Strong Convergence of Euler-Maruyama Schemes for McKean-Vlasov Stochastic Differential Equations under Local Lipschitz Conditions of State Variables

Yun Li* Xuerong Mao† Qingshuo Song‡ Fuke Wu§ George Yin¶

Abstract

This paper develops strong convergence of the Euler-Maruyama (EM) schemes for approximating McKean-Vlasov stochastic differential equations (SDEs). In contrast to the existing work, a novel feature is the use of a much weaker condition-local Lipschitzian in the state variable but under uniform linear growth assumption. To obtain the desired approximation, the paper first establishes the existence and uniqueness of solutions of the original McKean-Vlasov SDE using an Euler-like sequence of interpolations and partition of the sample space. Then, the paper returns to the analysis of the EM scheme for approximating solutions of McKean-Vlasov SDEs. A strong convergence theorem is established. Moreover, the convergence rates under global conditions are obtained.

Keywords. McKean-Vlasov SDE; one-sided local Lipschitz condition; local Lipschitz condition; interpolated Euler-like sequence; Euler-Maruyama scheme.

Mathematics Subject Classification (2010). 60H10, 65C35, 60J60.

Running title. EM Algorithms for McKean-Vlasov SDEs under Local Conditions

*School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan, Hubei, 430074, P.R. China, (li_yun@hust.edu.cn). The research of this author was supported in part by the National Natural Science Foundation of China (Grant No. 61873320).

†Department of Mathematics and Statistics, University of Strathclyde, Glasgow G1 1XH, UK, (x.mao@strath.ac.uk). The research of this author was supported in part by the Newton Fund, UK (NA160317, Royal Society-Newton Advanced Fellow-ship) and the Royal Society of Edinburgh (RSE1832).

‡Department of Mathematics, Worcester Polytechnic Institute, Worcester, MA 01609, USA, (qsong@wpi.edu).

§School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan, Hubei, 430074, P.R. China, (wufuke@hust.edu.cn). The research of this author was supported in part by the National Natural Science Foundation of China (Grant No. 61873320).

¶Department of Mathematics, University of Connecticut, Storrs, CT 06269-1009, USA, (gyin@uconn.edu). The research of this author was supported in part by the Air Force Office of Scientific Research.

1 Introduction

There has been increasing and resurgent interests to treat McKean-Vlasov stochastic differential equations (SDEs). The motivation stems from considerations of mean-field control, mean-field games, as well as complex networked systems; see [3, 15, 24]. The idea of using mean-field models has been around for a long time, one of the initial considerations was from statistical mechanics to replace "interactions of a large number of particles" in a "many body" problem (known to be notoriously difficult) by an average of the "bodies". The first mathematical treatment to justify such a replacement was [9]. The recent interest of mean-field models, mean-field controls, and mean-field games has stimulated much of the study on McKean-Vlasov SDEs. A distinct feature of such systems is the appearance of the probability laws in the coefficients of the resulting equations.

Consider a McKean-Vlasov SDE (also known as distribution-dependent SDE or mean-field SDE) of the form

$$dX(t) = b(t, X(t), \mathscr{L}(X(t)))dt + \sigma(t, X(t), \mathscr{L}(X(t)))dW(t), \quad X(0) = \xi, \tag{1.1}$$

where $\mathscr{L}(X(t))$ denotes the law or distribution of $X(t)$, $b : [0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d$ and $\sigma : [0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^{d\times d_1}$ are Borel measurable functions (see Section 2 for the definition of $\mathcal{P}_2(\mathbb{R}^d)$), ξ is an $\mathcal{F}_0$-measurable random variable satisfying $\mathbb{E}|\xi|^\beta < \infty$ for any $\beta > 0$, and $W(t)$ is a d_1-dimensional standard Brownian motion defined on the complete probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ with filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual conditions. In this paper, we aim to find numerical approximation of the solution of (1.1) under local conditions with respect to the state variable.

To numerically approximate the solution of (1.1), we first introduce the following intermediate system of stochastic differential equations

$$dX^{i,N}(t) = b(t, X^{i,N}(t), \mu^{X,N}_t)dt + \sigma(t, X^{i,N}(t), \mu^{X,N}_t)dW^i(t), \quad i = 1, 2, \dots, N, \tag{1.2}$$

where $(W^i, X^i(0))_{1\leq i\leq N}$ are independent copies of $(W, X(0))$, initial condition $X^{i,N}(0) = X^i(0)$, and $\mu^{X,N}_t := \frac{1}{N}\sum_{j=1}^N \delta_{X^{j,N}(t)}$ is the empirical measure of $(X^{j,N}(t))_{1\leq j\leq N}$. The above system in fact, stems from statistical physics and closely related to Dawson's original work. In our recent work [30], we treated switching diffusion systems with an additional random switching mechanism and obtained a law of large number type result, with emphasis on the probabilistic aspect of the problems. In this paper, we use (1.2) as a reference system to build numerical schemes. In the literature, (1.2) is sometimes referred to as an interacting particle system, whereas the recent interests in control systems theory, refer to (1.2) as a system with mean-field terms or a system with mean-field interactions. The essence is that as N is getting large, the solution of (1.2) approximates that of (1.1) in an appropriate sense, which is called the propagation of chaos.

To proceed, we present an EM numerical scheme to approximate the solution of (1.2), which in turn, approximates the solution of (1.1). We partition the time interval $[0, T]$ into n subintervals of equal length with $h_n := \frac{T}{n}$, and let $t^n_k = kh_n$ for any $k = 0, 1, \dots, n$. Compute the discrete approximation $X^{i,N,n}(t^n_k)$ of $X^{i,N}(t^n_k)$ by setting $X^{i,N,n}(0) = X^{i,N}(0)$ and defining

$$\begin{aligned} X^{i,N,n}(t^n_{k+1}) = X^{i,N,n}(t^n_k) \quad &+ \quad b(t^n_k, X^{i,N,n}(t^n_k), \mu^{X,N,n}_{t^n_k})h_n \\ &+ \quad \sigma(t^n_k, X^{i,N,n}(t^n_k), \mu^{X,N,n}_{t^n_k})\Delta W^{i,n}(k), \quad i = 1, 2, \dots, N, \end{aligned} \tag{1.3}$$

where $\mu^{X,N,n}_{t^n_k} := \frac{1}{N}\sum_{j=1}^N \delta_{X^{j,N,n}(t^n_k)}$ and $\Delta W^{i,n}(k) := W^i(t^n_{k+1}) - W^i(t^n_k)$. In this work, we demonstrate the strong convergence of the numerical scheme under rather weak conditions.

However, before the desired result can be obtained, we have to make sure that (1.1) is well behaved. The existence and uniqueness of the solution of McKean-Vlasov SDEs have been investigated intensively. The study of McKean-Vlasov SDEs was initiated in [25], inspired by the kinetic theory of Kac [17]. An illustration of the general theory of McKean-Vlasov SDEs and their particle approximations can be found in [38]. The existence and uniqueness of the weak solution to (1.1) were shown with constant diffusion and bounded drift coefficient in [16]. Moreover, the existence and uniqueness of the strong solution to (1.1) were established under the global Lipschitz condition by using the fixed point theorem on the space of continuous functions with values in $\mathcal{P}_2(\mathbb{R}^d)$, for example, in [1,7]. The existence and uniqueness were then obtained with one-sided global Lipschitz drift coefficient and global Lipschitz diffusion coefficient in [34,35,40]. Under measure-dependent Lyapunov condition and integrated Lyapunov condition, [12] proved the existence of a weak solution, together with the pathwise uniqueness. It is worth mentioning that the global conditions w.r.t. the state variable were used in [1,7,12,34,35,40].

Perhaps, a local condition was first introduced in [20], where the drift and diffusion coefficients satisfy a global Lipschitz condition w.r.t. the state variable but a local Lipschitz condition w.r.t. the measure. In this paper, we aim to treat (1.1) under local Lipschitz conditions w.r.t. the state variable. More precisely, for a neighborhood of radius R, we use the following form of conditions: Letting $|x| \vee |y| \leq R$, then

$$\begin{aligned}
&\langle x-y, b(t,x,\mu)-b(t,y,\mu)\rangle \leq L_R^{(1)}|x-y|^2 \text{ with } L_R^{(1)} \leq \alpha_1 \log R,\\
&\|\sigma(t,x,\mu)-\sigma(t,y,\mu)\| \leq L_R^{(2)}|x-y| \quad \text{with} \quad L_R^{(2)} \leq \sqrt{\alpha_2 \log R},
\end{aligned}$$

and α_1, α_2 are positive constants. Note that the logarithmic growth condition on Lipschitz constants were used in [11] and [41] to investigate the global flow for SDEs and the convergence rates of the EM schemes for SDEs, respectively.

Compared to the classical SDEs, the local conditions w.r.t. the state variable, together with the distribution-dependent coefficients create much difficulties. For the classical SDEs, it is well known that local Lipschitz type conditions ensure the existence and uniqueness of the local solution. This, together with some growth conditions such as the linear growth conditions, or monotone conditions, or the Khasminskii-type conditions, implies the non-explosion of the solution leading to existence and uniqueness of the global solution; see [26] and references therein. Nevertheless, the distribution-dependent coefficients introduce fundamental difficulties because their solution cannot be determined in a pathwise fashion. As a result, the standard truncation technique, the stopping time technique, and the Yamada-Watanabe principle, etc. cannot be applied directly. In addition, if the coefficients are locally Lipschitz, the method of the fixed point theorem generally fails. Inspired by [20], to overcome the difficulties, we use interpolated Euler-like sequence and partition of sample space. The key step of the argument is to prove the interpolated Euler-like sequence is Cauchy in a proper space. In addition, we also obtain the uniqueness of the solution.

Once the existence and uniqueness are established, we focus on numerical approximations for McKean-Vlasov SDEs under local Lipschitz conditions on drift and diffusion coefficients of the state variable. Although there were numerous results on strong convergence of numerical approximations for McKean-Vlasov SDEs, the conditions used are the global conditions w.r.t. the state variable. To mention just a few, an explicit Euler scheme was developed in [4] to handle a specific McKean-Vlasov SDE, but the convergence is established under the global conditions and constant diffusion coefficient. Reference [35] investigated the strong convergence of the tamed EM scheme for McKean-Vlasov SDEs, where the coefficients are of global w.r.t. the state variable. Later, the tamed Milstein scheme and adaptive EM scheme were developed for McKean-Vlasov SDEs; see [22, 23, 36]. In [2], the authors established the strong convergence of the EM scheme by the Zvonkin transformation when the drift coefficient is Hölder continuous w.r.t. the state variable.

Under non-Lipschitz conditions, the convergence rates of the EM schemes were obtained in [10]. However, empirical measure was not used in the reference system, thus the law is not computable. Under the conditions of measurability and linear growth of the coefficients, strong convergence of the EM scheme was proved in [42] using the Krylov estimate, whereas the law involved was not simulated. To the best of our knowledge, to date, there is no result concerning the strong convergence for McKean-Vlasov SDEs under local condition w.r.t. the state variable.

For the classical SDEs, the strong convergence of the EM scheme was proved under local Lipschitz continuity and linear growth condition in [13]. Furthermore, the convergence rates of the EM scheme for the classical SDEs with the coefficients satisfying local Lipschitz continuity and linear growth condition were obtained in [41], in which the local Lipschitz constants satisfy the same logarithmic growth condition as that of this paper. Compared to the classical SDEs, the main difference here is the need to approximate distributions at every step. To proceed, the stochastic interacting particle system (1.2) is used as a bridge. In particular, this paper establishes the propagations of chaos under the local conditions w.r.t. the state variable.

The rest of the paper is arranged as follows. Section 2 introduces some preliminary preparations and states main results. Section 3 obtains the existence and uniqueness of the solution to (1.1) using an interpolated Euler-like sequence. Section 4 proves the strong convergence of the EM scheme for system (1.2). An example is given to illustrate our results in Section 5. Finally, an appendix containing the proofs of two lemmas is provided at the end of the paper.

2 Preliminaries and Main Results

Throughout the paper, unless otherwise specified, we use the following notation. Let $|\cdot|$ denote the Euclidean norm and $\langle \cdot, \cdot \rangle$ denote the scalar product in $\mathbb{R}^d$. For a matrix A, denote the Frobenius norm by $\|A\| = \sqrt{\mathrm{tr}[AA^\top]}$. Owing to technical reasons, we restrict ourselves to the following subspace of $\mathcal{P}(\mathbb{R}^d)$

$$\mathcal{P}_2(\mathbb{R}^d) := \Big\{ \mu \in \mathcal{P}(\mathbb{R}^d) : [\mu]_2 := \int_{\mathbb{R}^d} |x|^2 \mu(dx) < \infty \Big\}.$$

Note that for any $x \in \mathbb{R}^d$, the Dirac measure δ_x belongs to $\mathcal{P}_2(\mathbb{R}^d)$. Moreover, $\mathcal{P}_2(\mathbb{R}^d)$ is a Polish space under the L^2-Wasserstein distance

$$W_2(\mu, \nu) := \inf_{\pi \in \mathcal{C}(\mu,\nu)} \Big(\int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 \pi(dx, dy) \Big)^{\frac{1}{2}}, \quad \mu, \nu \in \mathcal{P}_2(\mathbb{R}^d),$$

where $\mathcal{C}(\mu, \nu)$ is the collection of all couplings for μ and ν. In other words, $\pi \in \mathcal{C}(\mu, \nu)$ is a probability measure on $\mathbb{R}^d \times \mathbb{R}^d$ such that $\pi(\cdot \times \mathbb{R}^d) = \mu(\cdot)$ and $\pi(\mathbb{R}^d \times \cdot) = \nu(\cdot)$. In particular, if $\mu = \mathscr{L}(X)$ and $\nu = \mathscr{L}(Y)$ are the distributions of random variables X and Y respectively, then

$$W_2(\mu, \nu)^2 \le \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 \mathscr{L}((X, Y))(dx, dy) = \mathbb{E}|X - Y|^2,$$

in which $\mathscr{L}((X, Y))$ represents the joint distribution of random vector (X, Y); see [5, 7, 32, 39].

Denote by $L^2(\Omega;\mathbb{R}^d)$ the set of random variables X with values in $\mathbb{R}^d$ satisfying $\mathbb{E}|X|^2 < \infty$. Consider a terminal time $T < \infty$, and denote by $C([0,T];\mathbb{R}^d)$ the collection of continuous functions on $[0,T]$ with values in $\mathbb{R}^d$, endowed with the supremum norm. Denote by $L^2(\Omega;C([0,T];\mathbb{R}^d))$ the family of random variables X with values in $C([0,T];\mathbb{R}^d)$ which satisfy $\mathbb{E}[\sup_{0\le t\le T}|X(t)|^2] < \infty$. Then, $L^2(\Omega;C([0,T];\mathbb{R}^d))$ is a Banach space under the norm

$$\|X\|_{L^2} := \Big(\mathbb{E}\Big[\sup_{0\le t\le T}|X(t)|^2\Big]\Big)^{\frac{1}{2}}.$$

Throughout this paper, C, C_p, and C_N denote positive constants which may be different in different places, where the subscript p or N is used to emphasize that the constant depends on p or N.

Following [7,40], let us first recall the definition of a strong solution to the McKean-Vlasov SDE.

Definition 2.1. An $\mathbb{R}^d$-valued stochastic process $(X(t))_{0\le t\le T}$ is a unique solution to equation (1.1), if it satisfies the following properties:

(i) $(X(t))_{0\le t\le T}$ is $\{\mathcal{F}_t\}$-adapted and continuous;

(ii) $\displaystyle\int_0^T (|b(t,X(t),\mathscr{L}(X(t)))| + \|\sigma(t,X(t),\mathscr{L}(X(t)))\|^2)dt < \infty, \quad \mathbb{P}\text{-a.s.};$

(iii) $\displaystyle X(t) = \xi + \int_0^t b(s,X(s),\mathscr{L}(X(s)))ds + \int_0^t \sigma(s,X(s),\mathscr{L}(X(s)))dW(s), t\in[0,T], \mathbb{P}\text{-a.s.};$

(iv) If $(\bar{X}(t))_{0\le t\le T}$ is another solution with $\bar{X}(0) = \xi$, then $(\bar{X}(t))_{0\le t\le T}$ and $(X(t))_{0\le t\le T}$ are indistinguishable, that is, $\mathbb{P}\{X(t) = \bar{X}(t) \text{ for all } 0\le t\le T\} = 1$.

Remark 2.1. If for any initial condition ξ satisfying $\mathbb{E}|\xi|^2 < \infty$, equation (1.1) has a unique solution $(X(t))_{0\le t\le T}$ satisfying $\mathbb{E}|X(t)|^2 < \infty$ for any $t\in[0,T]$, then we say equation (1.1) has a unique solution in $L^2(\Omega;\mathbb{R}^d)$, which implies that $\mu_t = \mathscr{L}(X(t)) \in \mathcal{P}_2(\mathbb{R}^d)$.

We need the following assumptions.

(H1) (One-sided local Lipschitz condition on the drift coefficient with respect to the state variable) For each integer $R\ge 3$, there exists a positive constant $L_R^{(1)} \le \alpha_1 \log R$ for some constant α_1 such that for any $t\in[0,T]$, $x,y\in\mathbb{R}^d$ with $|x|\vee|y|\le R$, and $\mu\in\mathcal{P}_2(\mathbb{R}^d)$,

$$\langle x-y, b(t,x,\mu) - b(t,y,\mu)\rangle \le L_R^{(1)}|x-y|^2.$$

(H2) (Local Lipschitz condition on the diffusion coefficient with respect to the state variable) For each integer $R\ge 3$, there is a positive constant $L_R^{(2)} \le \sqrt{\alpha_2\log R}$ for some constant α_2 such that for any $t\in[0,T]$, $x,y\in\mathbb{R}^d$ with $|x|\vee|y|\le R$, and $\mu\in\mathcal{P}_2(\mathbb{R}^d)$,

$$\|\sigma(t,x,\mu) - \sigma(t,y,\mu)\| \le L_R^{(2)}|x-y|.$$

(H3) (Global Lipschitz condition on the measure) There exists a positive constant L such that for any $t\in[0,T]$, $x\in\mathbb{R}^d$, and $\mu,\nu\in\mathcal{P}_2(\mathbb{R}^d)$,

$$|b(t,x,\mu) - b(t,x,\nu)| \vee \|\sigma(t,x,\mu) - \sigma(t,x,\nu)\| \le LW_2(\mu,\nu).$$

(H4) (Joint continuity of the state and measure) For any $t \in [0,T]$, $b(t,\cdot,\cdot)$ is continuous on $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$.

(H5) (Linear growth condition on the state and measure) There exists a positive constant K_1 such that for any $t \in [0,T]$, $x \in \mathbb{R}^d$, and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$,

$$|b(t,x,\mu)| \vee \|\sigma(t,x,\mu)\| \leq K_1(1+|x|+W_2(\mu,\delta_0)),$$

where δ_0 denotes the Dirac measure at 0.

It was mentioned in the Introduction that the local conditions w.r.t. the state variable and the distribution-dependent coefficients may create trouble in the investigation of McKean-Vlasov SDEs. To overcome the difficulties, we utilize interpolated Euler-like sequence and partition of sample space to establish the existence and uniqueness of the solution to equation (1.1).

Theorem 2.2. *Let Assumptions* (H1)–(H5) *hold. Then, for any $\mathcal{F}_0$-measurable random variable ξ satisfying $\mathbb{E}|\xi|^\beta < \infty$ for any $\beta > 0$, equation* (1.1) *has a unique solution $X(t) \in L^2(\Omega;\mathbb{R}^d)$ for $t \in [0,T]$ with the initial value $X(0) = \xi$. Moreover, for any $p > 0$, this solution satisfies*

$$\mathbb{E}\Big[\sup_{0\leq t\leq T}|X(t)|^p\Big] < \infty. \tag{2.1}$$

We shall prove Theorem 2.2 by using interpolated Euler-like sequence in Section 3. One can also refer to [33] for another method.

Due to distribution-dependence in (1.1), we consider its approximation by stochastic interacting particle system (1.2). For $N \geq 1$ and $i = 1,2,\ldots,N$, let $(W^i, X^i(0))_{1\leq i\leq N}$ be independent copies of $(W, X(0))$. Consider the following non-interacting particle system associated with (1.1)

$$dX^i(t) = b(t,X^i(t),\mathscr{L}(X^i(t)))dt + \sigma(t,X^i(t),\mathscr{L}(X^i(t)))dW^i(t), \quad i=1,2,\ldots,N, \tag{2.2}$$

with the initial condition $X^i(0)$. According to Theorem 2.2, one has $\mathscr{L}(X^i(t)) = \mathscr{L}(X(t))$, $i = 1,2,\ldots,N$. Furthermore, let us define $[t]_n = t^n_k$ for all $t \in [t^n_k, t^n_{k+1})$ and $k = 0,1,\ldots,n-1$, then the continuous-time version of EM scheme (1.3) can be written as,

$$\begin{aligned} X^{i,N,n}(t) = X^{i,N,n}(0) \quad &+ \int_0^t b([s]_n, X^{i,N,n}([s]_n), \mu^{X,N,n}_{[s]_n})ds \\ &+ \int_0^t \sigma([s]_n, X^{i,N,n}([s]_n), \mu^{X,N,n}_{[s]_n})dW^i(s), \quad i=1,2,\ldots,N. \end{aligned} \tag{2.3}$$

For the EM scheme (2.3), it is necessary to estimate the difference of the functions $b(\cdot,x,\mu), \sigma(\cdot,x,\mu)$ at different times. To this end, we assume the Hölder continuity of b and σ with respect to t.

(H6) (Hölder continuity in time with exponent α) There exist constants $K_2 > 0$, $\rho > 1$, and $\alpha \in (0,1]$ such that for any $x \in \mathbb{R}^d$, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, and $t_1, t_2 \in [0,T]$,

$$\begin{aligned} &|b(t_1,x,\mu) - b(t_2,x,\mu)| \vee \|\sigma(t_1,x,\mu) - \sigma(t_2,x,\mu)\| \\ &\leq K_2(1+|x|^\rho + W_2(\mu,\delta_0)^\rho)|t_1-t_2|^\alpha. \end{aligned}$$

Theorem 2.3. *Let Assumptions* (H1)–(H6) *hold. Then, the EM scheme* (2.3) *converges to the non-interacting particle system* (2.2)*, that is,*

$$\lim_{N\to\infty}\lim_{n\to\infty}\sup_{1\leq i\leq N}\mathbb{E}\Big[\sup_{0\leq t\leq T}|X^i(t) - X^{i,N,n}(t)|^2\Big] = 0.$$

The proof of Theorem 2.3 is in Section 4.

Remark 2.4. For the classical SDEs whose coefficients are independent of distribution, under the global Lipschitz condition, the strong convergence rate of the EM scheme is $O((1/n)^{1/2})$; see [19,26]. Under a local Lipschitz condition, in general, the strong convergence can be obtained and the convergence rate requires additional conditions [13]. When the Lipschitz constants are of the order $O(\log R)$ for the drift coefficient and $O(\sqrt{\log R})$ for the diffusion coefficient, [41] obtained that the same convergence rate as the global Lipschitz condition. However, for McKean-Vlasov SDEs, we need to approximate the distribution $\mathscr{L}(X(t))$ for each $t \geq 0$ by the empirical measure. To do this, the interacting particle system (1.2) is used as a bridge. Because the use of the local conditions w.r.t. the state variable, both the propagation of chaos and the EM scheme associated with particle system (1.2) cannot be used to prove the rate of convergence. Thus the convergence rate of the numerical algorithms cannot be obtained. On the other hand, even without using the stochastic interacting particle system as a bridge, making use of the time-discretization about the McKean-Vlasov SDE, we still cannot prove the rate of convergence by using the same method as [41], because the restriction of the stopping time arguments.

Remark 2.5. When the drift coefficient b and the diffusion coefficient σ satisfy the global conditions w.r.t. the state variable, the strong convergence and the convergence rate were obtained with respect to N and step size for McKean-Vlasov SDEs in [2,4,14,35]. In this paper, under the local conditions w.r.t. the state variable, Theorem 2.3 provides a strong convergence result. If the global conditions hold ($L_R^{(1)}$ and $L_R^{(2)}$ are independent of R), according to the proofs of Lemma 4.3, Lemma 4.4, and Theorem 2.3, we can obtain the convergence rate with respect to N and step size. In other words, our proofs recover the rate of convergence results under the global Lipschitz conditions w.r.t. the state variable.

3 Existence and Uniqueness of Solutions

To prove Theorem 2.2, we use interpolated Euler-like sequence (only with respect to distributions, it is different from the classical EM scheme) with equidistant partitions of $[0, T]$. Then we show that this Euler-like sequence is Cauchy in $L^2(\Omega; C([0, T]; \mathbb{R}^d))$. The completeness of $L^2(\Omega; C([0, T]; \mathbb{R}^d))$ enables us to conclude that there is an $X : \Omega \to C([0, T]; \mathbb{R}^d)$, which is indeed the desired solution to (1.1).

For any integer $n \geq 1$, recall that $h_n = \frac{T}{n}$, $t_k^n = kh_n$ for $k = 0, 1, \ldots, n$. Let $X^{(n)}(0) = \xi$. In what follows, we define $X^{(n)}(t)$ step-by-step on the intervals $[0, t_1^n]$, $(t_1^n, t_2^n]$, $\ldots$, $(t_{n-1}^n, T]$. First, for $t \in [0, t_1^n]$, we consider the following classical SDE

$$dX^{(n)}(t) = b(t, X^{(n)}(t), \mu_0^{(n)})dt + \sigma(t, X^{(n)}(t), \mu_0^{(n)})dW(t),\ X^{(n)}(0) = \xi \in L^2(\Omega; \mathbb{R}^d), \tag{3.1}$$

where $\mu_0^{(n)} = \mathscr{L}(X^{(n)}(0)) = \mathscr{L}(\xi)$. We start by showing the existence and uniqueness of the solution to (3.1). Note that Assumptions (H1) and (H2) imply that

$$\langle x - y, b(t, x, \mu_0^{(n)}) - b(t, y, \mu_0^{(n)}) \rangle \leq L_R^{(1)} |x - y|^2$$

and

$$\|\sigma(t, x, \mu_0^{(n)}) - \sigma(t, y, \mu_0^{(n)})\| \leq L_R^{(2)} |x - y|,$$

for any $t \in [0, t_1^n]$, x, $y \in \mathbb{R}^d$ with $|x| \vee |y| \le R$, and that Assumption (H5) gives

$$\begin{aligned} |b(t, x, \mu_0^{(n)})| \vee \|\sigma(t, x, \mu_0^{(n)})\| &\le K_1(1 + |x| + W_2(\mu_0^{(n)}, \delta_0)) \\ &\le K_1\big[1 + (\mathbb{E}|X^{(n)}(0)|^2)^{\frac{1}{2}}\big](1 + |x|), \end{aligned}$$

for any $t \in [0, t_1^n]$ and $x \in \mathbb{R}^d$. These, together with the existence and uniqueness of the solution to the classical SDEs, imply that (3.1) has a unique solution on $[0, t_1^n]$; see [31, Theorem 3.1.1, p.44] for details. Moreover, by virtue of the linear growth condition, for any $p > 0$, there is a positive constant C such that

$$\mathbb{E}\Big[\sup_{0 \le t \le t_1^n} |X^{(n)}(t)|^p\Big] \le C(1 + \mathbb{E}|X^{(n)}(0)|^p). \tag{3.2}$$

For the proof of estimate (3.2), see [26, Theorem 4.4, p.61] for details. Next, for $t \in [t_1^n, t_2^n]$, we consider the following classical SDE

$$dX^{(n)}(t) = b(t, X^{(n)}(t), \mu_{t_1^n}^{(n)})dt + \sigma(t, X^{(n)}(t), \mu_{t_1^n}^{(n)})dW(t), \quad X^{(n)}(t_1^n) \in L^2(\Omega; \mathbb{R}^d), \tag{3.3}$$

where $\mu_{t_1^n}^{(n)} = \mathscr{L}(X^{(n)}(t_1^n))$. According to Assumptions (H1), (H2), and (H5) as well as estimate (3.2), the existence and uniqueness of the solution to (3.3) can be obtained by repeating the above procedure and using $(X^{(n)}(t_1^n), \mu_{t_1^n}^{(n)})$ in place of $(X^{(n)}(0), \mu_0^{(n)})$. Besides, for any $p > 0$,

$$\mathbb{E}\Big[\sup_{t_1^n \le t \le t_2^n} |X^{(n)}(t)|^p\Big] \le C\big(1 + \mathbb{E}|X^{(n)}(t_1^n)|^p\big).$$

Therefore, we can define $X^{(n)} : [0, T] \to L^2(\Omega; \mathbb{R}^d)$ inductively. In addition, for any $k = 0, 1, \dots, n-1$ and $p > 0$, we have

$$\mathbb{E}\Big[\sup_{t_k^n \le t \le t_{k+1}^n} |X^{(n)}(t)|^p\Big] \le C\big(1 + \mathbb{E}|X^{(n)}(t_k^n)|^p\big).$$

Moreover, for any $n \ge 1$ and $p > 0$, one has

$$\begin{aligned} \mathbb{E}\Big[\sup_{0 \le t \le T} |X^{(n)}(t)|^p\Big] &= \mathbb{E}\Big[\max_{0 \le k \le n-1} \sup_{t_k^n \le t \le t_{k+1}^n} |X^{(n)}(t)|^p\Big] \\ &\le \sum_{k=0}^{n-1} \mathbb{E}\Big[\sup_{t_k^n \le t \le t_{k+1}^n} |X^{(n)}(t)|^p\Big] < \infty, \end{aligned} \tag{3.4}$$

which implies that $X^{(n)} \in L^2(\Omega; C([0, T]; \mathbb{R}^d))$. It is worthwhile to mention that the bound in (3.4) exists but may depend on n. However, for our purpose, it is necessary to show that the bound in (3.4) is independent of n. To this end, recall that $[t]_n = t_k^n$ for all $t \in [t_k^n, t_{k+1}^n)$ and $k = 0, 1, \dots, n-1$, then for any $t \in [0, T]$,

$$X^{(n)}(t) = \xi + \int_0^t b(s, X^{(n)}(s), \mu_{[s]_n}^{(n)})ds + \int_0^t \sigma(s, X^{(n)}(s), \mu_{[s]_n}^{(n)})dW(s). \tag{3.5}$$

By exploiting this type of expression, we can prove the following lemma which indicates that the bound in (3.4) is independent of n.

Lemma 3.1. *Let Assumptions* (H1), (H2), *and* (H5) *hold. Then, for any $p > 0$, there exists a positive constant C_p, which is independent of n, such that*

$$\mathbb{E}\Big[\sup_{0 \le t \le T} |X^{(n)}(t)|^p\Big] \le C_p.$$

Lemma 3.2. *Assume that Assumptions* (H1), (H2), *and* (H5) *hold. Then, for any $p \geq 2$, and $s,t \in [0,T]$,*

$$\sup_{n\geq 1} \mathbb{E}|X^{(n)}(t) - X^{(n)}(s)|^p \leq C_p|t-s|^{\frac{p}{2}},$$

where C_p is independent of n.

The proofs of the above two lemmas are relegated to the appendix. With these two lemmas at hand, we proceed to prove an important lemma which indicates that $\{X^{(n)}\}_{n\geq 1}$ is a Cauchy sequence in $L^2(\Omega; C([0,T];\mathbb{R}^d))$.

Lemma 3.3. *Let Assumptions* (H1)–(H3), *and* (H5) *hold. The interpolated Euler-like sequence $\{X^{(n)}\}_{n\geq 1}$ is Cauchy in $L^2(\Omega; C([0,T];\mathbb{R}^d))$, that is,*

$$\|X^{(n)} - X^{(m)}\|_{L^2} = \Big(\mathbb{E}\Big[\sup_{0\leq t\leq T} |X^{(n)}(t) - X^{(m)}(t)|^2\Big]\Big)^{\frac{1}{2}} \to 0, \ \ \textit{as } n,m \to \infty.$$

Proof. For any $n,m \geq 1$, by (3.5) and using the Itô formula, we arrive at

$$|X^{(n)}(t) - X^{(m)}(t)|^2 = I_1(t) + I_2(t) + I_3(t), \tag{3.6}$$

where

$$\begin{aligned}
I_1(t) &= \int_0^t 2\langle X^{(n)}(s) - X^{(m)}(s), b(s, X^{(n)}(s), \mu^{(n)}_{[s]_n}) - b(s, X^{(m)}(s), \mu^{(m)}_{[s]_m})\rangle ds,\\
I_2(t) &= \int_0^t \|\sigma(s, X^{(n)}(s), \mu^{(n)}_{[s]_n}) - \sigma(s, X^{(m)}(s), \mu^{(m)}_{[s]_m})\|^2 ds,\\
I_3(t) &= \int_0^t 2\langle X^{(n)}(s) - X^{(m)}(s), \sigma(s, X^{(n)}(s), \mu^{(n)}_{[s]_n}) - \sigma(s, X^{(m)}(s), \mu^{(m)}_{[s]_m}) dW(s)\rangle.
\end{aligned}$$

For a sufficiently large $R > 0$, define

$$\Omega^{n,m}(R) = \Big\{\omega \in \Omega : \sup_{0\leq t\leq T} |X^{(n)}(t)| \vee \sup_{0\leq t\leq T} |X^{(m)}(t)| \geq R\Big\}.$$

Consequently, by the Chebyshev inequality and Lemma 3.1, we get that for any $q > 0$,

$$\mathbb{P}(\Omega^{n,m}(R)) \leq \frac{1}{R^q}\Big(\mathbb{E}\Big[\sup_{0\leq t\leq T} |X^{(n)}(t)|^q\Big] + \mathbb{E}\Big[\sup_{0\leq t\leq T} |X^{(m)}(t)|^q\Big]\Big) \leq \frac{2C_q}{R^q}, \tag{3.7}$$

where $R > 0$ and $q > 0$ are to be chosen later. Moreover, it follows from (3.7) that the estimate of $\mathbb{P}(\Omega^{n,m}(R))$ is independent of n and m. Therefore, the selected R and q are also independent of n and m. For any $t \in [0,T]$, by Assumptions (H1), (H3), and (H5), the Hölder inequality, we obtain

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{0\leq s\leq t} I_1(s)\Big]\\
&\leq 2\mathbb{E}\Big[\Big(\sup_{0\leq s\leq t}\int_0^s \langle X^{(n)}(r) - X^{(m)}(r), b(r, X^{(n)}(r), \mu^{(n)}_{[r]_n})\\
&\qquad - b(r, X^{(m)}(r), \mu^{(n)}_{[r]_n})\rangle dr\Big) \cdot I_{\Omega\setminus\Omega^{n,m}(R)}\Big]\\
&\quad + 2\mathbb{E}\Big[\Big(\sup_{0\leq s\leq t}\int_0^s \langle X^{(n)}(r) - X^{(m)}(r), b(r, X^{(n)}(r), \mu^{(n)}_{[r]_n})
\end{aligned}$$

$$
\begin{aligned}
&\qquad -b(r, X^{(m)}(r), \mu^{(n)}_{[r]_n})\rangle dr\Big) \cdot I_{\Omega^{n,m}(R)}\Big] \\
&\quad +2\mathbb{E}\Big[\int_0^t |X^{(n)}(s) - X^{(m)}(s)||b(s, X^{(m)}(s), \mu^{(n)}_{[s]_n}) - b(s, X^{(m)}(s), \mu^{(m)}_{[s]_m})|ds\Big] \\
&\le (2L_R^{(1)} + 2K_1 + L) \cdot \int_0^t \mathbb{E}|X^{(n)}(s) - X^{(m)}(s)|^2 ds \\
&\quad +8K_1\mathbb{E}\Big[\int_0^t \big(1 + |X^{(n)}(s)|^2 + |X^{(m)}(s)|^2 + W_2(\mu^{(n)}_{[s]_n}, \delta_0)^2\big)ds \cdot I_{\Omega^{n,m}(R)}\Big] \\
&\quad +3L\int_0^t \big(W_2(\mu^{(n)}_{[s]_n}, \mu^{(n)}_s)^2 + W_2(\mu^{(n)}_s, \mu^{(m)}_s)^2 + W_2(\mu^{(m)}_s, \mu^{(m)}_{[s]_m})^2\big)ds \\
&\le 2(L_R^{(1)} + K_1 + 2L) \cdot \int_0^t \mathbb{E}|X^{(n)}(s) - X^{(m)}(s)|^2 ds \\
&\quad +8TK_1\Big(\mathbb{E}\Big[\sup_{0\le s\le T}\big(1 + |X^{(n)}(s)|^2 + |X^{(m)}(s)|^2 + W_2(\mu^{(n)}_s, \delta_0)^2\big)^2\Big]\Big)^{\frac{1}{2}}\big(\mathbb{P}(\Omega^{n,m}(R))\big)^{\frac{1}{2}} \\
&\quad +3L\int_0^t \big(\mathbb{E}|X^{(n)}(s) - X^{(n)}([s]_n)|^2 + \mathbb{E}|X^{(m)}(s) - X^{(m)}([s]_m)|^2\big)ds \\
&\le 2(L_R^{(1)} + K_1 + 2L) \cdot \int_0^t \mathbb{E}|X^{(n)}(s) - X^{(m)}(s)|^2 ds + 16TK_1\sqrt{1+3C}\big(\mathbb{P}(\Omega^{n,m}(R))\big)^{\frac{1}{2}} \\
&\quad +3L\int_0^t \big(\mathbb{E}|X^{(n)}(s) - X^{(n)}([s]_n)|^2 + \mathbb{E}|X^{(m)}(s) - X^{(m)}([s]_m)|^2\big)ds. \qquad (3.8)
\end{aligned}
$$

By Assumptions (H2), (H3), and (H5), using the same technique as (3.8) was proved, we get

$$
\begin{aligned}
\mathbb{E}\Big[\sup_{0\le s\le t} I_2(s)\Big] &\le \mathbb{E}\Big[\int_0^t \|\sigma(s, X^{(n)}(s), \mu^{(n)}_{[s]_n}) - \sigma(s, X^{(m)}(s), \mu^{(m)}_{[s]_m})\|^2 ds\Big] \\
&\le 2\big[(L_R^{(2)})^2 + 3L^2\big] \cdot \int_0^t \mathbb{E}|X^{(n)}(s) - X^{(m)}(s)|^2 ds \\
&\quad +48TK_1^2\sqrt{1+3C}\big(\mathbb{P}(\Omega^{n,m}(R))\big)^{\frac{1}{2}} \\
&\quad +6L^2\int_0^t \big(\mathbb{E}|X^{(n)}(s) - X^{(n)}([s]_n)|^2 + \mathbb{E}|X^{(m)}(s) - X^{(m)}([s]_m)|^2\big)ds. \qquad (3.9)
\end{aligned}
$$

Exploiting the Burkholder-Davis-Gundy inequality, the Young inequality, and (3.9), we then obtain

$$
\begin{aligned}
&\mathbb{E}\Big[\sup_{0\le s\le t} I_3(s)\Big] \\
&\le 6\mathbb{E}\Big(\int_0^t |X^{(n)}(s) - X^{(m)}(s)|^2 \cdot \|\sigma(s, X^{(n)}(s), \mu^{(n)}_{[s]_n}) - \sigma(s, X^{(m)}(s), \mu^{(m)}_{[s]_m})\|^2 ds\Big)^{\frac{1}{2}} \\
&\le \frac{1}{2}\mathbb{E}\Big[\sup_{0\le s\le t}|X^{(n)}(s) - X^{(m)}(s)|^2\Big] \\
&\quad +36\big[(L_R^{(2)})^2 + 3L^2\big] \cdot \int_0^t \mathbb{E}|X^{(n)}(s) - X^{(m)}(s)|^2 ds + 864TK_1^2\sqrt{1+3C}\big(\mathbb{P}(\Omega^{n,m}(R))\big)^{\frac{1}{2}} \\
&\quad +108L^2\int_0^t \big(\mathbb{E}|X^{(n)}(s) - X^{(n)}([s]_n)|^2 + \mathbb{E}|X^{(m)}(s) - X^{(m)}([s]_m)|^2\big)ds. \qquad (3.10)
\end{aligned}
$$

Substituting (3.8)–(3.10) into (3.6) yields that

$$
\mathbb{E}\Big[\sup_{0\le s\le t}|X^{(n)}(s) - X^{(m)}(s)|^2\Big]
$$

$$
\begin{aligned}
&\le \big\{4(L_R^{(1)}+K_1+2L)+76\big[\big(L_R^{(2)}\big)^2+3L^2\big]\big\}\cdot\int_0^t \mathbb{E}|X^{(n)}(s)-X^{(m)}(s)|^2ds\\
&\quad +32TK_1(1+57K_1)\sqrt{1+3C}\big(\mathbb{P}(\Omega^{n,m}(R))\big)^{\frac{1}{2}}\\
&\quad +6L(1+38L)\int_0^t(\mathbb{E}|X^{(n)}(s)-X^{(n)}([s]_n)|^2+\mathbb{E}|X^{(m)}(s)-X^{(m)}([s]_m)|^2)ds. \qquad (3.11)
\end{aligned}
$$

In addition, for any $s\in[0,T]$, the result in Lemma 3.2 implies that

$$
\mathbb{E}|X^{(n)}(s)-X^{(n)}([s]_n)|^2\le Ch_n \quad \text{and} \quad \mathbb{E}|X^{(m)}(s)-X^{(m)}([s]_m)|^2\le Ch_m.
$$

These, together with (3.11), imply

$$
\begin{aligned}
&\mathbb{E}\Big[\sup_{0\le s\le t}|X^{(n)}(s)-X^{(m)}(s)|^2\Big]\\
&\le \big\{4(L_R^{(1)}+K_1+2L)+76\big[\big(L_R^{(2)}\big)^2+3L^2\big]\big\}\int_0^t \mathbb{E}|X^{(n)}(s)-X^{(m)}(s)|^2ds\\
&\quad +32TK_1(1+57K_1)\sqrt{1+3C}\big(\mathbb{P}(\Omega^{n,m}(R))\big)^{\frac{1}{2}}+6CTL(1+38L)(h_n+h_m).
\end{aligned}
$$

Using the Grönwall inequality and noting $L_R^{(1)}\le\alpha_1\log R$, $L_R^{(2)}\le\sqrt{\alpha_2\log R}$, and (3.7), we derive that

$$
\begin{aligned}
&\mathbb{E}\Big[\sup_{0\le s\le T}|X^{(n)}(s)-X^{(m)}(s)|^2\Big]\\
&\le 32TK_1(1+57K_1)\sqrt{1+3C}e^{4(K_1+2L)T+228L^2T}\frac{\sqrt{2C_q}}{R^{\frac{q}{2}}}R^{4(\alpha_1+19\alpha_2)T}\\
&\quad +6CTL(1+38L)e^{4(K_1+2L)T+228L^2T}R^{4(\alpha_1+19\alpha_2)T}(h_n+h_m)\\
&\le C\frac{\sqrt{C_q}}{R^{\frac{q}{2}}}R^{4(\alpha_1+19\alpha_2)T}+CR^{4(\alpha_1+19\alpha_2)T}(h_n+h_m). \qquad (3.12)
\end{aligned}
$$

We note that R is independent of n and m, h_n and h_m converge to 0 as $n,m\to\infty$. Letting $n,m\to\infty$, it follows from (3.12) that

$$
\lim_{n,m\to\infty}\mathbb{E}\Big[\sup_{0\le s\le T}|X^{(n)}(s)-X^{(m)}(s)|^2\Big]\le C\frac{\sqrt{C_q}}{R^{\frac{q}{2}}}R^{4(\alpha_1+19\alpha_2)T}. \qquad (3.13)
$$

Taking $q>0$ sufficiently large for $q>8(\alpha_1+19\alpha_2)T$, the right-hand side of (3.13) converges to 0 as $R\to\infty$. That is, for any $\varepsilon>0$, we can choose $R:=R(\varepsilon)>0$ sufficiently large, such that the right-hand side of (3.13) is less than ε. Therefore, the arbitrariness of ε implies that

$$
\lim_{n,m\to\infty}\mathbb{E}\Big[\sup_{0\le s\le T}|X^{(n)}(s)-X^{(m)}(s)|^2\Big]=0,
$$

which completes the proof. □

With these lemmas at hand, we can proceed to prove Theorem 2.2.

Proof of Theorem 2.2. We divide it into the following three steps.

Step 1: Existence. Note that $L^2(\Omega; C([0,T];\mathbb{R}^d))$ is complete. According to Lemma 3.3, there exists an $\{\mathcal{F}_t\}$-adapted continuous process $(X(t))_{0\le t\le T}$ such that

$$\mathbb{E}\Big[\sup_{0\le t\le T}|X^{(n)}(t)-X(t)|^2\Big]\to 0,\ \text{as } n\to\infty. \tag{3.14}$$

Note that $X^{(n)}(0)=\xi$ for any $n\ge 1$. Then, it follows from (3.14) that $X(0)=\xi$, $\mathbb{P}$-a.s. and

$$\mathbb{E}\Big[\sup_{0\le t\le T}|X(t)|^2\Big]\le C. \tag{3.15}$$

To prove that $(X(t))_{0\le t\le T}$ satisfies equation (1.1), we take the limit in (3.5). By the path continuity and (3.14), we need only prove that the right-hand side of (3.5) converges in probability to

$$\xi+\int_0^t b(s,X(s),\mu_s)ds+\int_0^t\sigma(s,X(s),\mu_s)dW(s),$$

in which $\mu_s=\mathscr{L}(X(s))$ for any $s\in[0,T]$. On the one hand, by Lemma 3.2 and (3.14), we have

$$\begin{aligned}
&\lim_{n\to\infty}\sup_{0\le s\le t}W_2(\mu^{(n)}_{[s]_n},\mu_s)^2\\
&\le 2\lim_{n\to\infty}\sup_{0\le s\le T}\mathbb{E}|X^{(n)}(s)-X^{(n)}([s]_n)|^2+2\lim_{n\to\infty}\mathbb{E}\Big[\sup_{0\le s\le T}|X^{(n)}(s)-X(s)|^2\Big]\\
&\le 2C\lim_{n\to\infty}h_n=0.
\end{aligned}$$

On the other hand, for any $s\in[0,T]$,

$$\lim_{n\to\infty}\mathbb{E}|X^{(n)}(s)-X(s)|^2\le\lim_{n\to\infty}\mathbb{E}\Big[\sup_{0\le s\le T}|X^{(n)}(s)-X(s)|^2\Big]=0,$$

which implies that there exists a subsequence (still indexed by n for notational simplicity) such that for any $s\in[0,T]$ and almost all $\omega\in\Omega$,

$$X^{(n)}(s,\omega)\to X(s,\omega),\ \text{as } n\to\infty.$$

Hence, for any $s\in[0,T]$, by Assumption (H4) and the joint continuity of $\sigma(s,\cdot,\cdot)$, we can derive that for any $s\in[0,T]$ and almost all $\omega\in\Omega$,

$$b(s,X^{(n)}(s,\omega),\mu^{(n)}_{[s]_n})\to b(s,X(s,\omega),\mu_s),\ \text{as } n\to\infty \tag{3.16}$$

and

$$\sigma(s,X^{(n)}(s,\omega),\mu^{(n)}_{[s]_n})\to\sigma(s,X(s,\omega),\mu_s),\ \text{as } n\to\infty. \tag{3.17}$$

Furthermore, according to Assumption (H5) and Lemma 3.1, for any $A\in\mathcal{F}$,

$$\begin{aligned}
\sup_{n\ge1}\mathbb{E}\big[|b(s,X^{(n)}(s),\mu^{(n)}_{[s]_n})|\cdot I_A\big] &\le K_1\sup_{n\ge1}\mathbb{E}\big[(1+|X^{(n)}(s)|+W_2(\mu^{(n)}_{[s]_n},\delta_0))\cdot I_A\big]\\
&\le K_1\sup_{n\ge1}\big[\mathbb{E}(1+|X^{(n)}(s)|+W_2(\mu^{(n)}_{[s]_n},\delta_0))^2\big]^{\frac12}\big(\mathbb{P}(A)\big)^{\frac12}\\
&\le K_1\sqrt{3(1+2C)}\big(\mathbb{P}(A)\big)^{\frac12}
\end{aligned}\tag{3.18}$$

and

$$\sup_{n\geq 1}\mathbb{E}\big[\|\sigma(s,X^{(n)}(s),\mu_{[s]_n}^{(n)})\|^2\cdot I_A\big]\leq 3K_1^2\sqrt{3(1+2C)}\big(\mathbb{P}(A)\big)^{\frac{1}{2}}. \tag{3.19}$$

Consequently, by the dominated convergence (see [8, Theorem 4.5.4, p.101] or [37, Theorem 4, p.188]) combined with (3.16) and (3.18), one obtains that for any $s\in[0,T]$,

$$\lim_{n\to\infty}\mathbb{E}|b(s,X^{(n)}(s),\mu_{[s]_n}^{(n)})-b(s,X(s),\mu_s)|=0.$$

Similarly, in view of (3.17) and (3.19), we get

$$\lim_{n\to\infty}\mathbb{E}\|\sigma(s,X^{(n)}(s),\mu_{[s]_n}^{(n)})-\sigma(s,X(s),\mu_s)\|^2=0.$$

Next, for any $t\in[0,T]$, Assumption (H5), together with (3.15) and Lemma 3.1, implies that

$$\begin{aligned}
&\sup_{n\geq 1}\sup_{s\in[0,t]}\mathbb{E}\big[|b(s,X^{(n)}(s),\mu_{[s]_n}^{(n)})-b(s,X(s),\mu_s)|+\|\sigma(s,X^{(n)}(s),\mu_{[s]_n}^{(n)})-\sigma(s,X(s),\mu_s)\|^2\big]\\
&\quad\leq K_1\sup_{n\geq 1}\sup_{s\in[0,t]}\mathbb{E}[2+|X^{(n)}(s)|+|X(s)|+W_2(\mu_{[s]_n}^{(n)},\delta_0)+W_2(\mu_s,\delta_0)]\\
&\qquad+6K_1^2\sup_{n\geq 1}\sup_{s\in[0,t]}\mathbb{E}[2+|X^{(n)}(s)|^2+|X(s)|^2+W_2(\mu_{[s]_n}^{(n)},\delta_0)^2+W_2(\mu_s,\delta_0)^2]\\
&\quad\leq 2K_1\sqrt{5(1+C)}+12K_1^2(1+2C)\leq C.
\end{aligned}$$

It suffices to show that

$$\Big([0,t]\ni s\mapsto\mathbb{E}|b(s,X^{(n)}(s),\mu_{[s]_n}^{(n)})-b(s,X(s),\mu_s)|\Big)_{n\geq 1}$$

and

$$\Big([0,t]\ni s\mapsto\mathbb{E}\|\sigma(s,X^{(n)}(s),\mu_{[s]_n}^{(n)})-\sigma(s,X(s),\mu_s)\|^2\Big)_{n\geq 1}$$

are both uniformly integrable on $[0,t]$. Moreover, by virtue of the dominated convergence (see [8, Theorem 4.5.4, p.101] or [37, Theorem 4, p.188]), we arrive at

$$\begin{aligned}
&\lim_{n\to\infty}\mathbb{E}\Big|\int_0^t(b(s,X^{(n)}(s),\mu_{[s]_n}^{(n)})-b(s,X(s),\mu_s))ds\Big|\\
&\quad\leq\lim_{n\to\infty}\int_0^t\mathbb{E}|b(s,X^{(n)}(s),\mu_{[s]_n}^{(n)})-b(s,X(s),\mu_s)|ds=0,
\end{aligned}$$

and

$$\begin{aligned}
&\lim_{n\to\infty}\mathbb{E}\Big|\int_0^t(\sigma(s,X^{(n)}(s),\mu_{[s]_n}^{(n)})-\sigma(s,X(s),\mu_s))dW(s)\Big|^2\\
&\quad=\lim_{n\to\infty}\int_0^t\mathbb{E}\|\sigma(s,X^{(n)}(s),\mu_{[s]_n}^{(n)})-\sigma(s,X(s),\mu_s)\|^2ds=0.
\end{aligned}$$

Therefore, by Definition 2.1 and (3.15), $X(t)\in L^2(\Omega;\mathbb{R}^d)$ is a solution to (1.1) for $t\in[0,T]$.

Step 2: Estimate (2.1). Let $X(t)\in L^2(\Omega;\mathbb{R}^d)$ be any solution to equation (1.1) for $t\in[0,T]$. Then we note that $\mathbb{E}|X(t)|^2<\infty$ for any $t\in[0,T]$. By the Lyapunov inequality, we prove that for any $p\geq 2$,

$$\mathbb{E}\Big[\sup_{0\leq t\leq T}|X(t)|^p\Big]<\infty.$$

For our goal, it is imperative to prove that the bound of $\mathbb{E}|X(t)|^2$ is independent of t. Therefore, exploiting the Hölder inequality, one gets

$$|X(t)|^2 \leq 3\Big[|\xi|^2 + T\int_0^t |b(s,X(s),\mu_s)|^2 ds + \Big|\int_0^t \sigma(s,X(s),\mu_s)dW(s)\Big|^2\Big].$$

For any $t \in [0,T]$, by the Burkholder-Davis-Gundy inequality, and then using Assumption (H5), we can calculate that

$$\begin{aligned}\mathbb{E}|X(t)|^2 &\leq 3\mathbb{E}|\xi|^2 + 9K_1^2(T+4)\mathbb{E}\Big[\int_0^t \big(1+|X(s)|^2 + W_2(\mu_s,\delta_0)^2\big)ds\Big] \\ &\leq 3\mathbb{E}|\xi|^2 + 2\cdot 9K_1^2(T+4)\int_0^t \big(1+\mathbb{E}|X(s)|^2\big)ds. \end{aligned} \tag{3.20}$$

Applying the Grönwall inequality to (3.20), we have

$$\sup_{0\leq t\leq T}\mathbb{E}|X(t)|^2 \leq (1+3\mathbb{E}|\xi|^2)e^{2\cdot 9K_1^2(T+4)T}.$$

In what follow, for any $p \geq 2$, we calculate the pth moment of the solution $(X(t))_{0\leq t\leq T}$. To this end, for every $R > 0$, define the stopping time

$$\tau_R = \inf\{t\in[0,T] : |X(t)| \geq R\} \wedge T.$$

Using similar technique as (3.20) was proved, we compute that for any $t \in [0,T]$,

$$\begin{aligned}&\mathbb{E}\Big[\sup_{0\leq s\leq t\wedge\tau_R}|X(s)|^p\Big] \\ &\leq 3^{p-1}\mathbb{E}|\xi|^p + 9^{p-1}K_1^p(T^{p-1}+M_pT^{\frac{p-2}{2}})\mathbb{E}\Big[\int_0^{t\wedge\tau_R}\big(1+|X(s)|^p + W_2(\mu_s,\delta_0)^p\big)ds\Big] \\ &\leq 3^{p-1}\mathbb{E}|\xi|^p + 9^{p-1}K_1^p(T^p+M_pT^{\frac{p}{2}})(1+3\mathbb{E}|\xi|^2)^{\frac{p}{2}}e^{9K_1^2(T+4)Tp} \\ &\quad +9^{p-1}K_1^p(T^{p-1}+M_pT^{\frac{p-2}{2}})\int_0^t\Big(1+\mathbb{E}\Big[\sup_{0\leq r\leq s\wedge\tau_R}|X(r)|^p\Big]\Big)ds,\end{aligned}$$

where $M_p = [p^{p+1}/2(p-1)^{p-1}]^{\frac{p}{2}}$. Note that $\tau_R \uparrow T$ a.s. Then, employing the Grönwall inequality and the Fatou lemma, we arrive at

$$\mathbb{E}\Big[\sup_{0\leq s\leq T}|X(s)|^p\Big] \leq \liminf_{R\to\infty}\mathbb{E}\Big[\sup_{0\leq s\leq T\wedge\tau_R}|X(s)|^p\Big] \leq C_p < \infty.$$

Therefore, the required assertion follows.

Step 3: Uniqueness. Assume that $(X(t))_{0\leq t\leq T}$ and $(\bar{X}(t))_{0\leq t\leq T}$ are two solutions to (1.1). Then, by (2.1), for any $p > 0$, there exists positive constant C_p such that

$$\mathbb{E}\Big[\sup_{0\leq t\leq T}|X(t)|^p\Big] \leq C_p, \quad \mathbb{E}\Big[\sup_{0\leq t\leq T}|\bar{X}(t)|^p\Big] \leq C_p. \tag{3.21}$$

First, for a sufficiently large $R > 0$, define

$$\Omega(R) = \Big\{\omega\in\Omega : \sup_{0\leq t\leq T}|X(t)| \vee \sup_{0\leq t\leq T}|\bar{X}(t)| \geq R\Big\}.$$

According to the Chebyshev inequality and (3.21), we can get that for any $q > 0$,

$$\mathbb{P}(\Omega(R)) \leq \frac{1}{R^q}\Big(\mathbb{E}\Big[\sup_{0\leq t\leq T}|X(t)|^q\Big] + \mathbb{E}\Big[\sup_{0\leq t\leq T}|\bar{X}(t)|^q\Big]\Big) \leq \frac{2C_q}{R^q}, \tag{3.22}$$

where $R > 0$ and $q > 0$ are to be chosen later. Next, using the Itô formula, we have

$$\begin{aligned} |X(t) - \bar{X}(t)|^2 \quad &= \quad \int_0^t 2\langle X(s) - \bar{X}(s), b(s, X(s), \mu_s) - b(s, \bar{X}(s), \nu_s)\rangle ds \\ &\quad + \int_0^t \|\sigma(s, X(s), \mu_s) - \sigma(s, \bar{X}(s), \nu_s)\|^2 ds \\ &\quad + \int_0^t 2\langle X(s) - \bar{X}(s), \sigma(s, X(s), \mu_s) - \sigma(s, \bar{X}(s), \nu_s) dW(s)\rangle \\ &=: \quad J_1(t) + J_2(t) + J_3(t), \end{aligned} \tag{3.23}$$

where $\nu_s = \mathscr{L}(\bar{X}(s))$ for any $s \in [0, T]$. Consequently, the uniqueness can be obtained by applying the same technique as Lemma 3.3 was proved. This completes the proof. □

Remark 3.4. We have established the existence and uniqueness of the solution for equation (1.1). It will be interesting to examine such equations with random coefficients and an additional random switching process. According to our method, if we consider the random coefficients $b(t, x, \mu, \tilde{\omega})$, $\sigma(t, x, \mu, \tilde{\omega})$ (with $\tilde{\omega} \in \tilde{\Omega}$), then we can still obtain the strong solution to equation (1.1) for all $\tilde{\omega} \in \tilde{\Omega}$; see [29] and [30] for the related works.

4 Convergence of the EM Algorithm

In this section, let us finish the proof of Theorem 2.3. To proceed, we first need to give some auxiliary lemmas. The following lemma establishes the well-posedness of the stochastic interacting particle system (1.2).

Lemma 4.1. *Suppose that Assumptions* (H1)–(H3)*, and* (H5) *hold. Then, for any $N \geq 1$, there exists a unique solution $(X^{i,N}(t))_{0\leq t\leq T}$ with the initial value $X^i(0)$ to* (1.2)*.*

Proof. For $\mathbf{x} := (x_1^\top, x_2^\top, \ldots, x_N^\top)^\top \in \mathbb{R}^{dN}$, $t \in [0, T]$, set

$$B(t, \mathbf{x}) = \Big(b\Big(t, x_1, \frac{1}{N}\sum_{j=1}^N \delta_{x_j}\Big)^\top, b\Big(t, x_2, \frac{1}{N}\sum_{j=1}^N \delta_{x_j}\Big)^\top, \ldots, b\Big(t, x_N, \frac{1}{N}\sum_{j=1}^N \delta_{x_j}\Big)^\top\Big)^\top$$

and

$$\Sigma(t, \mathbf{x}) = \begin{pmatrix} \sigma\Big(t, x_1, \frac{1}{N}\sum_{j=1}^N \delta_{x_j}\Big) & 0 & \cdots & 0 \\ 0 & \sigma\Big(t, x_2, \frac{1}{N}\sum_{j=1}^N \delta_{x_j}\Big) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma\Big(t, x_N, \frac{1}{N}\sum_{j=1}^N \delta_{x_j}\Big) \end{pmatrix},$$

where 0 represents a $d \times d_1$ null matrix. Then, (1.2) can be rewritten as

$$d\mathbf{X}^N(t) = B(t, \mathbf{X}^N(t))dt + \Sigma(t, \mathbf{X}^N(t))d\mathbf{W}^N(t), \tag{4.1}$$

with the initial condition $\mathbf{X}^N(0) = (X^{1,N}(0)^\top, X^{2,N}(0)^\top, \ldots, X^{N,N}(0)^\top)^\top$, and $\mathbf{W}^N(t) = (W^1(t)^\top, W^2(t)^\top, \ldots, W^N(t)^\top)^\top$ is a d_1N-dimensional Brownian motion. Note that

$$W_2\Big(\frac{1}{N}\sum_{j=1}^N \delta_{x_j}, \frac{1}{N}\sum_{j=1}^N \delta_{y_j}\Big)^2 \le \frac{1}{N}\sum_{j=1}^N |x_j - y_j|^2.$$

Therefore, on the one hand, it follows from Assumptions (H1)–(H3) that for any $t \in [0,T]$, and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{dN}$ with $|\mathbf{x}| \vee |\mathbf{y}| \le R$,

$$\begin{aligned}
\langle \mathbf{x} - \mathbf{y}, B(t,\mathbf{x}) - B(t,\mathbf{y})\rangle &= \sum_{i=1}^N \Big\langle x_i - y_i, b\Big(t, x_i, \frac{1}{N}\sum_{j=1}^N \delta_{x_j}\Big) - b\Big(t, y_i, \frac{1}{N}\sum_{j=1}^N \delta_{y_j}\Big)\Big\rangle \\
&\le L_R^{(1)} \sum_{i=1}^N |x_i - y_i|^2 + L\sum_{i=1}^N |x_i - y_i| \cdot W_2\Big(\frac{1}{N}\sum_{j=1}^N \delta_{x_j}, \frac{1}{N}\sum_{j=1}^N \delta_{y_j}\Big) \\
&\le (L_R^{(1)} + L)|\mathbf{x} - \mathbf{y}|^2,
\end{aligned}$$

and

$$\begin{aligned}
\|\Sigma(t,\mathbf{x}) - \Sigma(t,\mathbf{y})\| &= \Big(\sum_{i=1}^N \Big\|\sigma\Big(t, x_i, \frac{1}{N}\sum_{j=1}^N \delta_{x_j}\Big) - \sigma\Big(t, y_i, \frac{1}{N}\sum_{j=1}^N \delta_{y_j}\Big)\Big\|^2\Big)^{\frac{1}{2}} \\
&\le \Big(2(L_R^{(2)})^2 \sum_{i=1}^N |x_i - y_i|^2 + 2L^2 \sum_{i=1}^N W_2\Big(\frac{1}{N}\sum_{j=1}^N \delta_{x_j}, \frac{1}{N}\sum_{j=1}^N \delta_{y_j}\Big)^2\Big)^{\frac{1}{2}} \\
&\le \Big(2(L_R^{(2)})^2 |\mathbf{x} - \mathbf{y}|^2 + 2L^2 \sum_{j=1}^N |x_j - y_j|^2\Big)^{\frac{1}{2}} \\
&\le \sqrt{2(L_R^{(2)})^2 + 2L^2} \cdot |\mathbf{x} - \mathbf{y}|.
\end{aligned}$$

On the other hand, Assumption (H5) implies that for any $t \in [0,T]$, and $\mathbf{x} \in \mathbb{R}^{dN}$,

$$\begin{aligned}
|B(t,\mathbf{x})|^2 \vee \|\Sigma(t,\mathbf{x})\|^2 &= \sum_{i=1}^N \Big|b\Big(t, x_i, \frac{1}{N}\sum_{j=1}^N \delta_{x_j}\Big)\Big|^2 \vee \sum_{i=1}^N \Big\|\sigma\Big(t, x_i, \frac{1}{N}\sum_{j=1}^N \delta_{x_j}\Big)\Big\|^2 \\
&\le K_1^2 \sum_{i=1}^N \Big(1 + |x_i| + W_2\Big(\frac{1}{N}\sum_{j=1}^N \delta_{x_j}, \delta_0\Big)\Big)^2 \\
&\le 3K_1^2 \sum_{i=1}^N \Big(1 + |x_i|^2 + \frac{1}{N}\sum_{j=1}^N |x_j|^2\Big) \\
&\le 6K_1^2 N(1 + |\mathbf{x}|^2).
\end{aligned} \tag{4.2}$$

These, together with the existence and uniqueness of the solution to the classical SDEs, imply that (4.1) has a unique solution on $[0,T]$ for any N; see [31, Theorem 3.1.1, p.44] for details. □

Note that the continuous-time EM scheme to (4.1) is given by

$$\mathbf{X}^{N,n}(t) = \mathbf{X}^{N,n}(0) + \int_0^t B([s]_n, \mathbf{X}^{N,n}([s]_n))ds + \int_0^t \Sigma([s]_n, \mathbf{X}^{N,n}([s]_n))d\mathbf{W}^N(s),$$

where $\mathbf{X}^{N,n}(0) = (X^{1,N,n}(0)^\top, X^{2,N,n}(0)^\top, \ldots, X^{N,N,n}(0)^\top)^\top$. By virtue of the linear growth condition (4.2), for any $p \geq 2$, there is a positive constant C_N such that

$$\mathbb{E}\Big[\sup_{0\leq t\leq T} |\mathbf{X}^{N,n}(t)|^p\Big] \leq C_N. \tag{4.3}$$

The following lemma reveals the pth moments of the stochastic interacting particle system and the continuous-time EM scheme, defined by (1.2) and (2.3) respectively, are uniformly bounded.

Lemma 4.2. *Suppose that Assumptions* (H1)–(H3), *and* (H5) *hold. Then for any $p \geq 2$, there exists a positive constant C_p, which is independent of n and N, such that*

$$\sup_{1\leq i\leq N} \mathbb{E}\Big[\sup_{0\leq t\leq T} |X^{i,N}(t)|^p\Big] \vee \sup_{1\leq i\leq N} \mathbb{E}\Big[\sup_{0\leq t\leq T} |X^{i,N,n}(t)|^p\Big] \leq C_p.$$

This lemma can be established by using similar technique as Lemma 3.1, so we omit the proof.

The lemma below is concerned with the propagation of chaos for McKean-Vlasov SDEs under the local conditions w.r.t. the state variable.

Lemma 4.3. *Under Assumptions* (H1)–(H5),

$$\lim_{N\to\infty} \sup_{1\leq i\leq N} \mathbb{E}\Big[\sup_{0\leq t\leq T} |X^i(t) - X^{i,N}(t)|^2\Big] = 0.$$

Proof. For any $1 \leq i \leq N$ and $R > 0$, define the stopping time

$$\tau_R^i = \inf\{t \in [0,T] : |X^i(t)| \vee |X^{i,N}(t)| \geq R\}.$$

Then, by (1.2), (2.2), and using the Itô formula, we arrive at

$$|X^i(t\wedge\tau_R^i) - X^{i,N}(t\wedge\tau_R^i)|^2 = I_{1,R}(t) + I_{2,R}(t) + I_{3,R}(t), \tag{4.4}$$

where

$$\begin{aligned}
I_{1,R}(t) &= 2\int_0^{t\wedge\tau_R^i} \langle X^i(s) - X^{i,N}(s), b(s, X^i(s), \mu_s^i) - b(s, X^{i,N}(s), \mu_s^{X,N})\rangle ds,\\
I_{2,R}(t) &= \int_0^{t\wedge\tau_R^i} \|\sigma(s, X^i(s), \mu_s^i) - \sigma(s, X^{i,N}(s), \mu_s^{X,N})\|^2 ds,\\
I_{3,R}(t) &= \int_0^{t\wedge\tau_R^i} 2\langle X^i(s) - X^{i,N}(s), \sigma(s, X^i(s), \mu_s^i) - \sigma(s, X^{i,N}(s), \mu_s^{X,N}) dW(s)\rangle.
\end{aligned}$$

where $\mu_s^i = \mathscr{L}(X^i(s))$ for all $s \in [0,T]$. In order to take supremum over the time and the expectation, we need to estimate $\mathbb{E}\Big[\sup_{0\leq s\leq t} I_{i,R}(s)\Big]$, $i = 1,2,3$, respectively. For any $t \in [0,T]$, by virtue of Assumptions (H1), (H3), and the Young inequality, we obtain

$$\begin{aligned}
\mathbb{E}\Big[\sup_{0\leq s\leq t} I_{1,R}(s)\Big] \quad\leq\quad & 2\mathbb{E}\Big[\sup_{0\leq s\leq t}\int_0^{s\wedge\tau_R^i} \langle X^i(r) - X^{i,N}(r), b(r, X^i(r), \mu_r^i) - b(r, X^{i,N}(r), \mu_r^i)\rangle dr\Big]\\
&+2\mathbb{E}\Big[\int_0^{t\wedge\tau_R^i} |X^i(s) - X^{i,N}(s)| \cdot |b(s, X^{i,N}(s), \mu_s^i) - b(s, X^{i,N}(s), \mu_s^{X,N})| ds\Big]
\end{aligned}$$

$$
\begin{aligned}
&\leq \; 2L_R^{(1)}\mathbb{E}\Big[\int_0^{t\wedge\tau_R^i}|X^i(s)-X^{i,N}(s)|^2ds\Big]\\
&\quad +2L\mathbb{E}\Big[\int_0^{t\wedge\tau_R^i}|X^i(s)-X^{i,N}(s)|\cdot W_2(\mu_s^i,\mu_s^{X,N})ds\Big]\\
&\leq \; (2L_R^{(1)}+L)\mathbb{E}\Big[\int_0^{t\wedge\tau_R^i}|X^i(s)-X^{i,N}(s)|^2ds\Big]\\
&\quad +2L\mathbb{E}\Big[\int_0^{t\wedge\tau_R^i}\big(W_2(\mu_s^i,\tilde{\mu}_s^{X,N})^2+W_2(\tilde{\mu}_s^{X,N},\mu_s^{X,N})^2\big)ds\Big]\\
&\leq \; (2L_R^{(1)}+L)\cdot\int_0^t\mathbb{E}|X^i(s\wedge\tau_R^i)-X^{i,N}(s\wedge\tau_R^i)|^2ds\\
&\quad +2L\int_0^T\mathbb{E}[W_2(\mu_s,\tilde{\mu}_s^{X,N})^2]ds+2L\mathbb{E}\Big[\int_0^{t\wedge\tau_R^i}W_2(\tilde{\mu}_s^{X,N},\mu_s^{X,N})^2ds\Big],
\end{aligned}
\tag{4.5}
$$

where $\tilde{\mu}_s^{X,N}:=\frac{1}{N}\sum_{j=1}^N\delta_{X^j(s)}$ is the empirical measure of $(X^j(s))_{1\leq j\leq N}$ and we have used the fact that $\mu_s^i=\mu_s$. By Assumptions (H2) and (H3), using the same technique as (4.5) gives

$$
\begin{aligned}
\mathbb{E}\Big[\sup_{0\leq s\leq t}I_{2,R}(s)\Big] &\leq \mathbb{E}\Big[\int_0^{t\wedge\tau_R^i}\|\sigma(s,X^i(s),\mu_s^i)-\sigma(s,X^{i,N}(s),\mu_s^{X,N})\|^2ds\Big]\\
&\leq 2\big(L_R^{(2)}\big)^2\cdot\int_0^t\mathbb{E}|X^i(s\wedge\tau_R^i)-X^{i,N}(s\wedge\tau_R^i)|^2ds\\
&\quad +4L^2\int_0^T\mathbb{E}[W_2(\mu_s,\tilde{\mu}_s^{X,N})^2]ds+4L^2\mathbb{E}\Big[\int_0^{t\wedge\tau_R^i}W_2(\tilde{\mu}_s^{X,N},\mu_s^{X,N})^2ds\Big].
\end{aligned}
\tag{4.6}
$$

Exploiting the Burkholder-Davis-Gundy inequality, the Young inequality, and (4.6), we then obtain

$$
\begin{aligned}
&\mathbb{E}\Big[\sup_{0\leq s\leq t}I_{3,R}(s)\Big]\\
&\leq 6\mathbb{E}\Big(\int_0^{t\wedge\tau_R^i}|X^i(s)-X^{i,N}(s)|^2\|\sigma(s,X^i(s),\mu_s^i)-\sigma(s,X^{i,N}(s),\mu_s^{X,N})\|^2ds\Big)^{\frac{1}{2}}\\
&\leq \frac{1}{2}\mathbb{E}\Big[\sup_{0\leq s\leq t}|X^i(s\wedge\tau_R^i)-X^{i,N}(s\wedge\tau_R^i)|^2\Big]+36\big(L_R^{(2)}\big)^2\cdot\int_0^t\mathbb{E}|X^i(s\wedge\tau_R^i)-X^{i,N}(s\wedge\tau_R^i)|^2ds\\
&\quad +72L^2\int_0^T\mathbb{E}[W_2(\mu_s,\tilde{\mu}_s^{X,N})^2]ds+72L^2\mathbb{E}\Big[\int_0^{t\wedge\tau_R^i}W_2(\tilde{\mu}_s^{X,N},\mu_s^{X,N})^2ds\Big].
\end{aligned}
\tag{4.7}
$$

Substituting (4.5)–(4.7) into (4.4) yields that

$$
\begin{aligned}
&\mathbb{E}\Big[\sup_{0\leq s\leq t}|X^i(s\wedge\tau_R^i)-X^{i,N}(s\wedge\tau_R^i)|^2\Big]\\
&\leq \big[2(2L_R^{(1)}+L)+76\big(L_R^{(2)}\big)^2\big]\cdot\int_0^t\mathbb{E}\Big[\sup_{0\leq r\leq s}|X^i(r\wedge\tau_R^i)-X^{i,N}(r\wedge\tau_R^i)|^2\Big]ds\\
&\quad +4L(1+38L)\int_0^T\mathbb{E}[W_2(\mu_s,\tilde{\mu}_s^{X,N})^2]ds+4L(1+38L)\mathbb{E}\Big[\int_0^{t\wedge\tau_R^i}W_2(\tilde{\mu}_s^{X,N},\mu_s^{X,N})^2ds\Big].
\end{aligned}
\tag{4.8}
$$

Now let us estimate $W_2(\tilde{\mu}_s^{X,N},\mu_s^{X,N})^2$. According to [35] and the definition of the L^2-Wasserstein distance,

$$
W_2(\tilde{\mu}_s^{X,N},\mu_s^{X,N})^2\leq\frac{1}{N}\sum_{j=1}^N|X^j(s)-X^{j,N}(s)|^2\cdot I_{\{\tau_R^j>s\}}+\frac{1}{N}\sum_{j=1}^N|X^j(s)-X^{j,N}(s)|^2\cdot I_{\{\tau_R^j\leq s\}}.
$$

Thus, according to Theorem 2.2 and Lemma 4.2, one arrives at

$$
\begin{aligned}
&\mathbb{E}\Big[\int_0^{t\wedge\tau_R^i} W_2(\tilde{\mu}_s^{X,N},\mu_s^{X,N})^2 ds\Big] \le \mathbb{E}\Big[\int_0^{t} W_2(\tilde{\mu}_s^{X,N},\mu_s^{X,N})^2 ds\Big]\\
&\le \int_0^t \mathbb{E}\Big[\sup_{0\le r\le s}|X^i(r\wedge\tau_R^i)-X^{i,N}(r\wedge\tau_R^i)|^2\Big]ds\\
&\quad +2\int_0^T (\mathbb{E}(|X^i(s)|^2+|X^{i,N}(s)|^2)^2)^{\frac{1}{2}}\cdot(\mathbb{P}(\tau_R^i\le s))^{\frac{1}{2}}ds\\
&\le \int_0^t \mathbb{E}\Big[\sup_{0\le r\le s}|X^i(r\wedge\tau_R^i)-X^{i,N}(r\wedge\tau_R^i)|^2\Big]ds + CT\cdot(\mathbb{P}(\tau_R^i\le T))^{\frac{1}{2}},
\end{aligned}
\tag{4.9}
$$

where we have used the fact that $(X^j-X^{j,N})_{1\le j\le N}$ are identically distributed. Applying Theorem 2.2 and Lemma 4.2 again, we can compute that for any $q\ge 2$,

$$
\begin{aligned}
\mathbb{P}(\tau_R^i\le T) &\le \mathbb{E}\Big[I_{\{\tau_R^i\le T\}}\cdot\frac{|X^i(\tau_R^i)|^q+|X^{i,N}(\tau_R^i)|^q}{R^q}\Big]\\
&\le \frac{1}{R^q}\Big(\mathbb{E}\Big[\sup_{0\le t\le T}|X^i(t)|^q\Big]+\mathbb{E}\Big[\sup_{0\le t\le T}|X^{i,N}(t)|^q\Big]\Big)\le\frac{2C_q}{R^q}.
\end{aligned}
$$

This, together with (4.8) and (4.9), we obtain

$$
\begin{aligned}
&\mathbb{E}\Big[\sup_{0\le s\le t}|X^i(s\wedge\tau_R^i)-X^{i,N}(s\wedge\tau_R^i)|^2\Big]\\
&\le \big[4L_R^{(1)}+76\big(L_R^{(2)}\big)^2+2L(3+76L)\big]\cdot\int_0^t\mathbb{E}\Big[\sup_{0\le r\le s}|X^i(r\wedge\tau_R^i)-X^{i,N}(r\wedge\tau_R^i)|^2\Big]ds\\
&\quad +4L(1+38L)\int_0^T\mathbb{E}[W_2(\mu_s,\tilde{\mu}_s^{X,N})^2]ds+CL(1+38L)T\cdot\frac{\sqrt{2C_q}}{R^{\frac{q}{2}}}.
\end{aligned}
$$

Note that $L_R^{(1)}\le\alpha_1\log R$ and $L_R^{(2)}\le\sqrt{\alpha_2\log R}$. Using the Grönwall inequality, yields

$$
\begin{aligned}
&\sup_{1\le i\le N}\mathbb{E}\Big[\sup_{0\le s\le T}|X^i(s\wedge\tau_R^i)-X^{i,N}(s\wedge\tau_R^i)|^2\Big]\\
&\le CR^{4(\alpha_1+19\alpha_2)T}\cdot\int_0^T\mathbb{E}[W_2(\mu_s,\tilde{\mu}_s^{X,N})^2]ds+C\frac{\sqrt{C_q}}{R^{\frac{q}{2}}}\cdot R^{4(\alpha_1+19\alpha_2)T}.
\end{aligned}
\tag{4.10}
$$

Thus, by the Young inequality, Theorem 2.2, Lemma 4.2, and (4.10), we have that for any $\gamma>0$,

$$
\begin{aligned}
&\sup_{1\le i\le N}\mathbb{E}\Big[\sup_{0\le s\le T}|X^i(s)-X^{i,N}(s)|^2\Big]\\
&\le \sup_{1\le i\le N}\mathbb{E}\Big[\sup_{0\le s\le T}|X^i(s)-X^{i,N}(s)|^2\cdot I_{\{\tau_R^i>T\}}\Big]+\sup_{1\le i\le N}\mathbb{E}\Big[\sup_{0\le s\le T}|X^i(s)-X^{i,N}(s)|^2\cdot I_{\{\tau_R^i\le T\}}\Big]\\
&\le \sup_{1\le i\le N}\mathbb{E}\Big[\sup_{0\le s\le T}|X^i(s\wedge\tau_R^i)-X^{i,N}(s\wedge\tau_R^i)|^2\Big]+C\gamma+\frac{1}{2\gamma}\sup_{1\le i\le N}\mathbb{P}(\tau_R^i\le T)\\
&\le CR^{4(\alpha_1+19\alpha_2)T}\cdot\int_0^T\mathbb{E}[W_2(\mu_s,\tilde{\mu}_s^{X,N})^2]ds+C\frac{\sqrt{C_q}}{R^{\frac{q}{2}}}\cdot R^{4(\alpha_1+19\alpha_2)T}+C\gamma+\frac{C_q}{\gamma R^q}.
\end{aligned}
$$

Then, by the result of [6, Lemma 1.9, p.16], one can obtain

$$
\mathbb{E}[W_2(\mu_s,\tilde{\mu}_s^{X,N})^2]\le 4[\mu_s]_2\le 4\mathbb{E}\Big[\sup_{0\le s\le T}|X(s)|^2\Big]<\infty
$$

and

$$\lim_{N\to\infty}\mathbb{E}[W_2(\mu_s,\tilde{\mu}_s^{X,N})^2]=0.$$

Note that R is independent of N. Letting $N\to\infty$, the Lebesgue dominated convergence theorem implies that

$$\lim_{N\to\infty}\sup_{1\le i\le N}\mathbb{E}\Big[\sup_{0\le s\le T}|X^i(s)-X^{i,N}(s)|^2\Big]\le C\frac{\sqrt{C_q}}{R^{\frac{q}{2}}}\cdot R^{4(\alpha_1+19\alpha_2)T}+C\gamma+\frac{C_q}{\gamma R^q}. \tag{4.11}$$

Letting $q\ge 2$ be sufficiently large for $q>8(\alpha_1+19\alpha_2)T$, the first and third terms in the right-hand side of (4.11) converge to 0 as $R\to\infty$. Furthermore, the arbitrariness of γ yields the desired conclusion. This completes the proof. □

The following lemma shows the strong convergence of the continuous-time EM scheme w.r.t. the stochastic interacting particle system (1.2).

Lemma 4.4. *Suppose that Assumptions* (H1)–(H6) *hold and $p\ge 2$. Then the EM scheme* (2.3) *converges to the stochastic interacting particle system* (1.2), *that is,*

$$\lim_{n\to\infty}\sup_{1\le i\le N}\mathbb{E}\Big[\sup_{0\le t\le T}|X^{i,N}(t)-X^{i,N,n}(t)|^p\Big]=0.$$

Proof. By (1.2), (2.3), and using the Itô formula, one derives that

$$|X^{i,N}(t)-X^{i,N,n}(t)|^p=\sum_{i=1}^4\Lambda_i(t), \tag{4.12}$$

where

$$\begin{aligned}
\Lambda_1(t)&=p\int_0^t|X^{i,N}(s)-X^{i,N,n}(s)|^{p-2}\langle X^{i,N}(s)-X^{i,N,n}(s),\\
&\qquad b(s,X^{i,N}(s),\mu_s^{X,N})-b([s]_n,X^{i,N,n}([s]_n),\mu_{[s]_n}^{X,N,n})\rangle ds,\\
\Lambda_2(t)&=\frac{p}{2}\int_0^t|X^{i,N}(s)-X^{i,N,n}(s)|^{p-2}\\
&\qquad\times\|\sigma(s,X^{i,N}(s),\mu_s^{X,N})-\sigma([s]_n,X^{i,N,n}([s]_n),\mu_{[s]_n}^{X,N,n})\|^2ds,\\
\Lambda_3(t)&=\frac{p(p-2)}{2}\int_0^t|X^{i,N}(s)-X^{i,N,n}(s)|^{p-4}\cdot|(X^{i,N}(s)-X^{i,N,n}(s))^\top\\
&\qquad\times(\sigma(s,X^{i,N}(s),\mu_s^{X,N})-\sigma([s]_n,X^{i,N,n}([s]_n),\mu_{[s]_n}^{X,N,n}))|^2ds,\\
\Lambda_4(t)&=p\int_0^t|X^{i,N}(s)-X^{i,N,n}(s)|^{p-2}\langle X^{i,N}(s)-X^{i,N,n}(s),\\
&\qquad\sigma(s,X^{i,N}(s),\mu_s^{X,N})-\sigma([s]_n,X^{i,N,n}([s]_n),\mu_{[s]_n}^{X,N,n})dW(s)\rangle.
\end{aligned}$$

For a sufficiently large $R>0$, define

$$\Omega^{i,N,n}(R)=\Big\{\omega\in\Omega:\sup_{0\le t\le T}|X^{i,N}(t)|\vee\sup_{0\le t\le T}|X^{i,N,n}(t)|\ge R\Big\}.$$

Consequently, by the Chebyshev inequality and Lemma 4.2, we get that for any $q\ge 2$,

$$\mathbb{P}(\Omega^{i,N,n}(R))\le\frac{1}{R^q}\Big(\mathbb{E}\Big[\sup_{0\le t\le T}|X^{i,N}(t)|^q\Big]+\mathbb{E}\Big[\sup_{0\le t\le T}|X^{i,N,n}(t)|^q\Big]\Big)\le\frac{2C_q}{R^q}, \tag{4.13}$$

where $R>0$ and $q\geq 2$ are to be chosen later. Moreover, it follows from (4.13) that the estimate of $\mathbb{P}(\Omega^{i,N,n}(R))$ is independent of i, n, and N. Therefore, the selected R and q are also independent of i, n, and N. For any $t\in[0,T]$, by Assumptions (H1), (H3), (H5), (H6), and the Hölder inequality, we obtain

$$
\begin{aligned}
&\mathbb{E}\Big[\sup_{0\leq s\leq t}\Lambda_1(s)\Big]\\
&\leq p\mathbb{E}\Big[\sup_{0\leq s\leq t}\int_0^s |X^{i,N}(r)-X^{i,N,n}(r)|^{p-2}\langle X^{i,N}(r)-X^{i,N,n}([r]_n),\\
&\qquad\qquad b(r,X^{i,N}(r),\mu_r^{X,N})-b([r]_n,X^{i,N,n}([r]_n),\mu_{[r]_n}^{X,N,n})\rangle dr\Big]\\
&\quad+p\mathbb{E}\Big[\sup_{0\leq s\leq t}\int_0^s |X^{i,N}(r)-X^{i,N,n}(r)|^{p-2}\langle X^{i,N,n}([r]_n)-X^{i,N,n}(r),\\
&\qquad\qquad b(r,X^{i,N}(r),\mu_r^{X,N})-b([r]_n,X^{i,N,n}([r]_n),\mu_{[r]_n}^{X,N,n})\rangle dr\Big]\\
&\leq p\mathbb{E}\Big[\Big(\sup_{0\leq s\leq t}\int_0^s |X^{i,N}(r)-X^{i,N,n}(r)|^{p-2}\langle X^{i,N}(r)-X^{i,N,n}([r]_n),\\
&\qquad\qquad b(r,X^{i,N}(r),\mu_r^{X,N})-b(r,X^{i,N,n}([r]_n),\mu_r^{X,N})\rangle dr\Big)\cdot I_{\Omega\setminus\Omega^{i,N,n}(R)}\Big]\\
&\quad+p\mathbb{E}\Big[\Big(\sup_{0\leq s\leq t}\int_0^s |X^{i,N}(r)-X^{i,N,n}(r)|^{p-2}\langle X^{i,N}(r)-X^{i,N,n}([r]_n),\\
&\qquad\qquad b(r,X^{i,N}(r),\mu_r^{X,N})-b(r,X^{i,N,n}([r]_n),\mu_r^{X,N})\rangle dr\Big)\cdot I_{\Omega^{i,N,n}(R)}\Big]\\
&\quad+p\mathbb{E}\Big[\int_0^t |X^{i,N}(s)-X^{i,N,n}(s)|^{p-2}\cdot|X^{i,N}(s)-X^{i,N,n}([s]_n)|\\
&\qquad\qquad\times|b(s,X^{i,N,n}([s]_n),\mu_s^{X,N})-b([s]_n,X^{i,N,n}([s]_n),\mu_s^{X,N})|ds\Big]\\
&\quad+p\mathbb{E}\Big[\int_0^t |X^{i,N}(s)-X^{i,N,n}(s)|^{p-2}\cdot|X^{i,N}(s)-X^{i,N,n}([s]_n)|\\
&\qquad\qquad\times|b([s]_n,X^{i,N,n}([s]_n),\mu_s^{X,N})-b([s]_n,X^{i,N,n}([s]_n),\mu_{[s]_n}^{X,N,n})|ds\Big]\\
&\quad+p\mathbb{E}\Big[\int_0^t |X^{i,N}(s)-X^{i,N,n}(s)|^{p-2}\cdot|X^{i,N,n}(s)-X^{i,N,n}([s]_n)|\\
&\qquad\qquad\times|b(s,X^{i,N}(s),\mu_s^{X,N})-b([s]_n,X^{i,N,n}([s]_n),\mu_{[s]_n}^{X,N,n})|ds\Big]\\
&\leq 4pL_R^{(1)}\cdot\int_0^t\mathbb{E}|X^{i,N}(s)-X^{i,N,n}(s)|^p ds+4L_R^{(1)}\cdot\int_0^t\mathbb{E}|X^{i,N,n}(s)-X^{i,N,n}([s]_n)|^p ds\\
&\quad+2K_1p\mathbb{E}\Big[\int_0^t |X^{i,N}(s)-X^{i,N,n}(s)|^{p-1}\\
&\qquad\qquad\times(1+|X^{i,N}(s)|+|X^{i,N,n}([s]_n)|+W_2(\mu_s^{X,N},\delta_0))ds\cdot I_{\Omega^{i,N,n}(R)}\Big]\\
&\quad+2K_1p\mathbb{E}\Big[\int_0^t |X^{i,N}(s)-X^{i,N,n}(s)|^{p-2}\cdot|X^{i,N,n}(s)-X^{i,N,n}([s]_n)|\\
&\qquad\qquad\times(1+|X^{i,N}(s)|+|X^{i,N,n}([s]_n)|+W_2(\mu_s^{X,N},\delta_0))ds\cdot I_{\Omega^{i,N,n}(R)}\Big]\\
&\quad+K_2p\mathbb{E}\Big[\int_0^t |X^{i,N}(s)-X^{i,N,n}(s)|^{p-1}(1+|X^{i,N,n}([s]_n)|^\rho+W_2(\mu_s^{X,N},\delta_0)^\rho)\cdot h_n^\alpha ds\Big]
\end{aligned}
$$

$$
\begin{aligned}
&+K_2p\mathbb{E}\Big[\int_0^t |X^{i,N}(s)-X^{i,N,n}(s)|^{p-2}\cdot|X^{i,N,n}(s)-X^{i,N,n}([s]_n)| \\
&\qquad\qquad \times(1+|X^{i,N,n}([s]_n)|^\rho+W_2(\mu_s^{X,N},\delta_0)^\rho)\cdot h_n^\alpha ds\Big] \\
&+Lp\mathbb{E}\Big[\int_0^t |X^{i,N}(s)-X^{i,N,n}(s)|^{p-2}\cdot|X^{i,N}(s)-X^{i,N,n}([s]_n)|\cdot W_2(\mu_s^{X,N},\mu_{[s]_n}^{X,N,n})ds\Big] \\
&+2K_1p\mathbb{E}\Big[\int_0^t |X^{i,N}(s)-X^{i,N,n}(s)|^{p-2}\cdot|X^{i,N,n}(s)-X^{i,N,n}([s]_n)| \\
&\qquad\qquad \times(1+|X^{i,N}(s)|+|X^{i,N,n}([s]_n)|+W_2(\mu_s^{X,N},\delta_0)+W_2(\mu_{[s]_n}^{X,N,n},\delta_0))ds\Big] \\
\le\ & [4pL_R^{(1)}+2(3K_1+K_2)(p-1)+(2^p+p-1)L]\cdot\int_0^t \mathbb{E}|X^{i,N}(s)-X^{i,N,n}(s)|^p ds \\
&+4K_1\int_0^t \mathbb{E}[(1+|X^{i,N}(s)|+|X^{i,N,n}([s]_n)|+W_2(\mu_s^{X,N},\delta_0))^p\cdot I_{\Omega^{i,N,n}(R)}]ds \\
&+2K_2\int_0^t \mathbb{E}[(1+|X^{i,N,n}([s]_n)|^\rho+W_2(\mu_s^{X,N},\delta_0)^\rho)^p]\cdot h_n^{\alpha p}ds \\
&+4K_1\int_0^t \mathbb{E}[(1+|X^{i,N}(s)|+|X^{i,N,n}([s]_n)|+W_2(\mu_s^{X,N},\delta_0)+W_2(\mu_{[s]_n}^{X,N,n},\delta_0))^{\frac{p}{2}} \\
&\qquad\qquad \times|X^{i,N,n}(s)-X^{i,N,n}([s]_n)|^{\frac{p}{2}}]ds \\
&+\big(4L_R^{(1)}+2K_1+K_2+2^pL\big)\cdot\int_0^t \mathbb{E}|X^{i,N,n}(s)-X^{i,N,n}([s]_n)|^p ds \\
\le\ & [4pL_R^{(1)}+2(3K_1+K_2)(p-1)+(2^p+p-1)L]\cdot\int_0^t \mathbb{E}|X^{i,N}(s)-X^{i,N,n}(s)|^p ds \\
&+C_p(L_R^{(1)}+1)\cdot\int_0^T \mathbb{E}|X^{i,N,n}(s)-X^{i,N,n}([s]_n)|^p ds+C_pK_2Th_n^{\alpha p} \\
&+C_p\cdot\int_0^T \big(\mathbb{E}|X^{i,N,n}(s)-X^{i,N,n}([s]_n)|^p\big)^{\frac{1}{2}}ds+C_pK_1T\big(\mathbb{P}(\Omega^{i,N,n}(R))\big)^{\frac{1}{2}}.
\end{aligned}
\tag{4.14}
$$

For any $s\in[0,T)$, there exists a $k=0,1,\dots,n-1$, such that $s\in[t_k^n,t_{k+1}^n)$. Hence, according to (2.3), Assumption (H5), and Lemma 4.2, we can calculate that for any $p\ge 2$,

$$
\begin{aligned}
&\mathbb{E}|X^{i,N,n}(s)-X^{i,N,n}([s]_n)|^p \\
&=\mathbb{E}\Big|\int_{t_k^n}^s b([r]_n,X^{i,N,n}([r]_n),\mu_{[r]_n}^{X,N,n})dr+\int_{t_k^n}^s \sigma([r]_n,X^{i,N,n}([r]_n),\mu_{[r]_n}^{X,N,n})dW^i(r)\Big|^p \\
&\le 2^{p-1}h_n^{p-1}\int_{t_k^n}^s \mathbb{E}|b([r]_n,X^{i,N,n}([r]_n),\mu_{[r]_n}^{X,N,n})|^p dr \\
&\quad +2^{p-1}M_p\mathbb{E}\Big(\int_{t_k^n}^s \|\sigma([r]_n,X^{i,N,n}([r]_n),\mu_{[r]_n}^{X,N,n})\|^2dr\Big)^{\frac{p}{2}} \\
&\le 6^{p-1}K_1^p h_n^{\frac{p-2}{2}}(h_n^{\frac{p}{2}}+M_p)\int_{t_k^n}^s \mathbb{E}[1+|X^{i,N,n}([r]_n)|^p+W_2(\mu_{[r]_n}^{X,N,n},\delta_0)^p]dr \\
&\le C_ph_n^{\frac{p}{2}}.
\end{aligned}
$$

This, together with (4.14), implies that

$$
\mathbb{E}\Big[\sup_{0\le s\le t}\Lambda_1(s)\Big]
$$

$$
\begin{aligned}
&\le [4pL_R^{(1)}+2(3K_1+K_2)(p-1)+(2^p+p-1)L]\cdot\int_0^t \mathbb{E}|X^{i,N}(s)-X^{i,N,n}(s)|^p ds\\
&\quad +C_pT(L_R^{(1)}+1)h_n^{\frac{p}{2}}+C_pTh_n^{\frac{p}{4}}+C_pK_1T\big(\mathbb{P}(\Omega^{i,N,n}(R))\big)^{\frac12}+C_pK_2Th_n^{\alpha p}. \qquad (4.15)
\end{aligned}
$$

By Assumptions (H2), (H3), (H5), and (H6), using the same technique as (4.15) was proved, we get

$$
\begin{aligned}
&\mathbb{E}\Big[\sup_{0\le s\le t}(\Lambda_2(s)+\Lambda_3(s))\Big]\\
&\le \frac{p(p-1)}{2}\mathbb{E}\Big[\int_0^t |X^{i,N}(s)-X^{i,N,n}(s)|^{p-2}\\
&\qquad\qquad \times\|\sigma(s,X^{i,N}(s),\mu_s^{X,N})-\sigma([s]_n,X^{i,N,n}([s]_n),\mu_{[s]_n}^{X,N,n})\|^2 ds\Big]\\
&\le \frac{3p(p-1)}{2}\mathbb{E}\Big[\int_0^t |X^{i,N}(s)-X^{i,N,n}(s)|^{p-2}\\
&\qquad\qquad \times\|\sigma(s,X^{i,N}(s),\mu_s^{X,N})-\sigma(s,X^{i,N,n}([s]_n),\mu_s^{X,N})\|^2 ds\cdot I_{\Omega\backslash\Omega^{i,N,n}(R)}\Big]\\
&\quad +\frac{3p(p-1)}{2}\mathbb{E}\Big[\int_0^t |X^{i,N}(s)-X^{i,N,n}(s)|^{p-2}\\
&\qquad\qquad \times\|\sigma(s,X^{i,N}(s),\mu_s^{X,N})-\sigma(s,X^{i,N,n}([s]_n),\mu_s^{X,N})\|^2 ds\cdot I_{\Omega^{i,N,n}(R)}\Big]\\
&\quad +\frac{3p(p-1)}{2}\mathbb{E}\Big[\int_0^t |X^{i,N}(s)-X^{i,N,n}(s)|^{p-2}\\
&\qquad\qquad \times\|\sigma(s,X^{i,N,n}([s]_n),\mu_s^{X,N})-\sigma([s]_n,X^{i,N,n}([s]_n),\mu_s^{X,N})\|^2 ds\Big]\\
&\quad +\frac{3p(p-1)}{2}\mathbb{E}\Big[\int_0^t |X^{i,N}(s)-X^{i,N,n}(s)|^{p-2}\\
&\qquad\qquad \times\|\sigma([s]_n,X^{i,N,n}([s]_n),\mu_s^{X,N})-\sigma([s]_n,X^{i,N,n}([s]_n),\mu_{[s]_n}^{X,N,n})\|^2 ds\Big]\\
&\le 6p(p-1)\big[\big(L_R^{(2)}\big)^2+3(K_1^2+K_2^2+L^2)\big]\cdot\int_0^t \mathbb{E}|X^{i,N}(s)-X^{i,N,n}(s)|^p ds\\
&\quad +C_p(p-1)T\big[\big(L_R^{(2)}\big)^2+L^2\big]h_n^{\frac{p}{2}}+C_p(p-1)K_1^2T\big(\mathbb{P}(\Omega^{i,N,n}(R))\big)^{\frac12}\\
&\quad +C_p(p-1)K_2^2Th_n^{\alpha p}. \qquad (4.16)
\end{aligned}
$$

By using the Burkholder-Davis-Gundy inequality, the Young inequality, and (4.16), we then obtain

$$
\begin{aligned}
\mathbb{E}\Big[\sup_{0\le s\le t}\Lambda_4(s)\Big] &\le 3p\mathbb{E}\Big(\int_0^t |X^{i,N}(s)-X^{i,N,n}(s)|^{2p-2}\\
&\qquad \times\|\sigma(s,X^{i,N}(s),\mu_s^{X,N})-\sigma([s]_n,X^{i,N,n}([s]_n),\mu_{[s]_n}^{X,N,n})\|^2 ds\Big)^{\frac12}\\
&\le \frac12\mathbb{E}\Big[\sup_{0\le s\le t}|X^{i,N}(s)-X^{i,N,n}(s)|^p\Big]+C_ppT\big[\big(L_R^{(2)}\big)^2+L^2\big]h_n^{\frac{p}{2}}\\
&\quad +60p^2\big[\big(L_R^{(2)}\big)^2+3(K_1^2+K_2^2+L^2)\big]\cdot\int_0^t \mathbb{E}|X^{i,N}(s)-X^{i,N,n}(s)|^p ds\Big]\\
&\quad +C_ppK_1^2T\big(\mathbb{P}(\Omega^{i,N,n}(R))\big)^{\frac12}+C_ppK_2^2Th_n^{\alpha p}. \qquad (4.17)
\end{aligned}
$$

Substituting (4.15)–(4.17) into (4.12) yields that

$$
\begin{aligned}
&\mathbb{E}\Big[\sup_{0\le s\le t}|X^{i,N}(s)-X^{i,N,n}(s)|^p\Big]\\
&\le \big[8pL_R^{(1)}+132p^2\big(L_R^{(2)}\big)^2+C_p\big]\cdot\int_0^t \mathbb{E}\Big[\sup_{0\le r\le s}|X^{i,N}(r)-X^{i,N,n}(r)|^p\Big]ds\\
&\quad +C_p\big[L_R^{(1)}+\big(L_R^{(2)}\big)^2+1\big]h_n^{\frac{p}{2}}+C_p\big(\mathbb{P}(\Omega^{i,N,n}(R))\big)^{\frac12}+C_ph_n^{\alpha p}+C_ph_n^{\frac{p}{4}}. \qquad (4.18)
\end{aligned}
$$

Note that $L_R^{(1)}\le \alpha_1\log R$, $L_R^{(2)}\le\sqrt{\alpha_2\log R}$, and (4.13). By using the Grönwall inequality to (4.18), one arrives at

$$
\begin{aligned}
&\sup_{1\le i\le N}\mathbb{E}\Big[\sup_{0\le s\le T}|X^{i,N}(s)-X^{i,N,n}(s)|^p\Big]\\
&\le C_p\big[L_R^{(1)}+\big(L_R^{(2)}\big)^2+1\big]R^{4p(2\alpha_1+33p\alpha_2)T}\cdot h_n^{\frac{p}{2}}+C_p\frac{\sqrt{C_q}}{R^{\frac{q}{2}}}R^{4p(2\alpha_1+33p\alpha_2)T}\\
&\quad +C_pR^{4p(2\alpha_1+33p\alpha_2)T}\cdot h_n^{\alpha p}+C_pR^{4p(2\alpha_1+33p\alpha_2)T}\cdot h_n^{\frac{p}{4}}. \qquad (4.19)
\end{aligned}
$$

Note that R is independent of n, N, and $h_n\to 0$ as $n\to\infty$. Then, taking the limit on the both sides of (4.19), we get

$$
\lim_{n\to\infty}\sup_{1\le i\le N}\mathbb{E}\Big[\sup_{0\le s\le T}|X^{i,N}(s)-X^{i,N,n}(s)|^p\Big]\le C_p\frac{\sqrt{C_q}}{R^{\frac{q}{2}}}R^{4p(2\alpha_1+33p\alpha_2)T}. \qquad (4.20)
$$

Letting $q\ge 2$ be sufficiently large for $q>8p(2\alpha_1+33p\alpha_2)T$, the right-hand side of (4.20) converges to 0 as $R\to\infty$. Therefore, the proof is complete. □

Proof of Theorem 2.3. Note that for any $t\in[0,T]$,

$$
|X^i(t)-X^{i,N,n}(t)|^2\le 2|X^i(t)-X^{i,N}(t)|^2+2|X^{i,N}(t)-X^{i,N,n}(t)|^2.
$$

Therefore, by virtue of Lemmas 4.3 and 4.4, we complete the proof of Theorem 2.3. □

Remark 4.5. If the drift b and the diffusion coefficient σ satisfy the global Lipschitz conditions w.r.t. the state variable, that is, $L_R^{(1)}$ and $L_R^{(2)}$ are independent of R (We can rewrite them as L_1 and L_2), then the strong convergence with the corresponding convergence rate with respect to N and step size can be obtained in Theorem 2.3. Indeed, according to the proofs of Lemmas 4.3 and 4.4,

$$
\begin{aligned}
&\sup_{1\le i\le N}\mathbb{E}\Big[\sup_{0\le s\le T}|X^i(s)-X^{i,N,n}(s)|^2\Big]\\
&\le 2\sup_{1\le i\le N}\mathbb{E}\Big[\sup_{0\le s\le T}|X^i(s)-X^{i,N}(s)|^2\Big]+2\sup_{1\le i\le N}\mathbb{E}\Big[\sup_{0\le s\le T}|X^{i,N}(s)-X^{i,N,n}(s)|^2\Big]\\
&\le Ce^{4(L_1+19L_2^2)T}\cdot\int_0^T\mathbb{E}[W_2(\mu_s,\tilde{\mu}_s^{X,N})^2]ds+C\frac{\sqrt{C_q}}{R^{\frac{q}{2}}}\cdot e^{4(L_1+19L_2^2)T}+C\gamma+\frac{C_q}{\gamma R^q}\\
&\quad +C\big(L_1+L_2^2+1\big)e^{16(L_1+33L_2^2)T}h_n+C\frac{\sqrt{C_q}}{R^{\frac{q}{2}}}e^{16(L_1+33L_2^2)T}+Ce^{16(L_1+33L_2^2)T}h_n^{2\alpha}\\
&\quad +Ce^{16(L_1+33L_2^2)T}h_n^{\frac12}.
\end{aligned}
$$

Letting $R \to \infty$ and using the arbitrariness of γ, one obtains

$$\sup_{1\le i\le N} \mathbb{E}\Big[\sup_{0\le s\le T} |X^i(s) - X^{i,N,n}(s)|^2\Big] \le C\int_0^T \mathbb{E}[W_2(\mu_s, \tilde{\mu}_s^{X,N})^2]ds + Ch_n + Ch_n^{2\alpha} + Ch_n^{\frac{1}{2}}.$$

Recall that for a fixed $s \in [0,T]$ and some $p > 4$, there exists a positive constant C depending only on T such that

$$\int_{\mathbb{R}^d} |x|^p \mu_s(dx) \le \mathbb{E}\Big[\sup_{0\le s\le T} |X(s)|^p\Big] \le C.$$

This, together with the Glivenko-Cantelli theorem [7, Theorem 5.8, p.362], implies that

$$\sup_{1\le i\le N} \mathbb{E}\Big[\sup_{0\le s\le T} |X^i(s) - X^{i,N,n}(s)|^2\Big] \le C \begin{cases} N^{-\frac{1}{2}} + h_n + h_n^{\frac{1}{2}} + h_n^{2\alpha}, & d < 4, \\ N^{-\frac{1}{2}} \log N + h_n^{\frac{1}{2}} + h_n + h_n^{2\alpha}, & d = 4, \\ N^{-\frac{2}{d}} + h_n + h_n^{\frac{1}{2}} + h_n^{2\alpha}, & d > 4. \end{cases}$$

Therefore, we can obtain the rate of convergence under the global Lipschitz conditions w.r.t. the state variable.

5 Example

Example 5.1. Assume that $\phi : \mathbb{R}^d \to \mathbb{R}^d$ and $\psi : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ are Borel measurable functions given by

$$\phi(x) = \begin{pmatrix} x_1 \sin\big((\log(1 + |x_1|^2 + |x_2|^2 + \cdots + |x_{d-1}|^2))^2\big) \\ x_2 \sin\big((\log(1 + |x_2|^2 + |x_3|^2 + \cdots + |x_d|^2))^2\big) \\ \vdots \\ x_d \sin\big((\log(1 + |x_d|^2 + |x_1|^2 + \cdots + |x_{d-2}|^2))^2\big) \end{pmatrix}$$

and

$$\psi(x) = \begin{pmatrix} x_1 \sin\Big(\big(\log(1+|x_1|^2)\big)^{\frac{3}{2}}\Big) & x_1 \sin\Big(\big(\log(1+|x_1|^2+|x_2|^2)\big)^{\frac{3}{2}}\Big) & \cdots & x_1 \sin\Big(\big(\log(1+|x_1|^2+|x_d|^2)\big)^{\frac{3}{2}}\Big) \\ x_2 \sin\Big(\big(\log(1+|x_2|^2+|x_1|^2)\big)^{\frac{3}{2}}\Big) & x_2 \sin\Big(\big(\log(1+|x_2|^2)\big)^{\frac{3}{2}}\Big) & \cdots & x_2 \sin\Big(\big(\log(1+|x_2|^2+|x_d|^2)\big)^{\frac{3}{2}}\Big) \\ \vdots & \vdots & & \vdots \\ x_d \sin\Big(\big(\log(1+|x_d|^2+|x_1|^2)\big)^{\frac{3}{2}}\Big) & x_d \sin\Big(\big(\log(1+|x_d|^2+|x_2|^2)\big)^{\frac{3}{2}}\Big) & \cdots & x_d \sin\Big(\big(\log(1+|x_d|^2)\big)^{\frac{3}{2}}\Big) \end{pmatrix},$$

respectively. Then, it is easily seen that for any $x \in \mathbb{R}^d$,

$$|\phi(x)| \le |x|, \qquad \|\psi(x)\| \le \sqrt{d}|x|. \tag{5.1}$$

Furthermore, ϕ and ψ are both continuously differentiable on $\mathbb{R}^d$, and for any $x \in \mathbb{R}^d$, we can calculate that

$$\|(\partial_x \phi)(x)\| \le \sqrt{d}\big(1 + 4\log\big(1 + |x|^2\big)\big), \tag{5.2}$$

and

$$\|(\partial_x \psi)(x)\| \le d\Big(1 + 3\sqrt{\log\big(1 + |x|^2\big)}\Big). \tag{5.3}$$

Next, we consider the case $d_1 = d$ and the following McKean-Vlasov SDE

$$dX(t) = b(X(t), \mathscr{L}(X(t)))dt + \sigma(X(t), \mathscr{L}(X(t)))dW(t), \qquad t \in [0,T], \tag{5.4}$$

where $b:\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\to\mathbb{R}^d$ and $\sigma:\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\to\mathbb{R}^{d\times d}$ are Borel measurable functions defined by

$$b(x,\mu)=B_1x+B_2\phi(x)+B_3\int_{\mathbb{R}^d}y\mu(dy)+B_4\varphi(\mu)$$

and

$$\sigma(x,\mu)=\Sigma_1\psi(x)+\Sigma_2\cdot\mathrm{diag}\Big\{\int_{\mathbb{R}^d}y\mu(dy)\Big\}+\Sigma_3\mathrm{diag}\{\varphi(\mu)\},$$

B_1, B_2, B_3, B_4, Σ_1, Σ_2, and Σ_3 are deterministic $d\times d$ matrices, $\mathrm{diag}\{x\}$ represents a diagonal matrix with diagonal elements x, and

$$\varphi(\mu)=\Big(\sin\Big(\int_{\mathbb{R}^d}y_1\mu(dy)\Big),\ldots,\sin\Big(\int_{\mathbb{R}^d}y_d\mu(dy)\Big)\Big)^\top.$$

This equation is nonlinear with respect to the measure μ. Assume that there exists a positive constant K such that

$$\max\{\|B_1\|,\|B_2\|,\|B_3\|,\|B_4\|,\|\Sigma_1\|,\|\Sigma_2\|,\|\Sigma_3\|\}\le K.$$

For any $\mathcal{F}_0$-measurable random variable ξ satisfying $\mathbb{E}|\xi|^\beta<\infty$, $\beta>0$, equation (5.4) has a unique solution $(X(t))_{0\le t\le T}$. To proceed, we first examine the conditions in Theorem 2.2 are satisfied. According to (5.1) and the expressions of b and σ, Assumption (H5) can be verified easily. Hence, we need only verify that b and σ satisfy Assumptions (H1)-(H4). Applying (5.2) and (5.3), for any integer $R\ge 3$, $x,y\in\mathbb{R}^d$ with $|x|\vee|y|\le R$, and $\mu\in\mathcal{P}_2(\mathbb{R}^d)$, one computes

$$\begin{aligned}
|b(x,\mu)-b(y,\mu)| &= |B_1x+B_2\phi(x)-(B_1y+B_2\phi(y))|\\
&= \Big|\int_0^1[B_1+B_2(\partial_x\phi)(y+\rho(x-y))](x-y)d\rho\Big|\\
&\le \Big[\sup_{|z|\le R}\|B_1+B_2(\partial_x\phi)(z)\|\Big]\cdot|x-y|\\
&\le \Big[\|B_1\|+\|B_2\|\cdot\Big(\sup_{|z|\le R}\sqrt{d}\big(1+4\log\big(1+|z|^2\big)\big)\Big)\Big]\cdot|x-y|\\
&\le \big[K+K\sqrt{d}\big(1+4\log\big(1+R^2\big)\big)\big]\cdot|x-y|\\
&=: L_R^{(1)}|x-y|,
\end{aligned}\tag{5.5}$$

and

$$\begin{aligned}
\|\sigma(x,\mu)-\sigma(y,\mu)\| &= \|\Sigma_1\psi(x)-\Sigma_1\psi(y)\|\\
&= \Big|\int_0^1[\Sigma_1(\partial_x\psi)(y+\rho(x-y))](x-y)d\rho\Big|\\
&\le \Big[\sup_{|z|\le R}\|\Sigma_1(\partial_x\psi)(z)\|\Big]\cdot|x-y|\\
&\le \|\Sigma_1\|\cdot\Big[\sup_{|z|\le R}d\Big(1+3\sqrt{\log\big(1+|z|^2\big)}\Big)\Big]\cdot|x-y|\\
&\le Kd\Big(1+3\sqrt{\log\big(1+R^2\big)}\Big)\cdot|x-y|\\
&=: L_R^{(2)}|x-y|.
\end{aligned}$$

Moreover, we have that $L_R^{(1)} \leq 26K\sqrt{d}\log R$ and $L_R^{(2)} \leq \sqrt{56K^2d^2 \cdot \log R}$. These imply that Assumptions (H1) and (H2) are satisfied. In addition, applying [22, Lemma 5], for any $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$, one obtains

$$\|\varphi(\mu) - \varphi(\nu)\| \leq \sqrt{d}W_2(\mu,\nu) \quad \text{and} \quad \|\mathrm{diag}\{\varphi(\mu)\} - \mathrm{diag}\{\varphi(\nu)\}\| \leq \sqrt{d}W_2(\mu,\nu).$$

Therefore, for any $x \in \mathbb{R}^d$, and $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$, one computes

$$\begin{aligned} |b(x,\mu) - b(x,\nu)| &= \Big|B_3 \int_{\mathbb{R}^d} y\mu(dy) + B_4\varphi(\mu) - B_3\int_{\mathbb{R}^d} z\nu(dz) - B_4\varphi(\nu)\Big| \\ &\leq \|B_3\| \cdot W_2(\mu,\nu) + \|B_4\| \cdot \|\varphi(\mu) - \varphi(\nu)\| \\ &\leq K(1+\sqrt{d}) \cdot W_2(\mu,\nu), \end{aligned}$$

and

$$\begin{aligned} \|\sigma(x,\mu) - \sigma(x,\nu)\| &= \Big\|\Sigma_2 \cdot \mathrm{diag}\Big\{\int_{\mathbb{R}^d} y\mu(dy)\Big\} + \Sigma_3 \mathrm{diag}\{\varphi(\mu)\} \\ &\quad -\Sigma_2 \cdot \mathrm{diag}\Big\{\int_{\mathbb{R}^d} z\nu(dz)\Big\} - \Sigma_3\mathrm{diag}\{\varphi(\nu)\}\Big\| \\ &\leq \|\Sigma_2\| \cdot W_2(\mu,\nu) + \|\Sigma_3\| \cdot \|\mathrm{diag}\{\varphi(\mu)\} - \mathrm{diag}\{\varphi(\nu)\}\| \\ &\leq K(1+\sqrt{d}) \cdot W_2(\mu,\nu). \end{aligned}$$

These imply that Assumptions (H3) and (H4) are satisfied. Furthermore, one verifies that b does not satisfy the one-sided global Lipschitz condition. However, by Theorem 2.2 in this paper, equation (5.4) has a unique solution $(X(t))_{0\leq t\leq T}$ with $X(0) = \xi$, and for any $p > 0$,

$$\mathbb{E}\Big[\sup_{0\leq t\leq T} |X(t)|^p\Big] < \infty.$$

By virtue of Theorem 2.3, the EM scheme of the stochastic interacting particle system converges to non-interacting particle system associated with (5.4).

A Appendix: Proofs of Two Lemmas

Proof of Lemma 3.1. We adopt the approach in [26] to treat the distribution-dependent case. By the Lyapunov inequality, for any $n \geq 1$ and $p \geq 2$, using the elementary inequality $|a+b+c|^p \leq 3^{p-1}(|a|^p + |b|^p + |c|^p)$ and the Hölder inequality, it follows from (3.5) that

$$\begin{aligned} |X^{(n)}(t)|^p &\leq 3^{p-1}\Big[|\xi|^p + \Big|\int_0^t b(s, X^{(n)}(s), \mu^{(n)}_{[s]_n})ds\Big|^p + \Big|\int_0^t \sigma(s, X^{(n)}(s), \mu^{(n)}_{[s]_n})dW(s)\Big|^p\Big] \\ &\leq 3^{p-1}\Big[|\xi|^p + T^{p-1}\int_0^t |b(s, X^{(n)}(s), \mu^{(n)}_{[s]_n})|^p ds + \Big|\int_0^t \sigma(s, X^{(n)}(s), \mu^{(n)}_{[s]_n})dW(s)\Big|^p\Big]. \end{aligned}$$

This, together with (3.4) and the Burkholder-Davis-Gundy inequality (see [18, Theorem 3.28, p.166] or [27, Theorem 7.3, p.40]), implies that for every $t \in [0,T]$,

$$\begin{aligned} &\mathbb{E}\Big[\sup_{0\leq s\leq t} |X^{(n)}(s)|^p\Big] \\ &\leq 3^{p-1}\mathbb{E}|\xi|^p + 3^{p-1}T^{p-1}\mathbb{E}\Big[\int_0^t |b(s, X^{(n)}(s), \mu^{(n)}_{[s]_n})|^p ds\Big] \end{aligned}$$

$$
\begin{aligned}
&+3^{p-1}M_pT^{\frac{p-2}{2}}\mathbb{E}\Big[\int_0^t \|\sigma(s,X^{(n)}(s),\mu^{(n)}_{[s]_n})\|^p ds\Big]\\
&\le 3^{p-1}\mathbb{E}|\xi|^p + 9^{p-1}K_1^p(T^{p-1}+M_pT^{\frac{p-2}{2}})\mathbb{E}\Big[\int_0^t \big(1+|X^{(n)}(s)|^p+\mathbb{E}|X^{(n)}([s]_n)|^p\big)ds\Big]\\
&\le 3^{p-1}\mathbb{E}|\xi|^p + 2\cdot 9^{p-1}K_1^p(T^{p-1}+M_pT^{\frac{p-2}{2}})\int_0^t \Big(1+\mathbb{E}\Big[\sup_{0\le r\le s}|X^{(n)}(r)|^p\Big]\Big)ds, \qquad \text{(A.1)}
\end{aligned}
$$

where $M_p = [p^{p+1}/2(p-1)^{p-1}]^{\frac{p}{2}}$. Applying the Grönwall inequality to (A.1) yields that

$$
\mathbb{E}\Big[\sup_{0\le s\le T}|X^{(n)}(s)|^p\Big] \le C_p.
$$

It is easy to see that the positive constant C_p is dependent on p, T, and initial condition ξ, but independent of n. Therefore, the desired assertion follows. □

Proof of Lemma 3.2. We adopt the approach in [26, Theorem 4.3, p.61] to treat distribution-dependent SDEs. Set $0 \le s \le t \le T$. By the Hölder inequality, one gets

$$
\begin{aligned}
|X^{(n)}(t)-X^{(n)}(s)|^p \;\le\;& 2^{p-1}(t-s)^{p-1}\int_s^t |b(r,X^{(n)}(r),\mu^{(n)}_{[r]_n})|^p dr\\
&+2^{p-1}\Big|\int_s^t \sigma(r,X^{(n)}(r),\mu^{(n)}_{[r]_n})dW(r)\Big|^p.
\end{aligned}
$$

Taking the expectation on both sides, by the Burkholder-Davis-Gundy inequality and using Assumption (H5), we can then derive that

$$
\begin{aligned}
\mathbb{E}|X^{(n)}(t)-X^{(n)}(s)|^p \;\le\;& 2^{p-1}(t-s)^{p-1}\mathbb{E}\Big[\int_s^t |b(r,X^{(n)}(r),\mu^{(n)}_{[r]_n})|^p dr\Big]\\
&+2^{p-1}M_p(t-s)^{\frac{p-2}{2}}\mathbb{E}\Big[\int_s^t \|\sigma(r,X^{(n)}(r),\mu^{(n)}_{[r]_n})\|^p dr\Big]\\
\le\;& 6^{p-1}K_1^p(T^{\frac{p}{2}}+M_p)(t-s)^{\frac{p-2}{2}}\mathbb{E}\Big[\int_s^t (1+|X^{(n)}(r)|^p+\mathbb{E}|X^{(n)}([r]_n)|^p)dr\Big]\\
\le\;& C_p(t-s)^{\frac{p}{2}}.
\end{aligned}
$$

This leads to the desired assertion and the proof is therefore complete. □

Acknowledgement

The authors would like to thank the editors and reviewers for their comments and suggestions.

References

[1] K. Bahlali, M. A. Mezerdi, B. Mezerdi, Stability of McKean-Vlasov stochastic differential equations and applications, *Stochastics and Dynamics*, 20 (1) (2020), https://doi.org/10.1142/S0219493720500070.

[2] J. Bao, X. Huang, Approximations of McKean-Vlasov stochastic differential equations with irregular coefficients, *J. Theoretical Probab.*, (2021), https://doi.org/10.1007/s10959-021-01082-9.

[3] A. Bensoussan, J. Frehse, P. Yam, *Mean Field Games and Mean Field Type Control Theory*, SpringerBriefs Math., Springer, New York, 2013.

[4] A. Budhiraja, W.-T. L. Fan, Uniform in time interacting particle approximations for nonlinear equations of Patlak-Keller-Segel type, *Electron. J. Probab.*, 22 (2017), 1–37.

[5] P. Cardaliaguet, Notes on mean field games, Notes from P.-L. Lions' lectures at Collège de France, 2013.

[6] R. Carmona, Lectures on BSDEs, *Stochastic Control, and Stochastic Differential Games with Financial Applications*, SIAM., 2016.

[7] R. Carmona, F. Delarue, *Probabilistic Theory of Mean Field Games with Applications I: Mean Field FBSDEs, Control, and Games*, Probability Theory and Stochastic Modelling, Springer, Cham, 2018.

[8] K. L. Chung, *A Course in Probability Theory*, Academic Press, 3rd Ed., 2001.

[9] D. A. Dawson, Critical dynamics and fluctuations for a mean-field model of cooperative behavior, *J. Statist. Phys.*, 31 (1) (1983), 29–85.

[10] X. Ding, H. Qiao, Euler-Maruyama approximations for stochastic McKean-Vlasov equations with non-Lipschitz coefficients, *J. Theoretical Probab.*, 34 (2021), 1408–1425.

[11] S. Fang, P. Imkeller, T. Zhang, Global flows for stochastic differential equations without global Lipschitz conditions, *Ann. Probab.*, 35 (1) (2007), 180–205.

[12] W. Hammersley, D. Šiška, Ł. Szpruch, McKean-Vlasov SDEs under measure dependent Lyapunov conditions, *Ann. Inst. H. Poincaré Probab. Statist.*, 57 (2) (2021), 1032–1057.

[13] D. J. Higham, X. Mao, A. M. Stuart, Strong convergence of Euler-type methods for nonlinear stochastic differential equations, *SIAM J. Numer. Anal.*, 40 (3) (2002), 1041–1063.

[14] X. Huang, F. Yang, Distribution-dependent SDEs with Hölder continuous drift and α-stable noise, *Numerical Algorithms*, 86 (2021), 813–831.

[15] M. Huang, R. P. Malhamé, P. E. Caines, Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle, *Commun. Inf. Syst.*, 6 (3) (2006), 221–252.

[16] Z. Hao, M. Röckner, X. Zhang, Euler scheme for density dependent stochastic differential equations, *J. Differential Equations*, 274 (2021), 996–1014.

[17] M. Kac, *Foudations of kinetic theory*, In Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III, pages 171-197. University of California Press, Berkeley and Los Angeles, 1956.

[18] I. Karatzas, S. E. Shreve, *Brownian Motion and Stochastic Calculus*, New York : Springer-Verlag, 2nd Ed., 1991.

[19] P. E. Kloeden, E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, Berlin, 1995.

[20] P. E. Kloeden, T. Lorenz, Stochastic differential equations with nonlocal sample dependence, *Stochastic Anal. Appl.*, 28 (6) (2010), 937–945.

[21] V. N. Kolokoltsov, *Nonlinear Markov Processes and Kinetic Equations*, Cambridge University Press, 2010.

[22] C. Kumar, Neelima, On explicit Milstein-type scheme for McKean-Vlasov stochastic differential equations with super-linear drift coefficient, arXiv:2004.01266v1.

[23] C. Kumar, Neelima, C. Reisinger, W. Stockinger, Well-posedness and tamed schemes for McKean-Vlasov equations with common noise, arXiv:2006.00463v1.

[24] J.-M. Lasry, P.-L. Lions, Jeux à champ moyen. I. Le cas stationnaire, *C. R. Acad. Sci. Paris*, 343 (2006), 619–625.

[25] H. P. McKean, Jr., A class of Markov processes associated with nonlinear parabolic equations. *Proc. Nat. Acad. Sci. USA*, 56 (6) (1966), 1907–1911.

[26] X. Mao, *Stochastic Differential Equations and Applications*, Horwood, Chichester, 2nd Ed., 2007.

[27] X. Mao, C. Yuan, *Stochastic Differential Equations with Markovian Switching*, Imperial College Press, 2006.

[28] Yu. S. Mishura, A. Yu. Veretennikov, Existence and uniqueness theorems for solutions of McKean-Vlasov stochastic equations, arXiv: 1603.02212v12.

[29] S. Mehri, M. Scheutzow, W. Stannat, B. Z. Zangeneh, Propagation of chaos for stochastic spatially structured neuronal networks with delay driven by jump diffusions, *Ann. Appl. Probab.*, 30 (1) (2020), 175-207.

[30] S. L. Nguyen, G. Yin, T. A. Hoang, On laws of large numbers for systems with mean-field interactions and Markovian switching, *Stochstic Process. Appl.*, 130 (1) (2020), 262–296.

[31] C. Prévôt, M. Röckner, *A Concise Course on Stochastic Partial Differential Equations*, Springer-Verlag, 2007.

[32] P. Ren, F.-Y. Wang, Space-distribution PDEs for path independent additive functionals of McKean-Vlasov SDEs, *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 23 (3) (2020), https://doi.org/10.1142/S0219025720500186.

[33] P. Ren, H. Tang, F.-Y. Wang, Distribution-path dependent nonlinear SPDEs with application to stochastic transport type equations, arXiv:2007.09188v3.

[34] G. dos Reis, W. Salkeld, J. Tugaut, Freidlin-Wentzell LDP in path space for McKean-Vlasov equations and the functional iterated logarithm law, *Ann. Appl. Probab.*, 29 (3) (2019), 1487–1540.

[35] G. dos Reis, S. Engelhardt, G. Smith, Simulation of McKean-Vlasov SDEs with super-linear growth, *IMA Journal of Numerical Analysis*, (2021), https://doi.org/10.1093/imanum/draa099.

[36] C. Reisinger, W. Stockinger, An adaptive Euler-Maruyama scheme for McKean-Vlasov SDEs with super-linear growth and application to the mean-field FitzHugh-Nagumo model, *J. Comput. Appl. Math.*, 400 (2022), https://doi.org/10.1016/j.cam.2021.113725.

[37] A. N. Shiryaev, *Probability*, Springer, 2nd Ed., 1989.

[38] A. S. Sznitman, *Topics in propagation of chaos*, Springer, 1991.

[39] C. Villani, *Optimal Transport, Old and New*, Springer-Verlag, 2009.

[40] F.-Y. Wang, Distribution dependent SDEs for Landau type equations, *Stochastic Process. Appl.*, 128 (2) (2018), 595–621.

[41] C. Yuan, X. Mao, A note on the rate of convergence of the Euler-Maruyama method for stochastic differential equations, *Stochastic Anal. Appl.*, 26 (2) (2008), 325–333.

[42] X. Zhang, A discretized version of Krylov's estimate and its applications, *Electron. J. Probab.*, 24 (2019), 1–17.