Path Dependent McKean-Vlasov SDEs with Hölder Continuous Diffusion*

Xing Huang [a)], **Xucheng Wang** [a)]

a)Center for Applied Mathematics, Tianjin University, Tianjin 300072, China

xinghuang@tju.edu.cn, wxc3018233022@163.com

December 31, 2023

Abstract

In this paper, the well-posedness for one-dimensional path dependent McKean-Vlasov SDEs with $\alpha(\alpha \geq \frac{1}{2})$-Hölder continuous diffusion is investigated. Moreover, the associated quantitative propagation of chaos in the sense of Wasserstein distance, total variation distance as well as relative entropy is studied.

AMS subject Classification: 60H10, 60H05, 65C35.
Keywords: Path dependent McKean-Vlasov SDE, Yamada-Watanabe approximation, Hölder continuous diffusion, Propagation of chaos.

1 Introduction

Distribution dependent SDEs can be used to characterize the nonlinear Fokker-Planck-Kolmogorov equations. They are also called McKean-Vlasov SDEs due to the pioneer work in [18]. On the other hand, McKean-Vlasov SDE can be viewed as the limit equation of a single particle in the mean field interacting particle system, which is related to the propagation of chaos [24], so it is also called mean field SDE. Recently, there are plentiful results on McKean-Vlasov SDEs. With respect to the well-posedness, one can refer to [1, 4, 5, 12, 13, 19, 23, 25] and references therein, see also [14] for the path dependent case with singular drifts. In [4, 5, 12, 13, 23], the diffusion is assumed to be uniformly elliptic. For the propagation of chaos, see [2, 3, 6, 8, 9, 11, 15, 17, 24, 27]. One can also refer to [10, 22, 26] for the long time behavior of mean field interacting particle system and McKean-Vlasov SDEs.

*Supported in part by NNSFC (12271398).

The aim of this paper is to investigate the well-posedness and propagation of chaos of one-dimensional path dependent McKean-Vlasov SDEs with $\alpha(\alpha \geq \frac{1}{2})$-Hölder continuous diffusion. With respect to the well-posedness, we do not assume that the diffusion is elliptic.

Throughout the paper, fix a constant $r > 0$. Let $\mathscr{C} = C([-r, 0]; \mathbb{R})$, the continuous map from $[-r, 0]$ to $\mathbb{R}$. For any $f \in C([-r, \infty); \mathbb{R})$, $t \geq 0$, define $f_t \in \mathscr{C}$ as $f_t(s) = f(t+s), s \in [-r, 0]$, which is called the segment process. Let $\mathscr{P}(\mathbb{R})$ be the set of all probability measures in $\mathbb{R}$ equipped with the weak topology. Define

$$\mathscr{P}_1(\mathbb{R}) = \big\{\mu \in \mathscr{P}(\mathbb{R}) : \mu(|\cdot|) < \infty\big\}.$$

It is well known that $\mathscr{P}_1(\mathbb{R})$ is a Polish space under the L^1-Wasserstein distance

$$\mathbb{W}_1(\mu, \nu) := \inf_{\pi \in \mathbf{C}(\mu,\nu)} \left(\int_{\mathbb{R}\times\mathbb{R}} |x - y| \pi(\mathrm{d}x, \mathrm{d}y) \right), \quad \mu, \nu \in \mathscr{P}_1(\mathbb{R}),$$

where $\mathbf{C}(\mu, \nu)$ is the set of all couplings of μ and ν. By the adjoint formula, it holds

$$\mathbb{W}_1(\mu, \nu) = \sup_{\|f\|_{\mathrm{Lip}} \leq 1} |\mu(f) - \nu(f)|,$$

where

$$\|f\|_{\mathrm{Lip}} := \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|}.$$

Recall that for two probability measures μ, ν on some measurable space $(E, \mathscr{E})$, the entropy and total variation distance are defined as follows:

$$\mathrm{Ent}(\nu|\mu) := \begin{cases} \int_E (\log \frac{\mathrm{d}\nu}{\mathrm{d}\mu}) \mathrm{d}\nu, & \text{if } \nu \text{ is absolutely continuous with respect to } \mu, \\ \infty, & \text{otherwise,} \end{cases}$$

and

$$\|\mu - \nu\|_{var} := \sup_{\{f : |f(x)| \leq 1, x \in E\}} |\mu(f) - \nu(f)|.$$

By Pinsker's inequality (see [20]),

$$\|\mu - \nu\|_{var}^2 \leq 2\mathrm{Ent}(\nu|\mu), \quad \mu, \nu \in \mathscr{P}(E), \tag{1.1}$$

here $\mathscr{P}(E)$ denotes all probability measures on $(E, \mathscr{E})$.

Let $W = (W(t))_{t \geq 0}$ be a one-dimensional standard Brownian motion on a complete filtration probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$. Consider the following one-dimensional path dependent McKean-Vlasov SDE:

$$\mathrm{d}X(t) = b(t, X(t), \mathscr{L}_{X(t)})\mathrm{d}t + B(t, X_t, \mathscr{L}_{X(t)})\mathrm{d}t + \sigma(t, X(t))\mathrm{d}W(t), \tag{1.2}$$

where $b : [0, \infty) \times \mathbb{R} \times \mathscr{P}(\mathbb{R}) \to \mathbb{R}$, $B : [0, \infty) \times \mathscr{C} \times \mathscr{P}(\mathbb{R}) \to \mathbb{R}$, $\sigma : [0, \infty) \times \mathbb{R} \to \mathbb{R}$ are measurable and X_0 is an $\mathscr{F}_0$-measurable $\mathscr{C}$-valued random variable. Define the uniform norm $\|\xi\|_\infty := \sup_{s\in[-r,0]} |\xi(s)|, \xi \in \mathscr{C}$.

Definition 1.1. A continuous process $(X(t))_{t\geq -r}$ on $\mathbb{R}$ is called a strong solution of (1.2) in $\mathscr{P}_1(\mathbb{R})$, if for any $t \geq 0$, $X(t)$ is $\mathscr{F}_t$-measurable, $\mathbb{E}\|X_t\|_\infty < \infty$, and $\mathbb{P}$-a.s.

$$X(t) = X(0) + \int_0^t (b(s, X(s), \mathscr{L}_{X(s)}) + B(s, X_s, \mathscr{L}_{X(s)}))\mathrm{d}s + \int_0^t \sigma(s, X(s))\mathrm{d}W(s), \quad t \geq 0.$$

Throughout the paper, we fix $T > 0$, and consider the solution on $[0, T]$.

The remainder of this paper is organized as follows: In Section 2, the strong well-posedness of path dependent classical SDEs is addressed by Yamada-Watanabe's approximation; In Section 3, the well-posedness and quantitative propagation of chaos for path dependent McKean-Vlasov SDEs are investigated.

2 Multi-dimensional Path Dependent Classical SDEs with Hölder Continuous Diffusion

For any $n \geq 1$, let $\mathscr{C}^n = C([-r, 0]; \mathbb{R}^n)$, the continuous map from $[-r, 0]$ to $\mathbb{R}^n$, equipped with the uniform norm. In this section, we study the well-posedness of multi-dimensional path dependent classical SDEs with Hölder continuous diffusion. The main tool is the Yamada-Watanabe approximation, see [16]. Before going on, we state a time nonhomogeneous version of [21, Theorem 2.3]. More precisely, consider path dependent SDE on $\mathbb{R}^m$:

$$\mathrm{d}X(t) = f(t, X_t)\mathrm{d}t + g(t, X_t)\mathrm{d}\bar{W}(t), \quad X_0 = \xi \in \mathscr{C}^m, \tag{2.1}$$

here $f : [0, T] \times \mathscr{C}^m \to \mathbb{R}^m$, $g : [0, T] \times \mathscr{C}^m \to \mathbb{R}^m \otimes \mathbb{R}^d$ and $\bar{W}(t)$ is a d-dimensional standard Brownian motion on a complete filtration probability space $(\bar{\Omega}, \bar{\mathscr{F}}, \{\bar{\mathscr{F}}_t\}_{t\in[0,T]}, \bar{\mathbb{P}})$.

Proposition 2.1. *Assume that f and g are bounded on bounded sets. Suppose that for any $t \in [0, T]$, $f(t, \cdot)$ and $g(t, \cdot)$ are continuous in $(\mathscr{C}^m, \|\cdot\|_\infty)$ and there exists a constant $K \in \mathbb{R}$ such that*

$$2\langle f(t, \xi) - f(t, \eta), \xi(0) - \eta(0)\rangle + \|g(t, \xi) - g(t, \eta)\|_{HS}^2 \leq K\|\xi - \eta\|_\infty^2, \quad t \in [0, T], \xi, \eta \in \mathscr{C}^m.$$

Then (2.1) *has a unique non-explosive strong solution on* $[0, T]$.

Since the proof of Proposition 2.1 is completely the same with that of [21, Theorem 2.3], we omit it here.

The following stochastic Grönwall lemma comes from [21, Lemma 5.2], which is crucial in the proof of the main result of this section.

Lemma 2.2. *Let Z be a continuous adapted non-negative stochastic process which satisfies the inequality*

$$Z(t) \leq K \int_0^t \sup_{u\in[0,s]} Z(u)\mathrm{d}s + M(t) + C, \quad t \geq 0,$$

where $C \geq 0$, $K > 0$ and M is a continuous local martingale with $M(0) = 0$. Then for any $p \in (0,1)$, there exist finite constants $c_1(p), c_2(p)$ (not depending on K, C, T and M) such that

$$\mathbb{E}(\sup_{s\in[0,t]} |Z(s)|^p) \leq C^p c_1(p) \mathrm{e}^{c_2(p)Kt}, \quad t \geq 0.$$

For $x \in \mathbb{R}^d$, we denote x_i as the i-th component of x, that is $x = (x_1, x_2, \cdots, x_d)$. Consider

$$\mathrm{d}X(t) = F(t, X(t))\mathrm{d}t + H(t, X_t)\mathrm{d}t + G(t, X(t))\mathrm{d}\bar{W}(t), \quad X_0 = \xi \in \mathscr{C}^d, \tag{2.2}$$

where $F = (F_1, F_2, \cdots, F_d) : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d, H = (H_1, H_2, \cdots, H_d) : [0,T] \times \mathscr{C}^d \to \mathbb{R}^d$, $G : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$. We make the following assumptions.

(A1) F is locally bounded in $[0,T] \times \mathbb{R}^d$. For any $t \in [0,T]$, $F(t, \cdot)$ is continuous, and there exists a constant $K_1 \geq 0$ such that

$$(F_i(t,x) - F_i(t,y))\mathrm{sgn}(x_i - y_i) \leq K_1|x - y|, \quad t \in [0,T], x, y \in \mathbb{R}^d, 1 \leq i \leq d,$$

where $\mathrm{sgn}(\cdot)$ means the sign function.

(A2) There exist m real valued functions $(G_i)_{1\leq i\leq d}$ on $[0,T] \times \mathbb{R}$ such that

$$G(t,x) = \mathrm{diag}(G_1(t,x_1), G_2(t,x_2), \cdots, G_1(t,x_d)), \;\; x = (x_1, x_2, \cdots, x_d) \in \mathbb{R}^d, t \in [0,T].$$

Moreover, there exist constants $(\alpha_i)_{1\leq i\leq d} \subset [\frac{1}{2}, 1]$ and $K_2 \geq 0$ such that

$$|G_i(t,z) - G_i(t,\bar{z})| \leq K_2|z - \bar{z}|^{\alpha_i}, \quad |G_i(t,0)| \leq K_2, \quad z, \bar{z} \in \mathbb{R}, t \in [0,T], 1 \leq i \leq d.$$

(A3) There exists a constant $K_3 \geq 0$ such that

$$|H(t,\xi) - H(t,\eta)| \leq K_3\|\xi - \eta\|_\infty, \quad |H(t,0)| \leq K_3, \quad \xi, \eta \in \mathscr{C}^d, t \in [0,T].$$

Now, we provide the main result in this section.

Theorem 2.3. *Assume* (**A1**)-(**A3**). *Then for any $\xi \in \mathscr{C}^d$, (2.2) has a unique strong solution $(X^\xi(t))_{t\in[-r,T]}$ with initial value $X_0^\xi = \xi$ and for any $q \in (0,2]$, there exists a constant $C(T,q) > 0$ such that*

$$\mathbb{E} \sup_{t\in[-r,T]} |X^\xi(t)|^q \leq C(T,q)(1 + \|\xi\|_\infty^q). \tag{2.3}$$

Moreover, for any $p \in (0,1)$,

$$\mathbb{E} \sup_{t\in[-r,T]} |X^\xi(t) - X^\eta(t)|^p \leq C(p,T)\|\xi - \eta\|_\infty^p \tag{2.4}$$

for some constant $C(p,T) > 0$.

Proof. **Step 1. Existence of the strong solution.**

For $\varepsilon \in (0,1)$, note $\int_{\varepsilon/\mathrm{e}^{\frac{1}{\varepsilon}}}^{\varepsilon} \frac{\varepsilon}{x}\mathrm{d}x = 1$, so there exists a continuous function $\psi_\varepsilon : [0,\infty) \to [0,\infty)$ with support $[\varepsilon/\mathrm{e}^{\frac{1}{\varepsilon}}, \varepsilon]$ such that

$$0 \le \psi_\varepsilon(x) \le \frac{2\varepsilon}{x}, \quad x \in [\varepsilon/\mathrm{e}^{\frac{1}{\varepsilon}}, \varepsilon], \quad \int_{\varepsilon/\mathrm{e}^{\frac{1}{\varepsilon}}}^{\varepsilon} \psi_\varepsilon(u)\mathrm{d}u = 1.$$

Let

$$V_\varepsilon(x) := \int_0^{|x|}\int_0^y \psi_\varepsilon(z)\mathrm{d}z\mathrm{d}y, \quad x \in \mathbb{R}.$$

Then $V_\varepsilon \in C^2$,

$$|x| - \varepsilon \le V_\varepsilon(x) \le |x|, \quad 0 \le \mathrm{sgn}(x)V_\varepsilon'(x) \le 1, \quad x \in \mathbb{R}, \tag{2.5}$$

and

$$0 \le V_\varepsilon''(x) \le \frac{2\varepsilon}{|x|}\mathbf{1}_{[\varepsilon/\mathrm{e}^{\frac{1}{\varepsilon}},\varepsilon]}(|x|), \qquad x \in \mathbb{R}. \tag{2.6}$$

Let $\rho \in C_0^\infty(\mathbb{R})$ with $\rho \ge 0$ and $\int_{\mathbb{R}} \rho(x)\mathrm{d}x = 1$ be supported in $[-1,1]$. For any $n \ge 1$, define $\rho^n(x) = n\rho(nx), x \in \mathbb{R}$ and let

$$G_i^n(t,\cdot) = G_i(t,\cdot) * \rho^n, \quad 1 \le i \le d, t \in [0,T]$$

and

$$G^n(t,x) = \mathrm{diag}(G_1^n(t,x_1), G_2^n(t,x_2), \cdots, G_d^n(t,x_d)), \quad x = (x_1,x_2,\cdots,x_d) \in \mathbb{R}^d, t \in [0,T].$$

(A2) implies that

$$\begin{aligned}
&\lim_{n\to\infty} \sup_{t\in[0,T],x\in\mathbb{R}} |G_i^n(t,x) - G_i(t,x)| \\
&\le \lim_{n\to\infty} \sup_{t\in[0,T],x\in\mathbb{R}} \int_{\mathbb{R}} \rho^n(y)|G_i(t,x-y) - G_i(t,x)|\mathrm{d}y \\
&\le K_2 \lim_{n\to\infty} \int_{\mathbb{R}} |y|^{\alpha_i}\rho^n(y)\mathrm{d}y \\
&= K_2 \lim_{n\to\infty} \int_{\mathbb{R}} \frac{1}{n^{\alpha_i}}|x|^{\alpha_i}\rho(x)\mathrm{d}x = 0, \quad 1 \le i \le d
\end{aligned} \tag{2.7}$$

and for any $n \ge 1$, there exists a constant $L_n \ge 0$ such that

$$|G^n(t,x) - G^n(t,y)| \le L_n|x-y|, \quad |G^n(t,0)| \le 2\sqrt{d}K_2, \quad t \in [0,T], x,y \in \mathbb{R}^d, \tag{2.8}$$

where for the second inequality, it is sufficient to note that

$$|G_i^n(t,0)| \le \int_{\mathbb{R}} |G_i(t,-y)|n\rho(ny)\mathrm{d}y \le K_2\int_{\mathbb{R}} |y|^{\alpha_i}n\rho(ny)\mathrm{d}y + K_2 \le 2K_2, \quad 1 \le i \le d.$$

According to Proposition 2.1, it follows from **(A1)**, **(A3)** and (2.8) that for any $n \geq 1$, the SDE

$$\mathrm{d}X^n(t) = F(t, X^n(t))\mathrm{d}t + H(t, X_t^n)\mathrm{d}t + G^n(t, X^n(t))\mathrm{d}\bar{W}(t), \quad X_0^n = \xi \tag{2.9}$$

has a unique non-explosive strong solution. Moreover, in view of the second inequality of (2.8) and

$$|G_i^n(t, z) - G_i^n(t, \bar{z})| \leq K_2|z - \bar{z}|^{\alpha_i}, \quad t \in [0, T], z, \bar{z} \in \mathbb{R}, 1 \leq i \leq d, n \geq 1, \tag{2.10}$$

this together with **(A1)** and **(A3)** implies that there exists a constant $C(T) > 0$ such that

$$\sup_{n\geq 1} \mathbb{E} \sup_{t\in[-r,T]} |X^n(t)|^2 \leq C(T)(1 + \|\xi\|_\infty^2). \tag{2.11}$$

In fact, by **(A1)**, **(A3)**, (2.10) and the second inequality in (2.8), it holds

$$\begin{aligned} &\langle F(t, x), x\rangle \leq C(T)(1 + |x|^2), \quad |H(t, \xi)| \leq C(T)(1 + \|\xi\|_\infty), \\ &|G^n(t, x)| \leq C(T)(1 + |x|), \quad x \in \mathbb{R}^d, t \in [0, T], \xi \in \mathscr{C}^d, n \geq 1. \end{aligned} \tag{2.12}$$

For each integer $N \geq 1$, define $\tau_N^n = \inf\big\{t \in [0, T] : |X^n(t)| \geq N\big\}$ and $\inf \emptyset = \infty$ by convention. By Itô's formula, we derive from (2.12) that

$$\begin{aligned} |X^n(t \wedge \tau_N^n)|^2 \leq \xi(0)^2 + C + C \int_0^{t\wedge\tau_N^n} \sup_{u\in[-r,s]} |X^n(u)|^2 \mathrm{d}s \\ + 2\int_0^{t\wedge\tau_N^n} \langle X^n(s), G^n(s, X^n(s))\mathrm{d}\bar{W}(s)\rangle \end{aligned}$$

for some constant $C > 0$. Applying BDG's inequality, (2.12) and Grönwall's ineuqality, it is standard to derive (2.11). For any $1 \leq i \leq d$, let $X^{n,i}$ be the i-th component of X^n. For any $m, n \geq 1$, it follows from Itô's formula that

$$\begin{aligned} V_\varepsilon(X^{m,i}(t) - X^{n,i}(t)) = &\int_0^t V_\varepsilon'(X^{m,i}(s) - X^{n,i}(s))(F_i(s, X^m(s)) - F_i(s, X^n(s)))\mathrm{d}s \\ &+ \int_0^t V_\varepsilon'(X^{m,i}(s) - X^{n,i}(s))(H_i(s, X_s^m) - H_i(s, X_s^n))\mathrm{d}s \\ &+ \frac{1}{2}\int_0^t V_\varepsilon''(X^{m,i}(s) - X^{n,i}(s))(G_i^m(s, X^{m,i}(s)) - G_i^n(s, X^{n,i}(s)))^2\mathrm{d}s \\ &+ \int_0^t V_\varepsilon'(X^{m,i}(s) - X^{n,i}(s))(G_i^m(s, X^{m,i}(s)) - G_i^n(s, X^{n,i}(s)))\mathrm{d}\bar{W}^i(s). \end{aligned}$$

This combined with (2.5) implies that for any $1 \leq i \leq d$,

$$\sum_{i=1}^d |X^{m,i}(t) - X^{n,i}(t)| - d\varepsilon$$

$$
\begin{aligned}
&\le \int_0^t \sum_{i=1}^d V_\varepsilon'(X^{m,i}(s)-X^{n,i}(s))(F_i(s,X^m(s))-F_i(s,X^n(s)))\mathrm{d}s\\
&+\int_0^t \sum_{i=1}^d V_\varepsilon'(X^{m,i}(s)-X^{n,i}(s))(H_i(s,X^m_s)-H_i(s,X^n_s))\mathrm{d}s\\
&+\frac{1}{2}\int_0^t \sum_{i=1}^d V_\varepsilon''(X^{m,i}(s)-X^{n,i}(s))(G_i^m(s,X^{m,i}(s))-G_i^n(s,X^{n,i}(s)))^2\mathrm{d}s\\
&+\int_0^t \sum_{i=1}^d V_\varepsilon'(X^{m,i}(s)-X^{n,i}(s))(G_i^m(s,X^{m,i}(s))-G_i^n(s,X^{n,i}(s)))\mathrm{d}\bar{W}^i(s)\\
&=: J_1^{m,n}+J_2^{m,n}+J_3^{m,n}+J_4^{m,n},\quad t\in[0,T].
\end{aligned}
\tag{2.13}
$$

By (2.5) and **(A1)**, we conclude that

$$
\begin{aligned}
J_1^{m,n} &= \int_0^t \sum_{i=1}^d \Big[V_\varepsilon'(X^{m,i}(s)-X^{n,i}(s))\mathrm{sgn}(X^{m,i}(s)-X^{n,i}(s))\\
&\qquad\qquad \times (F_i(s,X^m(s))-F_i(s,X^n(s)))\mathrm{sgn}(X^{m,i}(s)-X^{n,i}(s))\Big]\mathrm{d}s\\
&\le K_1 d\int_0^t |X^m(s)-X^n(s)|\mathrm{d}s,\quad t\in[0,T].
\end{aligned}
$$

(2.5) and **(A3)** yield that

$$
J_2^{m,n} \le K_3 d\int_0^t \|X^m_s-X^n_s\|_\infty \mathrm{d}s,\quad t\in[0,T].
$$

Note that (2.6) and (2.10) derive

$$
\begin{aligned}
J_3^{m,n} &\le K_2^2\int_0^t \sum_{i=1}^d \varepsilon^{2\alpha_i}\mathrm{d}s+\int_0^t\sum_{i=1}^d \mathrm{e}^{\frac{1}{\varepsilon}}(G_i^m(s,X^{n,i}(s))-G_i^n(s,X^{n,i}(s)))^2\mathrm{d}s\\
&\le tK_2^2\sum_{i=1}^d \varepsilon^{2\alpha_i}+t\mathrm{e}^{\frac{1}{\varepsilon}}\sum_{i=1}^d \sup_{s\in[0,t],z\in\mathbb{R}}|G_i^m(s,z)-G_i^n(s,z)|^2\\
&\le tK_2^2\sum_{i=1}^d \varepsilon^{2\alpha_i}+2t\mathrm{e}^{\frac{1}{\varepsilon}}\sum_{i=1}^d \sup_{s\in[0,t],z\in\mathbb{R}}|G_i^m(s,z)-G_i(s,z)|^2\\
&\qquad + 2t\mathrm{e}^{\frac{1}{\varepsilon}}\sum_{i=1}^d \sup_{s\in[0,t],z\in\mathbb{R}}|G_i^n(s,z)-G_i(s,z)|^2\\
&=: H^{m,n,\varepsilon}(t),\quad t\in[0,T].
\end{aligned}
$$

So, (2.13) and Lemma 2.2 imply that for any $p\in(0,1)$, there exist constants $c_1(p),c_2(p)>0$ such that

$$
\mathbb{E}\sup_{t\in[0,T]}|X^m(t)-X^n(t)|^p \le c_1(p)\mathrm{e}^{c_2(p)(K_1+K_3)dT}(H^{m,n,\varepsilon}(T)+d\varepsilon)^p.
$$

Thanks to (2.7), we derive

$$\limsup_{m,n\to\infty}|H^{m,n,\varepsilon}(T)|\le TK_2^2\sum_{i=1}^d\varepsilon^{2\alpha_i}.$$

So, for $p\in(0,1)$, it holds

$$\limsup_{m,n\to\infty}\mathbb{E}\sup_{t\in[0,T]}|X^m(t)-X^n(t)|^p\le c_1(p)\mathrm{e}^{c_2(p)(K_1+K_3)dT}(TK_2^2\sum_{i=1}^d\varepsilon^{2\alpha_i}+d\varepsilon)^p.$$

Letting $\varepsilon\to0$, we conclude that for any $p\in(0,1)$,

$$\limsup_{m,n\to\infty}\mathbb{E}\sup_{t\in[0,T]}|X^m(t)-X^n(t)|^p=0.$$

So, there exists a continuous stochastic process $\{\bar X(t)\}_{t\in[-r,T]}$ satisfying $\bar X_0=\xi$ and

$$\limsup_{n\to\infty}\mathbb{E}\sup_{t\in[0,T]}|X^n(t)-\bar X(t)|^p=0. \tag{2.14}$$

This yields that there exists a subsequence $\{n_k\}_{k\ge1}$ such that $\mathbb{P}$-a.s.

$$\lim_{k\to\infty}\sup_{t\in[0,T]}|X^{n_k}(t)-\bar X(t)|=0,\quad \sup_{k\ge1}\sup_{t\in[-r,T]}(|X^{n_k}(t)|+|\bar X(t)|)<\infty. \tag{2.15}$$

Moreover, (2.11) and Fatou's Lemma imply

$$\mathbb{E}\sup_{t\in[-r,T]}|\bar X(t)|^2\le C(T)(1+\|\xi\|_\infty^2).$$

So, by the local boundedness of F,H, the second inequality in (2.15), we conclude that $\mathbb{P}$-a.s.

$$\sup_{k\ge1}\sup_{s\in[0,T]}(|F(s,X^{n_k}(s))|+|H(s,X_s^{n_k})|+|F(s,\bar X(s))|+|H(s,\bar X_s)|)<\infty.$$

which together with the continuity of $F(s,\cdot),H(s,\cdot)$, the first equality in (2.15) and the dominated convergence theorem implies that $\mathbb{P}$-a.s.

$$\lim_{k\to\infty}\sup_{t\in[0,T]}\left|\int_0^t(F(s,X^{n_k}(s))+H(s,X_s^{n_k}))\mathrm{d}s-\int_0^t(F(s,\bar X(s))+H(s,\bar X_s))\mathrm{d}s\right|=0.$$

Moreover, by Markov's inequality, BDG's inequality, (2.7), (2.10) and (2.14), for any $\varepsilon>0$ and $p\in(0,1)$, we have

$$\begin{aligned}&\limsup_{k\to\infty}\mathbb{P}\left(\sup_{t\in[0,T]}\left|\int_0^t[G^{n_k}(s,X^{n_k}(s))-G(s,\bar X(s))]\mathrm{d}\bar W(s)\right|\ge\varepsilon\right)\\&\le\limsup_{k\to\infty}\frac{1}{\varepsilon^p}\mathbb{E}\sup_{t\in[0,T]}\left|\int_0^t[G^{n_k}(s,X^{n_k}(s))-G(s,\bar X(s))]\mathrm{d}\bar W(s)\right|^p\end{aligned}$$

$$
\begin{aligned}
&\le c(p)\limsup_{k\to\infty}\frac{1}{\varepsilon^p}\mathbb{E}\left(\int_0^T\sum_{i=1}^d[G_i^{n_k}(s,X^{n_k,i}(s))-G_i(s,\bar{X}^i(s))]^2\mathrm{d}s\right)^{\frac{p}{2}}\\
&\le c(p)T^{\frac{p}{2}}\frac{1}{\varepsilon^p}\limsup_{k\to\infty}\left(\sum_{i=1}^d\sup_{s\in[0,T],x\in\mathbb{R}}|G_i^{n_k}(s,x)-G_i(s,x)|^2\right)^{\frac{p}{2}}\\
&+c(p)T^{\frac{p}{2}}K_2^p\frac{1}{\varepsilon^p}\limsup_{k\to\infty}\mathbb{E}\left(\sup_{s\in[0,T]}\sum_{i=1}^d|X^{n_k,i}(s)-\bar{X}^i(s)|^{2\alpha_i}\right)^{\frac{p}{2}}=0.
\end{aligned}
$$

Therefore, replacing n by n_k in (2.9) and letting $k\to\infty$, it holds $\mathbb{P}$-a.s.

$$
\bar{X}(t)=\int_0^t F(s,\bar{X}(s))\mathrm{d}s+\int_0^t H(s,\bar{X}_s)\mathrm{d}s+\int_0^t G(s,\bar{X}(s))\mathrm{d}\bar{W}(s),\quad t\in[0,T].
$$

This means that $\{\bar{X}(t)\}_{t\in[-r,T]}$ is a strong solution to (2.2).

Step 2. Uniqueness of the strong solution.

Let $X^\xi(t)$ be the solution to (2.2) with initial value $\xi\in\mathscr{C}^d$. By the same argument to derive (2.11), we obtain

$$
\mathbb{E}\sup_{t\in[-r,T]}|X^\xi(t)|^2\le C(T)(1+\|\xi\|_\infty^2).
$$

So, Jensen's inequality implies (2.3). For any $1\le i\le d$, let $X^{\xi,i}$ be the i-th component of X^ξ. Applying Ito's formula, for any $1\le i\le d$, we have

$$
\begin{aligned}
V_\varepsilon(X^{\xi,i}(t)-X^{\eta,i}(t))&=V_\varepsilon(\xi^i(0)-\eta^i(0))\\
&+\int_0^t V_\varepsilon'(X^{\xi,i}(s)-X^{\eta,i}(s))\big\{F_i(s,X^\xi(s))-F_i(s,X^\eta(s))\big\}\mathrm{d}s\\
&+\int_0^t V_\varepsilon'(X^{\xi,i}(s)-X^{\eta,i}(s))\big\{H_i(s,X_s^\xi)-H_i(s,X_s^\eta)\big\}\mathrm{d}s\\
&+\frac{1}{2}\int_0^t V_\varepsilon''(X^{\xi,i}(s)-X^{\eta,i}(s))\big\{G_i(s,X^{\xi,i}(s))-G_i(s,X^{\eta,i}(s))\big\}^2\mathrm{d}s\\
&+\int_0^t V_\varepsilon'(X^{\xi,i}(s)-X^{\eta,i}(s))\big\{G_i(s,X^{\xi,i}(s))-G_i(s,X^{\eta,i}(s))\big\}\mathrm{d}\bar{W}^i(s).
\end{aligned}
$$

By **(A1)-(A3)**, (2.5) and (2.6), it holds

$$
\begin{aligned}
|X^\xi(t)-X^\eta(t)|&\le d\varepsilon+K_2^2T\sum_{i=1}^d\varepsilon^{2\alpha_i}+C(T)\|\xi-\eta\|_\infty\\
&+C\int_0^t\sup_{u\in[0,s]}|X^\xi(u)-X^\eta(u)|\mathrm{d}s+M_t,\quad t\in[0,T]
\end{aligned}
$$

for a martingale M_t and some constants $C, C(T) > 0$. Then for any $p \in (0,1)$, applying Lemma 2.2, we get

$$\mathbb{E}\sup_{t\in[0,T]} |X^\xi(t) - X^\eta(t)|^p \le c_1(p)\mathrm{e}^{c_2(p)T}\left(d\varepsilon + K_2^2 T\sum_{i=1}^d \varepsilon^{2\alpha_i} + C(T)\|\xi-\eta\|_\infty\right)^p.$$

Letting $\varepsilon \to 0$, we derive (2.4), which yields the uniqueness of the strong solution of (2.2). □

3 Path Dependent McKean-Vlasov SDEs with Hölder Continuous Diffusion

Throughout this section, we make the following assumption.

(H) Assume that the following conditions hold.

(**Hb**) b is locally bounded in $[0,T]\times\mathbb{R}\times(\mathscr{P}_1(\mathbb{R}),\mathbb{W}_1)$. For any $t\in[0,T]$, $b(t,\cdot,\cdot)$ is continuous in $\mathbb{R}\times(\mathscr{P}_1(\mathbb{R}),\mathbb{W}_1)$, and there exists a constant $K_b\ge 0$ such that for $x,y\in\mathbb{R}$ and $\mu,\nu\in\mathscr{P}_1(\mathbb{R})$,

$$[b(t,x,\mu)-b(t,y,\nu)]\mathrm{sgn}(x-y)\le K_b(\mathbb{W}_1(\mu,\nu)+|x-y|),\quad t\in[0,T].$$

(**H**σ) There exist constants $K_\sigma\ge 0$ and $\alpha\in[\frac{1}{2},1]$ such that

$$|\sigma(t,x)-\sigma(t,y)|\le K_\sigma|x-y|^\alpha,\quad |\sigma(t,0)|\le K_\sigma,\quad x,y\in\mathbb{R}, t\in[0,T].$$

(**HB**) There exists a constant $K_B\ge 0$ and a probability measure m on $[-r,0]$ such that for any $\xi,\eta\in\mathscr{C},\mu,\nu\in\mathscr{P}_1(\mathbb{R}), t\in[0,T]$,

$$|B(t,\xi,\mu)-B(t,\eta,\nu)|\le K_B\big\{\|\xi-\eta\|_{L^1(m)}+\mathbb{W}_1(\mu,\nu)\big\},\quad |B(t,0,\delta_0)|\le K_B,$$

here δ_0 is the Dirac measure at the point 0.

3.1 Well-posedness

Theorem 3.1. *Assume* **(H)**. *Then for any $X_0\in L^1(\Omega\to(\mathscr{C},\|\cdot\|_\infty);\mathscr{F}_0,\mathbb{P})$, (1.2) has a unique strong solution $(X(t))_{t\in[-r,T]}$ with initial value X_0 and there exists a constant $C(T)>0$ such that*

$$\mathbb{E}\sup_{t\in[0,T]}\|X_t\|_\infty\le C(T)(1+\mathbb{E}\|X_0\|_\infty). \tag{3.1}$$

Moreover, for two solutions $X(t)$ and $\tilde{X}(t)$,

$$\begin{aligned}&\sup_{t\in[0,T]}\mathbb{E}|X(t)-\tilde{X}(t)|\\ &\quad\le C(T)\mathbb{E}\left\{|X(0)-\tilde{X}(0)|+K_B\int_{-r}^0 m([-r,u])|X(u)-\tilde{X}(u)|\mathrm{d}u\right\}.\end{aligned} \tag{3.2}$$

Proof. For $\mu \in C([0,T];\mathscr{P}_1(\mathbb{R}))$, $x \in \mathbb{R}$ and $\xi \in \mathscr{C}$, let $b^\mu(t,x) = b(t,x,\mu_t)$, $B^\mu(t,\xi) = B(t,\xi,\mu_t)$. Consider

$$\mathrm{d}X^\mu(t) = b^\mu(t,X^\mu(t))\mathrm{d}t + B^\mu(t,X_t^\mu)\mathrm{d}t + \sigma(t,X^\mu(t))\mathrm{d}W(t), \quad t \in [0,T]. \tag{3.3}$$

By **(H)** and Theorem 2.3, (3.3) is strongly well-posed and let $\Phi_t(\mu) = \mathscr{L}_{X^\mu(t)}, t \in [0,T]$, where $(X^\mu(t))_{t\in[-r,T]}$ solves (3.3) with $X_0^\mu \in L^1(\Omega \to (\mathscr{C}, \|\cdot\|_\infty);\mathscr{F}_0,\mathbb{P})$. In view of **(Hb)**,

$$b(t,x,\mu)\mathrm{sgn}(x) \le C_0(T)(1+|x|+\mu(|\cdot|)), \quad t \in [0,T] \tag{3.4}$$

holds for some constant $C_0(T) > 0$. So, by the similar argument to derive (2.11), we get

$$\mathbb{E}(\sup_{s\in[-r,t]}|X^\mu(s)|^2|\mathscr{F}_0) \le C(T)^2\left(1+\|X_0^\mu\|_\infty^2+\int_0^t \mu_s(|\cdot|)^2\mathrm{d}s\right), \quad t \in [0,T], \tag{3.5}$$

which yields

$$\mathbb{E}(\sup_{s\in[-r,t]}|X^\mu(s)|) \le C(T)\left(1+\mathbb{E}\|X_0^\mu\|_\infty+\left(\int_0^t \mu_s(|\cdot|)^2\mathrm{d}s\right)^{\frac{1}{2}}\right), \quad t \in [0,T] \tag{3.6}$$

for some constant $C(T) \ge 0$. By Itô's formula, it follows that

$$\begin{aligned} V_\varepsilon(X^\mu(t)-X^\nu(t)) &= V_\varepsilon(X^\mu(0)-X^\nu(0)) \\ &\quad + \int_0^t V_\varepsilon'(X^\mu(s)-X^\nu(s))\big\{b(s,X^\mu(s),\mu_s)-b(s,X^\nu(s),\nu_s)\big\}\mathrm{d}s \\ &\quad + \int_0^t V_\varepsilon'(X^\mu(s)-X^\nu(s))\big\{B(s,X_s^\mu,\mu_s)-B(s,X_s^\nu,\nu_s)\big\}\mathrm{d}s \\ &\quad + \frac{1}{2}\int_0^t V_\varepsilon''(X^\mu(s)-X^\nu(s))\big\{\sigma(s,X^\mu(s))-\sigma(s,X^\nu(s))\big\}^2\mathrm{d}s \\ &\quad + \int_0^t V_\varepsilon'(X^\mu(s)-X^\nu(s))\big\{\sigma(s,X^\mu(s))-\sigma(s,X^\nu(s))\big\}\mathrm{d}W(s) \\ &=: I_{1,\varepsilon} + I_{2,\varepsilon}(t) + I_{3,\varepsilon}(t) + I_{4,\varepsilon}(t) + I_{5,\varepsilon}(t). \end{aligned} \tag{3.7}$$

Using (2.5), we get

$$I_{1,\varepsilon} \le |X^\mu(0)-X^\nu(0)|.$$

Moreover, it follows from (2.5) and **(Hb)** that

$$I_{2,\varepsilon}(t) \le K_b\int_0^t \big\{|X^\mu(s)-X^\nu(s)|+\mathbb{W}_1(\mu_s,\nu_s)\big\}\mathrm{d}s, \quad t \in [0,T].$$

By (2.5), **(HB)**, Fubini's theorem and $t+\theta \le t, \theta \in [-r,0]$, we arrive at

$$I_{3,\varepsilon}(t) \le K_B\int_0^t \big\{\|X_s^\mu - X_s^\nu\|_{L^1(m)}+\mathbb{W}_1(\mu_s,\nu_s)\big\}\mathrm{d}s$$

$$
\begin{aligned}
&= K_B \int_{-r}^{0} \left(\int_{\theta}^{t+\theta} |X^\mu(u) - X^\nu(u)| \mathrm{d}u \right) m(\mathrm{d}\theta) + K_B \int_0^t \mathbb{W}_1(\mu_s, \nu_s) \mathrm{d}s \\
&\le K_B \int_{-r}^{0} \left(\int_0^t |X^\mu(u) - X^\nu(u)| \mathrm{d}u \right) m(\mathrm{d}\theta) \\
&+ K_B \int_{-r}^{0} \left(\int_{\theta}^{0} |X^\mu(u) - X^\nu(u)| \mathrm{d}u \right) m(\mathrm{d}\theta) + K_B \int_0^t \mathbb{W}_1(\mu_s, \nu_s) \mathrm{d}s \\
&\le K_B \left(\int_0^t |X^\mu(u) - X^\nu(u)| \mathrm{d}u \right) \\
&+ K_B \int_{-r}^{0} m([-r,u]) |X^\mu(u) - X^\nu(u)| \mathrm{d}u + K_B \int_0^t \mathbb{W}_1(\mu_s, \nu_s) \mathrm{d}s, \quad t \in [0,T].
\end{aligned}
$$

Furthermore, by **(Hσ)**, (2.6) and using $\alpha \in [1/2, 1]$, we deduce

$$
I_{4,\varepsilon}(t) \le K_\sigma^2 T \varepsilon^{2\alpha}, \quad t \in [0,T].
$$

In addition, by (2.5), **(Hσ)** and (3.6), we have $\mathbb{E} I_{5,\varepsilon}(t) = 0$. Taking expectation in (3.7), using (2.5) and letting $\varepsilon \downarrow 0$, there exists a constant $C > 0$ such that

$$
\begin{aligned}
\mathbb{E}|X^\mu(t) - X^\nu(t)| &\le \mathbb{E}|X^\mu(0) - X^\nu(0)| + K_B \int_{-r}^{0} m([-r,u]) \mathbb{E}|X^\mu(u) - X^\nu(u)| \mathrm{d}u \\
&\quad + C \int_0^t \mathbb{E}|X^\mu(s) - X^\nu(s)| \mathrm{d}s + (K_b + K_B) \int_0^t \mathbb{W}_1(\mu_s, \nu_s) \mathrm{d}s.
\end{aligned} \tag{3.8}
$$

It follows from Grönwall's inequality that

$$
\begin{aligned}
\mathbb{E}|X^\mu(t) - X^\nu(t)| &\le \mathrm{e}^{Ct} \left\{ \mathbb{E}|X^\mu(0) - X^\nu(0)| + K_B \int_{-r}^{0} m([-r,u]) \mathbb{E}|X^\mu(u) - X^\nu(u)| \mathrm{d}u \right\} \\
&\quad + C(T) \int_0^t \mathbb{W}_1(\mu_s, \nu_s) \mathrm{d}s.
\end{aligned}
$$

So, when $X_0^\mu = X_0^\nu$, for $\lambda = 2C(T)$, we get

$$
\sup_{t \in [0,T]} \mathrm{e}^{-\lambda t} \mathbb{W}_1(\Phi_t(\mu), \Phi_t(\nu)) \le \frac{1}{2} \sup_{t \in [0,T]} \mathrm{e}^{-\lambda t} \mathbb{W}_1(\mu_t, \nu_t).
$$

Set

$$
E_\lambda := \left\{ \mu \in C([0,T]; \mathscr{P}_1(\mathbb{R})) : \mu_0 = \mathscr{L}_{X^\mu(0)} \right\}
$$

and equip it with the complete metric

$$
\rho(\mu, \nu) := \sup_{t \in [0,T]} \mathrm{e}^{-\lambda t} \mathbb{W}_1(\mu_t, \nu_t), \quad \mu, \nu \in E_\lambda.
$$

Then Φ is strictly contractive in E_λ. Consequently, the Banach fixed point theorem together with the definition of Φ implies that there exists a unique $\mu \in E_\lambda$ such that

$$
\Phi_t(\mu) = \mu_t = \mathscr{L}_{X^\mu(t)}, \quad t \in [0,T].
$$

Finally, taking $\mu_t = \mathscr{L}_{X^\mu(t)}$ in (3.6), (3.1) follows from Grönwall's inequality. Similarly, taking $\mu_t = \mathscr{L}_{X(t)}, \nu_t = \mathscr{L}_{\tilde{X}(t)}$, $X^\mu(t) = X(t)$, $X^\nu(t) = \tilde{X}(t)$ in (3.8), (3.2) holds by Grönwall's inequality.

□

3.2 Propagation of Chaos

Let $N \geq 1$ be an integer and $(X_0^i, W^i(t))_{1\leq i\leq N}$ be i.i.d. copies of $(X_0, W(t))$ with $\mathscr{F}_0$-measurable $\mathscr{C}$-valued random variable X_0. Consider

$$\mathrm{d}X^i(t) = b(t, X^i(t), \mathscr{L}_{X^i(t)})\mathrm{d}t + B(t, X_t^i, \mathscr{L}_{X^i(t)})\mathrm{d}t + \sigma(t, X^i(t))\mathrm{d}W^i(t), \quad 1 \leq i \leq N.$$

Let

$$\tilde{\mu}_t^N = \frac{1}{N}\sum_{j=1}^N \delta_{X^j(t)}. \tag{3.9}$$

Consider the stochastic N-interacting particle system:

$$\mathrm{d}X^{i,N}(t) = b(t, X^{i,N}(t), \hat{\mu}_t^N)\mathrm{d}t + B(t, X_t^i, \hat{\mu}_t^N)\mathrm{d}t + \sigma(t, X^{i,N}(t))\mathrm{d}W^i(t), \quad X_0^{i,N} = X_0^i, \tag{3.10}$$

where $\hat{\mu}_t^N$ is the empirical distribution corresponding to $X^{1,N}(t), \cdots, X^{N,N}(t)$, i.e.

$$\hat{\mu}_t^N := \frac{1}{N}\sum_{j=1}^N \delta_{X^{j,N}(t)}.$$

Applying Theorem 2.3, the well-posedness of the stochastic N-interacting particle system (3.10) can be proved in the following lemma.

Lemma 3.2. *Assume* **(H)** *and* $X_0^i \in L^1(\Omega \to (\mathscr{C}, \|\cdot\|_\infty); \mathscr{F}_0, \mathbb{P}), 1 \leq i \leq N$. *Then, for each* $N \geq 1$, (3.10) *admits a unique strong solution* $\{(X^{i,N}(t))_{t\in[-r,T]}\}_{1\leq i\leq N}$ *and*

$$\mathbb{E}\sup_{t\in[-r,T]} |X^{i,N}(t)| \leq C(T)\mathbb{E}\sqrt{\frac{1}{N}\sum_{i=1}^N (1 + \|X_0^i\|_\infty^2)}, \quad 1 \leq i \leq N \tag{3.11}$$

holds for some constant $C(T) > 0$.

Proof. For $x := (x_1, x_2, \cdots, x_N)^* \in \mathbb{R}^N$, $\xi := (\xi_1, \xi_2, \cdots, \xi_N)^* \in \mathscr{C}^N$, set $\tilde{\mu}_x^N = \frac{1}{N}\sum_{i=1}^N \delta_{x_i}$ and

$$\hat{b}(t,x) := \big(b(t, x_1, \tilde{\mu}_x^N), \cdots, b(t, x_N, \tilde{\mu}_x^N)\big)^*, \quad \hat{B}(t,\xi) := \big(B(t, \xi_1, \tilde{\mu}_{\xi(0)}^N), \cdots, B(t, \xi_N, \tilde{\mu}_{\xi(0)}^N)\big)^*,$$
$$\hat{\sigma}(t,x) := \mathrm{diag}\big(\sigma(t,x_1), \cdots, \sigma(t,x_N)\big), \quad \hat{W}(t) := \big(W^1(t), \cdots, W^N(t)\big)^*, \quad t \in [0,T].$$

Then it is clear that $(\hat{W}(t))_{t\in[0,T]}$ is an N-dimensional Brownian motion and (3.10) can be reformulated as

$$\mathrm{d}\hat{X}(t) = \hat{b}(t, \hat{X}(t))\mathrm{d}t + \hat{B}(t, \hat{X}_t)\mathrm{d}t + \hat{\sigma}(t, \hat{X}(t))\mathrm{d}\hat{W}(t), \quad \hat{X}_0 = (X_0^1, X_0^2, \cdots, X_0^N)^*. \tag{3.12}$$

Note that

$$\mathbb{W}_1\left(\frac{1}{N}\sum_{i=1}^N\delta_{x_i},\frac{1}{N}\sum_{i=1}^N\delta_{\tilde{x}_i}\right)\le\frac{1}{N}\sum_{i=1}^N|x_i-\tilde{x}_i|,\quad x_i,\tilde{x}_i\in\mathbb{R},1\le i\le N. \tag{3.13}$$

It is not difficult to see from **(Hb)**, **(HB)** and (3.13) that $\hat{b}$ is locally bounded in $[0,T]\times\mathbb{R}^N$, for any $t\in[0,T]$, $\hat{b}(t,\cdot)$ is continuous,

$$\begin{aligned}(\hat{b}_i(t,x)-\hat{b}_i(t,y))\mathrm{sgn}(x_i-y_i)&=(b(t,x_i,\tilde{\mu}_x^N)-b(t,y_i,\tilde{\mu}_y^N))\mathrm{sgn}(x_i-y_i)\\&\le K_b(|x_i-y_i|+\mathbb{W}_1(\tilde{\mu}_x^N,\tilde{\mu}_y^N))\\&\le K_b(|x_i-y_i|+\frac{1}{N}\sum_{i=1}^N|x_i-y_i|)\\&\le K_b(1+N^{-\frac{1}{2}})|x-y|,\quad x,y\in\mathbb{R}^N,1\le i\le N,\end{aligned} \tag{3.14}$$

and

$$\begin{aligned}|\hat{B}(t,\xi)-\hat{B}(t,\eta)|^2&\le\sum_{i=1}^N|B(t,\xi_i,\tilde{\mu}_{\xi(0)}^N)-B(t,\eta_i,\tilde{\mu}_{\eta(0)}^N)|^2\\&\le 2K_B^2\sum_{i=1}^N(\|\xi_i-\eta_i\|_\infty^2+\mathbb{W}_1(\tilde{\mu}_{\xi(0)}^N,\tilde{\mu}_{\eta(0)}^N)^2)\\&\le 2K_B^2\sum_{i=1}^N(\|\xi_i-\eta_i\|_\infty^2+|\xi_i(0)-\eta_i(0)|^2)\\&\le 4K_B^2\sum_{i=1}^N\|\xi_i-\eta_i\|_\infty^2,\quad\xi,\eta\in\mathscr{C}^N.\end{aligned} \tag{3.15}$$

So, (3.14), (3.15) and **(Hσ)** yield that **(A1)**-**(A3)** hold for $\hat{b},\hat{B},\hat{\sigma},N$ replacing F,H,G,d respectively. Therefore, according to Theorem 2.3, for each $N\ge1$, (3.12) and consequently (3.10) admits a unique strong solution $\{(X^{i,N}(t))_{t\in[-r,T]}\}_{1\le i\le N}$. Finally, by Itô's formula, (3.4), **(HB)** and **(Hσ)**, there exists a constant $C>0$ such that

$$\begin{aligned}|X^{i,N}(t)|^2\le|X^i(0)|^2+C\int_0^t\left[1+|X^{i,N}(s)|^2+\frac{1}{N}\sum_{j=1}^N|X^{j,N}(s)|^2\right]\mathrm{d}s\\+C\int_0^t\|X_s^{i,N}\|_{L^1(m)}^2\mathrm{d}s+\int_0^t2X^{i,N}(s)\sigma(s,X^{i,N}(s))\mathrm{d}W^i(s).\end{aligned}$$

Using the same argument to derive (3.5), we arrive at

$$\sum_{i=1}^N\mathbb{E}\left(\sup_{t\in[-r,T]}|X^{i,N}(t)|^2|\mathscr{F}_0\right)\le C_0(T)\sum_{i=1}^N\left(1+\|X_0^i\|_\infty^2\right)$$

for some constant $C_0(T)>0$. This together with Jensen's inequality with respect to conditional expectation implies that

$$
\begin{aligned}
&\mathbb{E}\sup_{t\in[-r,T]}|X^{i,N}(t)|\\
&=\frac{1}{N}\mathbb{E}\left\{\mathbb{E}\left(\sum_{i=1}^{N}\sup_{t\in[-r,T]}|X^{i,N}(t)||\mathscr{F}_0\right)\right\}\\
&\le\frac{1}{N}\mathbb{E}\left\{\left[\mathbb{E}\left(N\sum_{i=1}^{N}\sup_{t\in[-r,T]}|X^{i,N}(t)|^2|\mathscr{F}_0\right)\right]^{\frac{1}{2}}\right\}\\
&\le\frac{1}{\sqrt{N}}\mathbb{E}\left[C(T)\sum_{i=1}^{N}\left(1+\|X_0^i\|_\infty^2\right)\right]^{\frac{1}{2}}\\
&=\sqrt{C(T)}\mathbb{E}\sqrt{\frac{1}{N}\sum_{i=1}^{N}(1+\|X_0^i\|_\infty^2)}.
\end{aligned}
$$

So, we complete the proof. □

Finally, we give the quantitative propagation of chaos.

Theorem 3.3. *Assume that* $\mathbb{E}\|X_0^i\|_\infty^p<\infty$ *for some* $p>1$ *and* $p\neq 2$. *Let* $\mu_t=\mathscr{L}_{X^i(t)}$.

(1) Then there exists a constant $C(p,T)>0$ *depending only on* p,T *such that*

$$\sup_{t\in[0,T]}\mathbb{E}\big|X^i(t)-X^{i,N}(t)\big|\le C(p,T)(1+\big(\mathbb{E}\|X_0^i\|_\infty^p\big)^{\frac{1}{p}})(N^{-1/2}+N^{-\frac{p-1}{p}}),\tag{3.16}$$

and consequently,

$$\sup_{t\in[0,T]}\mathbb{E}\mathbb{W}_1(\hat{\mu}_t^N,\mu_t)\le C(p,T)(1+\big(\mathbb{E}\|X_0^i\|_\infty^p\big)^{\frac{1}{p}})(N^{-1/2}+N^{-\frac{p-1}{p}}).\tag{3.17}$$

(2) If in addition, $\sigma^2\ge\delta$ *for some* $\delta>0$ *and there exists a constant* $K\ge 0$ *such that*

$$
\begin{aligned}
&|b(t,x,\mu)-b(t,x,\nu)|+|B(t,\xi,\mu)-B(t,\xi,\nu)|\\
&\qquad\le K(1\wedge\mathbb{W}_1(\mu,\nu)),\quad \mu,\nu\in\mathscr{P}_1(\mathbb{R}),t\in[0,T],x\in\mathbb{R},\xi\in\mathscr{C}.
\end{aligned}\tag{3.18}
$$

then there exists a constant $C(p,T)>0$ *depending only on* p,T *such that for any* $1\le k\le N$,

$$
\begin{aligned}
&\sup_{t\in[0,T]}\|\mathscr{L}_{(X^{1,N}(t),X^{2,N}(t),\cdots,X^{k,N}(t))}-\mu_t^{\otimes k}\|_{var}^2\\
&\le 2\sup_{t\in[0,T]}\mathrm{Ent}\left(\mu_t^{\otimes k}|\mathscr{L}_{(X^{1,N}(t),X^{2,N}(t),\cdots,X^{k,N}(t))}\right)\\
&\le kC(p,T)(1+\big(\mathbb{E}\|X_0^i\|_\infty^p\big)^{\frac{1}{p}})(N^{-\frac{1}{2}}+N^{-\frac{(p-1)}{p}}),
\end{aligned}
$$

where $\mu_t^{\otimes k}=\prod_{i=1}^k\mu_t$, *the* k*-independent product of* μ_t.

Proof. Applying Itô's formula, it holds

$$
\begin{aligned}
&V_\varepsilon(X^i(t)-X^{i,N}(t))\\
&=\int_0^t V_\varepsilon'(X^i(s)-X^{i,N}(s))\big(b(s,X^i(s),\hat\mu_s^N)-b(s,X^{i,N}(s),\mu_s)\big)\mathrm{d}s\\
&\quad+\int_0^t V_\varepsilon'(X^i(s)-X^{i,N}(s))\big(B(s,X_s^i,\hat\mu_s^N)-B(s,X_s^{i,N},\mu_s)\big)\mathrm{d}s\\
&\quad+\frac{1}{2}\int_0^t V_\varepsilon''(X^i(s)-X^{i,N}(s))\big(\sigma(s,X^i(s))-\sigma(s,X^{i,N}(s))\big)^2\mathrm{d}s\\
&\quad+\int_0^t V_\varepsilon'(X^i(s)-X^{i,N}(s))\big(\sigma(s,X^i(s))-\sigma(s,X^{i,N}(s))\big)\mathrm{d}W^i(s).
\end{aligned}
$$

By the same argument to derive (3.8) and adopting the triangle inequality for $\mathbb{W}_1$, we arrive at

$$
\mathbb{E}|X^i(t)-X^{i,N}(t)|\le C\int_0^t\{\mathbb{E}|X^i(s)-X^{i,N}(s)|+\mathbb{E}\mathbb{W}_1(\mu_s,\tilde\mu_s^N)+\mathbb{E}\mathbb{W}_1(\tilde\mu_s^N,\hat\mu_s^N)\}\mathrm{d}s,
$$

where $\tilde\mu^N$ was introduced in (3.9). By [7, Theorem 1], there exists a constant $C(p,T)>0$ such that

$$
\mathbb{E}\mathbb{W}_1(\mu_t,\tilde\mu_t^N)\le C(p,T)(1+(\mathbb{E}\|X_0^i\|_\infty^p)^{\frac{1}{p}})(N^{-1/2}+N^{-\frac{p-1}{p}}). \tag{3.19}
$$

As a result, it follows from (3.13) and (3.19) that

$$
\begin{aligned}
&\mathbb{E}|X^i(t)-X^{i,N}(t)|\\
&\le C_1\int_0^t\Big\{\mathbb{E}|X^i(s)-X^{i,N}(s)|+C(p,T)(1+(\mathbb{E}\|X_0^i\|_\infty^p)^{\frac{1}{p}})(N^{-1/2}+N^{-\frac{p-1}{p}})\Big\}\mathrm{d}s
\end{aligned}
$$

for some constant $C_1>0$. Consequently, we derive (3.16) by (3.1), (3.11) and Grönwall's inequality. Finally, note that

$$
\mathbb{W}_1(\hat\mu_s^N,\mu_s)\le\mathbb{W}_1(\hat\mu_s^N,\tilde\mu_s^N)+\mathbb{W}_1(\tilde\mu_s^N,\mu_s)\le\frac{1}{N}\sum_{i=1}^N|X^{i,N}(s)-X^i(s)|+\mathbb{W}_1(\tilde\mu_s^N,\mu_s),
$$

which together with (3.16) and (3.19) yields (3.17).

(2) Rewrite (3.10) as

$$
\mathrm{d}X^{i,N}(t)=b(t,X^{i,N}(t),\mu_t)\mathrm{d}t+B(t,X_t^{i,N},\mu_t)\mathrm{d}t+\sigma(t,X^{i,N}(t))\mathrm{d}\tilde W^i(t),\quad 1\le i\le k
$$

with

$$
\mathrm{d}\tilde W^i(t)=\tilde\Gamma^i(t)\mathrm{d}t+\mathrm{d}W^i(t),\quad 1\le i\le k
$$

and

$$
\tilde\Gamma^i(t)=\sigma(t,X^{i,N}(t))^{-1}[b(t,X^{i,N}(t),\hat\mu_t^N)-b(t,X^{i,N}(t),\mu_t)+B(t,X_t^{i,N},\hat\mu_t^N)-B(t,X_t^{i,N},\mu_t)].
$$

It follows from (3.18) and $\sigma^2 \geq \delta$ that there exists a constant $C > 0$ such that

$$|\tilde{\Gamma}^i(t)| \leq C(\mathbb{W}_1(\hat{\mu}_t^N, \mu_t) \wedge 1), \quad t \in [0,T], 1 \leq i \leq k. \tag{3.20}$$

Let

$$R_t^k = \exp\left\{-\sum_{i=1}^k \int_0^t \langle \tilde{\Gamma}^i(s), \mathrm{d}W^i(s)\rangle - \frac{1}{2}\sum_{i=1}^k \int_0^t |\tilde{\Gamma}^i(s)|^2 \mathrm{d}s\right\}, \quad t \in [0,T].$$

(3.20) and Girsanov's theorem imply that $\{R_t^k\}_{t\in[0,T]}$ is a martingale and $((\tilde{W}^i(t))_{1\leq i\leq k})_{t\in[0,T]}$ is a k-dimensional Brownian motion under $\mathbb{Q}_T^k = R_T^k\mathbb{P}$ and

$$\mathscr{L}_{(X^{1,N}(t), X^{2,N}(t), \cdots, X^{k,N}(t))}|\mathbb{Q}_T^k = \mu_t^{\otimes k}, \quad t \in [0,T]. \tag{3.21}$$

This implies that

$$\begin{aligned}\mu_t^{\otimes k}(f) &= \mathbb{E}[R_T^k f(X^{1,N}(t), X^{2,N}(t), \cdots, X^{k,N}(t))]\\ &= \mathbb{E}[R_t^k f(X^{1,N}(t), X^{2,N}(t), \cdots, X^{k,N}(t))], \quad f \in \mathscr{B}_b(\mathbb{R}^k), t \in [0,T].\end{aligned}$$

So, there exists a constant $C > 0$ such that

$$\begin{aligned}&\mathrm{Ent}(\mu_t^{\otimes k}|\mathscr{L}_{(X^{1,N}(t), X^{2,N}(t), \cdots, X^{k,N}(t))})\\ &= \mathbb{E}(R_t^k \log R_t^k) = \frac{1}{2}\sum_{i=1}^k \int_0^t \mathbb{E}^{\mathbb{Q}_T^k}|\tilde{\Gamma}^i(s)|^2\mathrm{d}s \leq C^2 k\int_0^t \mathbb{E}^{\mathbb{Q}_T^k}(\mathbb{W}_1(\hat{\mu}_s^N, \mu_s)\wedge 1)^2 \mathrm{d}s, \quad t \in [0,T].\end{aligned}$$

This together with Pinsker's inequality (1.1) yields

$$\begin{aligned}&\|\mu_t^{\otimes k} - \mathscr{L}_{(X^{1,N}(t), X^{2,N}(t), \cdots, X^{k,N}(t))}\|_{var}^2\\ &\leq 2\mathrm{Ent}(\mu_t^{\otimes k}|\mathscr{L}_{(X^{1,N}(t), X^{2,N}(t), \cdots, X^{k,N}(t))})\\ &\leq 2C^2 k\int_0^t \mathbb{E}^{\mathbb{Q}_T^k}(\mathbb{W}_1(\hat{\mu}_s^N, \mu_s))\mathrm{d}s.\end{aligned}$$

The proof is finished by (3.21) and (3.19). □

Remark 3.4. *For quantitative propagation of chaos, one can refer to [11] and references therein for the convolution type distribution dependent SDEs. Since we only assume that the drift is Lipschitz continuous under L^1-Wasserstein distance and the estimate in [7, Theorem 1] for the convergence rate of empirical distribution of i.i.d. random variables plays crucial role, the order of the quantitative propagation of chaos may be not optimal.*

References

[1] M. Bauer, T. M-Brandis, F. Proske, *Strong Solutions of Mean-Field Stochastic Differential Equations with irregular drift,* Electron. J. Probab. 23(2018), 1-35.

[2] R. J. Berman, M. Önnheim, *Propagation of Chaos for a Class of First Order Models with Singular Mean Field Interactions,* SIAM J. Math. Anal. 51(2019), 159-196.

[3] R. Carmona, F. Delarue, *Probabilistic theory of mean field games with applications. I.* Mean field FBSDEs, control, and games. Probability Theory and Stochastic Modelling, 83. Springer, Cham, 2018.

[4] P.-E. Chaudru de Raynal, *Strong well-posedness of McKean-Vlasov stochastic differential equation with Hölder drift,* Stochastic Process. Appl. 130(2020), 79-107.

[5] P.-E. Chaudru de Raynal, N. Frikha, *Well-posedness for some non-linear diffusion processes and related pde on the Wasserstein space,* J. Math. Pures Appl. 159(2022), 1-167.

[6] G. dos Reis, S. Engelhardt, G. Smith, *Simulation of McKean-Vlasov SDEs with super linear growth,* IMA J. Numer. Anal. 42(2022), 874-922.

[7] N. Fournier, A. Guillin, *On the rate of convergence in Wasserstein distance of the empirical measure,* Probab. Theory Related Fields 162(2015), 707-738.

[8] N. Fournier, M. Hauray, *Propagation of Chaos for the Landau equation with moderately soft potentials,* Ann. Probab. 44(2016), 3581-3660.

[9] O. Guéant O, J.-M. Lasry, P.-L. Lions, Mean Field Games and Applications, *Paris-Princeton Lectures on Mathematical Finance 2010. Lecture Notes in Math.* 2003, Springer, Berlin, 205-266.

[10] A. Guillin, W. Liu, L. Wu, C. Zhang, *The kinetic Fokker-Planck equation with mean field interaction,* J. Math. Pures Appl. 150(2021),1-23.

[11] Z. Hao, M. Röckner, X. Zhang, *Strong convergence of propagation of chaos for McKean-Vlasov SDEs with singular interactions,* arXiv:2204.07952.

[12] X. Huang, F.-Y. Wang, *Distribution dependent SDEs with singular coefficients,* S-tochastic Process. Appl. 129(2019), 4747-4770.

[13] X. Huang, F.-Y. Wang, *McKean-Vlasov SDEs with drifts discontinuous under Wasserstein distance,* Discrete Contin. Dyn. Syst. 41(2021), 1667-1679.

[14] X. Huang, *Path-distribution dependent SDEs with singular coefficients,* Electron. J. Probab. 26(2021), 1-21.

[15] B. Jourdain, J. Reygner, *Propagation of chaos for rank-based interacting diffusions and long time behaviour of a scalar quasilinear parabolic equation,* Stoch. PDE: Anal. Comp. 1(2013), 455-506.

[16] N. Ikeda, S. Watanabe, *Stochastic differential equations and diffusion processes*, North-Holland Mathematical Library, 24. North-Holland Publishing Co., Amsterdam-New York; Kodansha, Ltd., Tokyo, 1981.

[17] D. Lacker, *On a strong form of propagation of chaos for McKean-Vlasov equations,* Electron. Commun. Probab. 23(2018), 1-11.

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

[19] Yu. S. Mishura, A. Yu. Veretennikov, *Existence and uniqueness theorems for solutions of McKean-Vlasov stochastic equations,* Theor. Probability and Math. Statist. 103(2020), 59-101.

[20] M. S. Pinsker, *Information and Information Stability of Random Variables and Processes,* Holden-Day, San Francisco, 1964.

[21] M. von Renesse, M. Scheutzow, *Existence and uniqueness of solutions of stochastic functional differential equations,* Random Oper. Stoch. Equ. 18(2010), 267-284.

[22] P. Ren, F.-Y. Wang, *Exponential convergence in entropy and Wasserstein for McKean-Vlasov SDEs,* Nonlinear Anal. 206(2021), 112259.

[23] M. Röckner, X. Zhang, *Well-posedness of distribution dependent SDEs with singular drifts,* Bernoulli 27(2021), 1131-1158.

[24] A.-S. Sznitman, Topics in propagation of chaos, École d' Été de Prob. de Saint-Flour XIX-1989, 165-251, Lecture Notes in Math., 1464, Springer, Berlin, 1991.

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

[26] F.-Y. Wang, *Exponential Ergodicity for Singular Reflecting McKean-Vlasov SDEs,* arXiv:2108.03908.

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