EULER’S SCHEME OF MCKEAN-VLASOV SDES WITH NON-LIPSCHITZ COEFFICIENTS

ZHEN WANG, JIE REN AND YU MIAO

Abstract. In this paper, we show the strong well-posedness of Mckean-Vlasov SDEs with non-Lipschitz coefficients. Moreover, propagation of chaos and the convergence rate for Euler’s scheme of Mckean-Vlasov SDEs are also obtained. Keywords: Mckean-Vlasov SDEs, Propagation of chaos, Euler approximation, Non-Lipschitz.

1. Introduction

Let $(\Omega, \mathcal{F}, \mathcal{P}; (\mathcal{F}_t)_{t\geq 0})$ be a complete filtration probability space, endowed with a standard $d$-dimensional Brownain motion $(W_t)_{t\geq 0}$ on the probability space. Consider the following Mckean-Vlasov stochastic differential equations:

$$dX_t = \int_{\mathbb{R}^d} b(X_t, y) \mu(dy)dt + \int_{\mathbb{R}^d} \sigma(X_t, y) \mu(dy)dW_t, \quad X_0 = \xi,$$  \hspace{1cm} (1.1)

where $\mu$ be a probability measure on $\mathbb{R}^d$, the initial value $X_0 = \xi$ is a $\nu$-distributed, and $\mathcal{F}_0$-measurable $\mathbb{R}^d$-valued random variable and the coefficients

$$b : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d, \quad \sigma : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d \otimes \mathbb{R}^d$$

are Borel measurable functions.

Mckean-Vlasov SDEs were initiated by Mckean [28] using Kac’s formalism of molecular chaos [25], and have attracted substantial research interests(cf. [33]). Röckner and Zhang [31] showed the strong well-posedness of the above SDE under a singular distribution dependent drift and a constant matrix $\sigma$. Recently, Mishura and Veretennikov [30] established weak and strong existence and uniqueness results for multi-dimensional stochastic McKean-Vlasov equations if $b$ and $\sigma$ are of linear growth in $x$ and the diffusion matrix $\sigma$ is uniformly non-degenerate. Up to now, there are many works devoted to the study of Mckean-Vlasov SDEs(see [8, 22, 27] and references therein). Moreover, the non-degenerate assumption on $\sigma$ is usually required when $\sigma$ is non-constant.

This work is partially supported by an NNSFC grant of China (No.11901154) and (No. 11971154).
We note that if
\[ \bar{b}(x, \mu) := \int_{\mathbb{R}^d} b(x, y) \mu(dy), \quad \bar{\sigma}(x, \mu) := \int_{\mathbb{R}^d} \sigma(x, y) \mu(dy), \]
then Eq. (1.1) can be rewritten as
\[ dX_t = \bar{b}(X_t, \mu_t) \, dt + \bar{\sigma}(X_t, \mu_t) \, dW_t, \quad X_0 = \xi, \tag{1.2} \]
where \( \mu_t \) is the law of \( X_t \), and
\[ \bar{b} : \mathbb{R}^d \times \mathcal{P}_p(\mathbb{R}^d) \to \mathbb{R}^d, \quad \bar{\sigma} : \mathbb{R}^d \times \mathcal{P}_p(\mathbb{R}^d) \to \mathbb{R}^d \otimes \mathbb{R}^d \]
are Borel measurable functions. Here let \( \mathcal{P}_p(\mathbb{R}^d) \) denote a space of probability measures on \( \mathbb{R}^d \) with \( p \)-th moment \( (p \geq 1) \), that is, \( \mu(\cdot, \cdot) := \int_{\mathbb{R}^d} |x|^p \mu(dx) < \infty \). What’s more, Eq. (1.2) is also called mean-field SDE or distribution dependent SDE in the literature, which naturally appears in the studies of interacting systems and mean-field games (e.g. \([9, 10, 11, 25, 28, 33, 34] \)).

Meanwhile, numerous results are developed under irregular conditions, based on Eq. (1.2). Wang \([35]\) studied strong well-posedness of distribution dependent SDEs with one-sided Lipschitz continuous drift \( \bar{b} \) and Lipschitz-continuous diffusion \( \bar{\sigma} \). For additive noise case, existence and uniqueness of strong solutions with irregular drift were established in \([6]\). Huang and Wang \([21]\) proved the existence and uniqueness of strong solutions under non-degenerate noise, under integrability conditions on distribution dependent coefficients. Ding and Qiao \([16]\) obtained the strong well-posedness under non-Lipschitz coefficients by weak existence and pathwise uniqueness. In \([29]\), the unique strong solution was constructed by effective approximation procedures, without using the famous Yamada-Watanabe theorem. So far, related topics for Eq. (1.2) have been considerably investigated, such as Harnack inequality \([21, 35]\), ergodicity \([18, 32]\), and Feynman-Kac formulae \([7, 12, 32]\) and so on. In general, the model of (1.1) has comparatively more widespread applications than (1.2), which don’t require the coefficients depend on the law of the solution, such as \( b(x - y) \).

In the present paper, we make the following non-Lipschitz assumptions without non-degeneracy diffusion:

\textbf{(H1)} For any \( x, y \in \mathbb{R}^d \), and for some \( c_0 > 0 \),
\[ |b(x, y)| + |\sigma(x, y)| \leq c_0 (1 + |x| + |y|). \]

\textbf{(H2)} The functions \( b, \sigma \) are Borel continuous measurable functions in \( \mathbb{R}^d \times \mathbb{R}^d \). For all \( x, y \in \mathbb{R}^d \), there are constants \( \lambda_1, \lambda_2 > 0 \) such that
\[ |b(x_1, y_1) - b(x_2, y_2)| \leq \lambda_1 (|x_1 - x_2| \gamma_1(|x_1 - x_2|) + |y_1 - y_2| \gamma_1(|y_1 - y_2|)), \]
\[ ||\sigma(x_1, y_1) - \sigma(x_2, y_2)|| \leq \lambda_2 (|x_1 - x_2|^2 \gamma_2(|x_1 - x_2|) + |y_1 - y_2|^2 \gamma_2(|y_1 - y_2|)), \]
where \( \gamma_i \) is a positive continuous function on \( \mathbb{R}^+ \), bounded on \([1, \infty)\) and satisfying
\[ \lim_{x \downarrow 0} \frac{\gamma_i(x)}{\log(x^{-1})} = \delta_i < \infty, \quad i = 1, 2. \tag{1.3} \]
Throughout this paper, we assume $E|\xi|^2 < \infty$ and all relevant stochastic differential equations admit unique solutions. One of the main results of this paper is stated as follows.

**Theorem 1.1.** Assume $(H_1)$ and $(H_2)$. For any $T > 0$, (1.1) has a unique strong solution $(X_t)_{t \geq 0}$ such that

$$
E \left[ \sup_{t \in [0,T]} |X_t|^2 \right] \leq C(1 + E|\xi|^2),
$$

where $C$ is a constant depending on $T$ and $c_0$.

**Remark 1.1.** In this paper, we state the existence of solutions to (1.1) using the standard method-Picard successive approximation. Meanwhile, there are other ways to show the existence of solutions to (1.1), by the well-known result for bounded measurable drift $b$ obtained in [30] and the estimate of uniformity in [36], which is established by Krylov’s estimate and Zvonkin’s technique, and whose way is obtained under non-degenerate coefficient. Thereby, for degenerate diffusion terms, the strong well-posedness of McKean-Vlasov SDEs is established with non-Lipschitz coefficients, in this article.

**Remark 1.2.** We also mention our conditions. Theorem 1.1 can not be covered by the recent work of [16] in which the distribution dependent coefficients are non-Lipschitz continuous in the spatial variables and Lipschitz continuous in law under the Wasserstein metric. However, in this paper, if $b(x, y)$ and $\sigma(x, y)$ are Lipschitz continuous in $y$, it implies the conditions of the corresponding distribution dependent coefficients in [16]. That is, the conclusion improves the result of [16, Theorem 3.1].

Suppose that the sequence of $\{\xi_i, i \in \mathbb{N}\}$ is independent identically distributed with a common distribution $\nu$ in $\mathbb{R}^d$ and $\{W^i, i \in \mathbb{N}\}$ is a sequence of independent $d$-dimensional Brownian motions. Consider the following non-interacting particle systems:

$$
dX_t^i = \int_{\mathbb{R}^d} b(X_t^i, y) \mu_{X_t^i}(dy) dt + \int_{\mathbb{R}^d} \sigma(X_t^i, y) \mu_{X_t^i}(dy) dW^i_t, \quad X_0^i = \xi_i,
$$

where $\mu_{X_t^i}$ stands for the law of $X_t^i$, and $\{X^i, i \in \mathbb{N}\}$ is a set of independent identically distributed stochastic processes with the common distribution $\mu_\cdot$.

Next, We use the discretized $N$-interacting particle to approximate the above non-interacting particle. Namely, consider $N$-interacting particle $X^{N,i}$ satisfying

$$
dX^{N,i}_t = \bar{b} \left( X^{N,i}_t, \mu_t^N \right) dt + \bar{\sigma} \left( X^{N,i}_t, \mu_t^N \right) dW^i_t, \quad X_0^{N,i} = \xi_i,
$$

where $\mu_t^N$ is the empirical measure of $\{X^{N,i}_t, i = 1, 2, \ldots, N\}$ defined by

$$
\mu_t^N := \frac{1}{N} \sum_{j=1}^N \delta_{X^{N,j}_t},
$$
where $\delta_x$ stands for a Dirac measure at point $x$. Thus, for $i = 1, 2, \ldots, N$ and for any $t \geq 0$, (1.6) is equivalent to

$$
dX_t^{N,i} = \frac{1}{N} \sum_{j=1}^{N} b \left( X_t^{N,i}, X_t^{N,j} \right) dt + \frac{1}{N} \sum_{j=1}^{N} \sigma \left( X_t^{N,i}, X_t^{N,j} \right) dW_t^i, \quad X_0^{N,i} = \xi_i. \quad (1.7)
$$

The so-called propagation of chaos is that the law of fixed particles $X_t^{N,i}$ tends to the distribution of independent particles $X_t^i$ solving (1.5) with same law when $N$ goes to $+\infty$. The study of chaos have already profound impacts on theoretical and applied contexts. In [23], there are more developments with an eye toward models arising in the physical sciences. By the mean field game theory, [13, 20] have inspired new theoretical developments and applications in engineering and economics.

The propagation of chaos for the interacting particle of McKean-Vlasov SDEs has been revealed under Lipschitz assumptions, see [33]. We mention that new particle representations for ergodic McKean-Vlasov SDEs were introduced in [1]. Propagation of chaos and convergence rate of Euler’s approximation were established by Bao and Huang [4], where the coefficients are Hölder continuous in $x$ and Lipschitz continuous in $\mu$. Especially in the case of nondegenerate diffusion terms, Zhang [39] also showed the propagation of chaos for Euler’s approximation with the linear growth coefficients. Recently, Dos Reis, Smith and Tankov [17] took some examples in many different fields with Lipschitz coefficients.

The following results state the propagation of chaos for stochastic interacting particle systems. In other words, the continuous time Euler’s scheme of stochastic $N$-interacting particle systems converges strongly to non-interacting particle systems.

\textbf{Theorem 1.2.} Suppose that (H$_1$) and (H$_2$) hold. Then for any $T > 0$,

$$
\lim_{N \to \infty} \sup_{i=1,2,\ldots,N} \mathbb{E} \left( \sup_{t \in [0,T]} \left| X_t^{N,i} - X_t^i \right|^2 \right) = 0. \quad (1.8)
$$

In particular,

$$
\sup_{i=1,2,\ldots,N} \mathbb{E} \left[ \sup_{t \leq T} \left| X_t^{N,i} - X_t^i \right|^2 \right] < \infty.
$$

Another goal of this paper is to get the corresponding overall convergence rate. To discretize (1.7) in time, for fixed $h \in (0, 1)$, the corresponding Euler’s scheme is

$$
dX_t^{h,N,i} = \frac{1}{N} \sum_{j=1}^{N} b \left( X_{t_h}^{h,N,i}, X_{t_h}^{h,N,j} \right) dt + \frac{1}{N} \sum_{j=1}^{N} \sigma \left( X_{t_h}^{h,N,i}, X_{t_h}^{h,N,j} \right) dW_t^i, \quad X_0^{h,N,i} = \xi_i. \quad (1.9)
$$
Theorem 1.3. Under the assumptions of Theorem 1.2, for $h \in (0, 1)$ sufficiently small and $\alpha \in (0, \frac{1}{2})$, there exist a constant $C$ independent of $h$ and $N$ such that

$$
\sup_{i=1,2,\ldots,N} \mathbb{E} \left( \sup_{t \in [0,T]} \left| X_t^{h,N,i} - X_t^i \right|^2 \right) < C \left( \frac{1}{N} + h^{2\alpha} \right). \tag{1.10}
$$

This paper is organized as follows. In Section 2, we show the main lemmas for later use. Section 3 yields the strong wellposedness of (1.1). In Section 4, we construct the propagation of chaos and the corresponding overall convergence rate.

Throughout the paper, $C$ with or without indices will denote different positive constants (depending on the indices), whose values may change from one place to another and not important.

2. Preliminary

First, we introduce a well-studied metric on the space of distributions known as the Wasserstein distance which allows us to consider $\mathcal{P}_p(\mathbb{R}^d)$ as a metric space.

For $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$, the $\mathbb{W}_p$-Wasserstein distance between $\mu$ and $\nu$ is defined by

$$
\mathbb{W}_p(\mu, \nu) = \inf_{\pi \in \mathcal{C}(\mu, \nu)} \left( \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p \pi(dx, dy) \right)^{\frac{1}{p}}, \tag{2.1}
$$

where $\mathcal{C}(\mu, \nu)$ is the set of all couplings of $\mu$ and $\nu$ on $\mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$, such that $\pi \in \mathcal{C}(\mu, \nu)$ if and only if $\pi(\cdot, \mathbb{R}^d) = \mu(\cdot)$ and $\pi(\mathbb{R}^d, \cdot) = \nu(\cdot)$.

The topology induced by Wasserstein metric coincides with the topology of weak convergence of measure together with the convergence of all moments of order up to $p$ in [15].

For later use, for $0 < \eta < 1/e$, we define a strictly increasing, and concave function:

$$
\rho_\eta(x) = \begin{cases} 
    x \log x^{-1}, & 0 < x \leq \eta, \\
    (\log \eta^{-1} - 1)x + \eta, & x > \eta.
\end{cases}
$$

The generalization of Gronwall-Bellman type inequality has been verified by Bihari [3]. It can be also found in [37].

Lemma 2.1. Let $g(s), q(s)$ be two strictly positive functions on $\mathbb{R}_+$ satisfying $g(0) < \eta$ and

$$
g(t) \leq g(0) + \int_0^t q(s) \rho_\eta(g(s)) ds, \quad t \geq 0.
$$

Then

$$
g(t) \leq g(0) e^{\exp\{-\int_0^t q(s) ds\}}.
$$
Lemma 2.2. Assume \((H_1)\) and \((H_2)\). Then there exists a constant \(C\) independent of \(N\), such that for all \(T > 0\), \(N \in \mathbb{N}\) and \(i = 1, 2, 3, \ldots, N\),
\[
\mathbb{E}
\left[
\sup_{t \in [0,T]} \left| X_t^{N,i} \right|^2
\right]
\leq C(1 + \mathbb{E}|\xi_i|^2).
\tag{2.2}
\]

Proof. By Itô’s formula, Hölder’s inequality and the linear growth of \(b\) and \(\sigma\), we have
\[
\mathbb{E}
\left[
\sup_{t \in [0,T]} \left| X_t^{N,i} \right|^2
\right] =
\leq \mathbb{E}|\xi_i|^2 + C \int_0^T \mathbb{E}
\left[
\sup_{r \in [0,s]} \left( \left| X_r^{N,i} \right|^2 + \frac{1}{N} \sum_{j=1}^N \left| X_r^{N,j} \right|^2 + 1 \right)
\right] ds,
\]
where Jensen’s inequality and the condition \((H_1)\) are used in the last inequality.

Summing over \(i\) on both sides, we have
\[
\sum_{i=1}^N \mathbb{E}
\left[
\sup_{t \in [0,T]} \left| X_t^{N,i} \right|^2
\right] \leq \sum_{i=1}^N \mathbb{E}|\xi_i|^2 + C \int_0^T \mathbb{E}
\left[
\sup_{r \in [0,s]} \left( \frac{1}{N} \sum_{i=1}^N \left| X_r^{N,i} \right|^2 \right)
\right] ds.
\]

By the symmetry, it holds that
\[
\sum_{i=1}^N \mathbb{E}
\left[
\sup_{t \in [0,T]} \left| X_t^{N,i} \right|^2
\right] \leq \sum_{i=1}^N \mathbb{E}|\xi_i|^2 + C \int_0^T \mathbb{E}
\left[
\sup_{r \in [0,s]} \left( 2 \left| X_r^{N,i} \right|^2 + 1 \right)
\right] ds.
\]

Using symmetry, and removing the summation symbol, then
\[
\mathbb{E}
\left[
\sup_{t \in [0,T]} \left| X_t^{N,i} \right|^2
\right] \leq \mathbb{E}|\xi_i|^2 + C \int_0^T \mathbb{E}
\left[
\sup_{r \in [0,s]} \left( 2 \left| X_r^{N,i} \right|^2 + 1 \right)
\right] ds.
\]

With the aid of Gronwall’s inequality, we can obtain
\[
\mathbb{E}
\left[
\sup_{t \in [0,T]} \left| X_t^{N,i} \right|^2
\right] \leq C(1 + \mathbb{E}|\xi_i|^2),
\]
where the constant \(C\) depends on \(T, c_0\) but independent of \(N\). \(\square\)

**Corollary 2.1.** Under conditions \((H_1)\) and \((H_2)\), for all \(T > 0\), \(N \in \mathbb{N}\) and \(i = 1, 2, 3, \ldots, N\), there exists a constant \(C\) independent of \(N\) such that
\[
\mathbb{E}
\left[
\sup_{t \in [0,T]} \left| X_t^{h,N,i} \right|^2
\right] \leq C(1 + \mathbb{E}|\xi_i|^2).
\tag{2.3}
\]
Proof. It is similar to the proof of Lemma 2.2 about main verified processes. Combining with
\[
\sup_{0 \leq r \leq s} |X_{t_r}^{h,N,i}|^2 \leq |X_{s}^{h,N,i}|^2,
\]
the conclusion is true. \(\square\)

**Lemma 2.3**. Under the hypothesis of Lemma 2.2, we have for any \(s, t \in [0, T]\),
\[
\mathbb{E} |X_{t}^{N,i} - X_{s}^{N,i}|^2 \leq C(1 + \mathbb{E} |\xi|)^2 (t - s) \tag{2.4} \]
and
\[
\mathbb{E} |X_{t}^{h,N,i} - X_{s}^{h,N,i}|^2 \leq C(1 + \mathbb{E} |\xi|)^2 (t - s), \tag{2.5} \]
where \(C\) is a constant depending on \(T\) and \(c_0\) but independent of \(N\).

**Proof.** By the Hölder’s inequality, BDG’s inequality and the assumption \((H_1)\), it holds that
\[
\mathbb{E} |X_{t}^{N,i} - X_{s}^{N,i}|^2 = \leq 2\mathbb{E} \left[ \left| \int_{s}^{t} \frac{1}{N} \sum_{j=1}^{N} b(X_{u}^{N,i}, X_{u}^{N,j}) \, du \right|^2 + \left| \int_{s}^{t} \frac{1}{N} \sum_{j=1}^{N} \sigma(X_{u}^{N,i}, X_{u}^{N,j}) \, dW_{u}^j \right|^2 \right] \leq 2 \int_{s}^{t} \mathbb{E} \left[ \frac{1}{N} \sum_{j=1}^{N} \left( |b(X_{u}^{N,i}, X_{u}^{N,j})|^2 + |\sigma(X_{u}^{N,i}, X_{u}^{N,j})|^2 \right) \right] \, du \leq 2c_0 \int_{s}^{t} \mathbb{E} \left[ \frac{1}{N} \sum_{j=1}^{N} \left( |X_{u}^{N,i}|^2 + |X_{u}^{N,j}|^2 + 1 \right) \right] \, du.
\]
Again summing and removing over \(i\) of both sides with using symmetry of above inequality, then
\[
\mathbb{E} |X_{t}^{N,i} - X_{s}^{N,i}|^2 \leq C \int_{s}^{t} \mathbb{E} \left( |X_{u}^{N,i}|^2 + 1 \right) \, du,
\]
where the constant \(C\) depends on \(c_0\).

At last, by Gronwall’s inequality, we conclude the desired estimate. In the similar way as above discussions, it holds that \(X_{t}^{h,N,i}\) satisfies Equation (2.5). \(\square\)

3. **Proof of Theorem 1.1**

Based on the conditions \((H_1)\) and \((H_2)\) for Eq. (1.1), we have
\[
|\bar{b}(x, \mu)| + |\bar{\sigma}(x, \mu)| \leq C \left( 1 + |x| + \mathbb{W}_1(\mu, \delta_0) \right).
\]
Using (1.3) and the definition of \(\rho_\eta\), there is a sufficiently small \(0 < \eta < 1/e\), for any \(x \in \mathbb{R}^\ast\), such that
\[
x_{\gamma_i}(x) \leq \rho_\eta(x), \tag{3.1} \]
and

\[ x^2 \gamma_1(x) \leq \rho_\eta(x^2). \tag{3.2} \]

By (H₁) and (H₂) and Jensen’s inequality, similarly as in the proof of [2, Lemma 1.3], we have

\[
|\bar{b}(x_1, \mu_1) - \bar{b}(x_2, \mu_2)| = \left| \int_{\mathbb{R}^d} b(x_1, y_1) \mu_1(dy_1) - \int_{\mathbb{R}^d} b(x_2, y_2) \mu_1(dy_1) \right|
+ \left| \int_{\mathbb{R}^d} b(x_1, y_2) \mu_1(dy_1) - \int_{\mathbb{R}^d} b(x_2, y_2) \mu_2(dy_2) \right|
\leq \lambda_1|x_1 - x_2| \gamma_1(|x_1 - x_2|) + \lambda_1 \int_{\mathbb{R}^d \times \mathbb{R}^d} \rho_\eta(|y_1 - y_2|) \pi(dy_1, dy_2)
\leq \lambda_1|x_1 - x_2| \gamma_1(|x_1 - x_2|) + \lambda_1 \rho_\eta(\mathbb{W}_1(\mu_1, \mu_2)).
\]

By the similar deduction to above it holds that

\[ \|\bar{\sigma}(x_1, \mu_1) - \bar{\sigma}(x_2, \mu_2)\|^2 \leq \lambda_2|x_1 - x_2|^2 \gamma_2(|x_1 - x_2|) + \lambda_2 \rho_\eta(\mathbb{W}_2(\mu_1, \mu_2)^2). \]

Especially, if \( b(x, y) \) and \( \sigma(x, y) \) are Lipschitz continuous in \( y \), the corresponding \( \bar{b}(x, \mu) \) and \( \bar{\sigma}(x, \mu) \) are Lipschitz continuous in \( \mu \). In this situation, by [16, Theorem 3.1], the weak existence of (1.2) and pathwise uniqueness imply the strong existence of (1.1). In this part, we will directly construct the strong solutions to (1.1) by the method of successive approximations under the more general conditions (H₁) and (H₂).

**Proof of Theorem 1.1.** Consider the following distribution-iterated SDEs: for any \( t \in [0, T] \) and for each \( k \geq 1 \),

\[
X^{(k)}_t = X^{(k)}_0 + \int_0^t \int_{\mathbb{R}^d} b(X^{(k)}_s, \omega^{(k-1)}_s) \mu_{k-1}(d\omega^{(k-1)}_s) ds
+ \int_0^t \int_{\mathbb{R}^d} \sigma(X^{(k)}_s, \omega^{(k-1)}_s) \mu_{k-1}(d\omega^{(k-1)}_s) dW_s, \tag{3.3}
\]

where \( X^{(k)}_0 = \xi \) and \( \mu_{k-1} \) stands for the law of \( X^{(k-1)}_s \).

By [37, Theorem 4.1], Eq. (3.3) has a unique solution \( (X^{(k)}_t)_{t \geq 0} \). Next, existence of solutions to (1.1) can be shown by verifying the Cauchy sequence.

With BDG’s and Hölder’s inequality, it follows that

\[
\mathbb{E}\left[ \sup_{t \in [0, T]} |X^{(1)}_t|^2 \right] \leq C(1 + \mathbb{E}|\xi|^2), \tag{3.4}
\]

where the detailed proof can be found in literature [4] or [16].

Indeed, this can be handled in same manner by using the triple \((X^{(k+1)}, X^{(k)}, \mu^{(k)})\) in lieu of \((X^{(1)}, \xi, \nu)\). Therefore, (3.3) still holds true for \( k + 1 \). That is, for each
By Itô’s formula, then

\[ \mathbb{E} \left[ \sup_{t \in [0,T]} |X_t^{(k)}|^2 \right] \leq C(1 + \mathbb{E}|\xi|^2). \] (3.5)  

For notation brevity, we set

\[ Z_t^{(k+1)} := X_t^{(k+1)} - X_t^{(k)}. \]

By Itô’s formula, then

\[
|Z_t^{(k+1)}|^2 = \int_0^t \int_{\mathbb{R}^d} Z_s^{(k+1)} \left[ b(X_s^{(k+1)}, \omega_s^{(k)}) \mu_k(d\omega_s^{(k)}) - b(X_s^{(k)}, \omega_s^{(k-1)}) \mu_{k-1}(d\omega_s^{(k-1)}) \right] ds \\
+ \int_0^t \int_{\mathbb{R}^d} Z_s^{(k+1)} \left[ \sigma(X_s^{(k+1)}, \omega_s^{(k)}) \mu_k(d\omega_s^{(k)}) - \sigma(X_s^{(k)}, \omega_s^{(k-1)}) \mu_{k-1}(d\omega_s^{(k-1)}) \right] dW_s \\
+ \frac{1}{2} \int_0^t \int_{\mathbb{R}^d} \| \sigma(X_s^{(k+1)}, \omega_s^{(k)}) \mu_k(d\omega_s^{(k)}) - \sigma(X_s^{(k)}, \omega_s^{(k-1)}) \mu_{k-1}(d\omega_s^{(k-1)}) \|^2 ds \\
=: I_1(t) + I_2(t) + I_3(t).
\]

By (H1), we show

\[
I_1(t) = \int_0^t \int_{\mathbb{R}^d} Z_s^{(k+1)} \left[ b(X_s^{(k+1)}, \omega_s^{(k)}) \mu_k(d\omega_s^{(k)}) - b(X_s^{(k)}, \omega_s^{(k-1)}) \mu_{k-1}(d\omega_s^{(k-1)}) \right] ds \\
+ \int_0^t \int_{\mathbb{R}^d} Z_s^{(k+1)} \left[ b(X_s^{(k+1)}, \omega_s^{(k-1)}) \mu_{k-1}(d\omega_s^{(k-1)}) - b(X_s^{(k)}, \omega_s^{(k-1)}) \mu_{k-1}(d\omega_s^{(k-1)}) \right] ds \\
\leq C \int_0^t \left[ |Z_s^{(k+1)}|^2 \gamma_1(|Z_s^{(k+1)}|) + Z_s^{(k+1)} \rho_\eta \left( \mathbb{W}_1(\mu_s^{(k)}, \mu_s^{(k-1)}) \right) \right] ds.
\]

Similarly, by (H2), we have

\[
I_3(t) \leq C \int_0^t \left[ |Z_s^{(k+1)}|^2 \gamma_2(|Z_s^{(k+1)}|) \\
+ \int_{\mathbb{R}^d \times \mathbb{R}^d} |\omega_s^{(k)} - \omega_s^{(k-1)}|^2 \gamma_2(|\omega_s^{(k)} - \omega_s^{(k-1)}|) \pi(d\omega_s^{(k)}, d\omega_s^{(k-1)}) \right] ds \\
\leq C \int_0^t \left[ \rho_\eta(|Z_s^{(k+1)}|^2) + \rho_\eta \left( \mathbb{W}_2 \left( \mu_s^{(k)}, \mu_s^{(k-1)} \right)^2 \right) \right] ds.
\]
By BDG’s inequality and Young’s inequality, we have

\[
\mathbb{E}\left(\sup_{t \in [0,T]} I_2(t)\right) \leq C \mathbb{E}\left(\int_0^T \int_{\mathbb{R}^d} \left[|Z_s^{(k+1)}|^2 \rho_\eta(|Z_s^{(k+1)}|) + |Z_s^{(k+1)}|^2 \rho_\eta\left(\mathbb{W}_2\left(\mu_s^{(k)}, \mu_s^{(k-1)}\right)^2\right)\right] \, ds\right)^{\frac{1}{2}}
\]

\[ \leq C \mathbb{E}\left(\sup_{t \in [0,T]} |Z_t^{(k+1)}|^2 \int_0^T \int_{\mathbb{R}^d} \left[\rho_\eta(|Z_s^{(k+1)}|) + \rho_\eta\left(\mathbb{W}_2\left(\mu_s^{(k)}, \mu_s^{(k-1)}\right)^2\right)\right] \, ds\right)^{\frac{1}{2}}
\]

\[ \leq C \mathbb{E}\left(\sup_{t \in [0,T]} \int_0^t \int_{\mathbb{R}^d} \left[\rho_\eta(|Z_s^{(k+1)}|) + \rho_\eta\left(\mathbb{W}_2\left(\mu_s^{(k)}, \mu_s^{(k-1)}\right)^2\right)\right] \, ds\right)
\]

\[ + \frac{1}{4} \mathbb{E}\left(\sup_{t \in [0,T]} |Z_t^{(k+1)}|^2\right)
\]

Then, with Jensen’s inequality and Young’s inequality,

\[
\mathbb{E}\left(\sup_{t \in [0,T]} |Z_t^{(k+1)}|^2\right) \leq C \int_0^T \rho_\eta\left(\mathbb{E}\left(\sup_{r \in [0,s]} |Z_r^{(k+1)}|^2\right)\right) \, dr
\]

\[ + C \int_0^T \rho_\eta^2\left(\mathbb{E}\left(\sup_{r \in [0,s]} \mathbb{W}_1(\mu_r^{(k)}, \mu_r^{(k-1)})\right) + \rho_\eta\left(\mathbb{W}_2(\mu_r^{(k)}, \mu_r^{(k-1)})^2\right)\right) \, dr
\]

\[ \leq C \int_0^T \rho_\eta\left(\mathbb{E}\left(\sup_{r \in [0,s]} |Z_r^{(k+1)}|^2\right)\right) \, dr
\]

\[ + C \int_0^T \rho_\eta^2\left(\mathbb{E}\left(\sup_{r \in [0,s]} |Z_r^{(k)}|^2\right)\right)^{\frac{1}{2}} + \rho_\eta\left(\mathbb{E}\left(\sup_{r \in [0,s]} |Z_r^{(k)}|^2\right)\right) \, dr,
\]

where the last inequality is established with properties of \(\mathbb{W}_p\)-Wasserstein distance, as following:

\[
\sup_{r \in [0,s]} \mathbb{W}_1(\mu_r^{(k)}, \mu_r^{(k-1)}) \leq \mathbb{E}\left(\sup_{r \in [0,s]} \left|X_r^{(k)} - X_r^{(k-1)}\right|^2\right)^{\frac{1}{2}}, \quad (3.6)
\]

and

\[
\sup_{r \in [0,s]} \mathbb{W}_2(\mu_r^{(k)}, \mu_r^{(k-1)})^2 \leq \mathbb{E}\left(\sup_{r \in [0,s]} \left|X_r^{(k)} - X_r^{(k-1)}\right|^2\right).
\]

Furthermore, with the aid of (3.4), there exists a constant \(a\) such that

\[
\mathbb{E}\left[\sup_{r \in [0,s]} |Z_r^{(1)}|^2\right] \leq a.
\]

Therefore, by iteration and Lemma 2.1, we have

\[
\mathbb{E}\left(\sup_{t \in [0,T]} |Z_t^{(k+1)}|^2\right) \leq \left[\rho_\eta(\sqrt{a}) + \rho_\eta(a)\right] \exp(-kCT), \quad (3.8)
\]
Then, there is an \((F_t)_{t \in [0,T]}\)-adapted continuous stochastic process \((X_t)_{t \in [0,T]}\) and 
\(\mu_t\) is the law of \(X_t\) such that
\[
\lim_{k \to \infty} \mathbb{E}\left[ \sup_{t \in [0,T]} |X_t^{(k+1)} - X_t|^2 \right] = 0,
\]
and
\[
\lim_{k \to \infty} \sup_{t \in [0,T]} \mathbb{W}_2 \left( \mu_t^{(k)}, \mu_r \right)^2 \leq \lim_{k \to \infty} \mathbb{E}\left[ \sup_{t \in [0,T]} |X_t^{(k)} - X_t|^2 \right] = 0.
\]

By \((H_1)\), we find
\[
\int_0^t \int_{\mathbb{R}^d} \left[ b(X_s^{(k+1)}, \omega_s^{(k)}) \mu_k(d\omega^{(k)}) - b(X_s, \omega_s) \mu(d\omega) \right] ds \\
\leq C \int_0^T \left[ \int_{\mathbb{R}^d} \left| \omega_s^{(k)} - \omega_s \right| \gamma_1 \left( \left| \omega_s^{(k)} - \omega_s \right| \right) \pi(d\omega_s^{(k)}, d\omega_s) \right] ds \\
+ C \int_0^T \left[ |X_s^{(k+1)} - X_s| \gamma_1 \left( |X_s^{(k+1)} - X_s| \right) \right] ds \\
\leq C \int_0^T \left[ \rho_\eta \left( \mathbb{W}_1(\mu_s^{(k)}, \mu_s) \right) + \rho_\eta (|X_s^{(k+1)} - X_s|) \right] ds.
\]

By dominated convergence theorem and (3.9) and (3.10), we obtain
\[
\lim_{k \to \infty} \mathbb{E}\left[ \int_0^T \int_{\mathbb{R}^d} \left[ b(X_s^{(k+1)}, \omega_s^{(k)}) \mu_k(d\omega^{(k)}) - b(X_s, \omega_s) \mu(d\omega) \right] ds \right] = 0.
\]

Similarly,
\[
\lim_{k \to \infty} \mathbb{E}\left[ \int_0^T \int_{\mathbb{R}^d} \left[ \sigma(X_s^{(k+1)}, \omega_s^{(k)}) \mu_k(d\omega^{(k)}) - \sigma(X_s, \omega_s) \mu(d\omega) \right] dW_s \right] = 0.
\]

By taking \(k \to \infty\) in the equation (3.3), we derive (by extracting a suitable subsequence) SDE (1.1).

On the other hand, we show the uniqueness of (1.1). We assume that \((X_1^1)_{t \geq 0}\) and \((X_2^2)_{t \geq 0}\) are solutions to (1.1) with the same initial value \(\xi\). By the similar way, one has
\[
\lim_{k \to \infty} \mathbb{E}\left[ \sup_{t \in [0,T]} |X_t^1 - X_t^2|^2 \right] \leq C \int_0^T \rho_\eta \left( \mathbb{E}\left[ \sup_{r \in [0,s]} |X_r^1 - X_r^2|^2 \right] \right) ds.
\]

Once again invoking Lemma 2.1 yields the uniqueness.
Finally, we intend to show (1.4). Applying Hölder’s inequality, Jensen’s inequality, BDG’s inequality and the condition \((H_1)\), we show
\[
E\left[\sup_{t \in [0,T]} |X_t|^2\right] \leq 3 \left\{ E|\xi|^2 + T \int_0^T E\left[ \sup_{r \in [0,s]} |b(X_r, y)\mu(dy)|^2 \right] ds \right. \\
+ \left. \int_0^T E\left[ \sup_{r \in [0,s]} |\sigma(X_r, y)\mu(dy)|^2 \right] ds \right\} \\
\leq C \left\{ 1 + E|\xi|^2 + T \int_0^T E\left[ \sup_{r \in [0,s]} |X_r|^2 \right] ds + \int_0^T \left[ \sup_{r \in [0,s]} \mathbb{W}_1(\mu_r, \delta_0)^2 ds \right] \right\} \\
\leq C(1 + E|\xi|^2) + C \int_0^T E\left( \sup_{r \in [0,s]} |X_r|^2 \right) ds.
\]
Together with Gronwall’s inequality, implies
\[
E\left[\sup_{t \in [0,T]} |X_t|\right] \leq C(1 + E|\xi|^2).
\]
So the proof is finished. \(\square\)

4. Proof of Theorem 1.2

Proof of Theorem 1.2. Under the conditions \((H_1)\) and \((H_2)\), combining with Eq.(1.7) and (1.5), we find that
\[
X_t^{N,i} - X_t^i = \int_0^t \left[ \frac{1}{N} \sum_{j=1}^N b(X_s^{N,i}, X_s^{N,j}) - b(X_s^i, y)\mu_X(dy) \right] ds \\
+ \int_0^t \left[ \frac{1}{N} \sum_{j=1}^N \sigma(X_s^{N,i}, X_s^{N,j}) - \sigma(X_s^i, y)\mu_X(dy) \right] dW_s^i \\
= \int_0^t \frac{1}{N} \sum_{j=1}^N \left[ \{ b(X_s^{N,i}, X_s^{N,j}) - b(X_s^{N,i}, X_s^{N,j}) + \tilde{b}(X_s^{N,i}, X_s^{N,j}) \} \right] ds \\
+ \int_0^t \frac{1}{N} \sum_{j=1}^N \left[ \sigma(X_s^{N,i}, X_s^{N,j}) - \sigma(X_s^{N,i}, X_s^{N,j}) + \tilde{\sigma}(X_s^{N,i}, X_s^{N,j}) \right] dW_s^i.
\]
For simplicity of notation, \(\tilde{b}(x, x')\) and \(\tilde{\sigma}(x, x')\) can be redefined:
\[
\tilde{b}(x, x') := b(x, x') - \int_{\mathbb{R}^d} b(x, y)\mu_x(dy) \tag{4.1}
\]
and
\[
\tilde{\sigma}(x, x') := \sigma(x, x') - \int_{\mathbb{R}^d} \sigma(x, y)\mu_x(dy). \tag{4.2}
\]
By Itô’s formula and Jensen’s inequality, we get

\[
\mathbb{E} \left[ \sup_{t \in [0,T]} \left| X_{t,n}^{i} - X_{t}^{i} \right|^2 \right] \\
\leq 2\lambda_1 \int_{0}^{T} \mathbb{E} \left( \sup_{r \in [0,s]} \left| X_{r,n}^{i} - X_{r}^{i} \right|^2 \gamma_1 \left( \left| X_{r,n}^{i} - X_{r}^{i} \right| \right) \right) ds \\
+ \frac{1}{N} \sum_{j=1}^{N} \mathbb{E} \left( \sup_{r \in [0,s]} \left| X_{r,n}^{j} - X_{r}^{j} \right|^2 \gamma_1 \left( \left| X_{r,n}^{j} - X_{r}^{j} \right| \right) \right) ds \\
+ \lambda_2 \int_{0}^{T} \mathbb{E} \left( \sup_{r \in [0,s]} \left| X_{r,n}^{i} - X_{r}^{i} \right|^2 \gamma_2 \left( \left| X_{r,n}^{i} - X_{r}^{i} \right| \right) \right) ds \\
+ \frac{1}{N} \sum_{j=1}^{N} \mathbb{E} \left( \sup_{r \in [0,s]} \left| X_{r,n}^{j} - X_{r}^{j} \right|^2 \gamma_2 \left( \left| X_{r,n}^{j} - X_{r}^{j} \right| \right) \right) ds + \int_{0}^{T} \mathbb{E} \left[ \sup_{r \in [0,s]} \left| X_{r,n}^{i} - X_{r}^{i} \right|^2 \right] ds \\
+ \int_{0}^{T} \mathbb{E} \left[ \frac{1}{N} \sum_{j=1}^{N} \tilde{b} \left( X_{r,n}^{i}, X_{r,n}^{j} \right)^2 \right] ds + \left( \frac{1}{N} \sum_{j=1}^{N} \tilde{\sigma} \left( X_{r,n}^{i}, X_{r,n}^{j} \right)^2 \right) ds.
\]

Using symmetry technique as in the proof of Lemma 2.2, we have

\[
\mathbb{E} \left[ \sup_{t \in [0,T]} \left| X_{t,n}^{i} - X_{t}^{i} \right|^2 \right] \\
\leq 4\lambda_1 \int_{0}^{T} \mathbb{E} \left( \sup_{r \in [0,s]} \left| X_{r,n}^{i} - X_{r}^{i} \right|^2 \gamma_1 \left( \left| X_{r,n}^{i} - X_{r}^{i} \right| \right) \right) ds \\
+ 2\lambda_2 \int_{0}^{T} \mathbb{E} \left( \sup_{r \in [0,s]} \left| X_{r,n}^{i} - X_{r}^{i} \right|^2 \gamma_2 \left( \left| X_{r,n}^{i} - X_{r}^{i} \right| \right) \right) ds \tag{4.3} \]

\[
+ \int_{0}^{T} \mathbb{E} \left( \sup_{r \in [0,s]} \left| X_{r,n}^{i} - X_{r}^{i} \right|^2 \right) ds + \int_{0}^{T} \mathbb{E} \left[ \sup_{r \in [0,s]} \left| \frac{1}{N} \sum_{j=1}^{N} \tilde{b} \left( X_{r,n}^{i}, X_{r,n}^{j} \right) \right|^2 \right] ds \\
+ \int_{0}^{T} \mathbb{E} \left[ \frac{1}{N} \sum_{j=1}^{N} \tilde{\sigma} \left( X_{r,n}^{i}, X_{r,n}^{j} \right) \right]^2 ds.
\]

And due to the centering of \( \tilde{b}(x,y) \) and \( \tilde{\sigma}(x,y) \) with to their second variable, then if \( j \neq k \) such that

\[
\mathbb{E} \left[ \tilde{b}(X_{s}^{i}, X_{s}^{j}) \tilde{b}(X_{s}^{i}, X_{s}^{k}) \right] = 0 \tag{4.4}\]

and

\[
\mathbb{E} \left[ \tilde{\sigma}(X_{s}^{i}, X_{s}^{j}) \tilde{\sigma}(X_{s}^{i}, X_{s}^{k}) \right] = 0. \tag{4.5}
\]
In light of Lemma 2.2, we can obtain

\[ E\left[ \sup_{r \in [0,s]} \left| \frac{1}{N} \sum_{j=1}^{N} \tilde{b}(X_r^i, X_r^j) \right|^2 \right] = \frac{1}{N^2} E\left[ \sup_{r \in [0,s]} \left| \sum_{j,k} \tilde{b}(X_r^i, X_r^j) \tilde{b}(X_r^i, X_r^k) \right| \right] \]

\[ \leq \frac{2}{N^2} \sum_{j=1}^{N} E\left[ \sup_{r \in [0,s]} \left| \tilde{b}(X_r^i, X_r^j) \right|^2 \right] \]

\[ \leq \frac{C}{N^2} \sum_{j=1}^{N} \left( 1 + E\left[ \sup_{r \in [0,s]} |X_r^i|^2 \right] + E\left[ \sup_{r \in [0,s]} |X_r^j|^2 \right] \right) \leq \frac{C}{N} \]

and

\[ E\left[ \sup_{r \in [0,s]} \left| \frac{1}{N} \sum_{j=1}^{N} \tilde{\sigma}(X_r^i, X_r^j) \right|^2 \right] \leq \frac{C}{N}. \]

Furthermore, based on (3.1)-(3.2) and (4.4)-(4.7), the above inequality (4.3) is

\[ E\left[ \sup_{t \in [0,T]} \left| X_{N,i}^t - X_{i}^t \right|^2 \right] \leq (2\lambda_1 + \lambda_2) \int_0^T \rho_\eta \left( E\left[ \sup_{r \in [0,s]} |X_r^{N,i} - X_r^i|^2 \right] \right) ds + \frac{C}{N}. \]

Applying Gronwall-Belmman’s inequality (Lemma 2.1), we find

\[ E\left[ \sup_{t \in [0,T]} \left| X_{N,i}^t - X_{i}^t \right|^2 \right] \leq \frac{C}{N}. \]

Therefore,

\[ \sup_{i=1,2,\ldots,N} N E\left[ \sup_{t \leq T} \left| X_{N,i}^t - X_{i}^t \right|^2 \right] < \infty. \]

And by (4.8), for \( N \) is sufficiently large, the conclusion (1.10) is established. \( \square \)

### 5. Proof of Theorem 1.3

**Lemma 5.1.** Under the assumptions of Theorem 1.3, for any \( T > 0 \) and some \( \alpha \in (0, \frac{1}{2}) \), there is a constant \( C(T, \mathbb{E}[|\xi|^2], \lambda_1, \lambda_2, c_0) > 0 \), we have

\[ \sup_{i=1,2,\ldots,N} E\left( \sup_{t \in [0,T]} |X_{N,i}^t - X_{h,N,i}^t|^2 \right) < C h^{2\alpha}. \]

for \( h \in (0, 1) \) sufficiently small.

**Proof.** Set

\[ Z_{h,i}^t := X_{h,N,i}^t - X_{N,i}^t \]
then

\[ Z_t^{h,i} = \int_0^t \frac{1}{N} \sum_{j=1}^N \left[ b \left( X_s^{N,i}, X_s^{N,j} \right) - b \left( X_{s_h}^{h,N,i}, X_{s_h}^{h,N,j} \right) \right] ds \]

\[ + \int_0^t \frac{1}{N} \sum_{j=1}^N \left[ \sigma \left( X_s^{N,i}, X_s^{N,j} \right) - \sigma \left( X_{s_h}^{h,N,i}, X_{s_h}^{h,N,j} \right) \right] dW_s^i. \]

Using Itô’s formula and taking the expectation, we find

\[
\mathbb{E} \left[ \sup_{t \in [0,T]} \left| Z_t^{h,i} \right|^2 \right] =: Z(T)
\]

\[
\leq \int_0^T \mathbb{E} \left[ \sup_{r \in [0,s]} \left| Z_r^{h,i} \right|^2 \left( 2\lambda_1 \gamma_1 \left( \left| Z_r^{h,i} \right| \right) + \lambda_2 \gamma_2 \left( \left| Z_r^{h,i} \right| \right) \right) \right] ds
\]

\[
+ \int_0^T \mathbb{E} \left\{ \sup_{r \in [0,s]} \frac{1}{N} \sum_{j=1}^N \left[ 2\lambda_1 \left| Z_r^{h,i} \right| \left| Z_r^{h,j} \right| \gamma_1 \left( \left| Z_r^{h,j} \right| \right) + \lambda_2 \left| Z_r^{h,j} \right|^2 \gamma_2 \left( \left| Z_r^{h,j} \right| \right) \right] \right\} ds
\]

\[
+ 2 \int_0^T \mathbb{E} \left[ \sup_{r \in [0,s]} \frac{1}{N} \sum_{j=1}^N \left| Z_r^{h,i} \right| \left| b \left( X_r^{h,N,i}, X_r^{h,N,j} \right) - b \left( X_{r_h}^{h,N,i}, X_{r_h}^{h,N,j} \right) \right| \right] ds
\]

\[
+ \int_0^T \mathbb{E} \left[ \sup_{r \in [0,s]} \frac{1}{N} \sum_{j=1}^N \left| \sigma \left( X_r^{h,N,i}, X_r^{h,N,j} \right) - \sigma \left( X_{r_h}^{h,N,i}, X_{r_h}^{h,N,j} \right) \right|^2 \right] ds.
\]

With the aid of summing and symmetry in a similar way as Lemma 2.2, we obtain

\[
Z(T) \leq \int_0^T \mathbb{E} \left[ \sup_{r \in [0,s]} \left| Z_r^{h,i} \right|^2 \left( 4\lambda_1 \gamma_1 \left( \left| Z_r^{h,i} \right| \right) + 2\lambda_2 \gamma_2 \left( \left| Z_r^{h,i} \right| \right) \right) \right] ds + \int_0^T \mathbb{E} \left[ \sup_{r \in [0,s]} \left| Z_r^{h,i} \right|^2 \right] ds
\]

\[
+ \int_0^T \mathbb{E} \left[ \sup_{r \in [0,s]} \frac{1}{N} \sum_{j=1}^N \left| b \left( X_r^{h,N,i}, X_r^{h,N,j} \right) - b \left( X_{r_h}^{h,N,i}, X_{r_h}^{h,N,j} \right) \right|^2 \right] ds
\]

\[
+ \int_0^T \mathbb{E} \left[ \sup_{r \in [0,s]} \frac{1}{N} \sum_{j=1}^N \left| \sigma \left( X_r^{h,N,i}, X_r^{h,N,j} \right) - \sigma \left( X_{r_h}^{h,N,i}, X_{r_h}^{h,N,j} \right) \right|^2 \right] ds
\]

By lemma 2.3 and the condition \( (H_2) \) and Jensen’s inequality, together with (3.1) and (3.2), there is a constant \( \tilde{C} = \tilde{C}(\lambda_1, \lambda_2) \) independent of \( h \), we have

\[
Z(T) \leq C \int_0^T \rho_\eta \left( Z(s) \right) ds + \left[ \rho_\eta (\tilde{C} h^2) + \rho_\eta^2 (\tilde{C} h) \right].
\]

And for enough small \( x \), it is easy to see that \( \rho_\eta^{1/2} (x) \leq x^\alpha (\alpha < \frac{1}{2}). \) Using Lemma 2.1, we obtain for some constant \( C(T, \mathbb{E}[\xi^2], \lambda_1, \lambda_2, c_0) > 0 \),

\[
\sup_{i=1,2,\ldots,N} \mathbb{E} \left( \sup_{t \in [0,T]} \left| X_t^{N,i} - X_t^{h,N,i} \right|^2 \right) < C h^{2\alpha},
\]
where $h$ is sufficiently small.

**Proof of Theorem 1.3.** By Theorem 1.2 and Lemma 5.1, the proof of Theorem 1.3 can be complete.

**Acknowledgement:** The authors would like to thank ZiMo Hao for quite useful suggestions about this paper.

**References**

1. Airachid, H.; Bossy, M.; Ricci, C.; Szpruch, L.: New particle representations for ergodic McKean-Vlasov SDEs. ESAIM Proc. S., 65 (2019), 68-83.
2. Alain-Sol, S.: Topics in propagation of chaos. École d’Été de Probabilités de Saint-Flour XIX-1989, 165-251, Lecture Notes in Math. 1464, Springer, Berlin, 1991.
3. Bihari, I.: A generalization of a lemma of Bellman and its application to uniqueness problem of differential equations, Acta. Math. Acad. Sci. Hungar., 7 (1956) 71-94.
4. Bao, J.; Huang, X.: Approximations of McKean-Vlasov Stochastic Differential Equations with Irregular Coefficients. Journal of Theoretical Probability, doi.org/10.1007/s10959-021-01082-9 (2021).
5. Bao, J.; Huang, X.; Yuan, C.: Convergence rate of Euler-Maruyama scheme for SDEs with Hölder-Dini continuous drifts. J. Theoret. Probab., 32 (2019), no. 2, 848-871.
6. Bauer, M.; Meyer-Brandis, T.; Prokske, F.: Strong solutions of mean-field stochastic differential equations with irregular drift. Electron. J. Probab., 23 (2018), 1-35.
7. Buckdahn, R.; Li, J.; Peng, S.; Rainer, C.: Mean-field stochastic differential equations and associated PDEs. Ann. Probab., 45 (2017), 824-878.
8. Chaudru de Raynal, P.E.: Strong well-posedness of McKean-Vlasov stochastic differential equation with Hölder drift. Stoch. Process Appl., 130 (2020), 79-107.
9. Carmona, R.; Delarue, F.: Probabilistic analysis of mean-field games. (English summary) SIAM J. Control Optim., 51 (2013), no. 4, 2705-2734.
10. Carmona, R.; Delarue, F.: Probabilistic theory of mean field games with applications. II. Mean field games with common noise and master equations. Probability Theory and Stochastic Modelling, 84. Springer, Cham, 2018. xxiv+697 pp.
11. Carrillo, J.A.; Gvalani, R.S.; Pavliotis, G.A.; Schlichting, A.: Long-time behaviour and phase transitions for the McKean-Vlasov equation on the torus. Arch. Ration. Mech. Anal., 235 (2020), 635-690.
12. Crisan, D.; McMurray, E.: Smoothing properties of McKean-Vlasov SDEs. Probab. Theor. Relat. Fields, 171 (2018), 97-148.
13. Chassagneux J.-F.; Szpruch L.; Tse A.: Weak quantitative propagation of chaos via differential calculus on the space of measures, arXiv preprint arXiv:1901.02550.
14. Daniel L.: Hierarchies, entropy, and quantitative propagation of chaos for mean field diffusions, arXiv:2105.02983v1.
15. Dobrushin, R. L.: Prescribing a system of random variables by conditional distributions, Th. Probab. and its Applic., 3 (1970), 469.
16. Ding, X.; Qiao, H.: Euler-Maruyama approximations for stochastic McKean-Vlasov equations with non-Lipschitz coefficients. J. Theoret. Probab., 34 (2021), no. 3, 1408-1425.
17. Dos Reis, G.; Smith, G.; Tankov, P.: Importance sampling for McKean-Vlasov SDEs, arXiv:1803.09320.
18. Eberle, A.; Guillin, A.; Zimmer, R.: Quantitative Harris-type theorems for diffusions and McKean-Vlasov processes. Trans. Am. Math. Soc., 371 (2019), 7135-7173.
[19] Govindan, T.E.; Ahmed, N.U.: On Yosida approximations of McKean-Vlasov type stochastic evolution equations. *Stoch. Anal. Appl.*, 33 (2015), no. 3, 383-398.

[20] Huang M.; Malhamé R.P.; Caines P.E.: Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle, *Communications in Information and Systems*, 6 (2006), no. 3, 221-252.

[21] Huang, X.; Wang, F.-Y.: Distribution dependent SDEs with singular coefficients. *Stoch. Process. Appl.*, 129(2019), 4747-4770.

[22] Huang, X.; Wang, F.-Y.: McKean-Vlasov SDEs with drifts discontinuous under Wasserstein distance. [arXiv:2002.06877v2].

[23] Jabin, P.-E.; Wang, Z.: Mean field limit for stochastic particle systems. Active particles. Vol. 1. Advances in theory, models, and applications, 379-402, *Model. Simul. Sci. Eng. Technol.*, Birkhäuser/Springer, Cham, 2017.

[24] Jabin, P.-E.; Wang, Z.: Quantitative estimates of propagation of chaos for stochastic systems with $W^{1,\infty}$ kernels. *Invent. Math.*, 214 (2018), no. 1, 523-591.

[25] Kac M.: Foundations of kinetic theory. *Proceedings of The third Berkeley symposium on mathematical statistics and probability*, 1954-1955, vol. III, pp. 171-197.

[26] Kac, M.: Foundations of kinetic theory. Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954-1955, vol. III, pp. 171-197. University of California Press, Berkeley and Los Angeles, Calif., 1956.

[27] Li, J.; Min, H.: Weak solutions of mean-field stochastic differential equations and application to zero-sum stochastic differential games. *SIAM J. Control Optim.*, 54 (2016), 1826-1858.

[28] McKean, H.P.: A class Of Markov processes associated with nonlinear parabolic equations. *Proc. Nat. Acad. Sci. U.S.A.*, 56 (1966), 1907-1911.

[29] Mezerdi, M.A.; Bahlali, K.; Khelfallah, N.; Mezerdi, B.: Approximation and generic properties of McKean-Vlasov stochastic equations with continuous coefficients, arXiv: 1909.13699.

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

[31] Röckner, M.; Zhang, X.: Well-posedness of distribution dependent SDEs with singular drifts. *Bernoulli*, 27 (2021), no. 2, 1131-1158.

[32] Ren, P.; Röckner, M.; Wang, F.-Y.: Linearization of Nonlinear Fokker-Planck Equations and Applications, [arXiv:1904.06795v3].

[33] Sznitman, A.S.: Topics in propagation of chaos. In École d’Eté de Pro. de Saint-Flour XIX-1989, *Lecture Notes in Math.*, pages 165-251.

[34] Vlasov, A.A.: The vibrational properties of an electron gas. *Sov. Phys., Usp.* 10 (1968), 721.

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

[36] Xia, P.; Xie, L.; Zhang, X.; Zhao, G.: $Lq(Lp)$-theory of stochastic differential equations. *Stochastic Process. Appl.*, 130 (2020), no. 8, 5188-5211.

[37] Zhang X.: Homeomorphic flows for multi dimensional SDEs with non-Lipschitz coefficients. *Stochastic Process. Appl.*, 115 (2005) 435-448.

[38] Zhang X.: Euler-Maruyama approximations for SDEs with non-Lipschitz coefficients and applications. *J. Math. Anal. Appl.*, 316 (2006), no. 2, 447-458.

[39] Zhang, X.: A discretized version of Krylov’s estimate and its applications. *Electron. J. Probab.*, 24 (2019), no. 131, 17 pp.

ZHEN WANG: COLLEGE OF MATHEMATICS AND INFORMATION STATISTICS, HENAN NORMAL UNIVERSITY, XINXIANG, HENAN, 453000, P.R.CHINA

Email address: wangzhen881025@163.com
Jie Ren: College of Mathematics and Information Statistics, Henan University of Economics and Law, Zhengzhou, Henan, 450000, P.R.China

Email address: 20130006@huel.edu.cn

Yu Miao: College of Mathematics and Information Statistics, Henan Normal University, Xinxiang, Henan, 453000, P.R.China

Email address: yumiao728@gmail.com