int_0^1 |G(\xi,\zeta)\varphi(\zeta)\phi(\xi)|\,\mathrm{d}\zeta\mathrm{d}\xi
\end{aligned}$$

$$\begin{aligned}
&\leqslant \liminf_{N\to\infty} \sup_{\|\varphi\|_{L^{\frac{r}{r-1}}}\leqslant 1} \sup_{\|\phi\|_{L^1}\leqslant 1} \int_0^1\int_0^1 |G_N(\xi,\zeta)\varphi(\zeta)\phi(\xi)|\,\mathrm{d}\zeta\mathrm{d}\xi \\
&= \liminf_{N\to\infty} \|G_N\|_{L^\infty L^r} < \infty,
\end{aligned}$$

where we used the fact that $L^\infty([0,1])$ is dense in $L^1([0,1])$ or $L^{\frac{r}{r-1}}([0,1])$ and the lower semi-continuity of supremum. Furthermore, by Lemma 4.17, we conclude that

$$\mathbb{E}\int_0^T |\Phi(G, K-K_\varepsilon, v)|\,\mathrm{d}t \lesssim \|K-K_\varepsilon\|_{L_q^p} \xrightarrow{\varepsilon\to 0} 0.$$

We thus arrive at

$$\mathbb{E}\int_0^T |J_3|\mathrm{d}t \xrightarrow{\varepsilon\to 0} 0, \quad \text{uniformly in } N. \tag{4.23}$$

For the last term J_4, since the convergence of $G_{N'}$ w.r.t the cut-norm is stronger than the convergence w.r.t the weak topology of $L^1([0,1]\times[0,1])$, we find that

$$\begin{aligned}
J_4 = &\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_0^1\int_0^1 \Big(G_{N'}(\xi,\zeta)-G(\xi,\zeta)\Big)\nabla_x\varphi(x,\xi)K_\varepsilon(x-y)f(y,\zeta)f(x,\xi)\mathrm{d}\zeta\mathrm{d}\xi\mathrm{d}x\mathrm{d}y \\
&\xrightarrow{N'\to\infty}_\varepsilon 0,
\end{aligned}$$

here we used to the fact that

$$\begin{aligned}
&\left\|\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \nabla_x\varphi(x,\xi)K_\varepsilon(x-y)f(y,\zeta)f(x,\xi)\mathrm{d}x\mathrm{d}y\right\|_{L^\infty([0,1]\times[0,1])} \\
\leqslant &\|\nabla_x\varphi\|_{L^\infty}\|K_\varepsilon\|_{L^\infty}\|f\|^2_{L^\infty([0,1];L^1(\mathbb{R}^d))} < \infty.
\end{aligned}$$

In conclusion, we have

$$\begin{aligned}
&\lim_{N\to\infty}\left|\int_0^T \Phi(G_N,K,\nu_N)-\Phi(G,K,f)\mathrm{d}t\right| \\
\leqslant &\lim_{N\to\infty}\left|\int_0^T J_1(\varepsilon,N',N)\mathrm{d}t\right| + \lim_{N'\to\infty}\lim_{N\to\infty}\left|\int_0^T J_2(\varepsilon,N',N)\mathrm{d}t\right| \\
&+ \lim_{\varepsilon\to 0}\sup_N\left|\int_0^T J_3(\varepsilon,N)\mathrm{d}t\right| + \lim_{N'\to\infty}\left|\int_0^T J_4(\varepsilon,N')\mathrm{d}t\right| \\
=&0,
\end{aligned} \tag{4.24}$$

in probability. Therefore, we obtain

$$\mathcal{K}_N \to \int_0^t\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_0^1\int_0^1 G(\xi,\zeta)\nabla_x\varphi(x,\xi)K_\varepsilon(x-y)f(y,\zeta)f(x,\xi)\mathrm{d}\zeta\mathrm{d}\xi\mathrm{d}x\mathrm{d}y\mathrm{d}s,$$

almost surely up to subsequences.

As to the stochastic term M_N, by calculating the quadratic variation, we find

$$\mathbb{E}\sup_{t\in[0,T]}|M_N(t)|^2 \lesssim \sum_{i=1}^N \mathbb{E}\int_0^T\left|\int_{\frac{i-1}{N}}^{\frac{i}{N}} \nabla_x\varphi(X_i,\xi)\mathrm{d}\xi\right|^2 \mathrm{d}t$$

$$
\begin{aligned}
&\lesssim \mathbb{E}\int_0^T \sum_{i=1}^N \frac{1}{N}\int_{\frac{i-1}{N}}^{\frac{i}{N}} |\nabla_x \varphi(X_i,\xi)|^2 \,\mathrm{d}\xi\mathrm{d}t \\
&\lesssim \frac{1}{N}\mathbb{E}\int_0^T \left\langle |\nabla_x\varphi|^2, \nu_N\right\rangle \mathrm{d}t \\
&\lesssim \frac{\|\nabla\varphi\|_{L^\infty}^2}{N} \to 0. \qquad (4.25)
\end{aligned}
$$

Thus the noise M_N vanishes almost surely as N goes to infinity, up to subsequences.

By the weak convergence of ν_N to f, combining (4.19), (4.24), and (4.25) yields the result.

4.4.3 Well-posedness of the mean-field pde

In this section, we show the uniqueness of solutions to the generalized mean-field PDE (1.25). The proof is based on the following mild formulation.

Lemma 4.21. *Each solution f to* (1.25) *in the sense of Definition 1.19 satisfies the mild formulation. That is, for any $0 \leqslant t \leqslant T$,*

$$
\begin{aligned}
&\Gamma_t * f(0,x,\xi) - f(t,x,\xi) \\
&= \int_0^t \nabla.\Gamma_{t-s} * \left(\int_{\mathbb{R}^d}\int_0^1 K(\cdot - y)G(\xi,\zeta)f(s,\cdot,\xi)f(s,\zeta,y)\mathrm{d}\zeta\mathrm{d}y\right)(x)\mathrm{d}s
\end{aligned}
$$

holds in the distributional sense.

Proof We start with the following statement: given any function $\varphi \in C_b([0,1])$ and ϕ of class $C^1([0,t],\mathcal{S}(\mathbb{R}^d))$ for $t \in [0,T]$, it holds that,

$$
\begin{aligned}
&\int_{\mathbb{R}^d}\int_0^1 \phi(t,x)\varphi(\xi)f(t,x,\xi)\mathrm{d}\xi\mathrm{d}x - \int_{\mathbb{R}^d}\int_0^1 \phi(0,x)\varphi(\xi)f(0,x,\xi)\mathrm{d}\xi\mathrm{d}x \\
&= \int_0^t\int_{\mathbb{R}^d}\int_0^1 \partial_s\phi(s,x)\varphi(\xi)f(s,x,\xi)\mathrm{d}\xi\mathrm{d}x\mathrm{d}s + \int_0^t\int_{\mathbb{R}^d}\int_0^1 \Delta_x\phi(s,x)\varphi(\xi)f(s,x,\xi)\mathrm{d}\xi\mathrm{d}x\mathrm{d}s \\
&\quad + \int_0^t\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_0^1\int_0^1 \varphi(\xi)\nabla_x\phi(s,x)\cdot K(x-y)G(\xi,\zeta)f(s,x,\xi)f(s,\zeta,y)\mathrm{d}\xi\mathrm{d}\zeta\mathrm{d}x\mathrm{d}y\mathrm{d}s.
\end{aligned}
$$

Like in (2.47), it is straightforward to check the statement for finite linear combinations of functions ϕ of the form $\phi(s,x) = \phi_1(s)\phi_2(x)$, where $\phi_1 \in C^\infty([0,t])$ and $\phi_2 \in C_c^\infty(\mathbb{R}^d)$. Then one can approximate functions in $C^1([0,t],C_c^\infty(\mathbb{R}^d))$ with such combinations uniformly along with derivatives up to the second order. Since $C_c^\infty(\mathbb{R}^d)$ is dense in $\mathcal{S}(\mathbb{R}^d)$ with respect to the C^2-norm, the statement extends to functions in $\phi \in C^1([0,t],\mathcal{S}(\mathbb{R}^d))$.

For every $\phi_0 \in \mathcal{S}(\mathbb{R}^d)$ and $0 \leqslant s \leqslant t$, fix $\phi(s) := \Gamma_{t-s} * \phi_0$, where $\Gamma.$ stands for the heat kernel on $\mathbb{R}^d$. This implies that ϕ solves the backward heat equation,

$$
\partial_s\phi(s) + \Delta\phi(s) = 0, \quad \phi(t) = \phi_0.
$$

Therefore, we obtain

$$
\int_{\mathbb{R}^d}\int_0^1 \phi_0(x)\varphi(\xi)f(t,x,\xi)\mathrm{d}\xi\mathrm{d}x
$$

$$
\begin{aligned}
&= \int_{\mathbb{R}^d}\int_0^1 \Gamma_t * \phi_0(x)\varphi(\xi)f(0,x,\xi)\mathrm{d}\xi\mathrm{d}x \\
&\quad + \int_0^t\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_0^1\int_0^1 \varphi(\xi)\nabla_x\Gamma_{t-s} * \phi_0(x)\cdot K(x-y)G(\xi,\zeta)f(s,x,\xi)f(s,\zeta,y)\mathrm{d}\xi\mathrm{d}\zeta\mathrm{d}x\mathrm{d}y\mathrm{d}s \\
&= \int_{\mathbb{R}^d}\int_0^1 \phi_0(x)\varphi(\xi)\Gamma_t * f(0,x,\xi)\mathrm{d}\xi\mathrm{d}x - \\
&\quad \int_0^t\int_{\mathbb{R}^d}\int_0^1 \varphi(\xi)\phi_0(x)\nabla.\Gamma_{t-s} * \left(\int_{\mathbb{R}^d}\int_0^1 K(\cdot-y)G(\xi,\zeta)f(s,\cdot,\xi)f(s,\zeta,y)\mathrm{d}\zeta\mathrm{d}y\right)(x)\mathrm{d}\xi\mathrm{d}x\mathrm{d}s,
\end{aligned}
$$

for any $\phi_0 \in \mathcal{S}(\mathbb{R}^d)$ and $\varphi \in C_b([0,1])$. The result follows by extending test functions to $\mathcal{S}(\mathbb{R}^d \times [0,1])$.

Theorem 4.22. *There exists at most one solution to* (1.25) *in the sense of Definition 1.19.*

Proof Let $\kappa > 0$ be a positive number satisfying

$$
0 < \kappa < \min\left\{\frac{1}{p_2}, \frac{2}{d}(1-\frac{1}{r}), \frac{1}{2d}\Big(\frac{1}{r} + 1 - [\frac{d}{p_2} + \frac{2}{q_2} + \frac{1}{r}]\Big)\right\} < \frac{1}{d}. \tag{4.26}
$$

The condition $(\mathbf{K}'_\mathbf{r})$ implies that it happens either $\frac{d}{p_2} + \frac{2}{q_2} + \frac{1}{r} < 1$ or $r < \infty$, which together with $r > 1$ ensures the existence of κ.

Suppose that there exist two solutions f^1 and f^2 starting from the same initial data in $\mathcal{P}(\mathbb{R}^d \times [0,1])$. Since $\kappa < \frac{2}{d}(1-\frac{1}{r})$, we deduce from Corollary 4.18 that f^1 and f^2 belong to $L^{\frac{2}{d\kappa}}([0,T],(L^{\frac{r}{r-1}}[0,1],L^{\frac{1}{1-\kappa}}(\mathbb{R}^d)))$. Through the mild formulation given by Lemma 4.21, we compute the $L_\xi^{\frac{r}{r-1}}L_x^{\frac{1}{1-\kappa}}$-norm of $f^1 - f^2$. Firstly,

$$
\begin{aligned}
&\|f_t^1(\cdot,\xi) - f_t^2(\cdot,\xi)\|_{L^{\frac{1}{1-\kappa}}} \\
&\leqslant \int_0^t \left\|\nabla\Gamma_{t-s} * \left(\int_{\mathbb{R}^d}\int_0^1 G(\xi,\zeta)K(\cdot-y)f_s^1(y,\zeta)\mathrm{d}\zeta\mathrm{d}y\Big[f_s^1(\cdot,\xi) - f_s^2(\cdot,\xi)\Big]\right)\right\|_{L^{\frac{1}{1-\kappa}}}\mathrm{d}s \\
&\quad + \int_0^t \left\|\nabla\Gamma_{t-s} * \left(\int_{\mathbb{R}^d}\int_0^1 G(\xi,\zeta)K(\cdot-y)\Big[f_s^1(y,\zeta) - f_s^2(y,\zeta)\Big]f_s^2(\cdot,\xi)\mathrm{d}\zeta\mathrm{d}y\right)\right\|_{L^{\frac{1}{1-\kappa}}}\mathrm{d}s \\
&:= J_1(t,\xi) + J_2(t,\xi).
\end{aligned} \tag{4.27}
$$

Using Young's convolution inequality, we have

$$
\begin{aligned}
J_1(t,\xi) &\leqslant \int_0^t \|\nabla\Gamma_{t-s}\|_{L^{\frac{1}{1-\kappa^2}}}\left\|K * \tilde{f}_s^1(\cdot,\xi)\Big[f_s^1(\cdot,\xi) - f_s^2(\cdot,\xi)\Big]\right\|_{L^{\frac{1}{1-\kappa(1-\kappa)}}}\mathrm{d}s \\
&\leqslant \int_0^t \|\nabla\Gamma_{t-s}\|_{L^{\frac{1}{1-\kappa^2}}}\left\|K * \tilde{f}_s^1(\cdot,\xi)\right\|_{L^{\frac{1}{\kappa^2}}}\left\|f_s^1(\cdot,\xi) - f_s^2(\cdot,\xi)\right\|_{L^{\frac{1}{1-\kappa}}}\mathrm{d}s \\
&\lesssim \int_0^t (t-s)^{-\frac{1+d\kappa^2}{2}}\|K\|_{L^{p_2}}\left\|\tilde{f}_s^1(\cdot,\xi)\right\|_{L^p}\left\|f_s^1(\cdot,\xi) - f_s^2(\cdot,\xi)\right\|_{L^{\frac{1}{1-\kappa}}}\mathrm{d}s,
\end{aligned}
$$

where $\tilde{f}^1$ is defined by $\tilde{f}^1 := \int_0^1 G(\xi,\zeta)f^1(y,\zeta)\mathrm{d}\zeta$ and $\frac{1}{p} = 1 + \kappa^2 - \frac{1}{p_2}$ similar to (3.45). We further find that

$$
\left\|\tilde{f}^1(\cdot,\xi)\right\|_{L^p} = \left(\int_{\mathbb{R}^d}\left|\int_0^1 G(\xi,\zeta)f^1(y,\zeta)\mathrm{d}\zeta\right|^p\mathrm{d}y\right)^{\frac{1}{p}}
$$

$$\leqslant \left(\int_{\mathbb{R}^d} \left(\int_0^1 \left| G(\xi, \zeta) f^1(y, \zeta) \right| \mathrm{d}\zeta \right)^p \mathrm{d}y \right)^{\frac{1}{p}}$$
$$\leqslant \int_0^1 \left(\int_{\mathbb{R}^d} \left| G(\xi, \zeta) f^1(y, \zeta) \right|^p \mathrm{d}y \right)^{\frac{1}{p}} \mathrm{d}\zeta,$$

here we used Minkowski's integral inequality to find the last line. Then for almost every $\xi \in [0, 1]$, we obtain

$$\begin{aligned} \left\| \tilde{f}^1(\cdot, \xi) \right\|_{L^p} \leqslant & \int_0^1 |G(\xi, \zeta)| \| f^1(\cdot, \zeta) \|_{L^p} \mathrm{d}\zeta \\ \leqslant & \|G\|_{L^\infty L^r} \|f^1\|_{L_\xi^{\frac{r}{r-1}} L_x^p}. \end{aligned} \tag{4.28}$$

Therefore, we have the following estimate for $J_1(t, \xi)$,

$$J_1(t, \xi) \lesssim \int_0^t (t-s)^{-\frac{1+d\kappa^2}{2}} \|K\|_{L^{p_2}} \|f_s^1\|_{L_\xi^{\frac{r}{r-1}} L_x^p} \left\| f_s^1(\cdot, \xi) - f_s^2(\cdot, \xi) \right\|_{L^{\frac{1}{1-\kappa}}} \mathrm{d}s. \tag{4.29}$$

On the other hand, for $J_2(t, \xi)$, we find

$$\begin{aligned} J_2(t, \xi) \leqslant & \int_0^t (t-s)^{-\frac{1+d\kappa^2}{2}} \left\| K * \left[\tilde{f}_s^1 - \tilde{f}_s^2 \right](\cdot, \xi) f_s^2(\cdot, \xi) \right\|_{L^{\frac{1}{1-\kappa(1-\kappa)}}} \mathrm{d}s \\ \lesssim & \int_0^t (t-s)^{-\frac{1+d\kappa^2}{2}} \|K\|_{L^{p_2}} \left\| f_s^2(\cdot, \xi) \right\|_{L^p} \|\tilde{f}_s^1(\cdot, \xi) - \tilde{f}_s^2(\cdot, \xi)\|_{L^{\frac{1}{1-\kappa}}} \mathrm{d}s, \end{aligned}$$

where we used $(1 - \kappa(1 - \kappa))^{-1} < p$ following by $\kappa < \frac{1}{p_2}$. Using (4.28) with $\frac{1}{1-\kappa}$ replacing p, we have

$$J_2(t, \xi) \lesssim \int_0^t (t-s)^{-\frac{1+d\kappa^2}{2}} \|K\|_{L^{p_2}} \left\| f_s^2(\cdot, \xi) \right\|_{L^p} \|f_s^1 - f_s^2\|_{L_\xi^{\frac{r}{r-1}} L_x^{\frac{1}{1-\kappa}}} \mathrm{d}s. \tag{4.30}$$

Now that combining (4.29) and (4.30) yields that

$$\begin{aligned} & \|f_t^1 - f_t^2\|_{L_\xi^{\frac{r}{r-1}} L_x^{\frac{1}{1-\kappa}}} \\ \leqslant & \|J_1(t, \cdot)\|_{L^{\frac{r}{r-1}}} + \|J_2(t, \cdot)\|_{L^{\frac{r}{r-1}}} \\ \lesssim & \int_0^t (t-s)^{-\frac{1+d\kappa^2}{2}} \|K\|_{L^{p_2}} \left(\|f_s^1\|_{L_\xi^{\frac{r}{r-1}} L_x^p} + \|f_s^2\|_{L_\xi^{\frac{r}{r-1}} L_x^p} \right) \|f_s^1 - f_s^2\|_{L_\xi^{\frac{r}{r-1}} L_x^{\frac{1}{1-\kappa}}} \mathrm{d}s. \end{aligned}$$

Choose q, m such that $\frac{d}{p} + \frac{2}{q} = d$ and $\frac{1}{q_2} + \frac{1}{q} + \frac{d\kappa}{2} + \frac{1}{m} = 1$. Then applying Hölder's inequality then gives

$$\begin{aligned} & \|f_t^1 - f_t^2\|_{L_\xi^{\frac{r}{r-1}} L_x^{\frac{1}{1-\kappa}}} \\ \lesssim & \|K\|_{L_{q_2}^{p_2}} \left(\|f^1\|_{L_t^q L_\xi^{\frac{r}{r-1}} L_x^p} + \|f^2\|_{L_t^q L_\xi^{\frac{r}{r-1}} L_x^p} \right) \\ & \cdot \|f^1 - f^2\|_{L_t^{\frac{2}{d\kappa}} L_\xi^{\frac{r}{r-1}} L_x^{\frac{1}{1-\kappa}}} \left(\int_0^t (t-s)^{-\frac{1+d\kappa^2}{2} m} \mathrm{d}s \right)^{\frac{1}{m}} \\ \lesssim & \|f^1 - f^2\|_{L_t^{\frac{2}{d\kappa}} L_\xi^{\frac{r}{r-1}} L_x^{\frac{1}{1-\kappa}}}. \end{aligned}$$

The last inequality follows by Corllary 4.18 and $\frac{1}{m} > \frac{1+d\kappa^2}{2}$ (see (3.46)).

We thus arrive at

$$\|f_t^1 - f_t^2\|^{\frac{2}{d\kappa}}_{L_\xi^{\frac{r}{r-1}} L_x^{\frac{1}{1-\kappa}}} \lesssim \int_0^t \|f_s^1 - f_s^2\|^{\frac{2}{d\kappa}}_{L_\xi^{\frac{r}{r-1}} L_x^{\frac{1}{1-\kappa}}} \mathrm{d}s.$$

Therefore we obtain the $f_t^1 = f_t^2$ in $L^{\frac{r}{r-1}}([0,1], L^{\frac{1}{1-\kappa}}(\mathbb{R}^d))$ for a.e. $t \in (0,T)$ by applying Gronwall's inequality.

Now we finish the proofs of the theorems of this chapter.

Proof [Proof of Theorem 1.21] The result follows by combining Proposition 4.14, Proposition 4.20, and Theorem 4.22.

The differences between Theorem 1.21 and Theorem 1.20 are due to assumptions of convergences of graphons and initlal values.

Proof [Proof of Theorem 1.20] By Lemma 4.1 and Lemma 4.2, we know that, up to subsequences, there exists a limiting graphon G and permutations $\{\sigma_N : \{1,...,N\} \to \{1,...,N\}\}$, such that $\delta_\square(G_N, G) \to 0$ and $\|G_N^{\sigma_N} - G\|_\square \to 0$.

Therefore, extracting subsequences, reordering the particle systems (4.1), and applying Theorem 1.21 give that,

$$\sum_{i=1}^N \delta_{X_{\sigma_N(i)}(t)} 1_{[\frac{i-1}{N}, \frac{i}{N})}(\xi) \rightharpoonup f$$

in $C([0,T], \mathcal{P}(\mathbb{R}^d \times [0,1]))$ almost surely, where f solves (1.25) with the graphon G in the sense of Definition 1.19. Now we have

$$\mu_N(t) = \frac{1}{N}\sum_{i=1}^N \delta_{X_i(t)} = \frac{1}{N}\sum_{i=1}^N \delta_{X_{\sigma_N(i)(t)}} \rightharpoonup \int_0^1 f(t,\cdot,\xi)\mathrm{d}\xi$$

in $C([0,T], \mathcal{P}(\mathbb{R}^d))$ almost surely. This completes the proof.

Bibliography

[ACM19] Giovanni Alberti, Gianluca Crippa, and Anna Mazzucato. Exponential self-similar mixing by incompressible flows. *Journal of the American Mathematical Society*, 32(2):445–490, 2019.

[AGS08] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2008.

[AS16] Noga Alon and Joel H Spencer. *The probabilistic method*. John Wiley & Sons, 2016.

[BA94] Matania Ben-Artzi. Global solutions of two-dimensional Navier-Stokes and Euler equations. *Archive for rational mechanics and analysis*, 128(4):329–358, 1994.

[BBPS22a] Jacob Bedrossian, Alex Blumenthal, and Samuel Punshon-Smith. Almost-sure exponential mixing of passive scalars by the stochastic navier–stokes equations. *The Annals of Probability*, 50(1):241–303, 2022.

[BBPS22b] Jacob Bedrossian, Alex Blumenthal, and Samuel Punshon-Smith. The Batchelor spectrum of passive scalar turbulence in stochastic fluid mechanics at fixed reynolds number. *Communications on Pure and Applied Mathematics*, 75(6):1237–1291, 2022.

[BCCZ] Christian Borgs, Jennifer Chayes, Henry Cohn, and Yufei Zhao.

[BCD11] Hajer Bahouri, Jean-Yves Chemin, and Raphaël Danchin. *Fourier analysis and nonlinear partial differential equations*, volume 343. Springer Science & Business Media, 2011.

[BCN20] Gianmarco Bet, Fabio Coppini, and Francesca R Nardi. Weakly interacting oscillators on dense random graphs. *arXiv preprint arXiv:2006.07670*, 2020.

[BCW20] Erhan Bayraktar, Suman Chakraborty, and Ruoyu Wu. Graphon mean field systems. *arXiv preprint arXiv:2003.13180*, 2020.

[BFH18] Dominic Breit, Eduard Feireisl, and Martina Hofmanová. *Stochastically forced compressible fluid flows*, volume 3. Walter de Gruyter GmbH & Co KG, 2018.

[BFT15] Mireille Bossy, Olivier Faugeras, and Denis Talay. Clarification and complement to "Mean-field description and propagation of chaos in networks of

Hodgkin–Huxley and FitzHugh–Nagumo neurons". *The Journal of Mathematical Neuroscience (JMN)*, 5(1):1–23, 2015.

[BGSRS20] Thierry Bodineau, Isabelle Gallagher, Laure Saint-Raymond, and Sergio Simonella. Statistical dynamics of a hard sphere gas: fluctuating Boltzmann equation and large deviations. *arXiv preprint arXiv:2008.10403*, 2020.

[BGSRS22] Thierry Bodineau, Isabelle Gallagher, Laure Saint-Raymond, and Sergio Simonella. Long-time derivation at equilibrium of the fluctuating Boltzmann equation. *arXiv preprint arXiv:2201.04514*, 2022.

[BH77] Werner Braun and K Hepp. The Vlasov dynamics and its fluctuations in the 1/N limit of interacting classical particles. *Communications in mathematical physics*, 56(2):101–113, 1977.

[BJW20] Didier Bresch, Pierre-Emmanuel Jabin, and Zhenfu Wang. Mean-field limit and quantitative estimates with singular attractive kernels. *arXiv preprint arXiv:2011.08022*, 2020.

[BO19] Robert J Berman and Magnus Onnheim. Propagation of chaos for a class of first order models with singular mean field interactions. *SIAM Journal on Mathematical Analysis*, 51(1):159–196, 2019.

[Bog07] Vladimir I Bogachev. *Measure theory*, volume 2. Springer Science & Business Media, 2007.

[Bre94] Haïm Brezis. Remarks on the preceding paper by M. Ben-Artzi" Global solutions of two-dimensional Navier-Stokes and Euler equations". *Archive for Rational Mechanics and Analysis*, 128(4):359–360, 1994.

[BRS21] Vladimir I Bogachev, Michael Röckner, and Stanislav V Shaposhnikov. On the Ambrosio–Figalli–Trevisan superposition principle for probability solutions to Fokker–Planck–Kolmogorov equations. *Journal of Dynamics and Differential Equations*, 33(2):715–739, 2021.

[Car91] Eric A Carlen. Superadditivity of Fisher's information and logarithmic Sobolev inequalities. *Journal of Functional Analysis*, 101(1):194–211, 1991.

[CCH14] José Antonio Carrillo, Young-Pil Choi, and Maxime Hauray. The derivation of swarming models: mean-field limit and Wasserstein distances. In *Collective dynamics from bacteria to crowds*, pages 1–46. Springer, 2014.

[CF16] Zhen-Qing Chen and Wai-Tong Louis Fan. Fluctuation limit for interacting diffusions with partial annihilations through membranes. *Journal of Statistical Physics*, 164(4):890–936, 2016.

[CG19] Michele Coghi and Benjamin Gess. Stochastic nonlinear Fokker–Planck equations. *Nonlinear Analysis*, 187:259–278, 2019.

[CG22] Chenyang Chen and Hao Ge. Sample-path large deviation principle for a 2-D stochastic interacting vortex dynamics with singular kernel. *arXiv preprint arXiv:2205.11013*, 2022.

[Che17] Julien Chevallier. Fluctuations for mean-field interacting age-dependent Hawkes processes. *Electronic Journal of Probability*, 22, 2017.

[DDF88] Anthony D'Aristotile, Persi Diaconis, and David Freedman. On merging of probabilities. *The Indian Journal of Statistics, Series A*, pages 363–380, 1988.

[DE11] Paul Dupuis and Richard S Ellis. *A weak convergence approach to the theory of large deviations*, volume 902. John Wiley & Sons, 2011.

[DF37] Bruno De Finetti. La prévision: ses lois logiques, ses sources subjectives. In *Annales de l'institut Henri Poincaré*, volume 7, pages 1–68, 1937.

[Dob79] Roland L'vovich Dobrushin. Vlasov equations. *Functional Analysis and Its Applications*, 13(2):115–123, 1979.

[Dud18] Richard M Dudley. *Real analysis and probability*. CRC Press, 2018.

[Due16] Mitia Duerinckx. Mean-field limits for some Riesz interaction gradient flows. *SIAM Journal on Mathematical Analysis*, 48(3):2269–2300, 2016.

[Elw82] Kenneth David Elworthy. *Stochastic differential equations on manifolds*, volume 70. Cambridge University Press, 1982.

[Éme12] Michel Émery. *Stochastic calculus in manifolds*. Springer Science & Business Media, 2012.

[Eva98] Lawrence C Evans. Partial differential equations. *Graduate studies in mathematics*, 19(2), 1998.

[FGP11] Franco Flandoli, Massimiliano Gubinelli, and Enrico Priola. Full well-posedness of point vortex dynamics corresponding to stochastic 2D Euler equations. *Stochastic Processes and their Applications*, 121(7):1445–1463, 2011.

[FGV01] Gregory Falkovich, K Gawedzki, and Massimo Vergassola. Particles and fields in fluid turbulence. *Reviews of modern Physics*, 73(4):913, 2001.

[FH16] Nicolas Fournier and Maxime Hauray. Propagation of chaos for the Landau equation with moderately soft potentials. *The Annals of Probability*, 44(6):3581–3660, 2016.

[FHM14] Nicolas Fournier, Maxime Hauray, and Stéphane Mischler. Propagation of chaos for the 2D viscous vortex model. *Journal of the European Mathematical Society*, 16(7):1423–1466, 2014.

[FJ90] Noah E Friedkin and Eugene C Johnsen. Social influence and opinions. *Journal of Mathematical Sociology*, 15(3-4):193–206, 1990.

[FL21] Franco Flandoli and Dejun Luo. Mean field limit of point vortices with environmental noises to deterministic 2D Navier-Stokes equations. *arXiv preprint arXiv:2103.01497*, 2021.

[fla] Stochastic currents, author=Flandoli, Franco and Gubinelli, Massimiliano and Giaquinta, Mariano and Tortorelli, Vincenzo M, journal=Stochastic processes and their applications, volume=115, number=9, pages=1583–1601, year=2005, publisher=Elsevier.

[FM97] Begoña Fernandez and Sylvie Méléard. A Hilbertian approach for fluctuations on the McKean-Vlasov model. *Stochastic processes and their applications*, 71(1):33–53, 1997.

[FM07] Joaquin Fontbona and Miguel Martinez. Paths clustering and an existence result for stochastic vortex systems. *Journal of Statistical Physics*, 128(3):699–719, 2007.

[GK96] István Gyöngy and Nicolai Krylov. Existence of strong solutions for Itô's stochastic equations via approximations. *Probability theory and related fields*, 105(2):143–158, 1996.

[Hai14] Martin Hairer. A theory of regularity structures. *Inventiones mathematicae*, 198(2):269–504, 2014.

[HM14] Maxime Hauray and Stéphane Mischler. On Kac's chaos and related problems. *Journal of Functional Analysis*, 266(10):6055–6157, 2014.

[HMC+06] Minyi Huang, Roland P Malhamé, Peter E Caines, et al. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Communications in Information & Systems*, 6(3):221–252, 2006.

[Hoe94] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. In *The collected works of Wassily Hoeffding*, pages 409–426. Springer, 1994.

[HRZ22] Zimo Hao, Michael Röckner, and Xicheng Zhang. Strong convergence of propagation of chaos for McKean-Vlasov SDEs with singular interactions. *arXiv preprint arXiv:2204.07952*, 2022.

[HS55] Edwin Hewitt and Leonard J Savage. Symmetric measures on Cartesian products. *Transactions of the American Mathematical Society*, 80(2):470–501, 1955.

[Hsu02] Elton P Hsu. *Stochastic analysis on manifolds*. Number 38. American Mathematical Soc., 2002.

[Itô83] Kiyosi Itô. Distribution-valued processes arising from independent Brownian motions. *Mathematische Zeitschrift*, 182(1):17–33, 1983.

[Jab14] Pierre-Emmanuel Jabin. A review of the mean field limits for Vlasov equations. *Kinetic & Related Models*, 7(4):661, 2014.

[Jea15] J.H. Jeans. On the theory of star-streaming and the structure of the universe. *Monthly Notices of the Royal Astronomical Society*, 76:70–84, 1915.

[JH15] Pierre-Emmanuel Jabin and Maxime Hauray. Particles approximations of Vlasov equations with singular forces: Propagation of chaos. In *Annales scientifiques de l'École Normale Supérieure*, 2015.

[JM98] Benjamin Jourdain and Sylvie Méléard. Propagation of chaos and fluctuations for a moderate model with smooth initial data. In *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, volume 34, pages 727–766. Elsevier, 1998.

[JO04] Pierre-Emmanuel Jabin and Felix Otto. Identification of the dilute regime in particle sedimentation. *Communications in mathematical physics*, 250(2):415–432, 2004.

[JPS21] Pierre-Emmanuel Jabin, David Poyato, and Juan Soler. Mean-field limit of non-exchangeable systems. *arXiv preprint arXiv:2112.15406*, 2021.

[JW18] Pierre-Emmanuel Jabin and Zhenfu Wang. Quantitative estimates of propagation of chaos for stochastic systems with $W^{-1,\infty}$ kernels. *Inventiones mathematicae*, 214(1):523–591, 2018.

[Kan67] Mamoru Kanda. Regular points and Green functions in Markov processes. *Journal of the Mathematical Society of Japan*, 19(1):46–69, 1967.

[Kel17] John L Kelley. *General topology*. Courier Dover Publications, 2017.

[KS84] G Koumoullis and A Sapounakis. Two countability properties of sets of measures. *Michigan Mathematical Journal*, 31(1):31–47, 1984.

[KS21] Franziska Kuhn and Rene L. Schilling. Convolution inequalities for Besov and Triebel Lizorkin spaces, and applications to convolution semigroups. 2021.

[Kun97] Hiroshi Kunita. *Stochastic flows and stochastic differential equations*, volume 24. Cambridge university press, 1997.

[Kur14] Thomas Kurtz. Weak and strong solutions of general stochastic models. *Electronic Communications in Probability*, 19, 2014.

[Lac21] Daniel Lacker. Hierarchies, entropy, and quantitative propagation of chaos for mean field diffusions. *arXiv preprint arXiv:2105.02983*, 2021.

[Laz16] Dustin Lazarovici. The Vlasov-Poisson dynamics as the mean field limit of extended charges. *Communications in Mathematical Physics*, 347(1):271–289, 2016.

[LL07] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Japanese journal of mathematics*, 2(1):229–260, 2007.

[LLX20] Songzi Li, Xiang-Dong Li, and Yong-Xiao Xie. On the law of large numbers for the empirical measure process of generalized Dyson Brownian motion. *Journal of Statistical Physics*, 181(4):1277–1305, 2020.

[LLY19] Lei Li, Jian-Guo Liu, and Pu Yu. On the mean field limit for Brownian particles with Coulomb interaction in 3D. *Journal of Mathematical Physics*, 60(11):111501, 2019.

[Lov12] László Lovász. *Large networks and graph limits*, volume 60. American Mathematical Soc., 2012.

[LRW19] Daniel Lacker, Kavita Ramanan, and Ruoyu Wu. Local weak convergence for sparse networks of interacting processes. *arXiv preprint arXiv:1904.02585*, 2019.

[LS16] Eric Lucon and Wilhelm Stannat. Transition from gaussian to non-gaussian fluctuations for mean-field diffusions in spatial interaction. *The Annals of Applied Probability*, 26(6):3840–3909, 2016.

[Luc20] Eric Lucon. Quenched asymptotics for interacting diffusions on inhomogeneous random graphs. *Stochastic Processes and their Applications*, 130(11):6783–6842, 2020.

[McK67] HP McKean. Propagation of chaos for a class of non-linear parabolic equations. 1967 Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967) pp. 41–57 Air Force Office Sci. *Res., Arlington, Va*, 1967.

[McK75] HP McKean. Fluctuations in the kinetic theory of gases. *Communications on pure and applied mathematics*, 28(4):435–455, 1975.

[Mél96] Sylvie Méléard. Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models. In *Probabilistic models for nonlinear partial differential equations*, pages 42–95. Springer, 1996.

[MM13] Stéphane Mischler and Clément Mouhot. Kac's program in kinetic theory. *Inventiones mathematicae*, 193(1):1–147, 2013.

[MMN18] Song Mei, Andrea Montanari, and Phan-Minh Nguyen. A mean field view of the landscape of two-layer neural networks. *Proceedings of the National Academy of Sciences*, 115(33):E7665–E7671, 2018.

[MMW15] Stéphane Mischler, Clément Mouhot, and Bernt Wennberg. A new approach to quantitative propagation of chaos for drift, diffusion and jump processes. *Probability Theory and Related Fields*, 161(1):1–59, 2015.

[MP12] Carlo Marchioro and Mario Pulvirenti. *Mathematical theory of incompressible nonviscous fluids*, volume 96. Springer Science & Business Media, 2012.

[MW17] Jean-Christophe Mourrat and Hendrik Weber. Global well-posedness of the dynamic Phi4 model in the plane. *The Annals of Probability*, 45(4):2398–2476, 2017.

[NRS21] Quoc Hung Nguyen, Matthew Rosenzweig, and Sylvia Serfaty. Mean-field limits of Riesz-type singular flows with possible multiplicative transport noise. *arXiv preprint arXiv:2107.02592*, 2021.

[Nua06] David Nualart. *The Malliavin calculus and related topics*, volume 1995. Springer, 2006.

[OR19] Roberto I Oliveira and Guilherme H Reis. Interacting diffusions on random graphs with diverging average degrees: Hydrodynamics and large deviations. *Journal of Statistical Physics*, 176(5):1057–1087, 2019.

[ORS20] Roberto I Oliveira, Guilherme H Reis, and Lucas M Stolerman. Interacting diffusions on sparse graphs: hydrodynamics from local weak limits. *Electronic Journal of Probability*, 25:1–35, 2020.

[Osa85] Hirofumi Osada. A stochastic differential equation arising from the vortex problem. *Proceedings of the Japan Academy, Series A, Mathematical Sciences*, 61(10):333–336, 1985.

[Osa86] Hirofumi Osada. Propagation of chaos for the two dimensional Navier-Stokes equation. *Proceedings of the Japan Academy, Series A, Mathematical Sciences*, 62(1):8–11, 1986.

[Ros20] Matthew Rosenzweig. Mean-Field convergence of point vortices without regularity. *arXiv preprint arXiv:2004.04140*, 2020.

[Sei13] Christian Seis. Maximal mixing by incompressible fluid flows. *Nonlinearity*, 26(12):3279, 2013.

[Ser20] Sylvia Serfaty. Mean field limit for Coulomb-type flows. *Duke Mathematical Journal*, 169(15):2887–2935, 2020.

[SS00] Boris I Shraiman and Eric D Siggia. Scalar turbulence. *Nature*, 405(6787):639–646, 2000.

[Str00] Daniel W Stroock. *An introduction to the analysis of paths on a Riemannian manifold*. Number 74. American Mathematical Soc., 2000.

[SYY20] Jian Song, Jianfeng Yao, and Wangjun Yuan. High-dimensional limits of eigenvalue distributions for general Wishart process. *The Annals of Applied Probability*, 30(4):1642–1668, 2020.

[Szn] Alain-Sol Sznitman. Topics in propagation of chaos. In *Ecole d'été de probabilités de Saint-Flour XIX—1989, pages=165–251, year=1991, publisher=Springer*.

[Szn84] Alain-Sol Sznitman. Nonlinear reflecting diffusion process, and the propagation of chaos and fluctuations associated. *Journal of functional analysis*, 56(3):311–336, 1984.

[Tak85] Satoshi Takanobu. On the existence and uniqueness of SDE describing an n-particle system interacting via a singular potential. *Proceedings of the Japan Academy, Series A, Mathematical Sciences*, 61(9):287–290, 1985.

[Tan82] Hiroshi Tanaka. Fluctuation theory for Kac's one dimensional model of Maxwellian molecules. *The Indian Journal of Statistics, Series A*, pages 23–46, 1982.

[Tan83] Hiroshi Tanaka. Some probabilistic problems in the spatially homogeneous Boltzmann equation. In *Theory and application of random fields*, pages 258–267. Springer, 1983.

[Tan84] Hiroshi Tanaka. Limit theorems for certain diffusion processes with interaction. In *North-Holland Mathematical Library*, volume 32, pages 469–488. Elsevier, 1984.

[Tay13] Michael Taylor. *Partial differential equations II: Qualitative studies of linear equations*, volume 116. Springer Science & Business Media, 2013.

[TH81] Hiroshi Tanaka and Masuyuki Hitsuda. Central limit theorem for a simple diffusion model of interacting particles. *Hiroshima Mathematical Journal*, 11(2):415–423, 1981.

[Tre16] Dario Trevisan. Well-posedness of multidimensional diffusion processes with weakly differentiable coefficients. *Electronic Journal of Probability*, 21:1–41, 2016.

[Tri06] Hans Triebel. Theory of Function Spaces. III,. *Monographs in Mathematics*, 2006.

[Uch83] Kohei Uchiyama. Fluctuations of Markovian systems in Kac's caricature of a Maxwellian gas. *Journal of the Mathematical Society of Japan*, 35(3):477–499, 1983.

[Vil08] Cédric Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008.

[War00] Zellman Warhaft. Passive scalars in turbulent flows. *Annual Review of Fluid Mechanics*, 32(1):203–240, 2000.

[Wyn21] Dominic Wynter. Quantitative Propagation of Chaos for the Mixed-Sign Viscous Vortex Model on the Torus. *arXiv preprint arXiv:2106.05255*, 2021.

[WZZ21] Zhenfu Wang, Xianliang Zhao, and Rongchan Zhu. Gaussian fluctuations for interacting particle systems with singular kernels. *arXiv preprint arXiv:2105.13201*, 2021.

[WZZ22] Zhenfu Wang, Xianliang Zhao, and Rongchan Zhu. Mean-field limit of Non-exchangeable interacting diffusions with singular kernels. *arXiv preprint arXiv:2209.14002*, 2022.

[Zam98] Ram Zamir. A proof of the Fisher information inequality via a data processing argument. *IEEE Transactions on Information Theory*, 44(3):1246–1250, 1998.

[ZDE20] Michele Coti Zelati, Matias G Delgadino, and Tarek M Elgindi. On the relation between enhanced dissipation timescales and mixing rates. *Communications on Pure and Applied Mathematics*, 73(6):1205–1244, 2020.

[Zha10] Xicheng Zhang. Stochastic Volterra equations in Banach spaces and stochastic partial differential equation. *Journal of Functional Analysis*, 258(4):1361–1425, 2010.