

WELL-POSEDNESS AND AVERAGING PRINCIPLE FOR NON-GAUSSIAN MCKEAN–VLASOV STOCHASTIC DIFFERENTIAL EQUATIONS WITH LOCALLY LIPSCHITZ COEFFICIENTS

YING CHAO,* *Xi'an Jiaotong University*

JINQIAO DUAN,** *Great Bay University*

TING GAO,*** *Huazhong University of Science and Technology*

PINGYUAN WEI ,**** *Southeast University and Peking University*

Abstract

In this paper, we investigate a class of McKean–Vlasov stochastic differential equations (SDEs) with Lévy-type perturbations. We first establish the existence and uniqueness theorem for the solutions of the McKean–Vlasov SDEs by utilizing an Eulerlike approximation. Then, under suitable conditions, we demonstrate that the solutions of the McKean–Vlasov SDEs can be approximated by the solutions of the associated averaged McKean–Vlasov SDEs in the sense of mean square convergence. In contrast to existing work, a novel feature of this study is the use of a much weaker condition, locally Lipschitz continuity in the state variables, allowing for possibly superlinearly growing drift, while maintaining linearly growing diffusion and jump coefficients. Therefore, our results apply to a broader class of McKean–Vlasov SDEs.

Keywords: McKean–Vlasov stochastic differential equations; well-posedness; averaging principle; one-sided locally Lipschitz condition; Lévy-type perturbations

2020 Mathematics Subject Classification: Primary 60H10; 35Q83
Secondary 60G51; 34C29

1. Introduction

McKean–Vlasov stochastic differential equations (SDEs) have received significant attention in recent years due to their broad applications across various fields, such as stochastic control, stochastic games, and statistical physics. These equations were first introduced in [27], inspired by the kinetic theory of Kac [18], and differ from standard SDEs in that their coefficients additionally depend on the probability distribution of the solution process. In the literature, McKean–Vlasov SDEs are also referred to as mean-field SDEs, because they arise as the limits

Received 30 June 2025; 30 June 2025; accepted 30 June 2025.

* Postal address: School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China. Email: yingchao1993@xjtu.edu.cn

** Postal address: Department of Mathematics and Guangdong Provincial Key Laboratory of Mathematical and Neural Dynamical Systems, Great Bay University, Dongguan, Guangdong 523000, China. Email: duan@gbu.edu.cn

*** Postal address: School of Mathematics and Statistics and Center for Mathematical Sciences, Huazhong University of Science and Technology, Wuhan, Hubei 430074, China. Email: tgao0716@hust.edu.cn

**** Postal address: School of Mathematics, Southeast University, Nanjing 211189, China; Beijing International Center for Mathematical Research, Peking University, Beijing 100871, China. Email: weipingyuan@pku.edu.cn

of weakly interacting particle systems as the number of particles tends to infinity (so-called the propagation of chaos [39]).

In view of the development on the aforementioned McKean–Vlasov SDEs, the noise processes considered are primarily Gaussian. However, systems of practical relevance in physics and biology sometimes require modeling with non-Gaussian noise. This can be verified by some abrupt jumps in the individual particles and the related whole population. To reproduce the performance of these natural phenomena, it is appropriate to consider (non-Gaussian) Lévy-type perturbations [1, 11, 23]. In this paper, we focus on the following d -dimensional Lévy-type McKean–Vlasov SDE:

$$\begin{aligned} dX_\varepsilon(t) = & b\left(\frac{t}{\varepsilon}, X_\varepsilon(t-), \mathcal{L}_{X_\varepsilon(t)}\right) dt + \sigma\left(\frac{t}{\varepsilon}, X_\varepsilon(t-), \mathcal{L}_{X_\varepsilon(t)}\right) dW(t) \\ & + \int_U h\left(\frac{t}{\varepsilon}, X_\varepsilon(t-), \mathcal{L}_{X_\varepsilon(t)}, z\right) \tilde{N}(dt, dz), \quad X_\varepsilon(0) = x_0, \end{aligned} \quad (1.1)$$

for $t \in [0, T]$, with a small parameter $\varepsilon > 0$. Here, $\mathcal{L}_{X(t)}$ denotes the law of $X(t)$ at time t , and $W(t)$ is an m -dimensional standard Wiener process defined on the complete probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$, with $(\mathcal{F}_t)_{t \geq 0}$ satisfying the usual conditions. Let $(U, \mathcal{U}, \nu)$ be a σ -finite measure space with $U \subseteq \mathbb{R}^d \setminus \{0\}$, and let $N(dt, dz)$ be a Poisson random measure on $\mathbb{R}^+ \times U$ with intensity measure $\nu(dz)dt$, independent of $W(t)$. The compensated Poisson random measure is given by $\tilde{N}(dt, dz) = N(dt, dz) - \nu(dz)dt$. The precise assumptions on the coefficients $b: [0, T] \times \mathbb{R}^d \times M_2(\mathbb{R}^d) \rightarrow \mathbb{R}^d$, $\sigma: [0, T] \times \mathbb{R}^d \times M_2(\mathbb{R}^d) \rightarrow \mathbb{R}^{d \times m}$, and $h: [0, T] \times \mathbb{R}^d \times M_2(\mathbb{R}^d) \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ will be specified in later sections (see Section 2 for the definition of $M_2(\mathbb{R}^d)$). We also remark that $X_\varepsilon(t-)$ is the left limit at the point t , i.e. $X_\varepsilon(t-) = \lim_{s \uparrow t} X_\varepsilon(s)$.

The first aim of this paper is to consider the well-posedness of the McKean–Vlasov SDEs in the form of (1.1). Let us briefly review some previous works on the well-posedness of McKean–Vlasov SDEs with Brownian noise. Under the globally Lipschitz condition, the existence and uniqueness of strong solutions for McKean–Vlasov SDEs were obtained by using the fixed-point theorem, for example, in [2, 5]. Results for the case with a one-sided globally Lipschitz drift term and a globally Lipschitz diffusion term can be found in [10, 40]. To deal with the situation where the coefficients are locally Lipschitz with respect to (w.r.t.) the measure and globally Lipschitz w.r.t. the state variable, Kloeden and Lorenz [20] developed a method for constructing interpolated Eulerlike approximations. Recently, an extension to locally Lipschitz conditions w.r.t. the state variable under a uniform linear growth assumption was studied by Li *et al.* [22]; see also [9]. Moreover, Hong *et al.* [16] examined the strong and weak well-posedness of a class of McKean–Vlasov SDEs with the drift and diffusion coefficients fulfilling certain locally monotone conditions, whereas they need to impose additional structural assumptions on the coefficients to ensure a unique solution.

Unlike the case of Brownian noise, the study of McKean–Vlasov SDEs with Lévy noise is still in its infancy, although some interesting works are emerging [12–14, 17, 28]. In particular, Hao *et al.* [15] investigated a class of Lévy-type McKean–Vlasov SDEs satisfying global Lipschitz and linear growth conditions, established the existence and uniqueness of solutions, and explored their intrinsic link with nonlocal Fokker–Planck equations. The well-posedness results have been further developed for the case of superlinear drift, diffusion, and jump coefficients using the fixed-point theorem [28, 30]. Recently, Cavallazzi [6] has proven the strong well-posedness of McKean–Vlasov SDEs driven by Lévy process having a finite moment of order $\beta \in [1, 2]$ and under standard Lipschitz assumptions on the coefficients.

Motivated by previous works on the Brownian case as well as the Lévy case, in this paper, we aim to treat (1.1) only imposing locally Lipschitz conditions w.r.t. the state variable, allowing for a possibly superlinearly growing drift. We highlight that several essential difficulties arise. On the one hand, compared with classical SDEs, standard localization arguments cannot be applied directly due to the distribution-dependent coefficients. On the other hand, the non-Gaussian Lévy noise introduces challenges in both analytic and probabilistic aspects. Therefore, the results for classical SDEs (even with Lévy noise) or McKean–Vlasov SDEs with Brownian noise cannot be extended directly to McKean–Vlasov SDEs with Lévy noise. In this paper, we develop a Lévy-type technique of Eulerlike approximations to overcome the difficulties caused by the local conditions and distribution dependency. The crux of our method, which differs from the Brownian case [20, 22], lies in handling the drift terms under more general conditions as well as the jump terms.

Apart from the existence and uniqueness of solutions, we are further interested in establishing a stochastic averaging principle for (1.1) with drifts of polynomial growth under locally Lipschitz conditions w.r.t. the state variable. In fact, the averaging principle is a powerful method for extracting effective dynamics from complex systems arising in mechanics, mathematics, and other research areas. Since the pioneering work of Khasminskii [19], the averaging principle for usual SDEs has received significant attention and has stimulated much of the study in controls, stability analyses, and optimization methods. Although the problems considered take different forms (usually classified in terms of the noise or the conditions satisfied by their nonlinear terms), the essence behind the averaging method is to simplify dynamical systems and obtain approximate solutions to differential equations; see, e.g., [24, 33, 42]. Based on the idea of stochastic averaging, the second goal of this paper is to show that the solution of (1.1) converges to the following averaged equation (with $\tilde{X}(0) = x_0$) as ε tends to 0:

$$d\tilde{X}(t) = \bar{b}(\tilde{X}(t-), \mathcal{L}_{\tilde{X}(t)}) dt + \bar{\sigma}(\tilde{X}(t-), \mathcal{L}_{\tilde{X}(t)}) dW(t) + \int_U \bar{h}(\tilde{X}(t-), \mathcal{L}_{\tilde{X}(t)}, z) \tilde{N}(dt, dz), \quad (1.2)$$

in a certain sense, under appropriate averaging conditions. Here, $\bar{b}: \mathbb{R}^d \times M_2(\mathbb{R}^d) \rightarrow \mathbb{R}^d$, $\bar{\sigma}: \mathbb{R}^d \times M_2(\mathbb{R}^d) \rightarrow \mathbb{R}^{d \times m}$, and $\bar{h}: \mathbb{R}^d \times M_2(\mathbb{R}^d) \times U \rightarrow \mathbb{R}^d$ are Borel measurable functions. For more details on (1.2), see Section 3.

Again, we must point out that, compared with the case of classical SDEs, there are far fewer results on the averaging principle for McKean–Vlasov SDEs due to their distribution-dependent feature. Moreover, the existing studies on averaging principles for McKean–Vlasov SDEs primarily focus on the Brownian case [36, 41]. For some interesting results involving other types of noise, e.g., fractional Brownian noise, we refer to [37]. Nevertheless, to the best of the authors' knowledge, the averaging principle for McKean–Vlasov SDEs with Lévy noise has not yet been considered to date. This inspires us to establish an averaging principle.

The real-life applications of the Lévy-type McKean–Vlasov SDE (1.1) and its corresponding averaged equation (1.2) are not explored in this paper. Instead, we present an illustrative toy model in Example 4.1 and refer to [2, 5, 28] for discussions on potential applications of McKean–Vlasov SDEs with weak coefficient conditions in fields such as physics, finance, and population dynamics. To numerically approximate a solution of the McKean–Vlasov SDE in our setting, it is necessary to introduce an interacting particle system that is connected to the McKean–Vlasov SDE and is shown to converge to the true solution of the McKean–Vlasov SDE. This is popularly known as the propagation of chaos [39]. We present such a result in Appendix B. For more recent progress on propagation of chaos for jump processes, we refer to [6, 12–14, 17, 28, 30], and the references therein.

The rest of this paper is arranged as follows. In Section 2, we focus on investigating the existence and uniqueness of solutions for a class of McKean–Vlasov SDEs with Lévy-type perturbations. In Section 3, we prove an averaging principle for the solutions of the considered McKean–Vlasov SDEs. In Section 4, we present a specific example to illustrate the theoretical results of this paper. The details of the proof of Lemma 4 and the propagation of chaos result are postponed to Appendix B.

2. Well-posedness of Lévy-type McKean–Vlasov SDEs

We start with some notations used in the sequel. Let $|\cdot|$ and $\langle \cdot, \cdot \rangle$ be the Euclidean vector norm and the scalar product in $\mathbb{R}^d$, respectively. For a matrix A , we use the Frobenius norm defined as $\|A\| = \sqrt{\text{tr}[AA^T]}$, where A^T represents the transpose of the matrix A . Let $\mathcal{M}(\mathbb{R}^d)$ denote the space of all probability measures on $\mathbb{R}^d$ carrying the usual topology of weak convergence. Furthermore, for $p \geq 1$, let $\mathcal{M}_p(\mathbb{R}^d)$ represent the subspace of $\mathcal{M}(\mathbb{R}^d)$ as follows:

$$\mathcal{M}_p(\mathbb{R}^d) := \left\{ \mu \in \mathcal{M}(\mathbb{R}^d) : \mu(|\cdot|^p) := \int_{\mathbb{R}^d} |x|^p \mu(dx) < \infty \right\}.$$

For $\mu_1, \mu_2 \in \mathcal{M}_p(\mathbb{R}^d)$, the L^p -Wasserstein metric between μ_1 and μ_2 is defined as

$$W_p(\mu_1, \mu_2) := \inf_{\pi \in \mathcal{C}(\mu_1, \mu_2)} \left(\int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p \pi(dx, dy) \right)^{\frac{1}{p}},$$

where $\mathcal{C}(\mu_1, \mu_2)$ means the collection of all the probability measures whose marginal distributions are μ_1, μ_2 , respectively. Then $\mathcal{M}_p(\mathbb{R}^d)$ endowed with the above metric is a Polish space.

Let δ_x be the Dirac delta measure centered at the point $x \in \mathbb{R}^d$. A direct calculation shows that δ_x belongs to $\mathcal{M}_p(\mathbb{R}^d)$ for any $x \in \mathbb{R}^d$. Another remark is that if $\mu_1 = \mathcal{L}_X$ and $\mu_2 = \mathcal{L}_Y$ are the distributions of the random variables X and Y , respectively, then

$$(W_p(\mu_1, \mu_2))^p \leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p \mathcal{L}_{(X, Y)}(dx, dy) = \mathbb{E}|X - Y|^p,$$

where $\mathcal{L}_{(X, Y)}$ denotes the joint distribution of the random vector (X, Y) .

Given $T > 0$, let $D([0, T]; \mathbb{R}^d)$ be the collection of all càdlàg (i.e. right continuous with left limits) functions from $[0, T]$ to $\mathbb{R}^d$. Note that, at the endpoints of the closed interval $[0, T]$, we stipulate that an element in $D([0, T]; \mathbb{R}^d)$ is right continuous at 0 and has a left limit at T , respectively. For $1 \leq p < \infty$, we use $L^p(\Omega; \mathbb{R}^d)$ to denote the family of all $\mathbb{R}^d$ -valued random variables Y such that $\mathbb{E}|Y|^p < \infty$. Similarly, we denote by $L^p(\Omega; D([0, T]; \mathbb{R}^d))$ the subspace of all $D([0, T]; \mathbb{R}^d)$ -valued random variables X that satisfy $\mathbb{E}[\sup_{0 \leq t \leq T} |X(t)|^p] < \infty$. Then, we present the following proposition.

Proposition 1.

- (1) The space $D([0, T]; \mathbb{R}^d)$, equipped with the supremum norm, is a Banach space.
- (2) Let $p \in [1, \infty)$. The space $L^p(\Omega; D([0, T]; \mathbb{R}^d))$, equipped with the norm $\|X\|_{L^p} := (\mathbb{E}[\sup_{0 \leq t \leq T} |X(t)|^p])^{\frac{1}{p}}$, is also a Banach space.

Proof. (i) The proof is primarily based on the properties of càdlàg functions, as outlined on p. 140 of [1]. Let $B([0, T]; \mathbb{R}^d)$ denote the space of bounded functions from $[0, T]$

to $\mathbb{R}^d$. It is important to note that $B([0, T]; \mathbb{R}^d)$, when equipped with the supremum norm, is a Banach space [1, p. 6]. Referring to the property (4) in [1, p. 140], it follows that $D([0, T]; \mathbb{R}^d) \subset B([0, T]; \mathbb{R}^d)$. Hence, the Cauchy sequence $\{f_n\}$ of functions in $D([0, T]; \mathbb{R}^d)$ converges uniformly to some bounded function $f \in B([0, T]; \mathbb{R}^d)$. The desired result is then obtained by applying the property (6) in [1, p. 140], which states that the limit of a sequence of càdlàg functions on $[0, T]$ is itself càdlàg. (ii) The result follows directly from [32, Theorem 2.23 and Example 2.25], which establish the general completeness of L^p -spaces over Banach-valued random variables. For further references, see also [3, Theorem 4.1.3] and [4, Theorem 4.8 and Comment 4 in Chapter 4]. $\square$

We recall several useful inequalities that will be employed frequently throughout this paper. The first is Young's inequality, stated as

$$ab \leq \epsilon \frac{a^p}{p} + \epsilon^{-\frac{q}{p}} \frac{b^q}{q}, \quad \text{for all } \epsilon, a, b > 0, \text{ where } p > 1, \frac{1}{p} + \frac{1}{q} = 1. \quad (2.1)$$

Next, we list two elementary inequalities:

$$\left(\sum_{i=1}^k |a_i| \right)^l \leq \left(k \max_{1 \leq i \leq k} |a_i| \right)^l \leq k^l \sum_{i=1}^k |a_i|^l, \quad \text{for all } l > 0, a_i \in \mathbb{R}, k \in \mathbb{N}, \quad (2.2)$$

and

$$(a + b + c)^l \leq 3^{l-1}(|a|^l + |b|^l + |c|^l), \quad \text{for all } l \geq 1, a, b, c \in \mathbb{R}. \quad (2.3)$$

In addition, noting that stochastic integrals w.r.t. compensated Poisson random measures are local martingales, we require the following preparatory results to proceed with the analysis.

Proposition 2. Let $H: [0, T] \times U \rightarrow \mathbb{R}^d$ be a Borel measurable function satisfying $\int_0^t \int_U |H(s, z)|^2 \nu(dz) ds < \infty$, almost surely. Define the stochastic integral $I_t := \int_0^t \int_U H(s, z) \tilde{N}(ds, dz)$. Then, the following estimates hold.

(i) For any $p \geq 2$ and $0 \leq t \leq T$, there exists a constant $D_p > 0$ such that

$$\begin{aligned} \mathbb{E} \left(\sup_{0 \leq s \leq t} |I_s|^p \right) &\leq D_p \mathbb{E} \left[\left(\int_0^t \int_U |H(s, z)|^2 \nu(dz) ds \right)^{p/2} \right] \\ &\quad + D_p \mathbb{E} \left[\int_0^t \int_U |H(s, z)|^p \nu(dz) ds \right]. \end{aligned} \quad (2.4)$$

This result is commonly referred to as Kunita's first inequality [1, Theorem 4.4.23].

(ii) For any $1 \leq p \leq 2$ and $0 \leq t \leq T$, there exists a constant $K_p > 0$ such that

$$\mathbb{E} \left(\sup_{0 \leq s \leq t} |I_s|^p \right) \leq K_p \mathbb{E} \left[\left(\int_0^t \int_U |H(s, z)|^2 \nu(dz) ds \right)^{p/2} \right]. \quad (2.5)$$

Proof. We emphasize that this proposition can be viewed as a special instance of Novikov's result, which is rigorously established in [31, Theorem 1]; see also [21, Theorem 4.20] for applications of Novikov's result and its relation to variants of the Burkholder–Davis–Gundy (BDG) inequality. For the case $p \geq 2$, a proof of the conclusion (i) based on the BDG inequality

for local martingales is presented in [29, Lemma 1], whereas an alternative approach utilizing Itô's formula (applied to $x \mapsto x^p$) and Doob's martingale inequality can be found in [1, Theorem 4.4.23]. For the case $1 \leq p \leq 2$, the conclusion (ii) is stated in [7, Lemma 2.1] without proof. To ensure clarity for readers and maintain mathematical rigor, we provide a detailed proof for this case here. For convenience, we define the processes

$$A_t := \int_0^t \int_U |H(s, z)|^2 \nu(dz) ds$$

and

$$J_t := \int_0^t \int_U (A_s + \varepsilon)^{-\frac{2-p}{4}} H(s, z) \tilde{N}(ds, dz)$$

for $0 \leq t \leq T$, where $1 \leq p \leq 2$ and $\varepsilon > 0$ is a small parameter.

On the one hand, applying the integration by parts formula (also referred to Itô's product formula; see [1, Theorem 4.4.13]), we obtain

$$J_t \cdot (A_t + \varepsilon)^{\frac{2-p}{4}} = \int_0^t J_s d \left[(A_s + \varepsilon)^{\frac{2-p}{4}} \right] + \int_0^t \int_U H(s, z) \tilde{N}(ds, dz).$$

Noting that $(A_t + \varepsilon)^{\frac{2-p}{4}}$ is a nonnegative and nondecreasing process, we deduce the bound

$$|I_t| \leq |J_t| (A_t + \varepsilon)^{\frac{2-p}{4}} + \int_0^t |J_s| d \left[(A_s + \varepsilon)^{\frac{2-p}{4}} \right] \leq 2 \left(\sup_{0 \leq s \leq t} |J_s| \right) (A_t + \varepsilon)^{\frac{2-p}{4}}.$$

Since this estimate holds for all $t \geq 0$ and the right-hand side remains nondecreasing, it follows that

$$\sup_{0 \leq s \leq t} |I_s|^p \leq 2^p \left(\sup_{0 \leq s \leq t} |J_s| \right)^p (A_t + \varepsilon)^{\frac{(2-p)p}{4}}.$$

By Hölder inequality [26, p. 5], we further derive

$$\begin{aligned} \mathbb{E} \left(\sup_{0 \leq s \leq t} |I_s|^p \right) &\leq 2^p \mathbb{E} \left[\left(\sup_{0 \leq s \leq t} |J_s| \right)^p (A_t + \varepsilon)^{\frac{(2-p)p}{4}} \right] \\ &\leq 2^p \left[\mathbb{E} \left(\sup_{0 \leq s \leq t} |J_s| \right)^{p \cdot \frac{2}{p}} \right]^{\frac{p}{2}} \left[\mathbb{E} (A_t + \varepsilon)^{\frac{p(2-p)}{4} \cdot \frac{2}{2-p}} \right]^{\frac{2-p}{2}} \\ &= 2^p \left[\mathbb{E} \left(\sup_{0 \leq s \leq t} |J_s| \right)^2 \right]^{\frac{p}{2}} \left[\mathbb{E} (A_t + \varepsilon)^{\frac{p}{2}} \right]^{\frac{2-p}{2}}. \end{aligned} \quad (2.6)$$

On the other hand, applying the Itô isometry for integrals with respect to compensated Poisson random measures yields

$$\mathbb{E} |J_t|^2 = \mathbb{E} \int_0^t \int_U (A_s + \varepsilon)^{\frac{p-2}{2}} |H(s, z)|^2 \nu(dz) ds = \mathbb{E} \int_0^t (A_s + \varepsilon)^{\frac{p-2}{2}} dA_s \leq \frac{2}{p} \mathbb{E} (A_t + \varepsilon)^{\frac{p}{2}}.$$

By Doob's martingale inequality [1, Theorem 2.1.5], we then obtain

$$\mathbb{E} \left(\sup_{0 \leq s \leq t} |J_s|^2 \right) \leq 4 \mathbb{E} |J_t|^2 \leq \frac{8}{p} \mathbb{E} (A_t + \varepsilon)^{\frac{p}{2}}. \quad (2.7)$$

Combining (2.6) and (2.7), and using the fact that $\mathbb{E} (\sup_{0 \leq s \leq t} |J_s|)^2 = \mathbb{E} (\sup_{0 \leq s \leq t} |J_s|^2)$, we arrive at

$$\mathbb{E} \left(\sup_{0 \leq s \leq t} |I_s|^p \right) \leq \left(\frac{32}{p} \right)^{\frac{p}{2}} \mathbb{E} [\varepsilon + A_t]^{\frac{p}{2}}.$$

The result follows by letting $\varepsilon \rightarrow 0$. This completes the proof. $\square$

2.1. Formulation of the well-posedness results

This section is dedicated to establishing the existence and uniqueness theorem for the solutions of the d -dimensional Lévy-type McKean–Vlasov SDEs, i.e.

$$\begin{aligned} dX(t) = & b(t, X(t-), \mathcal{L}_{X(t)}) dt + \sigma(t, X(t-), \mathcal{L}_{X(t)}) dW(t) \\ & + \int_U h(t, X(t-), \mathcal{L}_{X(t)}, z) \tilde{N}(dt, dz) \end{aligned} \quad (2.8)$$

for $t \in [0, T]$ with initial condition $X(0) = x_0$. The functions b , σ , and h are defined as follows:

$$\begin{aligned} b: [0, T] \times \mathbb{R}^d \times M(\mathbb{R}^d) &\rightarrow \mathbb{R}^d, \quad \sigma: [0, T] \times \mathbb{R}^d \times M(\mathbb{R}^d) \rightarrow \mathbb{R}^{d \times m}, \\ h: [0, T] \times \mathbb{R}^d \times M(\mathbb{R}^d) \times U &\rightarrow \mathbb{R}^d, \end{aligned}$$

where b , σ , and h are Borel measurable functions. We now proceed by providing the precise definition of a solution to (2.8).

Definition 1. We say that (2.8) admits a unique strong solution if there exists an $\{\mathcal{F}_t\}_{0 \leq t \leq T}$ -adapted $\mathbb{R}^d$ -valued càdlàg stochastic process $(X(t))_{t \in [0, T]}$ such that

- (i) $X(t) = x_0 + \int_0^t b(s, X(s-), \mathcal{L}_{X(s)}) ds + \sigma(s, X(s-), \mathcal{L}_{X(s)}) dW(s) + \int_U h(s, X(s-), \mathcal{L}_{X(s)}, z) \tilde{N}(ds, dz)$, $t \in [0, T]$, $\mathbb{P}$ -almost surely;
- (ii) if $Y = (Y(t))_{t \in [0, T]}$ is another solution with $Y(0) = x_0$, then $\mathbb{P}(X(t) = Y(t) \text{ for all } t \in [0, T]) = 1$.

Assume that there exists a constant $\kappa \geq 2$ such that the following assumptions hold.

Assumption 1. (One-sided locally Lipschitz condition on the state variable.) For every $R > 0$, there exists a positive constant L_R such that for any $t \in [0, T]$, $x, y \in \mathbb{R}^d$ with $|x| \vee |y| \leq R$, and $\mu \in \mathcal{M}_2(\mathbb{R}^d)$,

$$\begin{aligned} & \langle x - y, b(t, x, \mu) - b(t, y, \mu) \rangle \vee \|\sigma(t, x, \mu) - \sigma(t, y, \mu)\|^2 \\ & \vee \int_U |h(t, x, \mu, z) - h(t, y, \mu, z)|^2 \nu(dz) \leq L_R |x - y|^2. \end{aligned}$$

Here, the symbol ' $\vee$ ' denotes the maximum of the multiple terms.

Assumption 2. (Globally 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 $\mu_1, \mu_2 \in \mathcal{M}_2(\mathbb{R}^d)$,

$$|b(t, x, \mu_1) - b(t, x, \mu_2)|^2 + \|\sigma(t, x, \mu_1) - \sigma(t, x, \mu_2)\|^2 \\ + \int_U |h(t, x, \mu, z) - h(t, y, \mu, z)|^2 \nu(dz) \leq LW_2^2(\mu_1, \mu_2).$$

Assumption 3. (Continuity.) For any $t \in [0, T]$, $b(t, \cdot, \cdot)$, $\sigma(t, \cdot, \cdot)$, and $\int_U h(t, \cdot, \cdot, z) \nu(dz)$ are continuous on $\mathbb{R}^d \times \mathcal{M}_2(\mathbb{R}^d)$.

Assumption 4. (One-sided linear and global linear growth condition.) There exists a positive constant K such that, for any $t \in [0, T]$, $x \in \mathbb{R}^d$, and $\mu \in \mathcal{M}_2(\mathbb{R}^d)$,

$$\langle x, b(t, x, \mu) \rangle \vee \|\sigma(t, x, \mu)\|^2 \vee \int_U |h(t, x, \mu, z)|^2 \nu(dz) \leq K(1 + |x|^2 + W_2^2(\mu, \delta_0)).$$

Assumption 5. (κ -order growth condition on the drift coefficient.) There exists a positive constant K_1 such that, for any $t \in [0, T]$, $x \in \mathbb{R}^d$, and $\mu \in \mathcal{M}_2(\mathbb{R}^d)$,

$$|b(t, x, \mu)|^2 \leq K_1(1 + |x|^\kappa + W_2^\kappa(\mu, \delta_0)).$$

Assumption 6. (r -order moment condition for the initial data.) Consider $x_0 \in L^r(\Omega; \mathbb{R}^d)$ for some $r \geq \max\{\kappa^2/2, 4\}$, i.e. $\mathbb{E}|x_0|^r < \infty$.

Assumption 7. (Additional growth conditions and Lipschitz type conditions on the jump coefficient h .) There exists a positive K_2 such that, for any $t \in [0, T]$, $x \in \mathbb{R}^d$, and $\mu \in \mathcal{M}_2(\mathbb{R}^d)$,

$$\int_U |h(t, x, \mu, z)|^r \nu(dz) \leq K_2(1 + |x|^r + W_2^r(\mu, \delta_0)).$$

In addition, if $\kappa > 2$, there exist constants $K_3, L' > 0$ such that, for any $t \in [0, T]$, $x, y \in \mathbb{R}^d$, and $\mu, \mu_1, \mu_2 \in \mathcal{M}_2(\mathbb{R}^d)$,

$$\int_U |h(t, x, \mu, z)|^\kappa \nu(dz) \leq K_3(1 + |x|^\kappa + W_2^\kappa(\mu, \delta_0)), \\ \int_U |h(t, x, \mu_1, z) - h(t, x, \mu_2, z)|^\kappa \nu(dz) \leq L' W_2^\kappa(\mu_1, \mu_2),$$

and for every $R > 0$, there exists a constant $L'_R > 0$ such that for any $t \in [0, T]$, $x, y \in \mathbb{R}^d$ with $|x| \vee |y| \leq R$, and $\mu \in \mathcal{M}_2(\mathbb{R}^d)$,

$$\int_U |h(t, x, \mu, z) - h(t, y, \mu, z)|^\kappa \nu(dz) \leq L'_R |x - y|^\kappa.$$

The main result of this section is stated as follows.

Theorem 1. (Well-posedness.) Let Assumptions 1–7 be satisfied. Then (2.8) admits a unique strong solution $(X(t))_{t \in [0, T]} \in L^\kappa(\Omega; \mathbb{R}^d)$ with the initial value $X(0) = x_0$, where $\kappa \geq 2$. Moreover, the following estimate holds:

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} |X(t)|^r \right] \leq C, \quad (2.9)$$

where $C := C(T, r, \mathbb{E}|x_0|^r)$ is a positive constant. Here, $r \geq \max\{\kappa^2/2, 4\}$.

Remark 1. We emphasize that the conditions in Assumptions 1–7 are carefully chosen, and the results in Theorem 1 are broadly applicable.

- (i) The one-sided locally Lipschitz condition in Assumption 1 is weaker than the classical locally Lipschitz condition. In fact, it is clear that the locally Lipschitz condition implies the one-sided locally Lipschitz condition (via the mean value inequality). However, the converse is false. For example, consider $b(t, x, \mu) = x^3 - x^{\frac{1}{3}} + t + \int_{\mathbb{R}} z\mu(dz)$ in $\mathbb{R}$. For $|x| \vee |y| \leq R$, we have

$$\begin{aligned} \langle x - y, b(t, x, \mu) - b(t, y, \mu) \rangle &= |x - y|^2 \left(x^2 + xy + y^2 \right) - (x - y) \left(x^{\frac{1}{3}} - y^{\frac{1}{3}} \right) \\ &\leq 3R^2 |x - y|^2, \end{aligned}$$

since $(x - y)(x^{\frac{1}{3}} - y^{\frac{1}{3}}) \geq 0$ for all x, y . Thus, b is one-sided locally Lipschitz but not locally Lipschitz.

- (ii) In contrast to the one-sided (globally) Lipschitz condition in the recent paper [30], which asserts that there exists a constant $C > 0$ such that for any $x, y \in \mathbb{R}^d$ and $\mu \in \mathcal{M}_2(\mathbb{R}^d)$,

$$\begin{aligned} \langle x - y, b(t, x, \mu) - b(t, y, \mu) \rangle + \|\sigma(t, x, \mu) - \sigma(t, y, \mu)\|^2 \\ + \int_U |h(t, x, \mu, z) - h(t, y, \mu, z)|^2 \nu(dz) \leq C|x - y|^2, \end{aligned}$$

the one-sided locally Lipschitz condition in Assumption 1 is expressed using the operation ‘ $\vee$ ’ instead of ‘ $+$ ’. This makes the condition in Assumption 1 weaker in some cases. For instance, consider b as a one-sided locally Lipschitz function and $\sigma = h = x$ with $\nu(U) < \infty$. In this case, Assumption 1 holds, but the one-sided (globally) Lipschitz condition in [30] is not satisfied.

- (iii) The result simplifies to the case of pure Brownian motion when $h \equiv 0$. In contrast to the Brownian motion model considered in [22], where the drift coefficient is required to satisfy a linear growth condition, the present framework imposes only a one-sided linear growth condition on the drift coefficient b . Furthermore, b is permitted to exhibit polynomial growth w.r.t. the state variable, as specified in Assumptions 4 and 5.
- (iv) Referring to [35, Theorem 25.3], when the jump coefficient h is a submultiplicative function with respect to z , the growth conditions in Assumption 7 can be interpreted as requiring that the jump measure $[\nu]_U$ has a bounded $|h|^r$ -moment or, equivalently, the associated Lévy motion has bounded $|h|^r$ -moments, for every $t \in [0, T]$, $x \in \mathbb{R}^d$, and $\mu \in \mathcal{M}_2(\mathbb{R}^d)$, where $r \geq \max\{\kappa^2/2, 4\}$ and $\kappa \geq 2$. In particular, the associated Lévy motion can be said to have bounded r -order moments when $h(t, x, \mu, z) = z$. We remark that while Brownian motion can be considered a 2-stable Lévy motion, our assumptions exclude applications involving jump measures associated with α -stable Lévy motions for $0 < \alpha < 2$. This exclusion arises because such α -stable Lévy motions process r -order moments only for $r < \alpha$; see [35, Example 25.10]. For a recent study addressing the strong well-posedness of McKean–Vlasov SDEs driven by Lévy noise with finite moments of order $\beta \in [1, 2]$, we refer to [6]. However, it should be pointed out that the assumptions in [6] regarding the coefficients with respect to both the space variable and the measure remain within the globally Lipschitz framework.

2.2. Euler-type approximation and auxiliary lemmas

A key aspect of our approach to proving Theorem 1 is the construction of an Eulerlike sequence for the McKean–Vlasov SDEs (2.8). Once we demonstrate that this sequence is Cauchy in an appropriate complete space (specifically, $L^k(\Omega; D([0, T]; \mathbb{R}^d))$, as is shown later), we can conclude that there exists a limiting process, which is indeed the desired solution to (2.8).

To this end, let $T > 0$ be given, and consider the equidistant partition of the interval $[0, T]$. For any integer $n \geq 1$, define $h_n = \frac{T}{n}$ and $t_k^n = kh_n$, $k = 0, 1, \dots, n$. For a fixed k ($0 \leq k \leq n-1$) and $t \in (t_k^n, t_{k+1}^n]$, we analyze the following approximation:

$$\begin{aligned} dX^{(n)}(t) = & b\left(t, X^{(n)}(t-), \mu_{t_k^n}^{(n)}\right) dt + \sigma\left(t, X^{(n)}(t-), \mu_{t_k^n}^{(n)}\right) dW(t) \\ & + \int_U h\left(t, X^{(n)}(t-), \mu_{t_k^n}^{(n)}, z\right) \tilde{N}(dt, dz), \end{aligned} \quad (2.10)$$

where $\mu_{t_k^n}^{(n)} = \mathcal{L}_{X^{(n)}(t_k^n)}$ denotes the law of $X^{(n)}(t_k^n)$. Observe that for each fixed k , if the initial value $X^{(n)}(t_k^n)$ and the distribution $\mu_{t_k^n}^{(n)}$ (at the left endpoint t_k^n) are known, then (2.10) reduces to a standard SDE that is independent of the law of $X^{(n)}(t)$. We now establish, by induction, the existence and uniqueness of the solution to (2.10).

In fact, for $k=0$ and $t \in [0, t_1^n]$, the distribution is $\mu_0^{(n)} = \mathcal{L}_{X^{(n)}(0)} = \mathcal{L}_{x_0}$. Applying Assumptions 1 and 4, we observe that the coefficients in (2.10) (with $k=0$) satisfy

$$\begin{aligned} & \left\langle x - y, b\left(t, x, \mu_0^{(n)}\right) - b\left(t, y, \mu_0^{(n)}\right) \right\rangle + \left\| \sigma\left(t, x, \mu_0^{(n)}\right) - \sigma\left(t, y, \mu_0^{(n)}\right) \right\|^2 \\ & + \int_U \left| h\left(t, x, \mu_0^{(n)}, z\right) - h\left(t, y, \mu_0^{(n)}, z\right) \right|^2 \nu(dz) \leq 3L_R |x - y|^2 \end{aligned}$$

and

$$\begin{aligned} & \left\langle x, b\left(t, x, \mu_0^{(n)}\right) \right\rangle + \left\| \sigma\left(t, x, \mu_0^{(n)}\right) \right\|^2 + \int_U \left| h\left(t, x, \mu_0^{(n)}, z\right) \right|^2 \nu(dz) \\ & \leq 3K \left(1 + |x|^2 + W_2^2\left(\mu_0^{(n)}, \delta_0\right)\right) \leq 3K \left(1 + |x|^2\right) \left(1 + \mathbb{E} \left|X^{(n)}(0)\right|^2\right). \end{aligned}$$

Referring to [25, Theorem 1.1], it admits a unique solution on $[0, t_1^n]$. Furthermore, by Assumption 5, it follows that for $r \geq \max\{\frac{\kappa^2}{2}, 4\}$, there exists a positive constant C such that

$$\mathbb{E} \left[\sup_{0 \leq t \leq t_1^n} \left| X^{(n)}(t) \right|^r \right] \leq C \left(1 + \mathbb{E} \left| X^{(n)}(0) \right|^r\right),$$

whose proof is quite similar to Lemma 1, and we omit the details here. Therefore, we can define $X^{(n)}(t_1^n)$ (which satisfies $\mathbb{E} |X^{(n)}(t_1^n)|^r < \infty$) and $\mu_{t_1^n}^{(n)} = \mathcal{L}_{X^{(n)}(t_1^n)}$.

For $k=1$ and $t \in (t_1^n, t_2^n]$, we can use $(X^{(n)}(t_1^n), \mu_{t_1^n}^{(n)})$ in place of $(X^{(n)}(0), \mu_0^{(n)})$ and repeat the above procedure. Inductively, for any $k=0, 1, \dots, n-1$ and $t \in (t_k^n, t_{k+1}^n]$, we obtain the existence and uniqueness of the solution to the SDE (2.10) as well as the corresponding estimate

$$\mathbb{E} \left[\sup_{t_k^n \leq t \leq t_{k+1}^n} \left| X^{(n)}(t) \right|^r \right] \leq C \left(1 + \mathbb{E} \left| X^{(n)}(t_k^n) \right|^r\right), \quad (2.11)$$

by similar arguments.

At this point, we define by $[t]_n = t_k^n$ for all $t \in (t_k^n, t_{k+1}^n]$, where $k = 0, 1, \dots, n-1$. Then, for $t \in [0, T]$, we introduce the following approximating SDE

$$\begin{aligned} dX^{(n)}(t) &= b\left(t, X^{(n)}(t-), \mu_{[t]_n}^{(n)}\right) dt + \sigma\left(t, X^{(n)}(t-), \mu_{[t]_n}^{(n)}\right) dW(t) \\ &\quad + \int_U h\left(t, X^{(n)}(t-), \mu_{[t]_n}^{(n)}, z\right) \tilde{N}(dt, dz), \end{aligned} \quad (2.12)$$

with the initial value $X^{(n)}(0) = x_0$, where $\mu_{[t]_n}^{(n)} = \mathcal{L}_{X^{(n)}([t]_n)}$. According to the previously presented procedures and results for (2.10), we conclude that there exists a unique solution to (2.12). In fact, for each $n \geq 1$ and $t \in [0, T]$, we can always find a certain k_* ($0 \leq k_* \leq n-1$) such that $t \in (t_{k_*}^n, t_{k_*+1}^n]$. Then, the solution to (2.12) can be written as

$$\begin{aligned} X^n(t) &= x_0 + \sum_{k=0}^{k_*} \int_{t_k^n}^{t_{k+1}^n \wedge t} b\left(s, X^{(n)}(s-), \mu_{t_k^n}^{(n)}\right) ds + \sigma\left(s, X^{(n)}(s-), \mu_{t_k^n}^{(n)}\right) dW(s) \\ &\quad + \int_U h\left(s, X^{(n)}(s-), \mu_{t_k^n}^{(n)}, z\right) \tilde{N}(ds, dz) \\ &= X^n(t_1^n) + \sum_{k=1}^{k_*} \int_{t_k^n}^{t_{k+1}^n \wedge t} b\left(s, X^{(n)}(s-), \mu_{t_k^n}^{(n)}\right) ds + \sigma\left(s, X^{(n)}(s-), \mu_{t_k^n}^{(n)}\right) dW(s) \\ &\quad + \int_U h\left(s, X^{(n)}(s-), \mu_{t_k^n}^{(n)}, z\right) \tilde{N}(ds, dz) \\ &\quad \dots \dots \\ &= X^n(t_{k_*}^n) + \int_{t_{k_*}^n}^t b\left(s, X^{(n)}(s-), \mu_{t_{k_*}^n}^{(n)}\right) ds + \sigma\left(s, X^{(n)}(s-), \mu_{t_{k_*}^n}^{(n)}\right) dW(s) \\ &\quad + \int_U h\left(s, X^{(n)}(s-), \mu_{t_{k_*}^n}^{(n)}, z\right) \tilde{N}(ds, dz), \end{aligned}$$

and it is well-defined based on the results for (2.10) with $k = 0, 1, \dots, k_*$. Moreover, we have the following estimate

$$\begin{aligned} \mathbb{E} \left[\sup_{0 \leq t \leq T} |X^{(n)}(t)|^r \right] &= \mathbb{E} \left[\max_{0 \leq k \leq n-1} \sup_{t_k^n \leq t \leq t_{k+1}^n} |X^{(n)}(t)|^r \right] \leq \sum_{k=0}^{n-1} \mathbb{E} \left[\sup_{t_k^n \leq t \leq t_{k+1}^n} |X^{(n)}(t)|^r \right] \\ &\leq C(n) < \infty. \end{aligned} \quad (2.13)$$

Under Assumption 6, which requires that the initial data x_0 satisfies $\mathbb{E}|x_0|^r < \infty$ with $r \geq \max\{\frac{\kappa^2}{2}, 4\} \geq \kappa$, we deduce that $X^{(n)} \in L^r(\Omega; D([0, T]; \mathbb{R}^d)) \subset L^\kappa(\Omega; D([0, T]; \mathbb{R}^d))$. Hence, the stochastic processes $\{X^{(n)}(t)\}_{n \geq 1}$ given by (2.12) form a sequence in $L^\kappa(\Omega; D([0, T]; \mathbb{R}^d))$. To demonstrate that this sequence is Cauchy, we require the following two auxiliary lemmas.

Lemma 1. (Uniform boundedness property.) *Under Assumptions 4, 6, and 7, for any $T > 0$, there exists a positive constant C_r (independent of n) such that*

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} |X^{(n)}(t)|^r \right] \leq C_r. \quad (2.14)$$

Proof. For $r \geq \max\{\frac{K}{2}, 4\}$ and $t \in [0, T]$, applying Itô's formula [1, Theorem 4.4.7] to $|x|^r$, along with the identity $v(dz)dt = N(dt, dz) - \tilde{N}(dt, dz)$, yields that

$$\begin{aligned} |X^{(n)}(t)|^r &= |x_0|^r + r \int_0^t |X^{(n)}(s-)|^{r-2} \left\langle X^{(n)}(s-), b\left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)}\right) \right\rangle ds \\ &\quad + \frac{r}{2} \int_0^t |X^{(n)}(s-)|^{r-2} \left\| \sigma(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)}) \right\|^2 ds \\ &\quad + \frac{r(r-2)}{2} \int_0^t |X^{(n)}(s-)|^{r-4} \left\| (X^{(n)}(s-))^T \sigma\left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)}\right) \right\|^2 ds \\ &\quad + r \int_0^t |X^{(n)}(s-)|^{r-2} \left\langle X^{(n)}(s-), \sigma\left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)}\right) dW(s) \right\rangle \\ &\quad + r \int_0^t \int_U |X^{(n)}(s-)|^{r-2} \left\langle X^{(n)}(s-), h\left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)}, z\right) \right\rangle \tilde{N}(ds, dz) \\ &\quad + \int_0^t \int_U \left[|X^{(n)}(s-)|^r + h\left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)}, z\right) \right]^r \\ &\quad - |X^{(n)}(s-)|^r - r |X^{(n)}(s-)|^{r-2} \cdot \left\langle X^{(n)}(s-), h\left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)}, z\right) \right\rangle \right] N(ds, dz). \end{aligned} \quad (2.15)$$

By virtue of Assumption 4, Young's inequality (2.1) (with $\epsilon = 1$, $p = \frac{r}{r-2}$, and $q = \frac{r}{2}$), Hölder inequality and the elementary inequality (2.3) (with $l = \frac{r}{2}$), one can estimate the second term of (2.15) by

$$\begin{aligned} &rK \int_0^t |X^{(n)}(s-)|^{r-2} \left(1 + |X^{(n)}(s-)|^2 + \mathbb{E} |X^{(n)}([s]_n -)|^2 \right) ds \\ &\leq (r-2)K \int_0^t |X^{(n)}(s-)|^r ds + 2K \int_0^t \left(1 + |X^{(n)}(s-)|^2 + \mathbb{E} |X^{(n)}([s]_n -)|^2 \right)^{\frac{r}{2}} ds \\ &\leq (r-2)K \int_0^t |X^{(n)}(s-)|^r ds + 2 \cdot 3^{\frac{r}{2}-1} K \int_0^t \left(1 + |X^{(n)}(s-)|^r + \mathbb{E} |X^{(n)}([s]_n -)|^r \right) ds. \end{aligned} \quad (2.16)$$

Analogously, the third and fourth terms of (2.15) can be estimated by

$$\frac{r-2}{2} K \int_0^t |X^{(n)}(s-)|^r ds + 3^{\frac{r}{2}-1} K \int_0^t \left(1 + |X^{(n)}(s-)|^r + \mathbb{E} |X^{(n)}([s]_n -)|^r \right) ds, \quad (2.17)$$

and

$$\frac{(r-2)^2}{2} K \int_0^t |X^{(n)}(s-)|^r ds + 3^{\frac{r}{2}-1} (r-2) K \int_0^t \left(1 + |X^{(n)}(s-)|^r + \mathbb{E} |X^{(n)}([s]_n -)|^r \right) ds, \quad (2.18)$$

respectively. Furthermore, note that the map $y \rightarrow |y|^r$ is of class C^2 and the remainder formula for $|y|^r$ gives

$$\begin{aligned} |y|^r - |b|^r - r|b|^{r-2} \langle b, y-b \rangle &\leq C_1 \int_0^1 |y-b|^2 |b + \theta(y-b)|^{r-2} d\theta \\ &\leq C_1 (|b|^{r-2} |y-b|^2 + |y-b|^r), \end{aligned} \quad (2.19)$$

for any $y, b \in \mathbb{R}^d$. The last term on the right-hand side of (2.15) can thus be estimated as

$$C_1 \int_0^t \int_U \left(\left| X^{(n)}(s-) \right|^{r-2} \left| h\left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)}, z\right) \right|^2 + \left| h\left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)}, z\right) \right|^r \right) N(ds, dz). \quad (2.20)$$

Denote the above upper-bound (2.20) by N_t . Substituting (2.16)–(2.20) into (2.15), taking the supremum over $[0, u]$ for $u \in [0, T]$ and then taking expectations gives that

$$\begin{aligned} \mathbb{E} \sup_{0 \leq t \leq u} |X^{(n)}(t)|^r &\leq \mathbb{E}|x_0|^r + \mathbb{E} \sup_{0 \leq t \leq u} |M_t| + \mathbb{E} \sup_{0 \leq t \leq u} |N_t| \\ &\quad + \frac{(r+1)(r-2)}{2} K \int_0^u \mathbb{E} \left| X^{(n)}(s-) \right|^r ds \\ &\quad + 3^{\frac{r}{2}-1} (r+1) K \int_0^u \left(1 + \mathbb{E} \left| X^{(n)}(s-) \right|^r + \mathbb{E} \left| X^{(n)}([s]_n -) \right|^r \right) ds, \end{aligned} \quad (2.21)$$

where

$$\begin{aligned} M_t &:= r \int_0^t \left| X^{(n)}(s-) \right|^{r-2} \left\langle X^{(n)}(s-), \sigma\left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)}\right) dW(s) \right\rangle \\ &\quad + r \int_0^t \int_U \left| X^{(n)}(s-) \right|^{r-2} \left\langle X^{(n)}(s-), h\left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)}, z\right) \right\rangle \tilde{N}(ds, dz) \end{aligned}$$

is indeed a local martingale. On the one hand, by the BDG inequality (for the Brownian case) [26, Theorem 7.3 in Chapter 1] and the inequality (2.5) (with $p = 1$) in Proposition 2, there exists a constant $C_2 > 0$ such that

$$\begin{aligned} \mathbb{E} \sup_{0 \leq t \leq u} |M_t| &\leq C_2 r \mathbb{E} \left[\int_0^u \left| X^{(n)}(s-) \right|^{2r-2} \left\| \sigma\left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)}\right) \right\|^2 ds \right]^{\frac{1}{2}} \\ &\quad + C_2 r \mathbb{E} \left[\int_0^u \int_U \left| X^{(n)}(s-) \right|^{2r-2} \left| h\left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)}, z\right) \right|^2 \nu(dz) ds \right]^{\frac{1}{2}} \\ &\leq C_2 r \mathbb{E} \left\{ \sup_{0 < t \leq u} \left| X^{(n)}(t-) \right|^{r-1} \left[\left(\int_0^u \left\| \sigma\left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)}\right) \right\|^2 ds \right)^{\frac{1}{2}} \right. \right. \\ &\quad \left. \left. + \left(\int_0^u \int_U \left| h\left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)}, z\right) \right|^2 \nu(dz) ds \right)^{\frac{1}{2}} \right] \right\}. \end{aligned}$$

Applying Assumption 4 yields

$$\begin{aligned} \mathbb{E} \sup_{0 \leq t \leq u} |M_t| &\leq C_2 r \mathbb{E} \left[\sup_{0 < t \leq u} \left| X^{(n)}(t-) \right|^{r-1} \cdot 2K \left(\int_0^u \sup_{0 < t \leq s} \left(1 + \left| X^{(n)}(t-) \right|^2 + \mathbb{E} \left| X^{(n)}([t]_n -) \right|^2 \right) ds \right)^{\frac{1}{2}} \right]. \end{aligned}$$

Then, due to Young's inequality (2.1) (with $\epsilon = \frac{1}{2C_2(r-1)}$, $p = \frac{r}{r-1}$, and $q = r$), Hölder inequality, the elementary inequality (2.3) (with $l = \frac{r}{2}$) and Lyapunov inequality, one can further conclude

$$\begin{aligned}
& \mathbb{E} \sup_{0 \leq t \leq u} |M_t| \\
& \leq \frac{1}{2} \mathbb{E} \sup_{0 < t \leq u} \left| X^{(n)}(t-) \right|^r + C_2^r (2(r-1))^{r-1} (2K)^r \mathbb{E} \\
& \quad \times \left[\int_0^u \sup_{0 < t \leq s} \left(1 + \left| X^{(n)}(t-) \right|^2 + \mathbb{E} \left| X^{(n)}([t]_n -) \right|^2 \right) ds \right]^{\frac{r}{2}} \\
& \leq \frac{1}{2} \mathbb{E} \sup_{0 < t \leq u} \left| X^{(n)}(t-) \right|^r + C_2^r (2(r-1))^{r-1} (2K)^r u^{\frac{r}{2}-1} \mathbb{E} \\
& \quad \times \left[\int_0^u \sup_{0 < t \leq s} \left(1 + \left| X^{(n)}(t-) \right|^2 + \mathbb{E} \left| X^{(n)}([t]_n -) \right|^2 \right)^{\frac{r}{2}} ds \right] \\
& \leq \frac{1}{2} \mathbb{E} \sup_{0 < t \leq u} \left| X^{(n)}(t-) \right|^r + C_2^r (2(r-1))^{r-1} (2K)^r (3u)^{\frac{r}{2}-1} \mathbb{E} \\
& \quad \times \int_0^u \sup_{0 < t \leq s} \left(1 + \left| X^{(n)}(t-) \right|^r + \mathbb{E} \left| X^{(n)}([t]_n -) \right|^r \right) ds. \tag{2.22}
\end{aligned}$$

On the other hand, utilizing Assumptions 4 and 7, along with Young's inequality (2.1) (with $\epsilon = 1$, $p = \frac{r}{r-2}$, and $q = \frac{r}{2}$), the elementary inequality (2.3) (with $l = \frac{r}{2}$), and Lyapunov's inequality, we obtain the following estimate for the supremum of $|N_t|$:

$$\begin{aligned}
\mathbb{E} \sup_{0 \leq t \leq u} |N_t| & \leq C_1 \mathbb{E} \int_0^u \int_U \left(\left| X^{(n)}(s-) \right|^{r-2} \left| h(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)}, z) \right|^2 \right. \\
& \quad \left. + \left| h(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)}, z) \right|^r \right) \nu(dz) ds \\
& \leq C_1 \mathbb{E} \int_0^u \left[K \left| X^{(n)}(s-) \right|^{r-2} \left(1 + \left| X^{(n)}(s-) \right|^2 + \mathbb{E} \left| X^{(n)}([s]_n -) \right|^2 \right) \right. \\
& \quad \left. + K_2 \left(1 + \left| X^{(n)}(s-) \right|^r + \left(\mathbb{E} \left| X^{(n)}([s]_n -) \right|^2 \right)^{\frac{r}{2}} \right) \right] ds \\
& \leq C_1 K \frac{r-2}{r} \int_0^u \mathbb{E} \sup_{0 < t \leq s} \left| X^{(n)}(t-) \right|^r ds + C_1 \left(\frac{2K}{r} 3^{\frac{r}{2}-1} + K_2 \right) \\
& \quad \times \int_0^u \left(1 + \mathbb{E} \sup_{0 < t \leq s} \left| X^{(n)}(t-) \right|^r + \mathbb{E} \sup_{0 < t \leq s} \left| X^{(n)}([t]_n -) \right|^r \right) ds. \tag{2.23}
\end{aligned}$$

Note that, by (2.13), we have

$$\mathbb{E} \sup_{0 < t \leq u} |X(t-)|^r \leq \mathbb{E} \sup_{0 \leq t \leq u} |X(t)|^r < \infty. \tag{2.24}$$

By combining all the estimates from (2.22) to (2.24) and applying Grönwall's inequality, we deduce from (2.21) that

$$\mathbb{E} \sup_{0 \leq t \leq T} \left| X^{(n)}(t) \right|^r \leq 2(1 + \mathbb{E}|x_0|^r) e^{\widehat{C}_r T} \leq C_r,$$

where $\widehat{C}_r = K(r-2) \frac{r^2+r+2C_1}{r} + 4 \cdot 3^{\frac{r}{2}-1} (r+1) + C_2^r 2^{r+1} (r-1)^{r-1} (2K)^r (3T)^{\frac{r}{2}-1} + 4C_1 \cdot \left(\frac{2K}{r} 3^{\frac{r}{2}-1} + K_2 \right)$. It is evident that the positive constant C_r depends on r , T , K , K_2 , and the initial condition x_0 , but is independent of n . Thus, the proof is complete. $\square$

Lemma 2. (Time Hölder continuity.) *Let Assumptions A4-A7 hold. For any initial condition $x_0 \in L^r(\Omega; \mathbb{R}^d)$ with $r \geq \kappa^2/2$, there exists a positive constant C_κ such that, for any $0 \leq s \leq t \leq T$ with $|t - s| \leq 1$,*

$$\sup_{n \geq 1} \mathbb{E} \left[\left| X^{(n)}(t) - X^{(n)}(s) \right|^\kappa \right] \leq C_\kappa |t - s|. \quad (2.25)$$

Proof. It follows from (2.12) that

$$\begin{aligned} X^{(n)}(t) - X^{(n)}(s) &= \int_s^t b(u, X^{(n)}(u-), \mu_{[u]_n}^{(n)}) du + \int_s^t \sigma(u, X^{(n)}(u-), \mu_{[u]_n}^{(n)}) dW(u) \\ &\quad + \int_s^t \int_U h(u, X^{(n)}(u-), \mu_{[u]_n}^{(n)}, z) \tilde{N}(du, dz). \end{aligned}$$

By taking expectations on both sides, one obtains

$$\begin{aligned} \mathbb{E} \left| X^{(n)}(t) - X^{(n)}(s) \right|^\kappa &\leq 3^{\kappa-1} \mathbb{E} \left[\left| \int_s^t b(u, X^{(n)}(u-), \mu_{[u]_n}^{(n)}) du \right|^\kappa \right] \\ &\quad + 3^{\kappa-1} \mathbb{E} \left[\left| \int_s^t \sigma(u, X^{(n)}(u-), \mu_{[u]_n}^{(n)}) dW(u) \right|^\kappa \right] \\ &\quad + 3^{\kappa-1} \mathbb{E} \left[\left| \int_s^t \int_U h(u, X^{(n)}(u-), \mu_{[u]_n}^{(n)}, z) \tilde{N}(du, dz) \right|^\kappa \right] \\ &=: B_1 + B_2 + B_3. \end{aligned} \quad (2.26)$$

We proceed by estimating B_1 , B_2 , and B_3 individually. To maintain clarity, we present only the core estimation steps for each term. By applying Hölder inequality, Assumption 5, Lyapunov inequality, and estimate (2.14), we derive the following bound for B_1 :

$$\begin{aligned} B_1 &\leq 3^{\kappa-1} (t-s)^{\kappa-1} \int_s^t \mathbb{E} \left[\left| b(u, X^{(n)}(u-), \mu_{[u]_n}^{(n)}) \right|^\kappa \right] du \\ &\leq 3^{\frac{3\kappa}{2}-2} K_1^{\frac{\kappa}{2}} (t-s)^{\kappa-1} \int_s^t \left(1 + \mathbb{E} \left| X^{(n)}(u-) \right|^{\frac{\kappa^2}{2}} + \left(\mathbb{E} \left| X^{(n)}([u]_n-) \right|^2 \right)^{\frac{\kappa^2}{4}} \right) du \\ &\leq 2 \cdot 3^{\frac{3\kappa}{2}-2} K_1^{\frac{\kappa}{2}} (t-s)^{\kappa-1} \int_s^t \left(1 + \mathbb{E} \sup_{0 \leq u \leq T} \left| X^{(n)}(u) \right|^{\frac{\kappa^2}{2}} \right) du \leq C_\kappa (t-s)^\kappa. \end{aligned}$$

For B_2 , using the BDG inequality, Hölder inequality, Assumption 4, and estimate (2.14), we have

$$\begin{aligned} B_2 &\leq 3^{\kappa-1} M_\kappa \mathbb{E} \left[\left| \int_s^t \left\| \sigma(u, X^{(n)}(u-), \mu_{[u]_n}^{(n)}) \right\|^2 du \right|^{\frac{\kappa}{2}} \right] \\ &\leq 3^{\kappa-1} M_\kappa (t-s)^{\frac{\kappa-2}{2}} \int_s^t \mathbb{E} \left[\left\| \sigma(u, X^{(n)}(u-), \mu_{[u]_n}^{(n)}) \right\|^\kappa \right] du \\ &\leq 2 \cdot 3^{\frac{3\kappa}{2}-2} M_\kappa K^{\frac{\kappa}{2}} (t-s)^{\frac{\kappa-2}{2}} \int_s^t \left(1 + \mathbb{E} \sup_{0 \leq u \leq T} \left| X^{(n)}(u) \right|^\kappa \right) du \leq C_\kappa (t-s)^{\frac{\kappa}{2}}, \end{aligned}$$

where $M_\kappa = [\kappa^{\kappa+1}/2(\kappa-1)^{\kappa-1}]^{\frac{\kappa}{2}}$. For B_3 , using Kunita's first inequality (i.e. the inequality (2.4) in Proposition 2), Hölder inequality, Assumptions 4 and 7, and estimate (2.14), we obtain

$$\begin{aligned}
B_3 &\leq 3^{\kappa-1} D \mathbb{E} \left[\left| \int_s^t \int_U \left| h \left(u, X^{(n)}(u-), \mu_{[u]_n}^{(n)}, z \right) \right|^2 v(dz) du \right|^{\frac{\kappa}{2}} \right] \\
&\quad + 3^{\kappa-1} D \int_s^t \mathbb{E} \left[\int_U \left| h \left(u, X^{(n)}(u-), \mu_{[u]_n}^{(n)}, z \right) \right|^{\kappa} v(dz) \right] du \\
&\leq 3^{\kappa-1} D (t-s)^{\frac{\kappa-2}{2}} \int_s^t \mathbb{E} \left[\left(\int_U \left| h \left(u, X^{(n)}(u-), \mu_{[u]_n}^{(n)}, z \right) \right|^2 v(dz) \right)^{\frac{\kappa}{2}} \right] du \\
&\quad + 3^{\kappa-1} D \int_s^t \mathbb{E} \left[\int_U \left| h \left(u, X^{(n)}(u-), \mu_{[u]_n}^{(n)}, z \right) \right|^{\kappa} v(dz) \right] du \\
&\leq 3^{\kappa-1} D K^{\frac{\kappa}{2}} (t-s)^{\frac{\kappa-2}{2}} \int_s^t \mathbb{E} \left[\left(1 + \left| X^{(n)}(u-) \right|^2 + \mathbb{E} X^{(n)}([u]_n -)^2 \right)^{\frac{\kappa}{2}} \right] du \\
&\quad + 3^{\kappa-1} D K_3 \int_s^t \mathbb{E} \left[\left(1 + \left| X^{(n)}(u-) \right|^{\kappa} + \left(\mathbb{E} X^{(n)}([u]_n -)^2 \right)^{\frac{\kappa}{2}} \right) \right] du \\
&\leq 2 \cdot 3^{\frac{3\kappa}{2}-2} D K^{\frac{\kappa}{2}} (t-s)^{\frac{\kappa-2}{2}} \int_s^t \left(1 + \mathbb{E} \sup_{0 \leq u \leq T} \left| X^{(n)}(u) \right|^{\kappa} \right) du \\
&\quad + 2 \cdot 3^{\kappa-1} D K_3 \int_s^t \left(1 + \mathbb{E} \sup_{0 \leq u \leq T} \left| X^{(n)}(u) \right|^{\kappa} \right) du \\
&\leq C_{\kappa} [(t-s) + (t-s)^{\frac{\kappa}{2}}],
\end{aligned}$$

where D is a positive constant dependent on κ . Consequently, the desired assertion follows by substituting the above estimates on B_1 , B_2 , and B_3 into (2.26) and then taking the supremum over n . $\square$

With Lemmas 1 and 2 established, we proceed to demonstrate the following result.

Lemma 3. (Cauchy sequences.) *The sequence $\{X^{(n)}(t)\}_{n \geq 1}$ given by (2.12) is a Cauchy sequence in $L^{\kappa}(\Omega; D([0, T]; \mathbb{R}^d))$. Specifically, for any $n, m \geq 1$, the following holds:*

$$\left\| X^{(n)} - X^{(m)} \right\|_{L^{\kappa}} = \left(\mathbb{E} \left[\sup_{0 \leq t \leq T} \left| X^{(n)}(t) - X^{(m)}(t) \right|^{\kappa} \right] \right)^{\frac{1}{\kappa}} \rightarrow 0, \quad \text{as } n, m \rightarrow \infty. \quad (2.27)$$

Proof. Note that, for $t \in [0, T]$, the difference between $X^{(n)}$ and $X^{(m)}$ satisfies the following equation:

$$\begin{aligned}
X^{(n)}(t) - X^{(m)}(t) &= \int_0^t \left[b \left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)} \right) - b \left(s, X^{(m)}(s-), \mu_{[s]_m}^{(m)} \right) \right] ds \\
&\quad + \int_0^t \left[\sigma \left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)} \right) - \sigma \left(s, X^{(m)}(s-), \mu_{[s]_m}^{(m)} \right) \right] dW(s) \\
&\quad + \int_0^t \int_U \left[h \left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)}, z \right) - h \left(s, X^{(m)}(s-), \mu_{[s]_m}^{(m)}, z \right) \right] \tilde{N}(ds, dz).
\end{aligned}$$

To facilitate the analysis, we define the stopping time:

$$\tau_R := \inf \left\{ t \in [0, T]: \left| X^{(n)}(t) \right| \vee \left| X^{(m)}(t) \right| > R \right\},$$

for each $R > 0$. The stopping time technique is employed here to ensure boundedness of the processes $X^{(n)}$ and $X^{(m)}$ up to τ_R , leveraging the fact that (2.12) describes a classical (nondistribution-dependent) SDE. It is clear that, for $0 \leq t \leq \tau_R \wedge T$, we have $|X^{(n)}(t-)| \leq R$ and $|X^{(m)}(t-)| \leq R$. Then, by De Morgan's law, we arrive at

$$\begin{aligned} \mathbb{E} \left[\sup_{0 \leq t \leq T} |X^{(n)}(t) - X^{(m)}(t)|^\kappa \right] &= \mathbb{E} \left[\sup_{0 \leq t \leq T} |X^{(n)}(t) - X^{(m)}(t)|^\kappa \mathbb{I}_{\{\tau_R > T\}} \right] \\ &\quad + \mathbb{E} \left[\sup_{0 \leq t \leq T} |X^{(n)}(t) - X^{(m)}(t)|^\kappa \mathbb{I}_{\{\tau_R \leq T\}} \right] \\ &=: J_1 + J_2, \end{aligned} \quad (2.28)$$

where $\mathbb{I}_A$ is the indicator function of the set A .

In the subsequent analysis, we estimate each term J_1 and J_2 on the right-hand side of (2.28).

(1) Estimation of the term J_1 . We note that

$$J_1 \leq \mathbb{E} \left[\sup_{0 \leq t \leq T} |X^{(n)}(t \wedge \tau_R) - X^{(m)}(t \wedge \tau_R)|^\kappa \right] < \infty, \quad (2.29)$$

where the finiteness of the term is guaranteed by Lemma 1. By applying Itô's formula, we obtain the following representation:

$$|X^{(n)}(t \wedge \tau_R) - X^{(m)}(t \wedge \tau_R)|^\kappa = J_{1,R}(t) + J_{2,R}(t) + J_{3,R}(t) + J_{4,R}(t) + J_{5,R}(t) + J_{6,R}(t),$$

where the individual terms $J_{i,R}$, for $i = 1, \dots, 6$, are given by

$$\begin{aligned} J_{1,R}(t) &= \kappa \int_0^{t \wedge \tau_R} |X^{(n)}(s-) - X^{(m)}(s-)|^{\kappa-2} \\ &\quad \cdot \left\langle X^{(n)}(s-) - X^{(m)}(s-), b \left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)} \right) - b \left(s, X^{(m)}(s-), \mu_{[s]_m}^{(m)} \right) \right\rangle ds, \\ J_{2,R}(t) &= \frac{\kappa}{2} \int_0^{t \wedge \tau_R} |X^{(n)}(s-) - X^{(m)}(s-)|^{\kappa-2} \left\| \sigma \left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)} \right) \right. \\ &\quad \left. - \sigma \left(s, X^{(m)}(s-), \mu_{[s]_m}^{(m)} \right) \right\|^2 ds, \\ J_{3,R}(t) &= \frac{\kappa(\kappa-2)}{2} \int_0^{t \wedge \tau_R} |X^{(n)}(s-) - X^{(m)}(s-)|^{\kappa-4} \\ &\quad \cdot \left\| \left(X^{(n)}(s-) - X^{(m)}(s-) \right)^\top \left(\sigma \left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)} \right) - \sigma \left(s, X^{(m)}(s-), \mu_{[s]_m}^{(m)} \right) \right) \right\|^2 ds, \\ J_{4,R}(t) &= \int_0^t \int_U \left[|X^{(n)}(s-) - X^{(m)}(s-) + h \left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)}, z \right) \right. \\ &\quad \left. - h \left(s, X^{(m)}(s-), \mu_{[s]_m}^{(m)}, z \right) \right|^\kappa \\ &\quad - |X^{(n)}(s-) - X^{(m)}(s-)|^\kappa - \kappa |X^{(n)}(s-) - X^{(m)}(s-)|^{\kappa-2} \\ &\quad \cdot \left\langle X^{(n)}(s-) - X^{(m)}(s-), h \left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)}, z \right) \right. \\ &\quad \left. - h \left(s, X^{(m)}(s-), \mu_{[s]_m}^{(m)}, z \right) \right\rangle \Big] N(ds, dz), \end{aligned}$$

$$\begin{aligned}
J_{5,R}(t) &= \kappa \int_0^{t \wedge \tau_R} \left| X^{(n)}(s-) - X^{(m)}(s-) \right|^{\kappa-2} \\
&\quad \cdot \left\langle X^{(n)}(s-) - X^{(m)}(s-), \sigma \left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)} \right) - \sigma \left(s, X^{(m)}(s-), \mu_{[s]_m}^{(m)} \right) dW(s) \right\rangle, \\
J_{6,R}(t) &= \kappa \int_0^{t \wedge \tau_R} \int_U \left| X^{(n)}(s-) - X^{(m)}(s-) \right|^{\kappa-2} \\
&\quad \cdot \left\langle X^{(n)}(s-) - X^{(m)}(s-), h \left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)}, z \right) \right. \\
&\quad \left. - h \left(s, X^{(m)}(s-), \mu_{[s]_m}^{(m)}, z \right) \right\rangle \tilde{N}(ds, dz).
\end{aligned}$$

In order to take the supremum over time and the expectation, we need to estimate $\mathbb{E} \left[\sup_{0 \leq t \leq u} J_{i,R}(t) \right]$ for $i = 1, \dots, 6$.

Note that the terms $J_{i,R}$ for $i = 1, 2, 3$ are standard Lebesgue integrals, and can be estimated in a similar manner. For any $u \in [0, T]$, applying Assumptions 1 and 2, we derive

$$\begin{aligned}
\mathbb{E} \left[\sup_{0 \leq t \leq u} J_{1,R}(t) \right] &\leq \kappa \mathbb{E} \left[\sup_{0 \leq t \leq u} \int_0^{t \wedge \tau_R} \left| X^{(n)}(s-) - X^{(m)}(s-) \right|^{\kappa-2} \right. \\
&\quad \cdot \left\langle X^{(n)}(s-) - X^{(m)}(s-), b \left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)} \right) \right. \\
&\quad \left. - b \left(s, X^{(m)}(s-), \mu_{[s]_m}^{(m)} \right) \right\rangle ds \Big] \\
&\quad + \kappa \mathbb{E} \left[\int_0^{u \wedge \tau_R} \left| X^{(n)}(s-) - X^{(m)}(s-) \right|^{\kappa-1} \right. \\
&\quad \cdot \left| b \left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)} \right) - b \left(s, X^{(m)}(s-), \mu_{[s]_m}^{(m)} \right) \right| ds \Big] \\
&\leq \kappa L_R \mathbb{E} \left[\int_0^{u \wedge \tau_R} \left| X^{(n)}(s-) - X^{(m)}(s-) \right|^{\kappa} ds \right] \\
&\quad + \kappa \sqrt{L} \mathbb{E} \left[\int_0^{u \wedge \tau_R} \left| X^{(n)}(s-) - X^{(m)}(s-) \right|^{\kappa-1} W_2 \left(\mu_{[s]_n}^{(n)}, \mu_{[s]_m}^{(m)} \right) ds \right].
\end{aligned}$$

By further applying Young's inequality (2.1) (with $\epsilon = 1$, $p = \frac{\kappa}{\kappa-1}$ and $q = \kappa$), we obtain

$$\begin{aligned}
&\mathbb{E} \left[\sup_{0 \leq t \leq u} J_{1,R}(t) \right] \\
&\leq \left(\kappa L_R + (\kappa - 1) \sqrt{L} \right) \mathbb{E} \left[\int_0^{u \wedge \tau_R} \left| X^{(n)}(s-) - X^{(m)}(s-) \right|^{\kappa} ds \right] \\
&\quad + 3^{\kappa-1} \sqrt{L} \int_0^{u \wedge \tau_R} \left[W_2^{\kappa} \left(\mu_{[s]_n}^{(n)}, \mu_s^{(n)} \right) + W_2^{\kappa} \left(\mu_s^{(n)}, \mu_s^{(m)} \right) + W_2^{\kappa} \left(\mu_s^{(m)}, \mu_{[s]_m}^{(m)} \right) \right] ds \\
&\leq \left[\kappa L_R + \sqrt{L} \left(\kappa - 1 + 3^{\kappa-1} \right) \right] \int_0^{u \wedge \tau_R} \mathbb{E} \left| X^{(n)}(s-) - X^{(m)}(s-) \right|^{\kappa} ds \\
&\quad + 3^{\kappa-1} \sqrt{L} \int_0^{u \wedge \tau_R} \left(\mathbb{E} \left| X^{(n)}(s-) - X^{(n)}([s]_n-) \right|^{\kappa} + \mathbb{E} \left| X^{(m)}(s-) - X^{(m)}([s]_m-) \right|^{\kappa} \right) ds.
\end{aligned} \tag{2.30}$$

Analogously, by Assumptions 1 and 2 and Young's inequality (2.1) (with $\epsilon = 1$, $p = \frac{\kappa}{\kappa-2}$, $q = \frac{\kappa}{2}$), we derive the following bounds for $J_{2,R}$ and $J_{3,R}$:

$$\begin{aligned} & \mathbb{E} \left[\sup_{0 \leq t \leq u} J_{2,R}(t) \right] \\ & \leq \left[\kappa L_R + L \left(\kappa - 2 + 2 \cdot 3^{\kappa-1} \right) \right] \int_0^{u \wedge \tau_R} \mathbb{E} \left| X^{(n)}(s-) - X^{(m)}(s-) \right|^\kappa ds \\ & \quad + 2L \cdot 3^{\kappa-1} \int_0^{u \wedge \tau_R} \left(\mathbb{E} \left| X^{(n)}(s-) - X^{(n)}([s]_n-) \right|^\kappa + \mathbb{E} \left| X^{(m)}(s-) - X^{(m)}([s]_m-) \right|^\kappa \right) ds, \end{aligned} \quad (2.31)$$

and

$$\begin{aligned} & \mathbb{E} \left[\sup_{0 \leq t \leq u} J_{3,R}(t) \right] \\ & \leq (\kappa - 2) \left[\kappa L_R + \left(\kappa - 2 + 2 \cdot 3^{\kappa-1} \right) L \right] \int_0^{u \wedge \tau_R} \mathbb{E} \left| X^{(n)}(s-) - X^{(m)}(s-) \right|^\kappa ds \\ & \quad + 2L \cdot 3^{\kappa-1} (\kappa - 2) \int_0^{u \wedge \tau_R} \left(\mathbb{E} \left| X^{(n)}(s-) - X^{(n)}([s]_n-) \right|^\kappa \right. \\ & \quad \left. + \mathbb{E} \left| X^{(m)}(s-) - X^{(m)}([s]_m-) \right|^\kappa \right) ds. \end{aligned} \quad (2.32)$$

As for the last three terms, we first use the remainder formula in (2.19) and Assumption 7 to obtain

$$\begin{aligned} & \mathbb{E} \left[\sup_{0 \leq t \leq u} J_{4,R}(t) \right] \\ & \leq C_1 \mathbb{E} \int_0^{u \wedge \tau_R} \int_U \left(\left| X^{(n)}(s-) - X^{(m)}(s-) \right|^{\kappa-2} \left| h \left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)}, z \right) \right. \right. \\ & \quad \left. \left. - h \left(s, X^{(m)}(s-), \mu_{[s]_m}^{(m)}, z \right) \right|^2 \right. \\ & \quad \left. + \left| h \left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)}, z \right) - h \left(s, X^{(m)}(s-), \mu_{[s]_m}^{(m)}, z \right) \right|^\kappa \right) \nu(dz) ds. \\ & \leq C_1 \left(L_R + \frac{(\kappa - 2 + 2 \cdot 3^{\kappa-1})L}{\kappa} \right) \int_0^{u \wedge \tau_R} \mathbb{E} \left| X^{(n)}(s-) - X^{(m)}(s-) \right|^\kappa ds \\ & \quad + 2C_1 L \cdot \frac{3^{\kappa-1}}{\kappa} \int_0^{u \wedge \tau_R} \left(\mathbb{E} \left| X^{(n)}(s-) - X^{(n)}([s]_n-) \right|^\kappa + \mathbb{E} \left| X^{(m)}(s-) - X^{(m)}([s]_m-) \right|^\kappa \right) ds \\ & \quad + C_1 2^{\kappa-1} \mathbb{E} \int_0^{u \wedge \tau_R} \int_U \left| h \left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)}, z \right) - h \left(s, X^{(m)}(s-), \mu_{[s]_n}^{(n)}, z \right) \right|^\kappa \nu(dz) ds \\ & \quad + C_1 2^{\kappa-1} \mathbb{E} \int_0^{u \wedge \tau_R} \int_U \left| h \left(s, X^{(m)}(s-), \mu_{[s]_n}^{(n)}, z \right) - h \left(s, X^{(m)}(s-), \mu_{[s]_m}^{(m)}, z \right) \right|^\kappa \nu(dz) ds \\ & \leq C_1 \left(L_R + \frac{(\kappa - 2 + 2 \cdot 3^{\kappa-1})L}{\kappa} + 2^{\kappa-1} L'_R + L' 6^{\kappa-1} \right) \int_0^{u \wedge \tau_R} \mathbb{E} \sup_{0 \leq t \leq s} \left| X^{(n)}(t) - X^{(m)}(t) \right|^\kappa ds \\ & \quad + 3^{\kappa-1} C_1 \left(\frac{2L}{\kappa} + 2^{\kappa-1} L' \right) \int_0^{u \wedge \tau_R} \left(\mathbb{E} \left| X^{(n)}(s-) - X^{(n)}([s]_n-) \right|^\kappa \right. \\ & \quad \left. + \mathbb{E} \left| X^{(m)}(s-) - X^{(m)}([s]_m-) \right|^\kappa \right) ds. \end{aligned} \quad (2.33)$$

We exploit the BDG inequality and Assumptions 2 and 3 to obtain

$$\begin{aligned}
& \mathbb{E} \left[\sup_{0 \leq t \leq u} J_{5,R}(t) \right] \\
& \leq \kappa \cdot \sqrt{32} \mathbb{E} \left(\int_0^{u \wedge \tau_R} \left| X^{(n)}(s-) - X^{(m)}(s-) \right|^{2\kappa-2} \left\| \sigma \left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)} \right) \right. \right. \\
& \quad \left. \left. - \sigma \left(s, X^{(m)}(s-), \mu_{[s]_m}^{(m)} \right) \right\|^2 ds \right)^{\frac{1}{2}} \\
& \leq 6\kappa \mathbb{E} \left[\sup_{0 \leq t \leq u} \left| X^{(n)}(t \wedge \tau_R) - X^{(m)}(t \wedge \tau_R) \right|^{\kappa-1} \left(\int_0^{u \wedge \tau_R} 2 \left\| \sigma \left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)} \right) \right. \right. \right. \\
& \quad \left. \left. - \sigma \left(s, X^{(m)}(s-), \mu_{[s]_m}^{(m)} \right) \right\|^2 \right. \\
& \quad \left. + 2 \left\| \sigma \left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)} \right) - \sigma \left(s, X^{(m)}(s-), \mu_{[s]_m}^{(m)} \right) \right\|^2 ds \right)^{\frac{1}{2}} \Bigg] \\
& \leq 6\kappa \mathbb{E} \left[\sup_{0 \leq t \leq u} \left| X^{(n)}(t \wedge \tau_R) - X^{(m)}(t \wedge \tau_R) \right|^{\kappa-1} \left(\int_0^{u \wedge \tau_R} 2L_R \left| X^{(n)}(s-) - X^{(m)}(s-) \right|^2 \right. \right. \\
& \quad \left. \left. + 2LW_2^2 \left(\mu_{[s]_n}^{(n)}, \mu_{[s]_m}^{(m)} \right) ds \right)^{\frac{1}{2}} \right].
\end{aligned}$$

Then owing to Young's inequality (2.1) (with $\epsilon = \frac{1}{24(\kappa-1)}$, $p = \frac{\kappa}{\kappa-1}$, and $q = \kappa$), Hölder inequality, the elementary inequality (2.2) (with $k = 2$ and $l = \frac{\kappa}{2}$), the elementary inequality (2.3) (with $l = \kappa$), and Lyapunov inequality, we further have

$$\begin{aligned}
& \mathbb{E} \left[\sup_{0 \leq t \leq u} J_{5,R}(t) \right] - \frac{1}{4} \mathbb{E} \left[\sup_{0 \leq t \leq u} \left| X^{(n)}(t \wedge \tau_R) - X^{(m)}(t \wedge \tau_R) \right|^\kappa \right] \\
& \leq 6^\kappa \cdot 2^{\frac{\kappa}{2}} (4(\kappa-1))^{\kappa-1} \mathbb{E} \left(\int_0^{u \wedge \tau_R} L_R \left| X^{(n)}(s-) - X^{(m)}(s-) \right|^2 + LW_2^2 \left(\mu_{[s]_n}^{(n)}, \mu_{[s]_m}^{(m)} \right) ds \right)^{\frac{\kappa}{2}} \\
& \leq 6^\kappa \cdot 2^{\frac{\kappa}{2}} (4(\kappa-1))^{\kappa-1} u^{\frac{\kappa}{2}-1} \mathbb{E} \left[\int_0^{u \wedge \tau_R} \left(L_R \left| X^{(n)}(s-) - X^{(m)}(s-) \right|^2 \right. \right. \\
& \quad \left. \left. + LW_2^2 \left(\mu_{[s]_n}^{(n)}, \mu_{[s]_m}^{(m)} \right) \right)^{\frac{\kappa}{2}} ds \right] \\
& \leq 12^\kappa \cdot (4(\kappa-1))^{\kappa-1} u^{\frac{\kappa}{2}-1} \mathbb{E} \left[\int_0^{u \wedge \tau_R} \left(L_R^{\frac{\kappa}{2}} \left| X^{(n)}(s-) - X^{(m)}(s-) \right|^\kappa \right. \right. \\
& \quad \left. \left. + L^{\frac{\kappa}{2}} W_2^\kappa \left(\mu_{[s]_n}^{(n)}, \mu_{[s]_m}^{(m)} \right) \right) ds \right] \\
& \leq 12^\kappa \cdot (4(\kappa-1))^{\kappa-1} u^{\frac{\kappa}{2}-1} \left(L_R^{\frac{\kappa}{2}} + L^{\frac{\kappa}{2}} 3^{\kappa-1} \right) \int_0^{u \wedge \tau_R} \mathbb{E} \left| X^{(n)}(s-) - X^{(m)}(s-) \right|^\kappa ds \\
& \quad + 12^\kappa \cdot (L)^{\frac{\kappa}{2}} (12(\kappa-1))^{\kappa-1} u^{\frac{\kappa}{2}-1} \int_0^{u \wedge \tau_R} \left(\mathbb{E} \left| X^{(n)}(s-) - X^{(n)}([s]_n-) \right|^\kappa \right. \\
& \quad \left. + \mathbb{E} \left| X^{(m)}(s-) - X^{(m)}([s]_m-) \right|^\kappa \right) ds.
\end{aligned} \tag{2.34}$$

Finally, we apply the inequality (2.5) (with $p = 1$) in Proposition 2 and Young's inequality (2.1) (with $\epsilon = \frac{1}{4D(\kappa-1)}$, $p = \frac{\kappa}{\kappa-1}$, and $q = \kappa$) to obtain

$$\begin{aligned}
& \mathbb{E} \left[\sup_{0 \leq t \leq u} J_{6,R}(t) \right] \\
& \leq \kappa D \mathbb{E} \left(\int_0^{u \wedge \tau_R} \int_U \left| X^{(n)}(s-) - X^{(m)}(s-) \right|^{2\kappa-2} \left| h \left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)}, z \right) \right. \right. \\
& \quad \left. \left. - h \left(s, X^{(m)}(s-), \mu_{[s]_m}^{(m)}, z \right) \right|^2 v(dz) ds \right)^{\frac{1}{2}} \\
& \leq \frac{1}{4} \mathbb{E} \left[\sup_{0 \leq t \leq u} \left| X^{(n)}(t \wedge \tau_R) - X^{(m)}(t \wedge \tau_R) \right|^\kappa \right] \\
& \quad + (2D)^\kappa \cdot (4(\kappa-1))^{\kappa-1} u^{\frac{\kappa}{2}-1} \left(L_R^{\frac{\kappa}{2}} + L^{\frac{\kappa}{2}} 3^{\kappa-1} \right) \int_0^{u \wedge \tau_R} \mathbb{E} \sup_{0 \leq t \leq s} \left| X^{(n)}(t) - X^{(m)}(t) \right|^\kappa ds \\
& \quad + (2D)^\kappa \cdot L^{\frac{\kappa}{2}} (12(\kappa-1))^{\kappa-1} u^{\frac{\kappa}{2}-1} \int_0^{u \wedge \tau_R} \left(\mathbb{E} \left| X^{(n)}(s-) - X^{(n)}([s]_n-) \right|^\kappa \right. \\
& \quad \left. + \mathbb{E} \left| X^{(m)}(s-) - X^{(m)}([s]_m-) \right|^\kappa \right) ds. \tag{2.35}
\end{aligned}$$

Substituting the estimates derived from (2.30)–(2.35) into (2.29) yields the inequality

$$\begin{aligned}
& \mathbb{E} \left[\sup_{0 \leq t \leq u} \left| X^{(n)}(t \wedge \tau_R) - X^{(m)}(t \wedge \tau_R) \right|^\kappa \right] \\
& \leq 2\widehat{M}_1(u) \int_0^u \mathbb{E} \sup_{0 \leq t \leq s} \left| X^{(n)}(t \wedge \tau_R) - X^{(m)}(t \wedge \tau_R) \right|^\kappa ds \\
& \quad + 2 \cdot 3^{\kappa-1} \widehat{M}_2(u) \int_0^{u \wedge \tau_R} \left(\mathbb{E} \left| X^{(n)}(s-) - X^{(n)}([s]_n-) \right|^\kappa + \mathbb{E} \left| X^{(m)}(s-) - X^{(m)}([s]_m-) \right|^\kappa \right) ds.
\end{aligned}$$

Here, $\widehat{M}_1(u) = \kappa^2 L_R + (\kappa-1+3^{\kappa-1})\sqrt{L} + (\kappa-1)(\kappa-2+2 \cdot 3^{\kappa-1})L + C_1 \left(L_R + \frac{(\kappa-2+2 \cdot 3^{\kappa-1})L}{\kappa} + 2\kappa^{-1}L'_R + 6^{\kappa-1}L' \right) + (12^\kappa + (2D)^\kappa)(4(\kappa-1))^{\kappa-1} \left(L_R^{\frac{\kappa}{2}} + L^{\frac{\kappa}{2}} 3^{\kappa-1} \right) u^{\frac{\kappa}{2}-1}$ and $\widehat{M}_2(u) = \sqrt{L} + 2L(\kappa-1) + C_1 \left(\frac{2L}{\kappa} + 2\kappa^{-1}L' \right) + (12^\kappa + (2D)^\kappa)L^{\frac{\kappa}{2}}(4(\kappa-1))^{\kappa-1}u^{\frac{\kappa}{2}-1}$. In addition, for any $t \in [0, T]$, the result in Lemma 2 implies that

$$\mathbb{E} \left| X^{(n)}(t) - X^{(n)}([t]_n) \right|^\kappa \leq Ch_n \quad \text{and} \quad \mathbb{E} \left| X^{(m)}(t) - X^{(m)}([t]_m) \right|^\kappa \leq Ch_m.$$

By these estimates, together with Grönwall's inequality, we conclude that

$$\begin{aligned}
J_1 & \leq \mathbb{E} \left[\sup_{0 \leq t \leq T} \left| X^{(n)}(t \wedge \tau_R) - X^{(m)}(t \wedge \tau_R) \right|^\kappa \right] \leq 2CT(h_n + h_m) \cdot 3^{\kappa-1} \widehat{M}_2(T) e^{2\widehat{M}_1(T) \cdot T} \\
& =: T(h_n + h_m)C(\kappa, T, L, L') e^{T \cdot C(\kappa, T, L, L_R, L', L'_R)}. \tag{2.36}
\end{aligned}$$

(2) Estimation of the term J_2 . With the aid of the Cauchy–Schwarz inequality, we have

$$\begin{aligned}
 J_2 &= \mathbb{E} \left[\sup_{0 \leq t \leq T} |X^{(n)}(t) - X^{(m)}(t)|^\kappa \mathbb{I}_{\{\tau_R \leq T\}} \right] \\
 &\leq \sqrt{\mathbb{E} \left(\sup_{0 \leq t \leq T} |X^{(n)}(t) - X^{(m)}(t)|^\kappa \right)^2} \sqrt{\mathbb{E} (\mathbb{I}_{\{\tau_R \leq T\}})^2} \\
 &\leq \sqrt{\mathbb{E} \left(\sup_{0 \leq t \leq T} |X^{(n)}(t) - X^{(m)}(t)|^{2\kappa} \right)} \sqrt{\mathbb{E} (\mathbb{I}_{\{\tau_R \leq T\}})} \\
 &\leq 2^{\kappa - \frac{1}{2}} \sqrt{\mathbb{E} \left[\sup_{0 \leq t \leq T} |X^{(n)}(t)|^{2\kappa} + \sup_{0 \leq t \leq T} |X^{(m)}(t)|^{2\kappa} \right]} \sqrt{\mathbb{P}(\tau_R \leq T)} \\
 &\leq C_1 \sqrt{\mathbb{P}(\tau_R \leq T)}.
 \end{aligned} \tag{2.37}$$

Here, the result of Lemma 1 has been utilized. Further, by employing the subadditivity of probability and invoking Lemma 1 once more, we can estimate

$$\begin{aligned}
 \mathbb{P}(\tau_R \leq T) &\leq \mathbb{E} \left(\mathbb{I}_{\{\tau_R \leq T\}} \frac{|X^{(n)}(\tau_R)|^4 + |X^{(m)}(\tau_R)|^4}{R^4} \right) \\
 &\leq \frac{1}{R^4} \left(\mathbb{E} \sup_{0 \leq t \leq T} |X^{(n)}(t)|^4 + \mathbb{E} \sup_{0 \leq t \leq T} |X^{(m)}(t)|^4 \right) \\
 &\leq \frac{C}{R^4}.
 \end{aligned}$$

By substituting this into (2.37), we further obtain

$$J_2 \leq \frac{C}{R^2}. \tag{2.38}$$

At this point, we can estimate (2.28) by combining (2.36) and (2.38) as follows:

$$\mathbb{E} \sup_{0 \leq t \leq T} |X^{(n)}(t) - X^{(m)}(t)|^\kappa \leq T(h_n + h_m)C(\kappa, T, L, L')e^{T \cdot C(\kappa, T, L, L_R, L', L'_R)} + \frac{C}{R^2}. \tag{2.39}$$

Note that R is independent of n and m , and $\frac{C}{R^2}$ converges to 0 as $R \rightarrow \infty$. For any given $\varepsilon > 0$, there exists a sufficiently large number $R(\varepsilon) > 0$, such that,

$$\frac{C}{R_*^2} < \frac{\varepsilon}{2},$$

when $R_* \geq R(\varepsilon)$. Since both h_n and h_m converge to 0 as $n, m \rightarrow \infty$, for the $\varepsilon > 0$ chosen previously, we have

$$T(h_n + h_m)C(\kappa, T, L, L')e^{T \cdot C(\kappa, T, L, L_{R_*}, L', L'_{R_*})} < \frac{\varepsilon}{2},$$

by letting $n, m \rightarrow \infty$. Consequently, we conclude that (2.27) holds. $\square$

2.3. Proof of Theorem 1

In this subsection, we turn to proving the main theorem in this section. The proof consists of three steps.

Step 1: Existence. Let $\{X^{(n)}(t)\}_{n \geq 1}$ be the Cauchy sequence in $L^\kappa(\Omega; D([0, T]; \mathbb{R}^d))$ given by (2.12). Keep in mind that the space $L^\kappa(\Omega; D([0, T]; \mathbb{R}^d))$, equipped with the norm $\|X\|_{L^\kappa} := (\mathbb{E} [\sup_{0 \leq t \leq T} |X(t)|^\kappa])^{\frac{1}{\kappa}}$, is a Banach space (see Proposition 1). Thus, there exists an $\{\mathcal{F}_t\}_{0 \leq t \leq T}$ -adapted $\mathbb{R}^d$ -valued càdlàg stochastic process $\{X(t)\}_{t \in [0, T]}$ with $X(0) = x_0$ and $\mu_t = \mathcal{L}_{X(t)}$ such that

$$\lim_{n \rightarrow \infty} \left[\mathbb{E} \left| X^{(n)}(t) - X(t) \right|^\kappa \right]^{\frac{1}{\kappa}} \leq \lim_{n \rightarrow \infty} \left[\mathbb{E} \sup_{0 \leq t \leq T} \left| X^{(n)}(t) - X(t) \right|^\kappa \right]^{\frac{1}{\kappa}} = 0. \quad (2.40)$$

We next prove that $\{X(t)\}_{t \in [0, T]}$ is a solution to (2.8). Indeed, the main idea is to show that the right-hand side of (2.12) converges in probability to

$$x_0 + \int_0^t b(s, X(s-), \mu_s) ds + \int_0^t \sigma(s, X(s-), \mu_s) dW(s) + \int_0^t \int_U h(s, X(s-), \mu_s, z) \tilde{N}(dz, ds),$$

by taking the limit on both sides of (2.12). Here $\mu_s = \mathcal{L}(X(s))$ for any $s \in [0, T]$.

First, it follows from (2.40) that there exists a subsequence (for notational simplicity, still indexed by n) such that, for all $s \in [0, T]$,

$$X^{(n)}(s, \omega) \rightarrow X(s, \omega), \quad \mathbb{P}\text{-almost surely.}$$

By applying Lemma 2, the Wasserstein distance between $\mu_{[s]_n}^{(n)}$ and μ_s satisfies

$$\begin{aligned} & \lim_{n \rightarrow \infty} \sup_{0 \leq s \leq t} W_2^\kappa \left(\mu_{[s]_n}^{(n)}, \mu_s \right) \\ & \leq 2^{\kappa-1} \lim_{n \rightarrow \infty} \sup_{0 \leq s \leq t} \mathbb{E} \left| X^{(n)}(s) - X^{(n)}([s]_n) \right|^\kappa + 2^{\kappa-1} \lim_{n \rightarrow \infty} \mathbb{E} \left[\sup_{0 \leq s \leq t} \left| X^{(n)}(s) - X(s) \right|^\kappa \right] \\ & \leq 2^{\kappa-1} C \lim_{n \rightarrow \infty} h_n = 0. \end{aligned} \quad (2.41)$$

Taking Assumption 3 into account, it follows immediately that, for all $s \in [0, T]$ and almost all $\omega \in \Omega$,

$$\begin{aligned} b \left(s, X^{(n)}(s), \mu_{[s]_n}^{(n)} \right) & \rightarrow b(s, X(s), \mu_s), \quad \sigma \left(s, X^{(n)}(s), \mu_{[s]_n}^{(n)} \right) \rightarrow \sigma(s, X(s), \mu_s), \\ \int_U h \left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)}, z \right) \nu(dz) & \rightarrow \int_U h(s, X(s-), \mu_s, z) \nu(dz), \end{aligned} \quad (2.42)$$

as $n \rightarrow \infty$.

Next, we claim that the sequences $\{b(s, X^{(n)}(s), \mu_{[s]_n}^{(n)})\}_{n \geq 1}$ and $\{\sigma(s, X^{(n)}(s), \mu_{[s]_n}^{(n)})\}_{n \geq 1}$ are uniformly integrable. In fact, from Assumptions 4 and 5 and Lemma 1, we obtain the following uniform boundedness,

$$\sup_{n \geq 1} \mathbb{E} |b(s, X^{(n)}(s), \mu_{[s]_n}^{(n)})| \leq \sqrt{3K_1} \sup_{n \geq 1} \mathbb{E} [1 + |X^{(n)}(s)|^{\frac{\kappa}{2}} + W_2^{\frac{\kappa}{2}}(\mu_{[s]_n}^{(n)}, \delta_0)] \leq \sqrt{3K_1}(1 + 2C),$$

$$\sup_{n \geq 1} \mathbb{E} \|\sigma(t, X^{(n)}(s), \mu_{[s]_n}^{(n)})\|^2 \leq K \sup_{n \geq 1} \mathbb{E} [1 + |X^{(n)}(s)|^2 + W_2^2(\mu_{[s]_n}^{(n)}, \delta_0)] \leq K(1 + 2C),$$

and the following uniform absolute continuity,

$$\begin{aligned} \sup_{n \geq 1} \mathbb{E} \left[\left| b \left(s, X^{(n)}(s), \mu_{[s]_n}^{(n)} \right) \right| \mathbb{I}_A \right] &\leq K_1 \sup_{n \geq 1} \left[\mathbb{E} \left(1 + |X^{(n)}(s)|^{\kappa} + W_2^{\kappa}(\mu_{[s]_n}^{(n)}, \delta_0) \right) \right]^{\frac{1}{2}} (\mathbb{P}(A))^{\frac{1}{2}} \\ &\leq K_1 \sqrt{1 + 2C} (\mathbb{P}(A))^{\frac{1}{2}} \rightarrow 0, \\ \sup_{n \geq 1} \mathbb{E} \left[\left\| \sigma \left(s, X^{(n)}(s), \mu_{[s]_n}^{(n)} \right) \right\|^2 \mathbb{I}_A \right] &\leq K \sup_{n \geq 1} \left[\mathbb{E} \left(1 + |X^{(n)}(s)|^2 + W_2^2(\mu_{[s]_n}^{(n)}, \delta_0) \right) \right]^{\frac{1}{2}} (\mathbb{P}(A))^{\frac{1}{2}} \\ &\leq K \sqrt{3(1 + 2C)} (\mathbb{P}(A))^{\frac{1}{2}} \rightarrow 0, \end{aligned}$$

when $\mathbb{P}(A) \rightarrow 0$. The uniform integrability of $\{b(s, X^{(n)}(s), \mu_{[s]_n}^{(n)})\}_{n \geq 1}$ and $\{\sigma(s, X^{(n)}(s), \mu_{[s]_n}^{(n)})\}_{n \geq 1}$ follows from [38, Lemma 3 in p. 190].

Hence, by applying the dominated convergence theorem [38, Theorem 4 on p. 188], together with (2.42), we obtain, for any $s \in [0, T]$,

$$\lim_{n \rightarrow \infty} \mathbb{E} \left| b \left(s, X^{(n)}(s), \mu_{[s]_n}^{(n)} \right) - b(s, X(s), \mu_s) \right| = 0, \quad (2.43)$$

$$\lim_{n \rightarrow \infty} \mathbb{E} \left\| \sigma \left(s, X^{(n)}(s), \mu_{[s]_n}^{(n)} \right) - \sigma(s, X(s), \mu_s) \right\|^2 = 0. \quad (2.44)$$

In addition, note that, following from (2.40),

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} |X(t)|^{\kappa} \right] \leq C.$$

We further have the following estimates based on Assumptions 4 and 5 and Lemma 1:

$$\begin{aligned} &\sup_{n \geq 1} \sup_{s \in [0, t]} \mathbb{E} \left| b \left(s, X^{(n)}(s), \mu_{[s]_n}^{(n)} \right) - b(s, X(s), \mu_s) \right| \\ &\leq \sqrt{3K_1} \sup_{n \geq 1} \sup_{s \in [0, t]} \mathbb{E} \left[2 + |X^{(n)}(s)|^{\frac{\kappa}{2}} + |X(s)|^{\frac{\kappa}{2}} + W_2^{\frac{\kappa}{2}}(\mu_{[s]_n}^{(n)}, \delta_0) + W_2^{\frac{\kappa}{2}}(\mu_s, \delta_0) \right] \\ &\leq 2\sqrt{3K_1}(1 + 2C), \end{aligned} \quad (2.45)$$

$$\begin{aligned} &\sup_{n \geq 1} \sup_{s \in [0, t]} \mathbb{E} \left\| \sigma \left(s, X^{(n)}(s), \mu_{[s]_n}^{(n)} \right) - \sigma(s, X(s), \mu_s) \right\|^2 \\ &\leq 2K \sup_{n \geq 1} \sup_{s \in [0, t]} \mathbb{E} \left[2 + |X^{(n)}(s)|^2 + |X(s)|^2 + W_2^{\kappa}(\mu_{[s]_n}^{(n)}, \delta_0) + W_2^{\kappa}(\mu_s, \delta_0) \right] \leq 4K(1 + 2C). \end{aligned} \quad (2.46)$$

For any $t \in [0, T]$, by applying the dominated convergence theorem in conjunction with (2.43) and (2.45), we eventually obtain

$$\begin{aligned} & \lim_{n \rightarrow \infty} \mathbb{E} \left| \int_0^t \left(b \left(s, X^{(n)}(s), \mu_{[s]_n}^{(n)} \right) - b(s, X(s), \mu_s) \right) ds \right| \\ & \leq \lim_{n \rightarrow \infty} \int_0^t \mathbb{E} \left| b \left(s, X^{(n)}(s), \mu_{[s]_n}^{(n)} \right) - b(s, X(s), \mu_s) \right| ds = 0. \end{aligned} \quad (2.47)$$

Similarly, in view of (2.44) and (2.46), we arrive at

$$\begin{aligned} & \lim_{n \rightarrow \infty} \mathbb{E} \left| \int_0^t \left(\sigma \left(s, X^{(n)}(s), \mu_{[s]_n}^{(n)} \right) - \sigma(s, X(s), \mu_s) \right) dW(s) \right|^2 \\ & = \lim_{n \rightarrow \infty} \int_0^t \mathbb{E} \left\| \sigma \left(s, X^{(n)}(s), \mu_{[s]_n}^{(n)} \right) - \sigma(s, X(s), \mu_s) \right\|^2 ds = 0. \end{aligned} \quad (2.48)$$

Finally, we examine the estimates for the integral w.r.t. the Poisson random measure. For any $u \in [0, T]$, it follows from the inequality (2.5) (with $p = 1$) in Proposition 2 and Assumptions 1 and 2 that

$$\begin{aligned} & \mathbb{E} \sup_{0 \leq u \leq t} \left| \int_0^u \int_U \left(h \left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)} \right) - h(s, X(s-), \mu_s, z) \right) \tilde{N}(ds, dz) \right| \\ & \leq D \mathbb{E} \left[\int_0^t \int_U \left| h \left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)} \right) - h(s, X(s-), \mu_s, z) \right|^2 \nu(dz) ds \right]^{\frac{1}{2}} \\ & \leq \sqrt{2D} \mathbb{E} \left[\int_0^t \int_U \left(\left| h \left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)} \right) - h(s, X(s-), \mu_{[s]_n}^{(n)}, z) \right|^2 \right. \right. \\ & \quad \left. \left. + \left| h \left(s, X(s-), \mu_{[s]_n}^{(n)} \right) - h(s, X(s-), \mu_s, z) \right|^2 \right) \nu(dz) ds \right]^{\frac{1}{2}} \\ & \leq \sqrt{2D} \mathbb{E} \left[\int_0^t \left| X^{(n)}(s-) - X(s-) \right|^2 ds + \int_0^t W_2^2 \left(\mu_{[s]_n}^{(n)}, \mu_s \right) ds \right]^{\frac{1}{2}} \\ & \leq 2D \mathbb{E} \left[\int_0^t \left| X^{(n)}(s-) - X(s-) \right|^2 ds \right]^{\frac{1}{2}} + 2D \mathbb{E} \left[\int_0^t W_2^2 \left(\mu_{[s]_n}^{(n)}, \mu_s \right) ds \right]^{\frac{1}{2}} \\ & \leq 2D\sqrt{T} \mathbb{E} \left[\sup_{0 \leq s \leq t} \left| X^{(n)}(s) - X(s) \right| \right] + 2D\sqrt{T} \sup_{0 \leq s \leq t} W_2 \left(\mu_{[s]_n}^{(n)}, \mu_s \right). \end{aligned}$$

By (2.40), (2.41), and Lyapunov inequality, we deduce that

$$\lim_{n \rightarrow \infty} \mathbb{E} \sup_{0 \leq u \leq t} \left| \int_0^u \int_U \left(h \left(s, X^{(n)}(s-), \mu_{[s]_n}^{(n)} \right) - h(s, X(s-), \mu_s, z) \right) \tilde{N}(ds, dz) \right| = 0. \quad (2.49)$$

As a consequence, by (2.47), (2.48), and (2.49), we conclude that the process $\{X(t)\}_{t \in [0, T]}$ is a strong solution to (2.8). This completes the proof of existence.

Step 2: Boundedness. For $t \in [0, T]$, let $X(t) \in L^k(\Omega; D([0, T]; \mathbb{R}^d))$ be a solution to (2.8). In the following, we estimate the r th moment of the solution $(X(t))_{t \in [0, T]}$, where $r \geq \max\{\frac{\kappa^2}{2}, 4\}$ and the initial value $X(0) = x_0$ satisfies $\mathbb{E}|x_0|^r < \infty$, as specified in Assumption 6.

For every $R > 0$, we define the stopping time

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

It is clear that $|X(t-)| \leq R$ for $0 \leq t \leq \pi_R$, and $\mathbb{E} \sup_{0 \leq t \leq u \wedge \pi_R} |X(t-)|^r < \infty$ for any $u \in [0, T]$. To derive an upper-bound for $\mathbb{E} \sup_{0 \leq t \leq u \wedge \pi_R} |X(t)|^r$, we employ the procedure similar to that in the proof of Lemma 1, where the case for $X^{(n)}(t)$ with $t \in [0, T]$ was considered. Specifically, for $r \geq \max\{\frac{\kappa^2}{2}, 4\}$ and $t \in [0, u \wedge \pi_R]$, and utilizing tools such as Itô's formula, as demonstrated in the proof of Lemma 1, we estimate that

$$\begin{aligned} \mathbb{E} \sup_{0 \leq t \leq u \wedge \pi_R} |X(t)|^r &\leq \mathbb{E}|x_0|^r + \frac{1}{2} \mathbb{E} \sup_{0 < t \leq u \wedge \pi_R} |X(t-)|^r + \widehat{D} \int_0^u \left(1 + \mathbb{E} \sup_{0 < t \leq s \wedge \pi_R} |X(t-)|^r \right) ds \\ &\quad + K(r-2) \frac{r^2 + r + 2C_1}{2r} \int_0^u \left(1 + \mathbb{E} \sup_{0 < t \leq s \wedge \pi_R} |X(t-)|^r \right) ds < \infty, \end{aligned} \quad (2.50)$$

with $\widehat{D} = 2 \cdot 3^{\frac{r}{2}-1} (r+1)K + C_2 2^r (r-1)^{r-1} (2K)^r (3T)^{\frac{r}{2}-1} + 2C_1 \cdot \left(\frac{2K}{r} 3^{\frac{r}{2}-1} + K_2 \right)$. Note that

$$\mathbb{E} \sup_{0 < t \leq u \wedge \pi_R} |X(t-)|^r \leq \mathbb{E} \sup_{0 \leq t \leq u \wedge \pi_R} |X(t)|^r.$$

Since $\pi_R \rightarrow T$, $\mathbb{P}$ -almost surely, we conclude the proof of the estimate (2.9) by applying Grönwall's inequality and the Fatou's lemma. Specifically, we have

$$\mathbb{E} \sup_{0 \leq t \leq T} |X(t)|^r \leq \liminf_{R \rightarrow \infty} \mathbb{E} \sup_{0 \leq t \leq T \wedge \pi_R} |X(t)|^r \leq C_r < \infty.$$

Step 3: Uniqueness. Let $X(t)$, $Y(t)$ be two solutions of (2.8) on the same probability space with $X(0) = Y(0)$. By (2.14), for a fixed $r \geq \max\{\kappa^2/2, 4\}$, there exists a positive constant C_r such that

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} |X(t)|^r \right] \leq C_r, \quad \mathbb{E} \left[\sup_{0 \leq t \leq T} |Y(t)|^r \right] \leq C_r.$$

For a sufficiently large $R > 0$, we define the stopping time

$$\bar{\tau}_R := \inf \{t \in [0, T]: |X(t)| \vee |Y(t)| > R\}.$$

To proceed, we compare $|X(t) - Y(t)|$ and $\bar{\tau}_R$ in this context with $|X^{(n)}(t) - X^{(m)}(t)|$ and τ_R , as introduced in the proof of Lemma 3. Clearly, the same method as used in the proof of Lemma 3 can be applied here, yielding the following estimate:

$$\begin{aligned} \mathbb{E} \left[\sup_{0 \leq t \leq T} |X(t) - Y(t)|^2 \right] &= \mathbb{E} \left[\sup_{0 \leq t \leq T} |X(t) - Y(t)|^2 \mathbb{I}_{\{\bar{\tau}_R > T\}} \right] \\ &\quad + \mathbb{E} \left[\sup_{0 \leq t \leq T} |X(t) - Y(t)|^2 \mathbb{I}_{\{\bar{\tau}_R \leq T\}} \right] \\ &\leq \mathbb{E} \left[\sup_{0 \leq t \leq T} |X(t \wedge \bar{\tau}_R) - Y(t \wedge \bar{\tau}_R)|^2 \right] + C_1 \sqrt{\mathbb{P}(\bar{\tau}_R \leq T)} \leq \frac{C}{R^2}. \end{aligned}$$

Letting $R \rightarrow \infty$ gives the uniqueness of the solution to (2.8).

This completes the proof of Theorem 1.

3. Stochastic averaging principle

In this section, we establish a stochastic averaging principle for the following stochastic integral equation

$$\begin{aligned} X_\varepsilon(t) = & x_0 + \int_0^t b\left(\frac{s}{\varepsilon}, X_\varepsilon(s-), \mathcal{L}_{X_\varepsilon(s)}\right) ds + \int_0^t \sigma\left(\frac{s}{\varepsilon}, X_\varepsilon(s-), \mathcal{L}_{X_\varepsilon(s)}\right) dW(s) \\ & + \int_0^t \int_U h\left(\frac{s}{\varepsilon}, X_\varepsilon(s-), \mathcal{L}_{X_\varepsilon(s)}, z\right) \tilde{N}(ds, dz), \quad t \in [0, T], \end{aligned} \quad (3.1)$$

where ε is a small positive parameter ($0 < \varepsilon \ll 1$). Assuming that (3.1) satisfies the conditions specified in Assumptions 1–7, the existence and uniqueness of its solution follow directly as a consequence of Theorem 1.

As mentioned in Section 1, our main goal is to demonstrate that the solution $(X_\varepsilon(t))_{t \in [0, T]}$ of (3.1) can be approximated by a simpler (or averaged) process in an appropriate sense. To proceed, we associate (3.1) with the following averaged McKean–Vlasov SDE:

$$\begin{aligned} \bar{X}(t) = & x_0 + \int_0^t \bar{b}(\bar{X}(s-), \mathcal{L}_{\bar{X}(s)}) ds + \int_0^t \bar{\sigma}(\bar{X}(s-), \mathcal{L}_{\bar{X}(s)}) dW(s) \\ & + \int_0^t \int_U \bar{h}(\bar{X}(s-), \mathcal{L}_{\bar{X}(s)}, z) \tilde{N}(ds, dz), \quad t \in [0, T], \end{aligned} \quad (3.2)$$

where $\bar{b}: \mathbb{R}^d \times M_2(\mathbb{R}^d) \rightarrow \mathbb{R}^d$, $\bar{\sigma}: \mathbb{R}^d \times M_2(\mathbb{R}^d) \rightarrow \mathbb{R}^{d \times m}$, and $\bar{h}: \mathbb{R}^d \times M_2(\mathbb{R}^d) \times U \rightarrow \mathbb{R}^d$ are Borel measurable functions. To ensure that (3.2) also admits a unique solution and to facilitate the application of stochastic averaging techniques, we impose specific averaging conditions. It is worth noting that these conditions differ slightly from the classical ones (see, e.g., [36, 42]) due to the distinct characteristics of the nonlinear terms involved in the equation.

Assumption 8. (Averaging conditions.) *There exist positive bounded functions (sometimes referred to as rate functions of convergence) φ_i , defined on $[0, T]$, with $\lim_{t \rightarrow \infty} \varphi_i(t) = 0$ for $i = 1, 2, 3$, such that*

$$\begin{aligned} \frac{1}{t} \int_0^t |b(s, x, \mu) - \bar{b}(x, \mu)|^2 ds &\leq \varphi_1(t) C_R^b (1 + |x|^2), \\ \frac{1}{t} \int_0^t \|\sigma(s, x, \mu) - \bar{\sigma}(x, \mu)\|^2 ds &\leq \varphi_2(t) C_R^\sigma (1 + |x|^2), \\ \frac{1}{t} \int_0^t \int_U |h(s, x, \mu, z) - \bar{h}(x, \mu, z)|^2 \nu(dz) ds &\leq \varphi_3(t) C_R^h (1 + |x|^2), \end{aligned}$$

respectively, for all $t \in [0, T]$, $x, y \in \mathbb{R}^d$ with $|x| \vee |y| \leq R$, and $\mu \in \mathcal{M}_2(\mathbb{R}^d)$. Here, C_R^b , C_R^σ , and C_R^h are positive constants.

Furthermore, if $\kappa > 2$, an additional condition is required.

Assumption 9. (Additional averaging conditions on the jump coefficients.) *There exists a positive bounded function φ , defined on $[0, T]$, with $\lim_{t \rightarrow \infty} \varphi(t) = 0$, such that*

$$\frac{1}{t} \int_0^t \int_U |h(s, x, \mu, z) - \bar{h}(x, \mu, z)|^r \nu(dz) ds \leq \varphi(t) C_R^h (1 + |x|^r),$$

$$\frac{1}{t} \int_0^t \int_U |h(s, x, \mu, z) - \bar{h}(x, \mu, z)|^k \nu(dz) ds \leq \varphi(t) C_R^h (1 + |x|^k),$$

respectively, for all $t \in [0, T]$, $x, y \in \mathbb{R}^d$ with $|x| \vee |y| \leq R$, and $\mu \in \mathcal{M}_2(\mathbb{R}^d)$.

The main theorem on the averaging principle for (3.1) is thus formulated as follows.

Theorem 2. (Averaging principle.) *Suppose that Assumptions 1–9 hold. Then, the following averaging principle holds:*

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \sup_{0 \leq t \leq T} |X_\varepsilon(t) - \bar{X}(t)|^2 = 0. \quad (3.3)$$

As a direct consequence of Theorem 2 and by applying the Chebyshev–Markov inequality, we have the following corollary.

Corollary 1. *The solution $X_\varepsilon(t)$ converges in probability to the averaged solution $\bar{X}(t)$. Specifically, for any $\delta > 0$,*

$$\mathbb{P}(\sup_{0 \leq t \leq T} |X_\varepsilon(t) - \bar{X}(t)| > \delta) \rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0.$$

Prior to establishing Theorem 2, it is necessary to address the well-posedness of the averaged equation (3.2). The following lemma ensures this property.

Lemma 4. *Under Assumptions 1–9, there exists a unique solution $\bar{X}(t)$ to the averaged equation (3.2).*

Proof. By Theorem 1, it suffices to verify that the coefficients functions $\bar{b}$, $\bar{\sigma}$, and $\bar{h}$ satisfy the conditions required for the existence and uniqueness of the solution. Note that both (3.1) and (3.2) share the same initial condition x_0 . The condition in Assumption 6 is directly satisfied. Regarding the conditions in Assumptions 1–5, we focus on the function $\bar{b}$, as similar arguments apply to the functions $\bar{\sigma}$ and $\bar{h}$. Finally, we verify that $\bar{h}$ satisfies the condition in Assumption 7. The details of these verifications are provided in Appendix A. $\square$

We now complete the proof of Theorem 2 as follows.

Proof of Theorem 2. For any $t \in [0, T]$, it follows from (3.1) and (3.2) that

$$\begin{aligned} X_\varepsilon(t) - \bar{X}(t) &= \int_0^t \left[b\left(\frac{s}{\varepsilon}, X_\varepsilon(s-), \mathcal{L}_{X_\varepsilon(s)}\right) - \bar{b}(\bar{X}(s-), \mathcal{L}_{\bar{X}(s)}) \right] ds \\ &\quad + \int_0^t \left[\sigma\left(\frac{s}{\varepsilon}, X_\varepsilon(s-), \mathcal{L}_{X_\varepsilon(s)}\right) - \bar{\sigma}(\bar{X}(s-), \mathcal{L}_{\bar{X}(s)}) \right] dW(s) \\ &\quad + \int_0^t \int_U \left[h\left(\frac{s}{\varepsilon}, X_\varepsilon(s-), \mathcal{L}_{X_\varepsilon(s)}, z\right) - \bar{h}(\bar{X}(s-), \mathcal{L}_{\bar{X}(s)}, z) \right] \tilde{N}(ds, dz). \end{aligned}$$

To handle the one-sided locally Lipschitz case, we introduce a stopping time η_R for each $R > 0$ defined as

$$\eta_R := \inf \{t \in [0, T] : |X_\varepsilon(t)| \vee |\bar{X}(t)| > R\}.$$

Using De Morgan's law, the following decomposition holds:

$$\begin{aligned} \mathbb{E} \left[\sup_{0 \leq t \leq T} |X_\varepsilon(t) - \bar{X}(t)|^2 \right] &= \mathbb{E} \left[\sup_{0 \leq t \leq T} |X_\varepsilon(t) - \bar{X}(t)|^2 \mathbb{I}_{\{\eta_R > T\}} \right] \\ &\quad + \mathbb{E} \left[\sup_{0 \leq t \leq T} |X_\varepsilon(t) - \bar{X}(t)|^2 \mathbb{I}_{\{\eta_R \leq T\}} \right] \\ &=: I_1 + I_2. \end{aligned}$$

We now proceed to estimate each term on the right-hand side of the equation above.

(1) Estimation of the term I_1 . We begin by bounding the term I_1 as follows

$$I_1 = \mathbb{E} \left[\sup_{0 \leq t \leq T} |X_\varepsilon(t) - \bar{X}(t)|^2 \mathbb{I}_{\{\eta_R > T\}} \right] \leq \mathbb{E} \left[\sup_{0 \leq t \leq T} |X_\varepsilon(t \wedge \eta_R) - \bar{X}(t \wedge \eta_R)|^2 \right] < \infty. \quad (3.4)$$

Here, we are effectively considering the process up to the stopping time η_R , which ensures that $|X_\varepsilon(t-)|$ and $|\bar{X}(t-)|$ are bounded by R for all $t \leq T$. By applying the Itô's formula, we obtain

$$|X_\varepsilon(t \wedge \eta_R) - \bar{X}(t \wedge \tau_R)|^2 = \sum_{i=1}^5 \Lambda_i(t),$$

where the terms Λ_i for $i = 1, \dots, 5$ are defined as follows:

$$\begin{aligned} \Lambda_1(t) &= 2 \int_0^{t \wedge \eta_R} \left\langle X_\varepsilon(s-) - \bar{X}(s-), b\left(\frac{s}{\varepsilon}, X_\varepsilon(s-), \mathcal{L}_{X_\varepsilon(s)}\right) - \bar{b}\left(\bar{X}(s-), \mathcal{L}_{\bar{X}(s)}\right) \right\rangle ds, \\ \Lambda_2(t) &= \int_0^{t \wedge \eta_R} \left\| \sigma\left(\frac{s}{\varepsilon}, X_\varepsilon(s-), \mathcal{L}_{X_\varepsilon(s)}\right) - \bar{\sigma}\left(\bar{X}(s-), \mathcal{L}_{\bar{X}(s)}\right) \right\|^2 ds, \\ \Lambda_3(t) &= \int_0^{t \wedge \eta_R} \int_U \left| h\left(\frac{s}{\varepsilon}, X_\varepsilon(s-), \mathcal{L}_{X_\varepsilon(s)}, z\right) - \bar{h}\left(\bar{X}(s-), \mathcal{L}_{\bar{X}(s)}, z\right) \right|^2 N(ds, dz), \\ \Lambda_4(t) &= 2 \int_0^{t \wedge \eta_R} \left\langle X_\varepsilon(s-) - \bar{X}(s-), \sigma\left(\frac{s}{\varepsilon}, X_\varepsilon(s-), \mathcal{L}_{X_\varepsilon(s)}\right) \right. \\ &\quad \left. - \bar{\sigma}\left(\bar{X}(s-), \mathcal{L}_{\bar{X}(s)}\right) dW(s) \right\rangle, \\ \Lambda_5(t) &= 2 \int_0^{t \wedge \eta_R} \left\langle X_\varepsilon(s-) - \bar{X}(s-), h\left(\frac{s}{\varepsilon}, X_\varepsilon(s-), \mathcal{L}_{X_\varepsilon(s)}, z\right) \right. \\ &\quad \left. - \bar{h}\left(\bar{X}(s-), \mathcal{L}_{\bar{X}(s)}, z\right) \right\rangle \tilde{N}(ds, dz). \end{aligned}$$

By taking the supremum over $[0, u]$ for $u \in [0, T]$ and then taking expectations, we can now estimate $\mathbb{E} \left[\sup_{0 \leq t \leq u} \Lambda_i(t) \right]$ for $i = 1, \dots, 5$, respectively. In view of Assumptions 1, 2, and 8, we obtain

$$\begin{aligned} &\mathbb{E} \left[\sup_{0 \leq t \leq u} \Lambda_1(t) \right] \\ &\leq 2 \mathbb{E} \left[\sup_{0 \leq t \leq u} \int_0^{t \wedge \eta_R} \left\langle X_\varepsilon(s-) - \bar{X}(s-), b\left(\frac{s}{\varepsilon}, X_\varepsilon(s-), \mathcal{L}_{X_\varepsilon(s)}\right) \right. \right. \end{aligned}$$

$$\begin{aligned}
& -b\left(\frac{s}{\varepsilon}, \bar{X}(s-), \mathcal{L}_{X_\varepsilon(s)}\right)\bigg\rangle ds\bigg] \\
& + 2\mathbb{E}\left[\sup_{0\leq t\leq u}\int_0^{t\wedge\eta_R}|X_\varepsilon(s-)-\bar{X}(s-)|\cdot\left|b\left(\frac{s}{\varepsilon}, \bar{X}(s-), \mathcal{L}_{X_\varepsilon(s)}\right)\right.\right. \\
& \quad \left.\left.-b\left(\frac{s}{\varepsilon}, \bar{X}(s-), \mathcal{L}_{\bar{X}(s)}\right)\right|\right] ds\bigg] \\
& + 2\mathbb{E}\left[\sup_{0\leq t\leq u}\int_0^{t\wedge\eta_R}|X_\varepsilon(s-)-\bar{X}(s-)|\cdot\left|b\left(\frac{s}{\varepsilon}, \bar{X}(s-), \mathcal{L}_{\bar{X}(s)}\right)\right.\right. \\
& \quad \left.\left.-\bar{b}\left(\bar{X}(s-), \mathcal{L}_{\bar{X}(s)}\right)\right|\right] ds\bigg] \\
& \leq 2L_R\mathbb{E}\left[\int_0^{u\wedge\eta_R}|X_\varepsilon(s-)-\bar{X}(s-)|^2 ds\right] \\
& \quad + 2\sqrt{L}\mathbb{E}\left[\int_0^{u\wedge\eta_R}|X_\varepsilon(s-)-\bar{X}(s-)|\cdot W_2\left(\mathcal{L}_{X_\varepsilon(s)}, \mathcal{L}_{\bar{X}(s)}\right) ds\right] \\
& \quad + 2\mathbb{E}\left[\int_0^{u\wedge\eta_R}|X_\varepsilon(s-)-\bar{X}(s-)|\cdot\left|b\left(\frac{s}{\varepsilon}, \bar{X}(s-), \mathcal{L}_{\bar{X}(s)}\right)\right.\right. \\
& \quad \left.\left.-\bar{b}\left(\bar{X}(s-), \mathcal{L}_{\bar{X}(s)}\right)\right|\right] ds\bigg] \\
& \leq (2L_R + 2\sqrt{L} + 1)\mathbb{E}\left[\int_0^{u\wedge\eta_R}|X_\varepsilon(s-)-\bar{X}(s-)|^2 ds\right] \\
& \quad + u\mathbb{E}\left[\frac{\varepsilon}{u\wedge\eta_R}\int_0^{\frac{u\wedge\eta_R}{\varepsilon}}\left|b\left(s, \bar{X}(s\varepsilon-), \mathcal{L}_{\bar{X}(s\varepsilon)}\right)-\bar{b}\left(\bar{X}(s\varepsilon-), \mathcal{L}_{\bar{X}(s\varepsilon)}\right)\right|^2 ds\right] \\
& \leq (2L_R + 2\sqrt{L} + 1)\int_0^u\mathbb{E}\sup_{0\leq t\leq s}|X_\varepsilon(t\wedge\eta_R)-\bar{X}(t\wedge\eta_R)|^2 ds + uC_R^b\varphi_1\left(\frac{u\wedge\eta_R}{\varepsilon}\right) \\
& \quad \times\left(1+\mathbb{E}\sup_{0\leq t\leq u}|\bar{X}(t)|^2\right) \\
& \leq (2L_R + 2\sqrt{L} + 1)\int_0^u\mathbb{E}\sup_{0\leq t\leq s}|X_\varepsilon(t\wedge\eta_R)-\bar{X}(t\wedge\eta_R)|^2 ds + uC_R^b\cdot C\varphi_1\left(\frac{u\wedge\eta_R}{\varepsilon}\right).
\end{aligned} \tag{3.5}$$

Here, we have used the fact that for each $u \in [0, T]$,

$$\left(\mathbb{E}\sup_{0\leq t\leq u}|\bar{X}(t)|^2\right)^{\frac{1}{2}}\leq(\mathbb{E}\sup_{0\leq t\leq u}|\bar{X}(t)|^r)^{\frac{1}{r}}<\infty,\quad\text{if }\mathbb{E}|X_\varepsilon(0)|^r<\infty.$$

By applying Assumptions 1, 2, 8, and employing techniques analogous to those used in deriving the estimate for Λ_1 , we establish the following bound for Λ_2 :

$$\begin{aligned}
& \mathbb{E}\left[\sup_{0\leq t\leq u}\Lambda_2(t)\right] \\
& \leq 3\mathbb{E}\left[\sup_{0\leq t\leq u}\int_0^{t\wedge\eta_R}\left\|\sigma\left(\frac{s}{\varepsilon}, X_\varepsilon(s-), \mathcal{L}_{X_\varepsilon(s)}\right)-\sigma\left(\frac{s}{\varepsilon}, \bar{X}(s-), \mathcal{L}_{X_\varepsilon(s)}\right)\right\|^2 ds\right] \\
& \quad + 3\mathbb{E}\left[\sup_{0\leq t\leq u}\int_0^{t\wedge\eta_R}\left\|\sigma\left(\frac{s}{\varepsilon}, \bar{X}(s-), \mathcal{L}_{X_\varepsilon(s)}\right)-\sigma\left(\frac{s}{\varepsilon}, \bar{X}(s-), \mathcal{L}_{\bar{X}(s)}\right)\right\|^2 ds\right]
\end{aligned}$$

$$\begin{aligned}
& + 3\mathbb{E} \left[\sup_{0 \leq t \leq u} \int_0^{t \wedge \eta_R} \left\| \sigma \left(\frac{s}{\varepsilon}, \bar{X}(s-), \mathcal{L}_{\bar{X}(s)} \right) - \bar{\sigma} \left(\bar{X}(s-), \mathcal{L}_{\bar{X}(s)} \right) \right\|^2 ds \right] \\
& \leq 3L_R \mathbb{E} \left[\int_0^{u \wedge \eta_R} |X_\varepsilon(s-) - \bar{X}(s-)|^2 ds \right] + 3L \mathbb{E} \left[\int_0^{u \wedge \eta_R} W_2^2(\mathcal{L}_{X_\varepsilon(s)}, \mathcal{L}_{\bar{X}(s)}) ds \right] \\
& + 3\mathbb{E} \left[\int_0^{u \wedge \eta_R} \left\| \sigma \left(\frac{s}{\varepsilon}, \bar{X}(s-), \mathcal{L}_{\bar{X}(s)} \right) - \bar{\sigma} \left(\bar{X}(s-), \mathcal{L}_{\bar{X}(s)} \right) \right\|^2 ds \right] \\
& \leq 3(L_R + L) \int_0^u \mathbb{E} \sup_{0 \leq t \leq s} |X_\varepsilon(t \wedge \eta_R) - \bar{X}(t \wedge \eta_R)|^2 ds + 3uC_R^\sigma \cdot C\varphi_2 \left(\frac{u \wedge \eta_R}{\varepsilon} \right). \quad (3.6)
\end{aligned}$$

Similarly, using Assumptions 1, 2, and 8, we obtain the following estimate for Λ_3 :

$$\begin{aligned}
& \mathbb{E} \left[\sup_{0 \leq t \leq u} \Lambda_3(t) \right] \\
& \leq \mathbb{E} \left[\int_0^{t \wedge \eta_R} \int_U \left| h \left(\frac{s}{\varepsilon}, X_\varepsilon(s-), \mathcal{L}_{X_\varepsilon(s)}, z \right) - \bar{h} \left(X(s-), \mathcal{L}_{\bar{X}(s)}, z \right) \right|^2 \nu(dz) ds \right] \\
& \leq 3(L_R + L) \int_0^u \mathbb{E} \sup_{0 \leq t \leq s} |X_\varepsilon(t \wedge \eta_R) - \bar{X}(t \wedge \eta_R)|^2 ds + 3uC_R^h \cdot C\varphi_3 \left(\frac{u \wedge \eta_R}{\varepsilon} \right). \quad (3.7)
\end{aligned}$$

Next, we apply the BDG inequality, along with Young's inequality (2.1) (with $\epsilon = \frac{1}{24}$ and $p = q = 2$) and the estimate (3.6), to derive the following bound for Λ_4 :

$$\begin{aligned}
& \mathbb{E} \left[\sup_{0 \leq t \leq u} \Lambda_4(t) \right] \\
& \leq 2\sqrt{32}\mathbb{E} \left(\int_0^{u \wedge \eta_R} |X_\varepsilon(s-) - \bar{X}(s-)|^2 \cdot \left\| \sigma \left(\frac{s}{\varepsilon}, X_\varepsilon(s-), \mathcal{L}_{X_\varepsilon(s)} \right) \right. \right. \\
& \quad \left. \left. - \bar{\sigma} \left(\bar{X}(s-), \mathcal{L}_{\bar{X}(s)} \right) \right\|^2 ds \right)^{\frac{1}{2}} \\
& \leq 12\mathbb{E} \left[\sup_{0 \leq t \leq u} |X_\varepsilon(t \wedge \eta_R) - \bar{X}(t \wedge \eta_R)| \left(\int_0^{u \wedge \eta_R} \left\| \sigma \left(\frac{s}{\varepsilon}, X_\varepsilon(s-), \mathcal{L}_{X_\varepsilon(s)} \right) \right. \right. \right. \\
& \quad \left. \left. - \bar{\sigma} \left(\bar{X}(s-), \mathcal{L}_{\bar{X}(s)} \right) \right\|^2 ds \right)^{\frac{1}{2}} \right] \\
& \leq \frac{1}{4}\mathbb{E} \left[\sup_{0 \leq t \leq u} |X_\varepsilon(t \wedge \eta_R) - \bar{X}(t \wedge \eta_R)|^2 \right] + 144\mathbb{E} \left[\int_0^{u \wedge \eta_R} \left\| \sigma \left(\frac{s}{\varepsilon}, X_\varepsilon(s-), \mathcal{L}_{X_\varepsilon(s)} \right) \right. \right. \\
& \quad \left. \left. - \bar{\sigma} \left(\bar{X}(s-), \mathcal{L}_{\bar{X}(s)} \right) \right\|^2 ds \right] \\
& \leq \frac{1}{4}\mathbb{E} \left[\sup_{0 \leq t \leq u} |X_\varepsilon(t \wedge \eta_R) - \bar{X}(t \wedge \eta_R)|^2 \right] + 432(L_R + L) \int_0^u \mathbb{E} \sup_{0 \leq t \leq s} |X_\varepsilon(t \wedge \eta_R) \\
& \quad - \bar{X}(t \wedge \eta_R)|^2 ds + 432uC_R^\sigma \cdot C\varphi_2 \left(\frac{u \wedge \eta_R}{\varepsilon} \right). \quad (3.8)
\end{aligned}$$

By applying the inequality (2.5) (with $p = 1$) from Proposition 2, Young's inequality (2.1) (with $\epsilon = \frac{1}{4D}$, $p = q = 2$) and the estimate (3.7), we arrive at the following

estimate for Λ_5 :

$$\begin{aligned}
 & \mathbb{E} \left[\sup_{0 \leq t \leq u} \Lambda_5(t) \right] \\
 & \leq 2D \mathbb{E} \left(\int_0^{u \wedge \eta_R} \int_U |X_\varepsilon(s-) - \bar{X}(s-)|^2 \left| h \left(\frac{s}{\varepsilon}, X_\varepsilon(s-), \mathcal{L}_{X_\varepsilon(s)}, z \right) \right. \right. \\
 & \quad \left. \left. - \bar{h} \left(\bar{X}(s-), \mathcal{L}_{\bar{X}(s)}, z \right) \right|^2 \nu(dz) ds \right)^{\frac{1}{2}} \\
 & \leq \frac{1}{4} \mathbb{E} \left[\sup_{0 \leq t \leq u} |X_\varepsilon(t \wedge \eta_R) - \bar{X}(t \wedge \eta_R)|^2 \right] \\
 & \quad + 4D^2 \mathbb{E} \left[\int_0^{u \wedge \eta_R} \int_U \left| h \left(\frac{s}{\varepsilon}, X_\varepsilon(s-), \mathcal{L}_{X_\varepsilon(s)}, z \right) - \bar{h} \left(\bar{X}(s-), \mathcal{L}_{\bar{X}(s)}, z \right) \right|^2 \nu(dz) ds \right] \\
 & \leq \frac{1}{4} \mathbb{E} \left[\sup_{0 \leq t \leq u} |X_\varepsilon(t \wedge \eta_R) - \bar{X}(t \wedge \eta_R)|^2 \right] \\
 & \quad + 12D^2(L_R + L) \int_0^u \mathbb{E} \sup_{0 \leq t \leq s} |X_\varepsilon(t \wedge \eta_R) - \bar{X}(t \wedge \eta_R)|^2 ds \\
 & \quad + 12D^2 u C_R^h \cdot C\varphi_3 \left(\frac{u \wedge \eta_R}{\varepsilon} \right). \tag{3.9}
 \end{aligned}$$

Finally, substituting the estimates (3.5)–(3.9) into the expression (3.4) for I_1 , and further utilizing Grönwall's inequality, we obtain

$$I_1 \leq \mathbb{E} \left[\sup_{0 \leq t \leq T} |X_\varepsilon(t \wedge \eta_R) - \bar{X}(t \wedge \eta_R)|^2 \right] \leq \widehat{N}_2 e^{\widehat{N}_1 T} \tag{3.10}$$

with $\widehat{N}_1 = 4(L_R + \sqrt{L}) + 2 + 12 \cdot (73 + 2D^2)(L_R + L)$ and $\widehat{N}_2 = 2TC_R^b \cdot C\varphi_1 \left(\frac{T \wedge \eta_R}{\varepsilon} \right) + 870TC_R^\sigma \cdot C\varphi_2 \left(\frac{T \wedge \eta_R}{\varepsilon} \right) + 6(1 + 4D^2)TC_R^\sigma C\varphi_3 \left(\frac{T \wedge \eta_R}{\varepsilon} \right)$.

- (2) Estimation of the term I_2 . Using the Cauchy–Schwarz inequality and Theorem 1, we deduce that

$$\begin{aligned}
 I_2 &= \mathbb{E} \left[\sup_{0 \leq t \leq T} |X_\varepsilon(t) - \bar{X}(t)|^2 \mathbb{I}_{\{\eta_R \leq T\}} \right] \leq \sqrt{\mathbb{E} \left(\sup_{0 \leq t \leq T} |X_\varepsilon(t) - \bar{X}(t)|^2 \right)^2} \sqrt{\mathbb{E} (\mathbb{I}_{\{\eta_R \leq T\}})^2} \\
 &\leq 2\sqrt{2} \sqrt{\mathbb{E} \left(\sup_{0 \leq t \leq T} |X_\varepsilon(t)|^4 + \sup_{0 \leq t \leq T} |\bar{X}(t)|^4 \right)} \sqrt{\mathbb{E} \left(\mathbb{I}_{\{\eta_R \leq T\}} \frac{|X_\varepsilon(\eta_R)|^4 + |\bar{X}(\eta_R)|^4}{R^4} \right)} \\
 &\leq \frac{2\sqrt{2}}{R^2} \left(\mathbb{E} \sup_{0 \leq t \leq T} |X_\varepsilon(t)|^4 + \mathbb{E} \sup_{0 \leq t \leq T} |\bar{X}(t)|^4 \right) \leq \frac{C}{R^2}. \tag{3.11}
 \end{aligned}$$

By combining the estimates on I_1 and I_2 , i.e. (3.10) and (3.11), we conclude that

$$\mathbb{E} \sup_{0 \leq t \leq T} |X_\varepsilon(t) - \bar{X}(t)|^2 \leq \widehat{N}_2 e^{\widehat{N}_1 T} + \frac{C}{R^2}.$$

Now, for any $\delta > 0$, we can choose $R > 0$ large enough such that $\frac{C}{R^2} < \frac{\delta}{2}$. In addition, by taking ε sufficiently small and using the averaging condition 8, we obtain that

$$\widehat{N}_2 e^{\widehat{N}_1 T} < \frac{\delta}{2}.$$

Thus, the arbitrariness of δ implies that $\mathbb{E} \sup_{0 \leq t \leq T} |X_\varepsilon(t) - \bar{X}(t)|^2$ converges to 0, as ε goes to 0. This completes the proof. $\square$

4. Example

In this section, we provide an illustrative example to demonstrate the theoretical results established in this paper. We highlight that the model (4.1) is carefully designed to satisfy all the conditions of our assumptions and to facilitate the explicit derivation of the corresponding averaged equation.

Example 1. Consider the following one-dimensional McKean–Vlasov SDE

$$\begin{aligned} dX_\varepsilon(t) = & \left[\left(X_\varepsilon(t-) - X_\varepsilon^3(t-) \right) \frac{\frac{t}{\varepsilon}}{1 + \frac{t}{\varepsilon}} + \mathbb{E} X_\varepsilon(t-) \right] dt \\ & + \left[X_\varepsilon(t-) \sin \left(\log^2 \left(1 + X_\varepsilon^2(t-) \right) \right) \frac{\frac{t}{\varepsilon}}{2 + \frac{t}{\varepsilon}} + \mathbb{E} X_\varepsilon(t-) \right] dW(t) \\ & + \int_U \left[X_\varepsilon(t-) \sin \left(\log^{\frac{3}{2}} \left(1 + X_\varepsilon^2(t-) \right) \right) \left(1 - e^{-\frac{t}{\varepsilon}} \right) + \mathbb{E} X_\varepsilon(t-) \right] \tilde{N}(dt, dz), \quad (4.1) \end{aligned}$$

with $t \in [0, T]$ and the initial condition $X_\varepsilon(0) = x_0$. Here, $W(t)$ is a scalar Wiener process, $U = \mathbb{R} \setminus \{0\}$, and ν is a finite measure with $\nu(U) = 1$. Define the following functions:

$$\begin{aligned} b(t, x, \mu) &= (x - x^3) \frac{t}{1+t} + \int_{\mathbb{R}} y \mu(dy), \quad \sigma(t, x, \mu) = \psi(x) \frac{t}{2+t} + \int_{\mathbb{R}} y \mu(dy), \\ h(t, x, \mu, z) &= \phi(x)(1 - e^{-t}) + \int_{\mathbb{R}} y \mu(dy), \end{aligned}$$

where $\psi(x) = x \sin(\log^2(1+x^2))$ and $\phi(x) = x \sin(\log^{\frac{3}{2}}(1+x^2))$ are continuously differentiable functions. For any $x \in \mathbb{R}$, we can show that

$$\begin{aligned} |\psi(x)| &\leq |x|, \quad |\phi(x)| \leq |x|, \quad |(\partial_x \psi)(x)| \leq 1 + 4 \log(1+x^2) \\ \text{and} \quad |(\partial_x \phi)(x)| &\leq 1 + 3\sqrt{\log(1+x^2)}. \end{aligned} \quad (4.2)$$

- (1) **Well-posedness.** To show that (4.1) has a unique solution $(X_\varepsilon(t))_{t \in [0, T]}$, we need to verify that the conditions in Theorem 1 are satisfied. For any $R > 0$, $x, y \in \mathbb{R}$ with $|x| \vee |y| \leq R$ and $\mu \in \mathcal{M}_2(\mathbb{R})$, we provide the following estimates:

$$\begin{aligned} (x-y)(b(t, x, \mu) - b(t, y, \mu)) &= (x-y)(x - x^3 - y + y^3) \frac{t}{1+t} \\ &\leq |x-y|^2 - |x-y|^2(x^2 + xy + y^2) \\ &\leq |x-y|^2(1 - xy) \leq (1 + R^2)|x-y|^2 =: L_R^1 |x-y|^2, \end{aligned}$$

$$\begin{aligned}
|\sigma(t, x, \mu) - \sigma(t, y, \mu)|^2 &= |\psi(x) - \psi(y)|^2 \left(\frac{t}{2+t} \right)^2 \\
&\leq \left| \int_0^1 (\partial_x \psi)(y + \theta(x-y))(x-y) d\theta \right|^2 \\
&\leq \left[\sup_{|z| \leq R} |(\partial_x \psi)(z)| \right]^2 |x-y|^2 \\
&\leq \left[\sup_{|z| \leq R} \left(1 + 4 \log(1+z^2) \right) \right]^2 |x-y|^2 \\
&\leq \left(1 + 4 \log(1+R^2) \right)^2 |x-y|^2 =: L_R^2 |x-y|^2,
\end{aligned}$$

$$\begin{aligned}
\int_U |h(t, x, \mu, z) - h(t, y, \mu, z)|^2 v(dz) &= |\phi(x) - \phi(y)|^2 (1 - e^{-t})^2 \\
&\leq \left| \int_0^1 (\partial_x \phi)(y + \theta(x-y))(x-y) d\theta \right|^2 \\
&\leq \left[\sup_{|z| \leq R} |(\partial_x \phi)(z)| \right]^2 \cdot |x-y|^2 \\
&\leq \left[\sup_{|z| \leq R} \left(1 + 3\sqrt{\log(1+z^2)} \right) \right]^2 |x-y|^2 \\
&\leq \left(1 + 3\sqrt{\log(1+R^2)} \right)^2 |x-y|^2 =: L_R^3 |x-y|^2.
\end{aligned}$$

These estimates imply that Assumption 1 is satisfied by denoting $L_R = \max\{L_R^1, L_R^2, L_R^3\}$. Next, we estimate the following for any $x \in \mathbb{R}$ and $\mu_1, \mu_2 \in \mathcal{M}_2(\mathbb{R})$

$$\begin{aligned}
&|b(t, x, \mu_1) - b(t, x, \mu_2)|^2 + \|\sigma(t, x, \mu_1) - \sigma(t, x, \mu_2)\|^2 \\
&+ \int_U |h(t, x, \mu_1, z) - h(t, x, \mu_2, z)|^2 v(dz) \\
&= 3 \left| \int_{\mathbb{R}} y \mu_1(dy) - \int_{\mathbb{R}} y \mu_2(dy) \right|^2 \leq 3W_2^2(\mu_1, \mu_2),
\end{aligned}$$

which shows that b, σ , and h satisfy Assumptions 2 and 3. Furthermore, using the bounds in (4.2) and the fact that $\frac{t}{1+t}$, $\frac{t}{2+t}$, and $1 - e^{-t}$ are bounded, we deduce that for any $x \in \mathbb{R}^d$ and $\mu \in \mathcal{M}_2(\mathbb{R})$,

$$\begin{aligned}
x \cdot b(t, x, \mu) &\leq x(x - x^3) \left(\frac{t}{1+t} \right) + x \int_{\mathbb{R}} y \mu(dy) \\
&\leq x^2 + \frac{1}{2}x^2 + \frac{1}{2} \left(\int_{\mathbb{R}} y \mu(dy) \right)^2 \leq 2 \left(1 + x^2 + W_2^2(\mu, \delta_0) \right), \\
|\sigma(t, x, \mu)|^2 &= \left| \psi(x) \frac{t}{2+t} + \int_{\mathbb{R}} y \mu(dy) \right|^2
\end{aligned}$$

$$\leq 2|\psi(x)|^2 + 2 \left(\int_{\mathbb{R}} y\mu(dy) \right)^2 \leq 2 \left(1 + |x|^2 + W_2^2(\mu, \delta_0) \right),$$

$$\begin{aligned} \int_U |h(t, x, \mu, z)|^2 \nu(dz) &= \left| \phi(x)(1 - e^{-t}) + \int_{\mathbb{R}} y\mu(dy) \right|^2 \leq 2|\phi(x)|^2 + 2 \left(\int_{\mathbb{R}} y\mu(dy) \right)^2 \\ &\leq 2 \left(1 + |x|^2 + W_2^2(\mu, \delta_0) \right). \end{aligned}$$

Thus, Assumption 4 holds. In addition, for any $x \in \mathbb{R}^d$ and $\mu \in \mathcal{M}_2(\mathbb{R})$, we have

$$\begin{aligned} |b(t, x, \mu)|^2 &\leq 2(x - x^3)^2 \left(\frac{t}{1+t} \right)^2 + 2 \left(\int_{\mathbb{R}} y\mu(dy) \right)^2 \leq 2x^2 - 4x^4 + 2x^6 + 2W_2^2(\mu, \delta_0) \\ &\leq 4 \left(1 + x^6 + W_2^6(\mu, \delta_0) \right). \end{aligned}$$

Thus, Assumption 5 holds with $\kappa = 6$. Finally, since $X_\varepsilon(0)$ is a constant, Assumption 6 (with $r \geq 18$) naturally holds. Due to the expression of h and the finiteness of ν , Assumption 7 can be easily verified using the same technique as Assumptions 1 and 2 were checked.

(2) **Averaging principle.** Define

$$\begin{aligned} \bar{b}(x, \mu) &= x - x^3 + \int_{\mathbb{R}} y\mu(dy), \quad \bar{\sigma}(x, \mu) = \psi(x) + \int_{\mathbb{R}} y\mu(dy), \\ \bar{h}(x, \mu, z) &= \phi(x) + \int_{\mathbb{R}} y\mu(dy). \end{aligned}$$

We can now verify that the averaging conditions in Assumptions 8 and 9 (with $\kappa = 6$ and $r \geq 18$) are satisfied:

$$\begin{aligned} \frac{1}{t} \int_0^t |b(s, x, \mu) - \bar{b}(x, \mu)|^2 ds &= \frac{1}{t} \int_0^t |x - x^3|^2 \left[1 - \frac{s}{1+s} \right]^2 ds = x^2(1 - x^2)^2 \frac{1}{1+t} \\ &\leq \varphi_1(t) C_R^b (1 + |x|^2) \end{aligned}$$

$$\begin{aligned} \frac{1}{t} \int_0^t |\sigma(s, x, \mu) - \bar{\sigma}(x, \mu)|^2 ds &= \frac{1}{t} \int_0^t \psi^2(x) \left[1 - \frac{s}{2+s} \right]^2 ds = \psi^2(x) \frac{2}{2+t} \\ &\leq \varphi_2(t) C_R^\sigma (1 + |x|^2) \end{aligned}$$

$$\begin{aligned} \frac{1}{t} \int_0^t \int_U |h(s, x, \mu, z) - \bar{h}(x, \mu, z)|^2 \nu(dz) ds &= \frac{1}{t} \int_0^t \phi^2(x) [1 - (1 - e^{-s})]^2 ds \\ &= \phi^2(x) \frac{1 - e^{-2t}}{2t} \leq \varphi_3(t) C_R^h (1 + |x|^2), \end{aligned}$$

and

$$\frac{1}{t} \int_0^t \int_U |h(s, x, \mu, z) - \bar{h}(x, \mu, z)|^l \nu(dz) ds = \frac{1}{t} \int_0^t \phi^l(x) [1 - (1 - e^{-s})]^l ds$$

$$= \phi^l(x) \frac{1 - e^{-lt}}{lt} \\ \leq \varphi(t) C_R^h \left(1 + |x|^l\right), \quad l = r \text{ or } \kappa,$$

for $x \in \mathbb{R}$ with $|x| \leq R$, where the functions $\varphi_1(t) = \frac{1}{1+t}$, $\varphi_2(t) = \frac{1}{2+t}$, $\varphi_3(t) = \frac{1-e^{-2t}}{2t}$, and $\varphi(t) = \frac{1-e^{-lt}}{lt}$ are continuous, positive, and bounded, with the property that $\lim_{t \rightarrow \infty} \varphi_i(t) = \lim_{t \rightarrow \infty} \varphi(t) = 0$, for $i = 1, 2, 3$.

Based on the discussion and the result of Theorem 2, the solution of (4.1) can be approximated by the following equation (for $t \in [0, T]$ and $\bar{X}(0) = x_0$)

$$d\bar{X}(t) = \left[\left(\bar{X}(t-) - \bar{X}^3(t-) \right) + \mathbb{E}\bar{X}(t-) \right] dt \\ + \left[\bar{X}(t-) \sin \left(\log^2 \left(1 + \bar{X}^2(t-) \right) \right) + \mathbb{E}\bar{X}(t-) \right] dW(t) \\ + \int_U \left[\bar{X}(t-) \sin \left(\log^{\frac{3}{2}} \left(1 + \bar{X}^2(t-) \right) \right) + \mathbb{E}\bar{X}(t-) \right] \tilde{N}(dt, dz), \quad (4.3)$$

in the sense of mean square.

We now carry out numerical simulations to compute the solutions of (4.1) and (4.3) with $x_0 = 1$, $T = 10$, $\varepsilon = 0.01$ and $x_0 = 1$, $T = 10$, $\varepsilon = 0.001$, respectively. Figure 1(a) and (b) illustrate the comparison between the solution $X_\varepsilon(t)$ of (4.1) and the averaged solution $\bar{X}(t)$ of (4.3). As shown, the solutions of the original equation and the averaged equation exhibit strong agreement. In addition, one can find that, for fixed sample points, the error $\sup_{0 \leq t \leq 10} |X_\varepsilon(t) - \bar{X}(t)|$ decreases when ε changes from 0.01 to 0.001. This observed behavior aligns with the predictions of the averaging principle stated in Theorem 2.

We remark that in our numerical simulations to approximate the McKean–Vlasov SDE (4.1) and (4.3), we use N -dimensional systems of interacting particles, which can be regarded as standard SDEs. This approach is based on the so-called propagation of chaos result (see Appendix B). Based on Proposition 3, we briefly introduce an Euler–Maruyama (EM) numerical scheme to approximate the solution of (B.2), which, in turn, serves as an approximation for the solution of the McKean–Vlasov SDE (2.8). To this end, we partition the time interval $[0, T]$ into n subintervals of equal length and define $t_k^n = kh_n$ for $k = 0, 1, \dots, n$, where $n \in \mathbb{N}$ and the step size is given by $h_n = \frac{T}{n}$. The EM scheme for the interacting particle system (B.2) is specified by the initial condition $X^{i,N,n}(0) = X^{i,N}(0)$ and the recurrence relation

$$X^{i,N,n}(t_{k+1}^n) = X^{i,N,n}(t_k^n) + b(t_k^n, X^{i,N,n}(t_k^n), \mu_{t_k^n}^{X,N,n})h_n + \sigma(t_k^n, X^{i,N,n}(t_k^n), \mu_{t_k^n}^{X,N,n})\Delta W^{i,n}(k) \\ + \int_{t_k^n}^{t_{k+1}^n} \int_U h(t_k^n, X^{i,N,n}(t_k^n), \mu_{t_k^n}^{X,N,n}, z) \tilde{N}^i(dt, dz), \quad i = 1, 2, \dots, N, \quad (4.4)$$

where $X^{i,N,n}(t_k^n)$ denotes the approximation of $X^{i,N}(t_k^n)$, $\mu_{t_k^n}^{X,N,n} = \frac{1}{N} \sum_{j=1}^N \delta_{X^{j,N,n}(t_k^n)}$ is the empirical measure, and $\Delta W^{i,n}(t) = W^i(t_{k+1}^n) - W^i(t_k^n)$ is the Brownian increment. To simulate the integrals w.r.t. the compensated Poisson random measure $\tilde{N}(dt, dz) = N(dt, dz) - \nu(dz) dt$, we also employ the technique of introducing a compound Poisson process $\int_U z \tilde{N}(t, dz)$, as detailed in [1, Section 4.3.2].

For this example, we simulate $N = 100$ particles with a time step 0.01, $T = 10$. Figure 2(a) and (b) depict the realizations of the interacting particle systems associated with the McKean–Vlasov SDEs (4.1) and (4.3), respectively, under the initial conditions $X_\varepsilon(0) = \bar{X}(0) = 1$ and

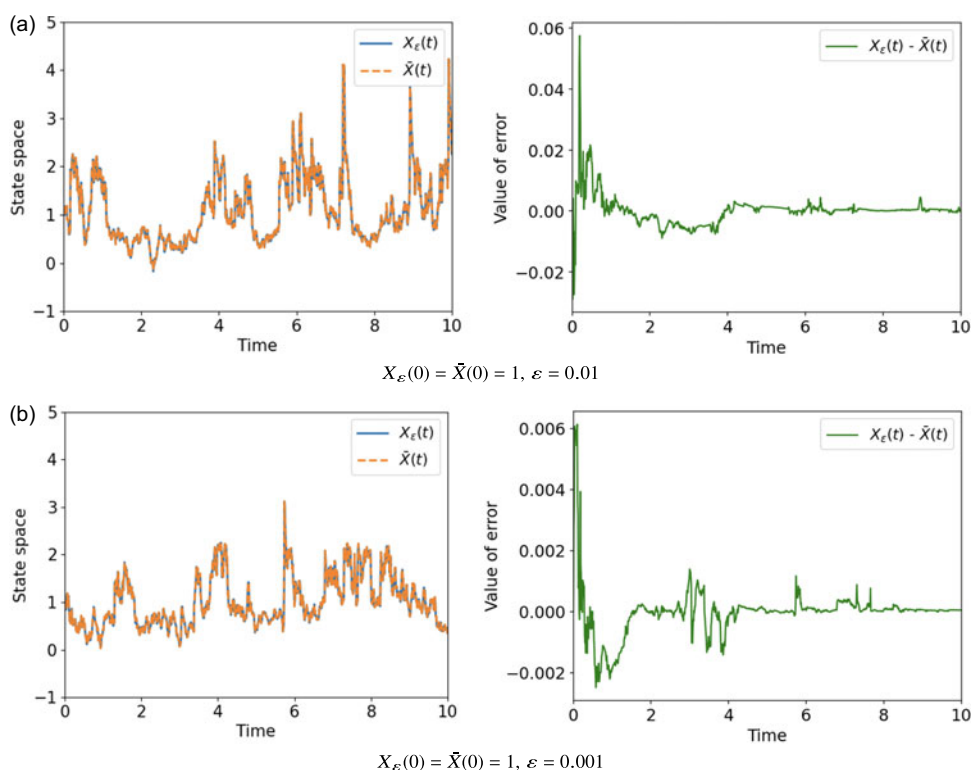


FIGURE 1. Comparison of the solution $X_\varepsilon(t)$ for (4.1) with the averaged solution $\tilde{X}(t)$ for (4.3): (a) $X_\varepsilon(0) = \tilde{X}(0) = 1$, $\varepsilon = 0.01$; (b) $X_\varepsilon(0) = \tilde{X}(0) = 1$, $\varepsilon = 0.001$.

$\varepsilon = 0.01$. Numerically, the Wasserstein distance between the distributions of the solutions to (4.1) (i.e. $\mathcal{L}_{X_\varepsilon(t)}$) and (4.3) (i.e. $\mathcal{L}_{\tilde{X}(t)}$) is approximated via the empirical distributions of the interacting particle systems, as illustrated in Figure 2(c).

Appendix A. Details of the proof for Lemma 4

Proof for Lemma 4. For any $t \in [0, T]$, $x, y \in \mathbb{R}^d$ and $\mu, \mu_1, \mu_2 \in \mathcal{M}_2(\mathbb{R}^d)$, we calculate successively that

$$\begin{aligned}
 \langle x - y, \bar{b}(x, \mu) - \bar{b}(y, \mu) \rangle &\leq |x - y|^2 + \frac{1}{2t} \int_0^t \left(|b(s, x, \mu) - \bar{b}(x, \mu)|^2 + |b(s, y, \mu) - \bar{b}(y, \mu)|^2 \right) \\
 &\quad \times ds + L_R |x - y|^2 \\
 &\leq \varphi_1(t) C_R^b (1 + |x|^2 + |y|^2) + \left(\frac{1}{2} + L_R \right) |x - y|^2, \\
 \langle x, \bar{b}(x, \mu) \rangle &\leq \frac{|x|^2}{2} + \frac{1}{2t} \int_0^t |b(s, x, \mu) - \bar{b}(x, \mu)|^2 ds + K(1 + |x|^2 + W_2^2(\mu, \delta_0)) \\
 &\leq \frac{1}{2} \varphi_1(t) C_R^b (1 + |x|^2) + \frac{3}{2} K(1 + |x|^2 + W_2^2(\mu, \delta_0)), \\
 |\bar{b}(x, \mu_1) - \bar{b}(x, \mu_2)|^2 &\leq \frac{3}{t} \int_0^t \left(|b(s, x, \mu_1) - \bar{b}(x, \mu_1)|^2 + |b(s, x, \mu_2) - \bar{b}(x, \mu_2)|^2 \right)
 \end{aligned}$$

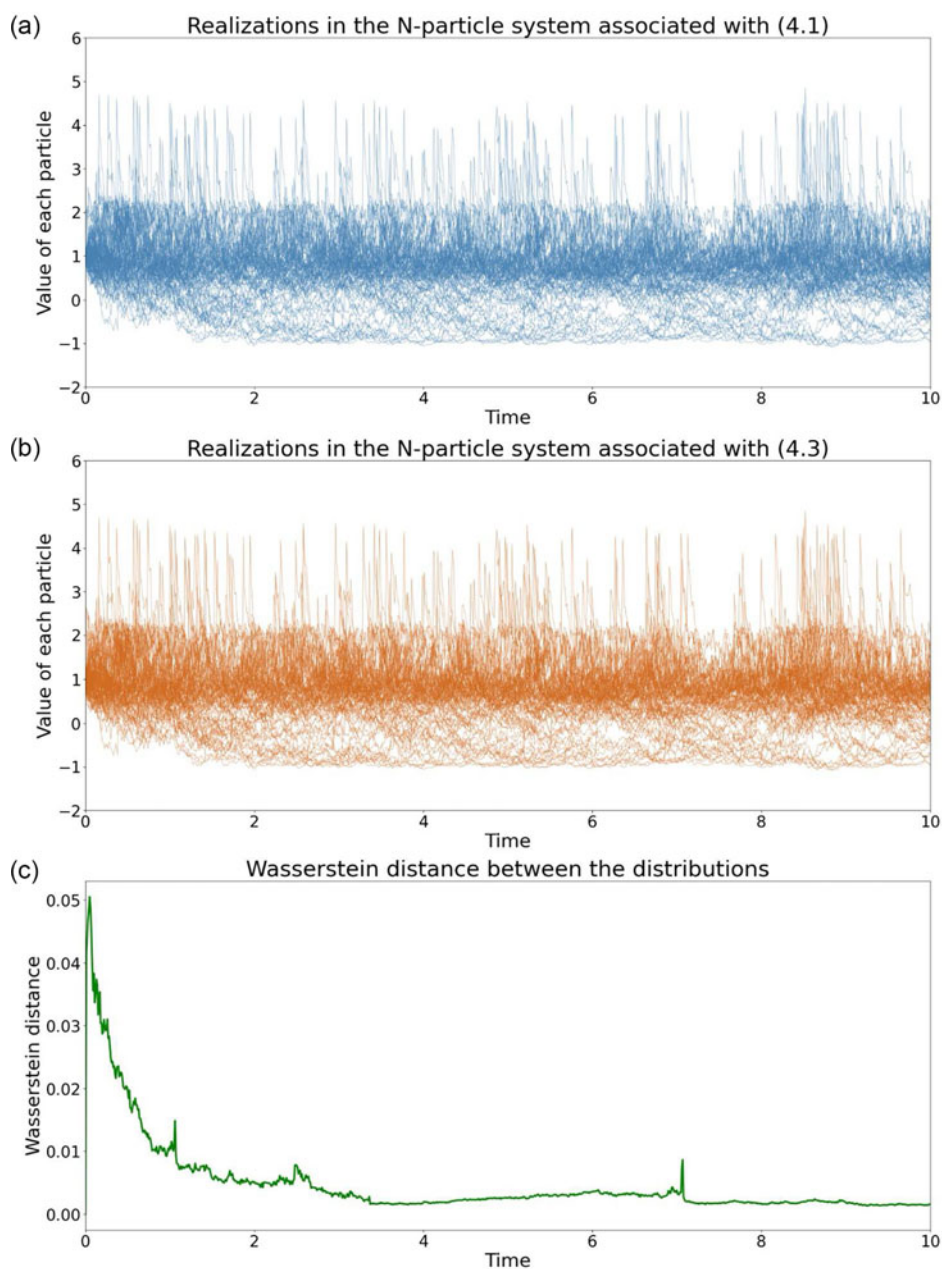


FIGURE 2. Comparison of the interacting N -particle systems associated with McKean–Vlasov SDEs (4.1) and (4.3), for $N = 100$ and $\varepsilon = 0.01$.

$$\begin{aligned}
& + |b(s, x, \mu_1) - b(s, x, \mu_2)|^2 \Big) \, ds \\
& \leq 6C_R^b \varphi_1(t)(1 + |x|^2) + 3LW_2^2(\mu_1, \mu_2), \\
|\bar{b}(x, \mu_1) - \bar{b}(y, \mu_2)|^2 & \leq \frac{3}{t} \int_0^t \Big(|b(s, x, \mu_1) - \bar{b}(x, \mu_1)|^2 + |b(s, y, \mu_2) - \bar{b}(y, \mu_2)|^2 \\
& + |b(s, x, \mu_1) - b(s, y, \mu_2)|^2 \Big) \, ds \\
& \leq 6C_R^b \varphi_1(t)(1 + |x|^2) + \frac{3}{t} \int_0^t |b(s, x, \mu_1) - b(s, y, \mu_2)|^2 \, ds, \\
|\bar{b}(x, \mu)|^2 & \leq 2 \left| \frac{1}{t} \int_0^t b(s, x, \mu) - \bar{b}(x, \mu) \, ds \right|^2 + 2 \left| \frac{1}{t} \int_0^t b(s, x, \mu) \, ds \right|^2 \\
& \leq 2C_R^b \varphi_1(t)(1 + |x|^2) + 2K(1 + |x|^\kappa + W_2^\kappa(\mu, \delta_0)).
\end{aligned}$$

Similar estimates hold for $\bar{\sigma}(x, \mu)$ and $\bar{h}(x, \mu, z)$. Let $t \rightarrow \infty$ in the above estimates, we conclude that the averaged equation (3.2) satisfies Assumptions 1–5.

We next check the extra conditions for $\bar{h}$ by calculating that (for $l = r$ or κ)

$$\begin{aligned}
\int_U |\bar{h}(x, \mu, z)|^l v(dz) & \leq \frac{2^{l-1}}{t} \int_0^t \int_U \Big(|h(s, x, \mu, z) - \bar{h}(x, \mu, z)|^l \\
& + |h(s, x, \mu, z)|^l \Big) v(dz) \, ds \\
& \leq 2^{l-1} C_R^h \varphi(t)(1 + |x|^l) + 2^{l-1} K(1 + |x|^l + W_2^l(\mu, \delta_0)), \\
\int_U |\bar{h}(x, \mu, z) - \bar{h}(y, \mu, z)|^\kappa v(dz) & \leq \frac{3^{\kappa-1}}{t} \int_0^t \int_U (|h(s, x, \mu, z) - \bar{h}(x, \mu, z)|^\kappa \\
& + |h(s, y, \mu, z) - \bar{h}(y, \mu, z)|^\kappa \\
& + |h(s, x, \mu, z) - h(s, y, \mu, z)|^\kappa) v(dz) \, ds \\
& \leq 3^{\kappa-1} \varphi(t) C_R^h (2 + |x|^\kappa + |y|^\kappa) + 3^{\kappa-1} L'_R |x - y|^\kappa, \\
\int_U |\bar{h}(x, \mu_1, z) - \bar{h}(x, \mu_2, z)|^\kappa v(dz) & \leq \frac{3^{\kappa-1}}{t} \int_0^t \int_U (|h(s, x, \mu_1, z) - \bar{h}(x, \mu_1, z)|^\kappa \\
& + |h(s, x, \mu_2, z) - \bar{h}(x, \mu_2, z)|^\kappa \\
& + |h(s, x, \mu_1, z) - h(s, x, \mu_2, z)|^\kappa) v(dz) \, ds \\
& \leq 2 \cdot 3^{\kappa-1} \varphi(t) C_R^h (1 + |x|^\kappa) + 3^{\kappa-1} L' W_2^\kappa(\mu_1, \mu_2).
\end{aligned}$$

Taking $t \rightarrow \infty$, we conclude that Assumption 7 holds. $\square$

Appendix B. Propagation of chaos

For $N \geq 1$ and $i = 1, 2, \dots, N$, let $(W^i, \tilde{N}^i, X^i(0))$ be independent copies of $(W, \tilde{N}, X(0))$. We introduce the noninteracting particle system associated with the McKean–Vlasov SDE (2.8). The state $X^i(t)$ of the particle i is given by

$$dX^i(t) = b(t, X^i(t-), \mathcal{L}_{X^i(t)}) \, dt + \sigma(t, X^i(t-), \mathcal{L}_{X^i(t)}) \, dW^i(t)$$

$$+ \int_U h(t, X^i(t-), \mathcal{L}_{X^i(t)}, z) \tilde{N}^i(dt, dz) \quad (\text{B.1})$$

for $t \in [0, T]$ with initial data $X^i(0)$. According to Theorem 1, we have $\mathcal{L}_{X^i(t)} = \mathcal{L}_{X(t)}$, for all $i = 1, 2, \dots, N$. Here, $X(t)$ is the solution of the McKean–Vlasov SDE (2.8) for $t \in [0, T]$ with initial data $X(0) = x_0$.

We also consider the associated interacting particle system

$$\begin{aligned} dX^{i,N}(t) &= b\left(t, X^{i,N}(t-), \mu_t^{X,N}\right) dt + \sigma\left(t, X^{i,N}(t-), \mu_t^{X,N}\right) dW^i(t) \\ &\quad + \int_U h\left(t, X^{i,N}(t-), \mu_t^{X,N}, z\right) \tilde{N}^i(dt, dz) \end{aligned} \quad (\text{B.2})$$

with initial data $X^{i,N}(0) = X^i(0)$, where $\mu_t^{X,N}$ is the empirical measure of N interacting particles given by $\mu_t^{X,N} = \frac{1}{N} \sum_{j=1}^N \delta_{X^{j,N}(t)}$. We proceed to establish and prove the propagation of chaos result. Furthermore, we note that compared with the existing literature on the Lévy case, particularly [6, 28, 30], the coefficient conditions in our framework are somewhat more relaxed, as discussed in Remark 1.

Proposition 3. (Propagation of chaos.) *Suppose Assumptions 1–7 hold and $r \geq 4$. Then, the interacting particle system (B.2) is well-posed and converges to the noninteracting particle system (B.1), that is,*

$$\lim_{N \rightarrow \infty} \sup_{1 \leq i \leq N} \sup_{0 \leq t \leq T} \mathbb{E} |X^i(t) - X^{i,N}(t)|^2 = 0. \quad (\text{B.3})$$

Proof. First, note that the interacting particle system $\{X^{i,N}\}_{1 \leq i \leq N}$ given in (B.2) can be regarded as a system of ordinary SDEs driven by Lévy noise, taking values in $\mathbb{R}^{d \times N}$. Thus, according to [25, Theorem 1.1], it has a unique càdlàg solution under Assumptions A1, A4, A5 such that

$$\sup_{1 \leq i \leq N} \mathbb{E} \sup_{0 \leq t \leq T} |X^{i,N}(t)|^4 \leq C,$$

for any $N \geq 1$, where $C > 0$ is independent of N .

To handle the one-sided locally Lipschitz case, for any $1 \leq i \leq N$ and $R > 0$, define the stopping time:

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

Then, by De Morgan's Law, we obtain

$$\begin{aligned} \sup_{0 \leq t \leq T} \mathbb{E} |X^i(t) - X^{i,N}(t)|^2 &\leq \sup_{0 \leq t \leq T} \mathbb{E} \left[|X^i(t) - X^{i,N}(t)|^2 \mathbb{I}_{\{\zeta_R > T\}} \right] \\ &\quad + \sup_{0 \leq t \leq T} \mathbb{E} \left[|X^i(t) - X^{i,N}(t)|^2 \mathbb{I}_{\{\zeta_R \leq T\}} \right] \\ &=: Q_1 + Q_2, \end{aligned}$$

where $\mathbb{I}_A$ is the indicator function of set A . Similarly to the proof of Theorem 2, we now estimate Q_1 and Q_2 respectively.

(1) Estimation of the term Q_1 . Note that

$$Q_1 = \sup_{0 \leq t \leq T} \mathbb{E} \left[|X^i(t) - X^{i,N}(t)|^2 \mathbb{I}_{\{\zeta_R > T\}} \right] \leq \sup_{0 \leq t \leq T} \mathbb{E} |X^i(t \wedge \zeta_R) - X^{i,N}(t \wedge \zeta_R)|^2 < \infty. \quad (\text{B.4})$$

By Itô's formula, we have

$$\begin{aligned}
 & \mathbb{E} |X^i(t \wedge \zeta_R) - X^{i,N}(t \wedge \zeta_R)|^2 \\
 &= 2\mathbb{E} \int_0^{t \wedge \zeta_R} \langle X^i(s-) - X^{i,N}(s-), b(s, X^i(s-), \mathcal{L}_{X^i(s)}) - b(s, X^{i,N}(s-), \mu_s^{X,N}) \rangle ds \\
 &\quad + \mathbb{E} \int_0^{t \wedge \zeta_R} \|\sigma(s, X^i(s-), \mathcal{L}_{X^i(s)}) - \sigma(s, X^{i,N}(s-), \mu_s^{X,N})\|^2 ds \\
 &\quad + \mathbb{E} \int_0^{t \wedge \zeta_R} \int_U |h(s, X^i(s-), \mathcal{L}_{X^i(s)}, z) - h(s, X^{i,N}(s-), \mu_s^{X,N}, z)|^2 \nu(dz) ds \\
 &=: Q_{1,R}(t) + Q_{2,R}(t) + Q_{3,R}.
 \end{aligned} \tag{B.5}$$

We now estimate $Q_{i,R}$ for $i = 1, 2, 3$ by Assumptions 1 and 2 and obtain that

$$\begin{aligned}
 Q_{1,R}(t) &\leq (2L_R + 1) \int_0^{t \wedge \zeta_R} \mathbb{E} |X^i(s-) - X^{i,N}(s-)|^2 ds + L \int_0^{t \wedge \zeta_R} \mathbb{E} W_2^2(\mathcal{L}_{X^i(s)}, \mu_s^{X,N}) ds, \\
 Q_{2,R}(t) &\leq 2L_R \int_0^{t \wedge \zeta_R} \mathbb{E} |X^i(s-) - X^{i,N}(s-)|^2 ds + 2L \int_0^{t \wedge \zeta_R} \mathbb{E} W_2^2(\mathcal{L}_{X^i(s)}, \mu_s^{X,N}) ds, \\
 Q_{3,R}(t) &\leq 2L_R \int_0^{t \wedge \zeta_R} \mathbb{E} |X^i(s-) - X^{i,N}(s-)|^2 ds + 2L \int_0^{t \wedge \zeta_R} \mathbb{E} W_2^2(\mathcal{L}_{X^i(s)}, \mu_s^{X,N}) ds,
 \end{aligned}$$

where

$$\int_0^{t \wedge \zeta_R} \mathbb{E} W_2^2(\mathcal{L}_{X^i(s)}, \mu_s^{X,N}) ds \leq 2 \int_0^{t \wedge \zeta_R} \mathbb{E} W_2^2(\mathcal{L}_{X^i(s)}, \mu_s^X) ds + 2 \int_0^{t \wedge \zeta_R} \mathbb{E} W_2^2(\mu_s^X, \mu_s^{X,N}) ds, \tag{B.6}$$

with $\mu_t^X = \frac{1}{N} \sum_{i=1}^N \delta_{X^i(t)}$ the empirical measure of N noninteracting particles. Then, by combining these estimates and applying Grönwall's inequality, we eventually have

$$Q_1 \leq \sup_{0 \leq t \leq T} \mathbb{E} |X^i(t \wedge \zeta_R) - X^{i,N}(t \wedge \zeta_R)|^2 \leq 10Le^{(6L_R+1+10L)T} \int_0^T \mathbb{E} W_2^2(\mathcal{L}_{X^i(s)}, \mu_s^X) ds.$$

(2) Estimation of the term Q_2 . Using the Cauchy–Schwarz inequality and Theorem 1, we deduce that

$$\begin{aligned}
 Q_2 &= \sup_{0 \leq t \leq T} \mathbb{E} \left[|X^i(t) - X^{i,N}(t)|^2 \mathbb{I}_{\{\zeta_R \leq T\}} \right] \leq \sup_{0 \leq t \leq T} \sqrt{\mathbb{E} (|X^i(t) - X^{i,N}(t)|^2)^2} \sqrt{\mathbb{E} (\mathbb{I}_{\{\zeta_R \leq T\}})^2} \\
 &\leq 2\sqrt{2} \sqrt{\mathbb{E} \left(\sup_{0 \leq t \leq T} |X^i(t)|^4 + \sup_{0 \leq t \leq T} |X^{i,N}(t)|^4 \right)} \sqrt{\mathbb{E} \left(\mathbb{I}_{\{\zeta_R \leq T\}} \frac{|X^i(\zeta_R)|^4 + |X^{i,N}(\zeta_R)|^4}{R^4} \right)} \\
 &\leq \frac{2\sqrt{2}}{R^2} \left(\mathbb{E} \sup_{0 \leq t \leq T} |X^i(t)|^4 + \mathbb{E} \sup_{0 \leq t \leq T} |X^{i,N}(t)|^4 \right) \leq \frac{C}{R^2}.
 \end{aligned} \tag{B.7}$$

With the estimations of Q_1 and Q_2 at hand, we conclude that

$$\sup_{1 \leq i \leq N} \sup_{0 \leq t \leq T} \mathbb{E} |X^i(t) - X^{i,N}(t)|^2 \leq 10Le^{(6L_R+1+10L)T} \int_0^T \sup_{1 \leq i \leq N} \mathbb{E} W_2^2(\mathcal{L}_{X^i(s)}, \mu_s^X) ds + \frac{C}{R^2}. \tag{B.8}$$

Note that, by [5, Theorem 5.8], we have the following estimate for the Wasserstein distance:

$$\mathbb{E}W_2^2(\mathcal{L}_{X^{i(s)}}, \mu_s^X) \leq C \begin{cases} N^{-1/2}, & \text{if } d < 4, \\ N^{-1/2} \ln(N), & \text{if } d = 4, \\ N^{-1/2}, & \text{if } d > 4. \end{cases}$$

Thus, we observe that the right-hand side of the estimate (B.8) converges to 0 as $N \rightarrow \infty$. The result follows, and the proof is complete. $\square$

Acknowledgements

The authors thank Prof. Yanjie Zhang for helpful discussions. The authors also thank the referees for their careful reading of the manuscript and invaluable comments, which were very useful in improving this paper.

Data availability statement

The datasets supporting the findings of this work are available from the corresponding author on reasonable request.

Funding statement

The research of Y. Chao was partially supported by NSFC grant 12101484, the Fundamental Research Funds for the Central Universities (xzy012025071), the Guangdong Provincial Key Laboratory of Mathematical and Neural Dynamical Systems (DSNS2025003), and NSFC grants 12271424 and 12371276. The research of J. Duan was partially supported by NSFC grant 12141107, the Guangdong Provincial Key Laboratory of Mathematical and Neural Dynamical Systems (2024B1212010004), the Cross-Disciplinary Research Team on Data Science and Intelligent Medicine (2023KCXTD054), and the Guangdong–Dongguan Joint Research Grant 2023A1515140016. The research of T. Gao was partially supported by the National Key R&D Program of China (2021ZD0201300) and NSFC grant 12401233. The research of P. Wei was partially supported by CPSF grants 2022TQ0009 and 2022M720264, the Jiangsu Provincial Scientific Research Center of Applied Mathematics (BK20233002), and the National Key R&D Program of China (2020YFA0712800).

Competing interest

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

- [1] APPLEBAUM, D. (2009). *Lévy Processes and Stochastic Calculus*, 2nd edn. Cambridge University Press, Cambridge.
- [2] BAHLALI, K., MEZERDI, M. A., and MEZERDI, B. (2020). Stability of McKean–Vlasov stochastic differential equations and applications. *Stoch. Dyn.* **20**(01), 2050007.
- [3] BOGACHEV, V. I. (2007). *Measure Theory. Vol. I, II*. Springer, Berlin.
- [4] BREZIS, H. (2011). *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer, New York.
- [5] CARMONA, R. and DELARUE, F. (2018). *Probabilistic Theory of Mean Field Games with Applications II*. Springer, Cham.
- [6] CAVALLAZZI, T. (2023). Well-posedness and propagation of chaos for Lévy-driven McKean–Vlasov SDEs under Lipschitz assumptions. Arxiv:[2301.08594](https://arxiv.org/abs/2301.08594).

- [7] DAREIOTIS, K., KUMAR, C., and SABANIS, S. (2016). On tamed Euler approximations of SDEs driven by Lévy noise with applications to delay equations. *SIAM J. Numer. Anal.* **54**(3), 1840–1872.
- [8] DELLACHERIE, C. and MEYER, P. A. (2018). *Probabilities and Potential B*. North-Holland, Amsterdam.
- [9] DING, X. and QIAO, H. (2021). Euler–Maruyama approximations for stochastic McKean–Vlasov equations with non-Lipschitz coefficients. *J. Theoret. Probab.* **34**(3), 1408–1425.
- [10] DOS REIS, G., ENGELHARDT, S. and SMITH, G. (2022). Simulation of McKean–Vlasov SDEs with super-linear growth. *IMA J. Numer. Anal.* **42**(1), 874–922.
- [11] DUAN, J. (2015). *An Introduction to Stochastic Dynamics*. Cambridge University Press, Cambridge.
- [12] FRIKHA, N., KONAKOV, V. and MENOZZI, S. (2021). Well-posedness of some non-linear stable driven SDEs. *Discrete Contin. Dyn. Syst.* **42**(2), 849–898.
- [13] FRIKHA, N. and LI, L. (2021). Well-posedness and approximation of some one-dimensional Lévy-driven non-linear SDEs. *Stochastic Process. Appl.* **132**, 76–107.
- [14] GRAHAM C. (1992). Nonlinear diffusion with jumps. *Ann. Inst. H. Poincaré Probab. Statist.* **28**(3), 393–402.
- [15] HAO, T. and LI, J. (2016). Mean-field SDEs with jumps and nonlocal integral-PDEs. *Nonlinear Diff. Equ. Appl.* **23**(2), 1–51.
- [16] HONG, W., HU, S. and LIU, W. (2022). McKean-Vlasov SDEs and SPDEs with locally monotone coefficients. [ArXiv:2205.04043](https://arxiv.org/abs/2205.04043).
- [17] JOURDAIN, B., MÉLÉARD S. and WOYCZYNSKI W. (2008). Nonlinear SDEs driven by Lévy processes and related PDEs. *ALEA Lat. Am. J. Probab. Math. Stat.* **4**, 1–29.
- [18] KAC, K. (1956). Foundations of kinetic theory. *Proceedings of The third Berkeley Symposium on Mathematical Statistics and Probability*, 171–197.
- [19] KHASMINSKII, R. (1968). On the averaging principle for stochastic differential Itô equation. *Kibernetika* **4**(3), 260–279.
- [20] KLOEDEN, P. E. and LORENZ, T. (2010). Stochastic differential equations with nonlocal sample dependence. *Stochastic Anal. Appl.* **28**(6), 937–945.
- [21] KÜHN, F. and SCHILLING, R. L. (2023). Maximal inequalities and some applications. *Probab. Surv.* **20**, 382–485.
- [22] LI, Y., MAO, X., SONG, Q., WU, F. and YIN, G. (2023). Strong convergence of Euler–Maruyama schemes for McKean–Vlasov stochastic differential equations under local Lipschitz conditions of state variables. *IMA J. Numer. Anal.* **43**(2), 1001–1035.
- [23] LIU, W., SONG, Y., ZHAI, J. and ZHANG, T. (2023). Large and moderate deviation principles for McKean–Vlasov SDEs with jumps. *Potential Anal.* **59**(3), 1141–1190.
- [24] MA, S. and KANG, Y. (2019). Periodic averaging method for impulsive stochastic differential equations with Lévy noise. *Appl. Math. Lett.* **93**, 91–97.
- [25] MAJKA, M. B. (2020). A note on existence of global solutions and invariant measures for jump sdes with locally one-sided Lipschitz drift. *Probab. Theory Stoch. Model.* **40**(1), 37–55.
- [26] MAO, X. (2007). *Stochastic Differential Equations and Applications*, Horwood, Chichester.
- [27] MCKEAN, H. P. (1966). A class of Markov processes associated with nonlinear parabolic equations. *Proc. Natl. Acad. Sci.* **56**(6), 1907–1911.
- [28] MEHRI, S., SCHEUTZOW, M., STANNAT, W. and ZANGENEH, B. Z. (2020). Propagation of chaos for stochastic spatially structured neuronal networks with delay driven by jump diffusions. *Ann. Appl. Probab.* **30**(1), 175–207.
- [29] MIKULEVICIUS, R. and PRAGARAUSKAS, H. (2012). On L_p -estimates of some singular integrals related to jump processes. *SIAM J. Math. Anal.* **44**(4), 2305–2328.
- [30] NEELIMA, S. B., KUMAR, C., DOS REIS, G. and REISINGER, C. (2020). Well-posedness and tamed Euler schemes for McKean–Vlasov equations driven by Lévy noise. [ArXiv:2010.08585](https://arxiv.org/abs/2010.08585).
- [31] NOVIKOV, A. A. (1975). On discontinuous martingales. *Theor. Probability Appl.* **20**(1), 11–26.
- [32] PAVLIOTIS, G. A., and ANDREW, S. (2008) *Multiscale Methods: Averaging and Homogenization*. Springer, New York.
- [33] PEI, B., XU, Y. and WU, J. (2020). Stochastic averaging for stochastic differential equations driven by fractional Brownian motion and standard Brownian motion. *Appl. Math. Lett.* **100**, 106006.
- [34] PROTTER, P. (2009). *Stochastic Integration and Differential Equations*, 2nd edn. Springer, New York.
- [35] SATO, K.-I. (2013). *Lévy processes and infinitely divisible distributions*, Revised edn. Cambridge University Press, Cambridge.
- [36] SHEN, G., SONG, J. and WU, J. (2022). Stochastic averaging principle for distribution dependent stochastic differential equations. *Appl. Math. Lett.* **125**, 107761.
- [37] SHEN, G., XIANG, J. and WU, J. (2022). Averaging principle for distribution dependent stochastic differential equations driven by fractional Brownian motion and standard Brownian motion. *J. Differ. Equ.* **321**, 381–414.
- [38] SHIRYAEV, A. N. (1989). *Probability*. Spring, New York.

- [39] SZNITMAN, A. S. (1991). Topics in propagation of chaos. Vol. 1464, *Lecture Notes in Mathematics*, 165–251. Springer, Berlin.
- [40] WANG, F. (2018). Distribution dependent SDEs for Landau type equations. *Stochastic Process. Appl.* **128**(2), 595–621.
- [41] XU, J., LIU, J. and MIAO, Y. (2021). Strong averaging principle for two-time-scale stochastic McKean-Vlasov equations. *Appl. Math. Opt.* **84**(1), 837–867.
- [42] XU, Y., DUAN, J. and XU, W. (2011). An averaging principle for stochastic dynamical systems with Lévy noise. *Phys. D* **240**(17), 1395–1401.