Journal of Computational and Applied Mathematics 441 (2024) 115682

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Convergence rate in $\mathcal{L}^p$ sense of tamed EM scheme for highly nonlinear neutral multiple-delay stochastic McKean–Vlasov equations

Shuaibin Gao [a], Qian Guo [a], Junhao Hu [b], Chenggui Yuan [c,*]

[a] Department of Mathematics, Shanghai Normal University, Shanghai 200234, China
[b] School of Mathematics and Statistics, South-Central University For Nationalities, Wuhan 430074, China
[c] Department of Mathematics, Swansea University, Bay Campus, Swansea, SA1 8EN, UK

ARTICLE INFO

Keywords:
The tamed Euler–Maruyama scheme
Neutral multiple-delay stochastic McKean-Vlasov equation
Strong convergence rate
Propagation of chaos

ABSTRACT

This paper focuses on the numerical scheme of highly nonlinear neutral multiple-delay stochastic McKean-Vlasov equation (NMSMVE) by virtue of the stochastic particle method. First, under general assumptions, the results about propagation of chaos in $\mathcal{L}^p$ sense are revealed, where the convergence rate loses a little due to the proof technique. Then the tamed Euler–Maruyama scheme to the corresponding particle system is established and the convergence rate in $\mathcal{L}^p$ sense is obtained. Furthermore, combining these two results gives the convergence error in $\mathcal{L}^p$ sense between the objective NMSMVE and numerical approximation, which is related to the particle number and step size. Finally, two numerical examples are provided to support the finding.

1. Introduction

The theories of stochastic McKean-Vlasov equations (SMVEs) have been investigated by plenty of scholars, since SMVEs appear in many research fields, such as biological systems, chemistry and mean-field games [1–3]. The salient feature of SMVEs is that the coefficients depend on the distributions of state variables, which brings difficulties to the research. SMVEs are also called distribution dependent stochastic differential equations (SDEs) or mean-field SDEs. Reviewing the pioneering works, SMVEs were studied by McKean in [4–6], which were inspired by [7]. The existence and uniqueness of SMVEs were discussed in [8–10]. As for other theories of SMVEs, we refer the readers to [11–16].

As so often is the case, the true solutions to SMVEs cannot be expressed explicitly. Hence, analyzing the numerical solutions is a common way to get the properties of the true solutions. However, the classical Euler–Maruyama (EM) scheme cannot be used to simulate SDEs with superlinear coefficients well [17]. By borrowing the ideas in [18–20], the tamed EM scheme for SMVEs with superlinear drift coefficients was proposed in [21]. The tamed Milstein scheme for SMVEs was established to improve the convergence rate in [22]. In addition, after the existence and uniqueness of the solution to SMVE with superlinear drift and diffusion coefficients were proven by using the new method, the tamed EM and Milstein schemes were also analyzed in [23].

When the time-delay is taken into consideration, the theories for SMVEs with delay were discussed in [24–27]. The neutral SMVEs with delay refer to a class of SMVEs which not only depend on the present, past state variables, but also contain derivatives with delay. The neutral SMVEs with delay were approximated by the tamed EM scheme in [28]. However, the equation form is limited. For example, the following scalar equation is not included in [28]:

$$d[Y(t)+Y^3(t-\rho)] = [-2Y(t)+Y(t-\rho)-2Y^5(t-\rho)+\mathbb{E}Y(t)]dt + [Y(t)+Y(t-\rho)]dB(t). \tag{1.1}$$

* Corresponding author.
E-mail address: C.Yuan@swansea.ac.uk (C. Yuan).

https://doi.org/10.1016/j.cam.2023.115682
Received 6 April 2023; Received in revised form 9 September 2023
Available online 11 November 2023

Here, ρ is the constant delay. Actually, some numerical schemes for neutral stochastic differential delay equations (NSDDEs), whose coefficients are not dependent of the distributions, also have this limitation, such as [29–33]. To overcome this shortcoming and relax the constraint of the delay variables, we use the techniques in [34,35] to approximate the highly nonlinear neutral SMVEs with delay.

In this paper, we focus on a class of highly nonlinear neutral multiple-delay stochastic McKean-Vlasov equations (NMSMVEs) of the form:

$$\begin{aligned} d[Y(t)-D(Y(t-\rho))] = {}& \alpha\left(Y(t), Y(t-\rho_2), \dots, Y(t-\rho_r), \mathbb{L}_{Y(t)}, \mathbb{L}_{Y(t-\rho_2)}, \dots, \mathbb{L}_{Y(t-\rho_r)}\right) dt \\ & + \beta\left(Y(t), Y(t-\rho_2), \dots, Y(t-\rho_r), \mathbb{L}_{Y(t)}, \mathbb{L}_{Y(t-\rho_2)}, \dots, \mathbb{L}_{Y(t-\rho_r)}\right) dB(t), \end{aligned} \tag{1.2}$$

on $t \in [0,T]$, where $\mathbb{L}_{Y(t-\rho_v)}$ is the law of Y at time $t-\rho_v$ for $v \in \mathbb{S}_r := \{1,2,\dots,r\}$. Moreover, for $v \in \mathbb{S}_r$, let $0 \le \rho_v \le \rho$. Here, $D : \mathbb{R}^d \to \mathbb{R}^d$, $\alpha : (\mathbb{R}^d)^r \times (\mathcal{P}_2(\mathbb{R}^d))^r \to \mathbb{R}^d$, $\beta : (\mathbb{R}^d)^r \times (\mathcal{P}_2(\mathbb{R}^d))^r \to \mathbb{R}^{d\times m}$.

In reality, the multiple-delay systems are very significant, and they turn up on many occasions [36–39]. Then in this paper, the tamed EM scheme is established for NMSMVEs (1.2). Additionally, the convergence rate in $\mathcal{L}^p$ sense is shown, which is differential from the previous papers [21,22,24,26,28,40,41], where only the convergence rate in $\mathcal{L}^2$ sense was given. After analyzing the existing results, we find that the key to overcome this problem is to obtain the propagation of chaos in $\mathcal{L}^p$ sense. By virtue of the theory in [42], we give this result in Theorems 3.1 and 3.2.

All in all, the main contributions of the present paper can be stated as follows.

- There is only one delay in [26,28], but we shall deal with multiple delays in (1.2). Moreover, the coefficients of (1.2) depend on the distributions of the delay variables.
- The form of the equations and the constraint of the delay variables are more general, which are allowed highly nonlinear.
- The requirement for the neutral term is also allowed highly nonlinear.
- The propagation of chaos in $\mathcal{L}^p$ sense is shown with the aid of the theory in [42]. Then the convergence rate in $\mathcal{L}^p$ sense of the tamed EM scheme is presented.

The organization for the rest of the paper is as follows. We simplify the NMSMVEs by the projection operator and give the moment boundedness of the true solution in Section 2. The propagation of chaos in $\mathcal{L}^p$ sense is shown in Section 3. In Section 4, the tamed EM scheme is established to approximate NMSMVEs. Section 5 contains a one-dimensional example and a two-dimensional example.

2. Preliminaries

Let $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\ge 0}, \mathbb{P}\right)$ be a complete probability space with a filtration $\{\mathcal{F}_t\}_{t\ge0}$ satisfying the usual conditions (i.e., it is increasing and right continuous while $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets). For $x \in \mathbb{R}^d$, let $|x|$ be its Euclidean norm. For the real numbers b_1, b_2, denote $b_1 \wedge b_2 = \min\{b_1, b_2\}$ and $b_1 \vee b_2 = \max\{b_1, b_2\}$. Let $\lfloor b_1 \rfloor$ be the largest integer that does not exceed b_1. For a set S, define $\mathbb{I}_S(x) = 1$ if $x \in S$ and $\mathbb{I}_S(x) = 0$ if $x \notin S$ (i.e., $\mathbb{I}_S$ is indicator function). Assume that $\mathcal{C} := \mathcal{C}([-\rho, 0]; \mathbb{R}^d)$ is the family of all continuous functions φ from $[-\rho, 0]$ to $\mathbb{R}^d$ with the norm $\|\varphi\| = \sup_{-\rho\le\theta\le0}|\varphi(\theta)|$. The probability expectation with respect to $\mathbb{P}$ is defined by $\mathbb{E}$. For $p \ge 1$, $\mathcal{L}^p := \mathcal{L}^p(\Omega, \mathcal{F}, \mathbb{P})$ is the set of random variables X with $\mathbb{E}|X|^p < \infty$. Let $B(t)$ be an m-dimensional Brownian motion on the probability space. Denote $\mathbb{R}_+ = [0, +\infty)$.

Let $\delta_y(\cdot)$ stand for the Dirac measure at point $y \in \mathbb{R}^d$. Assume that $\mathcal{P}(\mathbb{R}^d)$ is the family of all probability measures on $\mathbb{R}^d$. For $q \ge 1$, define

$$\mathcal{P}_q(\mathbb{R}^d) = \left\{ \mu \in \mathcal{P}(\mathbb{R}^d) : \left(\int_{\mathbb{R}^d} |y|^q \mu(dy) \right)^{1/q} < \infty \right\},$$

and set $\mathcal{W}_q(\mu) = \left(\int_{\mathbb{R}^d} |y|^q \mu(dy)\right)^{1/q}$ for any $\mu \in \mathcal{P}_q(\mathbb{R}^d)$. For $q \ge 1$, the Wasserstein distance of $\mu, \nu \in \mathcal{P}_q(\mathbb{R}^d)$ is defined by

$$\mathbb{W}_q(\mu, \nu) = \inf_{\pi \in \mathfrak{C}(\mu,\nu)} \left(\int_{\mathbb{R}^d\times\mathbb{R}^d} |y_1 - y_2|^q \pi(dy_1, dy_2) \right)^{1/q},$$

where $\mathfrak{C}(\mu, \nu)$ is the family of all couplings for μ, ν, i.e., $\pi(\cdot, \mathbb{R}^d) = \mu(\cdot)$ and $\pi(\mathbb{R}^d, \cdot) = \nu(\cdot)$.

We quote Lemma 2.3 in [8] as the following lemma.

Lemma 2.1. *For any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, we have $\mathbb{W}_2(\mu, \delta_0) = \mathcal{W}_2(\mu)$.*

Define the segment process $y_t = \{y(t+\theta) : -\rho \le \theta \le 0\}$ for any $t \ge 0$. Then $y_t \in \mathcal{C}$. In order to simplify the equation form, we introduce the following projection operator. Let $\Gamma_\theta(\varphi) : \mathcal{C} \to \mathbb{R}^d$, $\Gamma_\theta(\varphi) = \varphi(\theta)$ for $\varphi \in \mathcal{C}$ and $\theta \in [-\rho, 0]$. In addition, we set

$$\Gamma(\varphi) = (\Gamma_{\bar{s}_1}(\varphi), \Gamma_{\bar{s}_2}(\varphi), \dots, \Gamma_{\bar{s}_r}(\varphi)),$$

$$\mathbb{L}_{\Gamma(\varphi)} = (\mathbb{L}_{\Gamma_{\bar{s}_1}(\varphi)}, \mathbb{L}_{\Gamma_{\bar{s}_2}(\varphi)}, \dots, \mathbb{L}_{\Gamma_{\bar{s}_r}(\varphi)}),$$

for any $\varphi \in C$ and $\bar{s}_1, \bar{s}_2, \ldots, \bar{s}_r \in [-\rho, 0]$. Let $\bar{s}_1 = 0$, $\bar{s}_r = -\rho$ throughout the paper. For another, it should be noted that we can arrange the r-delays in (1.2) into a non subtractive sequence $\{\rho_1, \rho_2, \ldots, \rho_r\}$. Then set $\rho_1 = -\bar{s}_1 = 0$, $\rho_r = -\bar{s}_r = \rho$ and $\rho_v = -\bar{s}_v$, $v \in \{2, 3, \ldots, r-1\}$. Based on these notations, NMSMVE (1.2) can be rewritten as

$$d[Y(t) - D(Y(t-\rho))] = \alpha\left(\Gamma(Y_t), \mathbb{L}_{\Gamma(Y_t)}\right) dt + \beta\left(\Gamma(Y_t), \mathbb{L}_{\Gamma(Y_t)}\right) dB(t), \quad t \in [0, T], \tag{2.1}$$

with the initial value $Y_0 = \xi \in C^{\check{p}}_{\mathcal{F}_0}([-\rho, 0]; \mathbb{R}^d)$, where $C^{\check{p}}_{\mathcal{F}_0}([-\rho, 0]; \mathbb{R}^d)$ is the family of all $\mathcal{F}_0$-measurable C-valued random variables ξ with $\mathbb{E}\|\xi\|^{\check{p}} < \infty$ for any $\check{p} > 0$. Moreover, $D : \mathbb{R}^d \to \mathbb{R}^d$, $\alpha : \left(\mathbb{R}^d\right)^r \times \left(\mathcal{P}_2(\mathbb{R}^d)\right)^r \to \mathbb{R}^d$, $\beta : \left(\mathbb{R}^d\right)^r \times \left(\mathcal{P}_2(\mathbb{R}^d)\right)^r \to \mathbb{R}^{d\times m}$ are all Borel-measurable.

Remark 1. One can observe that

$$\begin{aligned}\Gamma(Y_t) =& (\Gamma_{\bar{s}_1}(Y_t), \Gamma_{\bar{s}_2}(Y_t), \ldots, \Gamma_{\bar{s}_r}(Y_t))\\ =& (Y_t(\bar{s}_1), Y_t(\bar{s}_2), \ldots, Y_t(\bar{s}_r))\\ =& (Y(t), Y(t+\bar{s}_2), \ldots, Y(t+\bar{s}_r))\\ =& (Y(t), Y(t-\rho_2), \ldots, Y(t-\rho)).\end{aligned}$$

The Theorem 1 in [42] is cited as the following theorem.

Theorem 2.2. *Let $\{X_n\}_{n\ge 1}$ be a sequence of independent and identically distributed (i.i.d.) random variables in $\mathbb{R}^d$ with the distribution $\mu \in \mathcal{P}_{\bar{p}}(\mathbb{R}^d)$ and define the empirical measure $\mu_N = \frac{1}{N}\sum_{j=1}^N \delta_{X_j}$ for $j \in \mathbb{S}_N$. Then for $\bar{p} > p \ge 2$, there exists a constant $C_{p,\bar{p},d}$ depending on p, $\bar{p}$, d, such that, for all $N \ge 1$,*

$$\mathbb{E}\left(\mathbb{W}_p^p(\mu_N, \mu)\right) \le C_{p,\bar{p},d}\begin{cases} N^{-1/2} + N^{-(\bar{p}-p)/\bar{p}}, & \text{if } p > d/2 \text{ and } \bar{p} \neq 2p,\\ N^{-1/2}\log(1+N) + N^{-(\bar{p}-p)/\bar{p}}, & \text{if } p = d/2 \text{ and } \bar{p} \neq 2p,\\ N^{-p/d} + N^{-(\bar{p}-p)/\bar{p}}, & \text{if } 2 \le p < d/2.\end{cases}$$

Then the special case is stated as the following corollary.

Corollary 2.3. *Assume that the settings in Theorem 2.2 hold. Then for $\bar{p} > 2p$ and $p \ge 2$, there exists a constant $C_{p,\bar{p},d}$ depending on p, $\bar{p}$, d, such that, for all $N \ge 1$,*

$$\mathbb{E}\left(\mathbb{W}_p^p(\mu_N, \mu)\right) \le C_{p,\bar{p},d}\begin{cases} N^{-1/2}, & \text{if } p > d/2,\\ N^{-1/2}\log(1+N), & \text{if } p = d/2,\\ N^{-p/d}, & \text{if } 2 \le p < d/2.\end{cases}$$

In order to make assumptions on coefficients, we denote $U_i : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, $i = 1, 2, 3$ and there exist constants $K_U > 0$ and $l_i \ge 1$ such that

$$0 \le U_i(\bar{x}, \hat{x}) \le K_U(1 + |\bar{x}|^{l_i} + |\hat{x}|^{l_i}), \tag{2.2}$$

for any $\bar{x}, \hat{x} \in \mathbb{R}^d$ and $i = 1, 2, 3$.

Let $x^{(r)} = \left(x_1, x_2, \ldots, x_r\right)$, $y^{(r)} = \left(y_1, y_2, \ldots, y_r\right)$ for x_i, $y_i \in \mathbb{R}^d$, $i \in \mathbb{S}_r$, and $\mu^{(r)} = \left(\mu_1, \mu_2, \ldots, \mu_r\right)$, $\nu^{(r)} = \left(\nu_1, \nu_2, \ldots, \nu_r\right)$ for μ_i, $\nu_i \in \mathcal{P}_2(\mathbb{R}^d)$, $i \in \mathbb{S}_r$.

Assumption 2.4. There exist positive constants K_{11}, K_{12}, K_{13} such that

$$\begin{aligned}&\left|\alpha(x^{(r)}, \mu^{(r)}) - \alpha(y^{(r)}, \mu^{(r)})\right| \le K_{11}\sum_{i=1}^r \left[U_1(x_i, y_i)|x_i - y_i|\right],\\ &\left|\alpha(x^{(r)}, \mu^{(r)}) - \alpha(x^{(r)}, \nu^{(r)})\right| \le K_{12}\sum_{i=1}^r \mathbb{W}_2(\mu_i, \nu_i),\\ &\left(x_1 - D(x_r) - y_1 + D(y_r)\right)^T\left(\alpha(x^{(r)}, \mu^{(r)}) - \alpha(y^{(r)}, \nu^{(r)})\right)\\ &\le K_{13}\left[\sum_{i=1}^{r-1}|x_i - y_i|^2 + U_2^2(x_r, y_r)|x_r - y_r|^2 + \sum_{i=1}^r \mathbb{W}_2^2(\mu_i, \nu_i)\right],\end{aligned}$$

for any $x^{(r)}, y^{(r)} \in (\mathbb{R}^d)^r$ and $\mu^{(r)}, \nu^{(r)} \in (\mathcal{P}_2(\mathbb{R}^d))^r$.

Assumption 2.5. There exists a constant $K_2 > 0$ such that

$$\begin{aligned}&\left|\beta(x^{(r)}, \mu^{(r)}) - \beta(y^{(r)}, \mu^{(r)})\right| \le K_2\left[\sum_{i=1}^{r-1}|x_i - y_i| + U_2(x_r, y_r)|x_r - y_r|\right],\\ &\left|\beta(x^{(r)}, \mu^{(r)}) - \beta(x^{(r)}, \nu^{(r)})\right| \le K_2\sum_{i=1}^r \mathbb{W}_2(\mu_i, \nu_i),\end{aligned}$$

for any $x^{(r)}, y^{(r)} \in (\mathbb{R}^d)^r$ and $\mu^{(r)}, \nu^{(r)} \in (\mathcal{P}_2(\mathbb{R}^d))^r$.

Assumption 2.6. $D(0) = 0$ and there exists a constant $K_3 > 0$ such that

$$\left|D(x_r) - D(y_r)\right| \le K_3 U_3(x_r, y_r)|x_r - y_r|,$$

for any $x_r,\ y_r \in \mathbb{R}^d$.

Assumption 2.7. There exists a constant $K_4 > 0$ such that for all positive $\check{p}$,

$$\mathbb{E}\left(\sup_{-\rho\le t_1,t_2\le 0} |\xi(t_1) - \xi(t_2)|^{\check{p}}\right) \le K_4|t_1 - t_2|^{\check{p}/2}.$$

Remark 2. In the numerical examples, for the third inequality in Assumption 2.4, we need to check the case $\mu^{(r)} = \nu^{(r)}$ as follows

$$\begin{aligned}&\left(x_1 - D(x_r) - y_1 + D(y_r)\right)^T \left(\alpha(x^{(r)}, \mu^{(r)}) - \alpha(y^{(r)}, \mu^{(r)})\right)\\ &\le C\left[\sum_{i=1}^{r-1}|x_i - y_i|^2 + U_2^2(x_r, y_r)|x_r - y_r|^2\right].\end{aligned} \tag{2.3}$$

In fact, one can see that

$$\begin{aligned}&\left(x_1 - D(x_r) - y_1 + D(y_r)\right)^T \left(\alpha(x^{(r)}, \mu^{(r)}) - \alpha(y^{(r)}, \nu^{(r)})\right)\\ &= \left(x_1 - D(x_r) - y_1 + D(y_r)\right)^T \left(\alpha(x^{(r)}, \mu^{(r)}) - \alpha(y^{(r)}, \mu^{(r)})\right)\\ &\quad + \left(x_1 - D(x_r) - y_1 + D(y_r)\right)^T \left(\alpha(y^{(r)}, \mu^{(r)}) - \alpha(y^{(r)}, \nu^{(r)})\right).\end{aligned}$$

Note that

$$\begin{aligned}&\left(x_1 - D(x_r) - y_1 + D(y_r)\right)^T \left(\alpha(y^{(r)}, \mu^{(r)}) - \alpha(y^{(r)}, \nu^{(r)})\right)\\ &\le (|x_1 - y_1|^2 + |D(x_r) - D(y_r)|^2) + \frac{1}{2}rK_{12}^2\sum_{i=1}^{r}\mathbb{W}_2^2(\mu_i, \nu_i)\\ &\le (|x_1 - y_1|^2 + K_3^2U_2^2(x_r, y_r)|x_r - y_r|^2) + \frac{1}{2}rK_{12}^2\sum_{i=1}^{r}\mathbb{W}_2^2(\mu_i, \nu_i),\end{aligned}$$

where in the last inequality, we have chosen sufficiently large l_2 such that $U_3(x_r, y_r) \le U_2(x_r, y_r)$ hold with the aid of (2.2) and Assumption 2.6. Then the third inequality in Assumption 2.4 holds by the above inequality and (2.3).

Moreover, by Assumptions 2.4–2.6, one can see there exist some constants $\bar{C}_1, \bar{C}_2, \bar{C}_3$ such that

$$\left|\alpha(x^{(r)}, \mu^{(r)})\right| \le \bar{C}_1\left(1 + \sum_{i=1}^{r}\left[U_1(x_i, 0)|x_i| + \mathcal{W}_2(\mu_i)\right]\right), \tag{2.4}$$

$$\left(x_1 - D(x_r)\right)^T \alpha(x^{(r)}, \mu^{(r)}) \le \bar{C}_2\left[1 + \sum_{i=1}^{r-1}|x_i|^2 + U_2^2(x_r, 0)|x_r|^2 + \sum_{i=1}^{r}\mathcal{W}_2^2(\mu_i)\right], \tag{2.5}$$

$$\left|\beta(x^{(r)}, \mu^{(r)})\right|^2 \le \bar{C}_3\left[1 + \sum_{i=1}^{r-1}|x_i|^2 + U_2^2(x_r, 0)|x_r|^2 + \sum_{i=1}^{r}\mathcal{W}_2^2(\mu_i)\right], \tag{2.6}$$

$$\left|D(x_r)\right| \le K_3U_3(x_r, 0)|x_r|, \tag{2.7}$$

for any $x^{(r)} \in (\mathbb{R}^d)^r$ and $\mu^{(r)} \in (\mathcal{P}_2(\mathbb{R}^d))^r$. In the rest of this paper, let $l_U = max\{l_1, l_2, l_3\}$ and $\xi \in C_{\mathcal{F}_0}^{\check{p}}([-\rho, 0]; \mathbb{R}^d)$ for any $\check{p} > 0$ for simplicity.

Theorem 2.8. *Let Assumptions 2.4–2.6 hold. Then there exists a unique strong solution $Y(t)$ to (2.1) and $Y(t)$ satisfies for any $\bar{p} > 0$ and $T > 0$,*

$$\mathbb{E}\left(\sup_{0\le t\le T}|Y(t)|^{\bar{p}}\right) \le C_{\check{C}_*, T, \|\xi\|, l_*, \bar{p}_j},$$

where the notations $\check{C}_*, l_*, \bar{p}_j$ are defined in the proof.

Proof. We divide the proof into three steps, and the techniques used here can be found in [28,34,43].

Step 1: For $t \in [0, T]$ and $n = 1, 2, \ldots$, let $Y_{(n)}(t)$ solve the distribution-interaction equation:

$$d[Y_{(n)}(t) - D(Y_{(n)}(t - \rho))] = \alpha\left(\Gamma_{(n)}(Y_t), \mathbb{L}_{\Gamma_{(n-1)}(Y_t)}\right)dt + \beta\left(\Gamma_{(n)}(Y_t), \mathbb{L}_{\Gamma_{(n-1)}(Y_t)}\right)dB(t), \tag{2.8}$$

with the initial value $Y_{(n)}(\theta) = \xi(\theta), \theta \in [-\rho, 0]$. In this step, we shall prove that (2.8) admits a unique solution $Y_{(n)}(t)$ and give the moment boundedness of $Y_{(n)}(t)$. Here, $\mathbb{L}_{\Gamma_{(n)}(Y_t)}$ is the law of $\Gamma_{(n)}(Y_t)$, where

$$\begin{aligned}\Gamma_{(n)}(Y_t) &= \left(\Gamma_{\bar{s}_1}(Y_{t,(n)}), \Gamma_{\bar{s}_2}(Y_{t,(n)}), \ldots, \Gamma_{\bar{s}_r}(Y_{t,(n)})\right)\\ &= \left(Y_{t,(n)}(\bar{s}_1), Y_{t,(n)}(\bar{s}_2), \ldots, Y_{t,(n)}(\bar{s}_r)\right)\\ &= \left(Y_{(n)}(t+\bar{s}_1), Y_{(n)}(t+\bar{s}_2), \ldots, Y_{(n)}(t+\bar{s}_r)\right)\\ &= \left(Y_{(n)}(t), Y_{(n)}(t-\rho_2), \ldots, Y_{(n)}(t-\rho)\right).\end{aligned}$$

Moreover,

$$\begin{cases} Y_{(0)}(\theta) = \xi(\theta), & \theta \in [-\rho, 0], \\ Y_{(0)}(t) = \xi(0), & t \in [0, T]. \end{cases}$$

When $n = 1$, the drift and diffusion coefficients of (2.8) depend on the state variable $Y_{(1)}(\cdot)$ and distribution of ξ. Then the NSDDE is independent of the distribution $\mathbb{L}_{Y_{(1)}(\cdot)}$, we obtain from [43] that it admits a unique solution $Y_{(1)}(t)$ under the given conditions. Since $\mathcal{W}_2^2(\mathbb{L}_{Y_{(0)}(s+\bar{s}_i)}) \leq \mathbb{E}\|\xi\|^2$ for $i \in \mathbb{S}_r$, the procedure to estimate the moment of $Y_{(1)}(t)$ is similar to Step 3, we omit the details here. For every integer $\tilde{N} \geq 1$, define

$$\tau_{\tilde{N},(1)} = T \wedge \inf\left\{t \in [0, T] : |Y_{(1)}(t)| \geq \tilde{N}\right\}.$$

We can get from Hölder's inequality, Gronwall's inequality and the Burkholder–Davis–Gundy (BDG) inequality that, for $\bar{p} \geq 2$ and $t \in [0, T]$,

$$\mathbb{E}\left(\sup_{0\leq u\leq t\wedge\tau_{\tilde{N},(1)}} \left|Y_{(1)}(u)\right|^{\bar{p}}\right) \leq \hat{C}_{T*} + \hat{C}_{T*}(T+1)\mathbb{E}\left(\sup_{0\leq u\leq t\wedge\tau_{\tilde{N},(1)}} \left|Y_{(1)}(u-\rho)\right|^{\bar{p}l_*}\right), \tag{2.9}$$

where $\hat{C}_{T*} = \hat{C}_* e^{T\hat{C}_*}$, $l_* = l_U + 1$, and

$$\begin{aligned}\hat{C}_* = &\left[(1+(1+r)\mathbb{E}\|\xi\|^{\bar{p}}) \vee (K_U^{\bar{p}} 2^{\bar{p}-1} + r) \vee \frac{\bar{p}(\bar{p}-1)}{2} \vee \bar{p}\right] \\ &\cdot \left(\left[(\bar{C}_2 + \bar{C}_3)2^{\bar{p}-2}(1 \vee K_3^{\bar{p}-2})\left(\frac{p-2}{p} \vee \frac{2}{p}\right)\right] \vee \left[\frac{(4\sqrt{2})^{\bar{p}}(2(\bar{p}-1))^{\bar{p}}\bar{C}_3^{\bar{p}}}{\bar{p}^{\bar{p}-2}}\right]\right).\end{aligned}$$

Define the sequence $\bar{p}_j$ by

$$\bar{p}_j = (2 - j + \lfloor\frac{T}{\rho}\rfloor)\bar{p} l_*^{1-j+\lfloor\frac{T}{\rho}\rfloor}, \quad j = 1, 2, \ldots, \lfloor\frac{T}{\rho}\rfloor + 1.$$

For $u \in [0, \rho]$, we derive from (2.9) that

$$\begin{aligned}\mathbb{E}\left(\sup_{0\leq u\leq \rho\wedge\tau_{\tilde{N},(1)}} \left|Y_{(1)}(u)\right|^{\bar{p}_1}\right) &\leq \hat{C}_{T*} + \hat{C}_{T*}(T+1)\mathbb{E}\left(\sup_{-\rho\leq u\leq 0} \left|Y_{(1)}(u)\right|^{\bar{p}_1 l_*}\right) \\ &\leq \hat{C}_{T*} + \hat{C}_{T*}(T+1)\mathbb{E}\|\xi\|^{\bar{p}_1 l_*}.\end{aligned}$$

Then for $u \in [0, 2\rho]$, using (2.9) with Hölder's inequality gives that

$$\mathbb{E}\left(\sup_{0\leq u\leq 2\rho\wedge\tau_{\tilde{N},(1)}} \left|Y_{(1)}(u)\right|^{\bar{p}_2}\right) \leq \hat{C}_{T*} + \hat{C}_{T*}(T+1)[\hat{C}_{T*} + \hat{C}_{T*}(T+1)\mathbb{E}\|\xi\|^{\bar{p}_1 l_*}]^{\frac{\bar{p}_2 l_*}{\bar{p}_1}}.$$

By induction, we derive that

$$\mathbb{E}\left(\sup_{0\leq u\leq\left[\left((\lfloor\frac{T}{\rho}\rfloor+1)\rho\right)\wedge\tau_{\tilde{N},(1)}\right]} \left|Y_{(1)}(u)\right|^{\bar{p}}\right) \leq C_{\hat{C}_{T*},T,\|\xi\|,l_*,\bar{p}_j},$$

where $j = 1, 2, \ldots, \lfloor\frac{T}{\rho}\rfloor + 1$. Then Fatou's lemma leads to

$$\mathbb{E}\left(\sup_{0\leq u\leq\left[(\lfloor\frac{T}{\rho}\rfloor+1)\rho\right]} \left|Y_{(1)}(u)\right|^{\bar{p}}\right) \leq C_{\hat{C}_{T*},T,\|\xi\|,l_*,\bar{p}_j}.$$

Now, we assume that

$$\mathbb{E}\left(\sup_{0\leq t\leq T} \left|Y_{(n-1)}(t)\right|^{\bar{p}}\right) \leq C_{\hat{C}_{T*},T,\|\xi\|,l_*,\bar{p}_j}.$$

Replacing $\left(\Gamma_{(1)}(Y_t), \mathbb{L}_{\Gamma_{(0)}(Y_t)}\right)$ by $\left(\Gamma_{(n)}(Y_t), \mathbb{L}_{\Gamma_{(n-1)}(Y_t)}\right)$ and repeating these procedures give that (2.8) admits a unique solution $Y_{(n)}(t)$ and it has the property that

$$\mathbb{E}\left(\sup_{0\le t\le T}\left|Y_{(n)}(t)\right|^{\bar{p}}\right) \le C_{\hat{C}_{T*},T,\|\xi\|,l_*,\bar{p}_j}.$$

Step 2: We mainly prove that the limits of the sequences $\{Y_{(n)}(\cdot)\}$ and $\{\mathbb{L}_{Y_{(n)}(\cdot)}\}$ (as $n \to \infty$) satisfy $\mathbb{P}$-a.s.

$$d[Y(t) - D(Y(t-\rho))] = \alpha\left(\Gamma(Y_t), \mathbb{L}_{\Gamma(Y_t)}\right) dt + \beta\left(\Gamma(Y_t), \mathbb{L}_{\Gamma(Y_t)}\right) dB(t), \quad t \in [0,T].$$

Then the uniqueness of the solution can be shown. This proof can be completed by using the Step 2 of Theorem 3.8 in [28] and the iterative technique, we omit it.

Step 3: We will give the moment boundedness of the solution $Y(t)$ in this step. Set $\bar{p} \ge 2$ first. For any $t \in [0,T]$, by Itô's formula, we have

$$\begin{aligned}
&|Y(t) - D(Y(t-\rho))|^{\bar{p}} - |\xi(0) - D(\xi(-\rho))|^{\bar{p}} \\
&\le \bar{p}\int_0^t |Y(s) - D(Y(s-\rho))|^{\bar{p}-2}\,(Y(s) - D(Y(s-\rho)))^T \alpha\left(\Gamma(Y_s), \mathbb{L}_{\Gamma(Y_s)}\right) ds \\
&+ \frac{\bar{p}(\bar{p}-1)}{2}\int_0^t |Y(s) - D(Y(s-\rho))|^{\bar{p}-2}\,|\beta\left(\Gamma(Y_s), \mathbb{L}_{\Gamma(Y_s)}\right)|^2 ds \\
&+ \bar{p}\int_0^t |Y(s) - D(Y(s-\rho))|^{\bar{p}-2}\,(Y(s) - D(Y(s-\rho)))^T \beta\left(\Gamma(Y_s), \mathbb{L}_{\Gamma(Y_s)}\right) dB(s) \\
&=: \bar{p}I_1(t) + \frac{\bar{p}(\bar{p}-1)}{2}I_2(t) + \bar{p}I_3(t).
\end{aligned}$$

Define $\tau_{\tilde{N}} = T \wedge inf\{t \in [0,T] : |Y(t)| \ge \tilde{N}\}$ for every integer $\tilde{N} \ge 1$. By (2.5), (2.6), Hölder's inequality and Young's inequality, we derive that

$$\begin{aligned}
&\mathbb{E}\left[\sup_{0\le s\le t\wedge\tau_{\tilde{N}}}(I_1(s) + I_2(s))\right] \\
&\le(\bar{C}_2 + \bar{C}_3)\mathbb{E}\int_0^{t\wedge\tau_{\tilde{N}}} |Y(s) - D(Y(s-\rho))|^{\bar{p}-2} \\
&\quad\cdot\left[1 + \sum_{i=1}^{r-1}|Y(s+\bar{s}_i)|^2 + U_2^2(Y(s-\rho),0)|Y(s-\rho)|^2 + \sum_{i=1}^{r}\mathcal{W}_2^2(\mathbb{L}_{Y(s+\bar{s}_i)})\right] ds \\
&\le(\bar{C}_2 + \bar{C}_3)2^{\bar{p}-2}(1 \vee K_3^{\bar{p}-2})\mathbb{E}\int_0^{t\wedge\tau_{\tilde{N}}}\left(|Y(s)|^{\bar{p}-2} + U_3^{\bar{p}-2}(Y(s-\rho),0)|Y(s-\rho)|^{\bar{p}-2}\right) \\
&\quad\cdot\left[1 + \sum_{i=1}^{r-1}|Y(s+\bar{s}_i)|^2 + U_2^2(Y(s-\rho),0)|Y(s-\rho)|^2 + \sum_{i=1}^{r}\mathcal{W}_2^2(\mathbb{L}_{Y(s+\bar{s}_i)})\right] ds \\
&\le(\bar{C}_2 + \bar{C}_3)2^{\bar{p}-2}(1 \vee K_3^{\bar{p}-2})\left(\frac{p-2}{p} \vee \frac{2}{p}\right)\mathbb{E}\int_0^{t\wedge\tau_{\tilde{N}}}\left[1 + \sum_{i=1}^{r-1}|Y(s+\bar{s}_i)|^{\bar{p}} + \|\xi\|^{\bar{p}}\right. \\
&\quad\left. + \left(U_2(Y(s-\rho),0) \vee U_3(Y(s-\rho),0)\right)^{\bar{p}}|Y(s-\rho)|^{\bar{p}} + \sum_{i=1}^{r}\mathcal{W}_2^p(\mathbb{L}_{Y(s+\bar{s}_i)})\right] ds \\
&\le(\bar{C}_2 + \bar{C}_3)2^{\bar{p}-2}(1 \vee K_3^{\bar{p}-2})\left(\frac{p-2}{p} \vee \frac{2}{p}\right)\mathbb{E}\int_0^{t\wedge\tau_{\tilde{N}}}\left[1 + \|\xi\|^{\bar{p}} + \sum_{i=1}^{r-1}|Y(s+\bar{s}_i)|^{\bar{p}}\right. \\
&\quad\left. + K_U^{\bar{p}}\left(1 + |Y(s-\rho)|^{l_2\vee l_3}\right)^{\bar{p}}|Y(s-\rho)|^{\bar{p}} + \sum_{i=1}^{r}\mathcal{W}_2^p(\mathbb{L}_{Y(s+\bar{s}_i)})\right] ds \\
&\le(\bar{C}_2 + \bar{C}_3)2^{\bar{p}-2}(1 \vee K_3^{\bar{p}-2})\left(\frac{p-2}{p} \vee \frac{2}{p}\right)\int_0^t\left[1 + \mathbb{E}\|\xi\|^{\bar{p}} + (r-1)\mathbb{E}\left(\sup_{0\le u\le s\wedge\tau_{\tilde{N}}}|Y(u)|^{\bar{p}}\right)\right. \\
&\quad\left. + K_U^{\bar{p}}2^{\bar{p}-1}\mathbb{E}\left(\sup_{0\le u\le s\wedge\tau_{\tilde{N}}}|Y(u)|^{\bar{p}}\right) + K_U^{\bar{p}}2^{\bar{p}-1}\mathbb{E}|Y(s-\rho)|^{l_U\bar{p}+\bar{p}} + r\mathbb{E}\left(\sup_{0\le u\le s\wedge\tau_{\tilde{N}}}|Y(u)|^{\bar{p}}\right)\right] ds \\
&\le(\bar{C}_2 + \bar{C}_3)2^{\bar{p}-2}(1 \vee K_3^{\bar{p}-2})[(1 + \mathbb{E}\|\xi\|^{\bar{p}}) \vee (K_U^{\bar{p}}2^{\bar{p}-1} + r)]\left(\frac{p-2}{p} \vee \frac{2}{p}\right) \\
&\quad\cdot\mathbb{E}\int_0^t\left(1 + \sup_{0\le u\le s\wedge\tau_{\tilde{N}}}|Y(u)|^{\bar{p}} + |Y(s-\rho)|^{(l_U+1)\bar{p}}\right) ds.
\end{aligned}$$

Using Young's inequality, Hölder's inequality and BDG's inequality gives that

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{0\le s\le t\wedge\tau_{\tilde N}} I_3(s)\right]\\
&\le 4\sqrt{2}\mathbb{E}\left[\int_0^{t\wedge\tau_{\tilde N}} |Y(s)-D(Y(s-\rho))|^{2\bar p-2}\left|\beta\left(\Gamma(Y_s),\mathbb{L}_{\Gamma(Y_s)}\right)\right|^2 ds\right]^{1/2}\\
&\le \frac{1}{2}\mathbb{E}\left(\sup_{0\le s\le t\wedge\tau_{\tilde N}}|Y(s)-D(Y(s-\rho))|^{\bar p}\right)\\
&+\frac{(4\sqrt{2})^{\bar p}(2(\bar p-1))^{\bar p-1}\bar C_3^{\bar p}}{\bar p^{\bar p-2}}\mathbb{E}\left(\int_0^{t\wedge\tau_{\tilde N}}\left[1+\sum_{i=1}^{r-1}|Y(s+\bar s_i)|^2+U_2^2(Y(s-\rho),0)|Y(s-\rho)|^2\right.\right.\\
&\left.\left.+\sum_{i=1}^{r}\mathcal{W}_2^2(\mathbb{L}_{Y(s+\bar s_i)})\right]ds\right)^{\bar p/2}\\
&\le \frac{1}{2}\mathbb{E}\left(\sup_{0\le s\le t\wedge\tau_{\tilde N}}|Y(s)-D(Y(s-\rho))|^{\bar p}\right)\\
&+\frac{(4\sqrt{2})^{\bar p}(2(\bar p-1))^{\bar p-1}\bar C_3^{\bar p}}{\bar p^{\bar p-2}}[(1+\mathbb{E}\|\xi\|^{\bar p})\vee(K_U^{\bar p}2^{\bar p-1}+r)]\mathbb{E}\int_0^t\left(1+\sup_{0\le u\le s\wedge\tau_{\tilde N}}|Y(u)|^{\bar p}+|Y(s-\rho)|^{(l_U+1)\bar p}\right)ds.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
&\mathbb{E}\left(\sup_{0\le u\le t\wedge\tau_{\tilde N}}|Y(u)|^{\bar p}\right)\\
&\le 2^{\bar p-1}\left[\mathbb{E}\left(\sup_{0\le u\le t\wedge\tau_{\tilde N}}|D(Y(u-\rho))|^{\bar p}\right)+\mathbb{E}\left(\sup_{0\le u\le t\wedge\tau_{\tilde N}}|Y(u)-D(Y(u-\rho))|^{\bar p}\right)\right]\\
&\le \check C_*\left[1+\mathbb{E}\int_0^t\left(\sup_{0\le u\le s\wedge\tau_{\tilde N}}|Y(u)|^{\bar p}\right)ds+\mathbb{E}\int_0^{t\wedge\tau_{\tilde N}}|Y(s-\rho)|^{(l_U+1)\bar p}ds+\mathbb{E}\left(\sup_{0\le u\le t\wedge\tau_{\tilde N}}|Y(u-\rho)|^{(l_U+1)\bar p}\right)\right],
\end{aligned}
$$

where

$$
\begin{aligned}
\check C_* =&\left[(1+\mathbb{E}\|\xi\|^{\bar p})\vee(K_U^{\bar p}2^{\bar p-1}+r)\vee\frac{\bar p(\bar p-1)}{2}\vee\bar p\right]\\
&\cdot\left(\left[(\bar C_2+\bar C_3)2^{\bar p-2}(1\vee K_3^{\bar p-2})\left(\frac{p-2}{p}\vee\frac{2}{p}\right)\right]\vee\left[\frac{(4\sqrt{2})^{\bar p}(2(\bar p-1))^{\bar p}\bar C_3^{\bar p}}{\bar p^{\bar p-2}}\right]\right).
\end{aligned}
$$

Thanks to Gronwall's inequality, we get that

$$
\mathbb{E}\left(\sup_{0\le u\le t\wedge\tau_{\tilde N}}|Y(u)|^{\bar p}\right)\le\check C_{T*}+\check C_{T*}(T+1)\mathbb{E}\left(\sup_{0\le u\le t\wedge\tau_{\tilde N}}|Y(u-\rho)|^{\bar p l_*}\right), \tag{2.10}
$$

where $\check C_{T*}=\check C_* e^{T\check C_*}$ and $l_*=l_U+1$. Define a sequence $\bar p_j$ by

$$
\bar p_j=(2-j+\lfloor\frac{T}{\rho}\rfloor)\bar p l_*^{1-j+\lfloor\frac{T}{\rho}\rfloor},\quad j=1,2,\dots,\lfloor\frac{T}{\rho}\rfloor+1.
$$

One can see that $\bar p_{j+1}l_*<\bar p_j$ for $j=1,2,\dots,\lfloor\frac{T}{\rho}\rfloor+1$ and $\bar p_{\lfloor\frac{T}{\rho}\rfloor+1}=\bar p$. For $u\in[0,\rho]$, (2.10) means that

$$
\mathbb{E}\left(\sup_{0\le u\le\rho\wedge\tau_{\tilde N}}|Y(u)|^{\bar p_1}\right)\le\check C_{T*}+\check C_{T*}(T+1)\mathbb{E}\left(\sup_{-\rho\le u\le 0}|Y(u)|^{\bar p_1 l_*}\right)\le\check C_{T*}+\check C_{T*}(T+1)\mathbb{E}\|\xi\|^{\bar p_1 l_*}.
$$

Then for $u\in[0,2\rho]$, using $\bar p_2 l_*<\bar p_1$ and (2.10) with Hölder's inequality gives that

$$
\begin{aligned}
\mathbb{E}\left(\sup_{0\le u\le 2\rho\wedge\tau_{\tilde N}}|Y(u)|^{\bar p_2}\right)&\le\check C_{T*}+\check C_{T*}(T+1)\mathbb{E}\left(\sup_{0\le u\le 2\rho\wedge\tau_{\tilde N}}|Y(u-\rho)|^{\bar p_2 l_*}\right)\\
&\le\check C_{T*}+\check C_{T*}(T+1)\left[\mathbb{E}\left(\sup_{0\le u\le 2\rho\wedge\tau_{\tilde N}}|Y(u-\rho)|^{\bar p_1}\right)\right]^{\frac{\bar p_2 l_*}{\bar p_1}}\\
&\le\check C_{T*}+\check C_{T*}(T+1)[\check C_{T*}+\check C_{T*}(T+1)\mathbb{E}\|\xi\|^{\bar p_1 l_*}]^{\frac{\bar p_2 l_*}{\bar p_1}}.
\end{aligned}
$$

By induction, we get that

$$\mathbb{E}\left(\sup_{0\le u\le\left[\left((\lfloor\frac{T}{\rho}\rfloor+1)\rho\right)\wedge\tau_{\tilde N}\right]}|Y(u)|^{\bar p}\right)\le C_{\check C_{T*},T,\|\xi\|,l_*,\bar p_j},$$

where $j=1,2,\ldots,\lfloor\frac{T}{\rho}\rfloor+1$. The Fatou lemma leads to

$$\mathbb{E}\left(\sup_{0\le u\le\left[(\lfloor\frac{T}{\rho}\rfloor+1)\rho\right]}|Y(u)|^{\bar p}\right)\le C_{\check C_{T*},T,\|\xi\|,l_*,\bar p_j}.$$

When $\bar p\in(0,2)$, the desired result follows by the Hölder inequality. □

3. Propagation of chaos

In this section, we will use the stochastic particle method in [44,45] to approximate NMSMVE (2.1). For any $i\in\mathbb{S}_N$, let (B^i,ξ^i) be independent copies of (B,ξ) and all (B^i,ξ^i) are i.i.d. Moreover, for $\xi^i\in C^{\check p}_{\mathcal{F}_0}([-\rho,0];\mathbb{R}^d)$ and any $\check p>0$, set $\mathbb{E}\left(\sup_{-\rho\le t_1,t_2\le 0}|\xi^i(t_1)-\xi^i(t_2)|^{\check p}\right)\le K_4|t_1-t_2|^{\check p/2}$. A non-interacting particle system is given by

$$d[Y^i(t)-D(Y^i(t-\rho))]=\alpha\left(\Gamma(Y^i_t),\mathbb{L}_{\Gamma(Y^i_t)}\right)dt+\beta\left(\Gamma(Y^i_t),\mathbb{L}_{\Gamma(Y^i_t)}\right)dB^i_t, \tag{3.1}$$

with the initial value ξ^i, where

$$\Gamma(Y^i_t)=(\Gamma_{\bar s_1}(Y^i_t),\ldots,\Gamma_{\bar s_r}(Y^i_t))=(Y^i(t),\ldots,Y^i(t-\rho)),$$
$$\mathbb{L}_{\Gamma(Y^i_t)}=(\mathbb{L}_{\Gamma_{\bar s_1}(Y^i_t)},\ldots,\mathbb{L}_{\Gamma_{\bar s_r}(Y^i_t)})=(\mathbb{L}_{Y^i(t)},\ldots,\mathbb{L}_{Y^i(t-\rho)}).$$

One can see that $\mathbb{L}_{\Gamma(Y^i_t)}=\mathbb{L}_{\Gamma(Y_t)}$, $i\in\mathbb{S}_N$. To deal with $\mathbb{L}_{\Gamma(Y^i_t)}$, we introduce the following interacting particle system of the form

$$d[Y^{i,N}(t)-D(Y^{i,N}(t-\rho))]=\alpha\left(\Gamma(Y^{i,N}_t),\mathbb{L}_{\Gamma(Y^N_t)}\right)dt+\beta\left(\Gamma(Y^{i,N}_t),\mathbb{L}_{\Gamma(Y^N_t)}\right)dB^i(t), \tag{3.2}$$

with the initial value ξ^i, where

$$\Gamma(Y^{i,N}_t)=(\Gamma_{\bar s_1}(Y^{i,N}_t),\ldots,\Gamma_{\bar s_r}(Y^{i,N}_t))=(Y^{i,N}(t),\ldots,Y^{i,N}(t-\rho)),$$

$$\mathbb{L}_{\Gamma(Y^N_t)}=(\mathbb{L}_{\Gamma_{\bar s_1}(Y^N_t)},\ldots,\mathbb{L}_{\Gamma_{\bar s_r}(Y^N_t)})=(\mathbb{L}_{Y^N(t)},\ldots,\mathbb{L}_{Y^N(t-\rho)}),$$

and

$$\mathbb{L}_{Y^N(t-\bar s_v)}(\cdot):=\frac{1}{N}\sum_{j=1}^N\delta_{Y^{j,N}(t-\bar s_v)}(\cdot),\quad v\in\mathbb{S}_r.$$

In the following of this paper, let $p\ge 2$. The theory of the propagation of chaos is stated as the following theorem.

Theorem 3.1. *Let Assumptions 2.4–2.6 hold and $(pl_U+\varepsilon)p<\varepsilon\bar p$ hold for $\varepsilon\in(0,1]$. Then there exists a constant C independent of N such that, for any $i\in\mathbb{S}_N$,*

$$\mathbb{E}\left(\sup_{0\le t\le T}|Y^i(t)-Y^{i,N}(t)|^p\right)\le C\begin{cases}(N^{-1/2})^{\lambda_{T,\rho,p}}, & \text{if } p>d/2,\\ [N^{-1/2}\log(1+N)]^{\lambda_{T,\rho,p}}, & \text{if } p=d/2,\\ (N^{-p/d})^{\lambda_{T,\rho,p}}, & \text{if } 2\le p<d/2,\end{cases}$$

where $\lambda_{T,\rho,p}=(\frac{p-\varepsilon}{p})^{\lfloor\frac{T}{\rho}\rfloor}$.

Proof. For any $i\in\mathbb{S}_N$ and $t\in[0,T]$, set

$$\Xi^i(t)=Y^i(t)-D(Y^i(t-\rho))-Y^{i,N}(t)+D(Y^{i,N}(t-\rho)).$$

Then using Itô's formula leads to

$$\begin{aligned}&|\Xi^i(t)|^p-|\Xi^i(0)|^p\\ &\le p\int_0^t|\Xi^i(s)|^{p-2}(\Xi^i(s))^T\left[\alpha\left(\Gamma(Y^i_s),\mathbb{L}_{\Gamma(Y^i_s)}\right)-\alpha\left(\Gamma(Y^{i,N}_s),\mathbb{L}_{\Gamma(Y^N_s)}\right)\right]ds\\ &+\frac{p(p-1)}{2}\int_0^t|\Xi^i(s)|^{p-2}\left|\beta\left(\Gamma(Y^i_s),\mathbb{L}_{\Gamma(Y^i_s)}\right)-\beta\left(\Gamma(Y^{i,N}_s),\mathbb{L}_{\Gamma(Y^N_s)}\right)\right|^2ds\\ &+p\int_0^t|\Xi^i(s)|^{p-2}(\Xi^i(s))^T\left[\beta\left(\Gamma(Y^i_s),\mathbb{L}_{\Gamma(Y^i_s)}\right)-\beta\left(\Gamma(Y^{i,N}_s),\mathbb{L}_{\Gamma(Y^N_s)}\right)\right]dB^i(s)\\ &=:J^i_1(t)+J^i_2(t)+J^i_3(t).\end{aligned}$$

For $\varepsilon \in (0,1]$, we get from Hölder's inequality, Young's inequality and Assumptions 2.4, 2.5 that

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{0\le s\le t}(J_1^i(s)+J_2^i(s))\right]\\
\le& C\mathbb{E}\int_0^t \left|\Xi^i(s)\right|^{p-2}\Bigg[\sum_{v=1}^{r-1}|Y^i(s+\bar{s}_v)-Y^{i,N}(s+\bar{s}_v)|^2\\
&+U_2^2(Y^i(s-\rho),Y^{i,N}(s-\rho))|Y^i(s-\rho)-Y^{i,N}(s-\rho)|^2+\sum_{v=1}^{r}\mathbb{W}_2^2(\mathbb{L}_{Y^i(s+\bar{s}_v)},\mathbb{L}_{Y^N(s+\bar{s}_v)})\Bigg]ds\\
\le& C\mathbb{E}\int_0^t|\Xi^i(s)|^pds+C\mathbb{E}\int_0^t\Bigg[\sum_{v=1}^{r-1}|Y^i(s+\bar{s}_v)-Y^{i,N}(s+\bar{s}_v)|^p\\
&+U_2^p(Y^i(s-\rho),Y^{i,N}(s-\rho))|Y^i(s-\rho)-Y^{i,N}(s-\rho)|^p+\sum_{v=1}^{r}\mathbb{W}_p^p(\mathbb{L}_{Y^i(s+\bar{s}_v)},\mathbb{L}_{Y^N(s+\bar{s}_v)})\Bigg]ds\\
\le& C\mathbb{E}\int_0^t\left|\Xi^i(s)\right|^pds+C\int_0^t\mathbb{E}\left(\sup_{0\le u\le s}|Y^i(u)-Y^{i,N}(u)|^p\right)ds\\
&+C\int_0^t\left[\mathbb{E}\left(1+|Y^i(s-\rho)|^{l_Up+\varepsilon}+|Y^{i,N}(s-\rho)|^{l_Up+\varepsilon}\right)^{p/\varepsilon}\right]^{\varepsilon/p}\\
&\quad\cdot\left[\mathbb{E}|Y^i(s-\rho)-Y^{i,N}(s-\rho)|^p\right]^{(p-\varepsilon)/p}ds\\
&+C\mathbb{E}\int_0^t\sum_{v=1}^{r}\mathbb{W}_p^p(\mathbb{L}_{Y^i(s+\bar{s}_v)},\mathbb{L}_{Y^N(s+\bar{s}_v)})ds\\
\le& C\mathbb{E}\int_0^t\left|\Xi^i(s)\right|^pds+C\int_0^t\mathbb{E}\left(\sup_{0\le u\le s}|Y^i(u)-Y^{i,N}(u)|^p\right)ds\\
&+C\int_0^t\left[\mathbb{E}|Y^i(s-\rho)-Y^{i,N}(s-\rho)|^p\right]^{(p-\varepsilon)/p}ds\\
&+C\mathbb{E}\int_0^t\sum_{v=1}^{r}\mathbb{W}_p^p(\mathbb{L}_{Y^i(s+\bar{s}_v)},\mathbb{L}_{Y^N(s+\bar{s}_v)})ds.
\end{aligned}
$$

By Assumption 2.5, BDG's inequality, Young's inequality and Hölder's inequality, we derive that

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{0\le s\le t}J_3^i(s)\right]\\
\le& C\mathbb{E}\left[\int_0^t|\Xi^i(s)|^{2p-2}\left|\beta\left(\Gamma(Y_s^i),\mathbb{L}_{\Gamma(Y_s^i)}\right)-\beta\left(\Gamma(Y_s^{i,N}),\mathbb{L}_{\Gamma(Y_s^N)}\right)\right|^2ds\right]^{1/2}\\
\le& \frac{1}{2}\mathbb{E}\left(\sup_{0\le s\le t}|\Xi^i(s)|^p\right)+C\mathbb{E}\Bigg[\int_0^t\Bigg[\sum_{v=1}^{r-1}|Y^i(s+\bar{s}_v)-Y^{i,N}(s+\bar{s}_v)|^2\\
&\quad+U_2^2(Y^i(s-\rho),Y^{i,N}(s-\rho))|Y^i(s-\rho)-Y^{i,N}(s-\rho)|^2+\sum_{v=1}^{r}\mathbb{W}_2^2(\mathbb{L}_{Y^i(s+\bar{s}_v)},\mathbb{L}_{Y^N(s+\bar{s}_v)})\Bigg]ds\Bigg]^{p/2}\\
\le& \frac{1}{2}\mathbb{E}\left(\sup_{0\le s\le t}|\Xi^i(s)|^p\right)+C\int_0^t\mathbb{E}\left(\sup_{0\le u\le s}|Y^i(u)-Y^{i,N}(u)|^p\right)ds\\
&+C\int_0^t\left[\mathbb{E}|Y^i(s-\rho)-Y^{i,N}(s-\rho)|^p\right]^{(p-\varepsilon)/p}ds+C\mathbb{E}\int_0^t\sum_{v=1}^{r}\mathbb{W}_p^p(\mathbb{L}_{Y^i(s+\bar{s}_v)},\mathbb{L}_{Y^N(s+\bar{s}_v)})ds.
\end{aligned}
$$

Thanks to Gronwall's inequality, we have

$$
\begin{aligned}
\mathbb{E}\left(\sup_{0\le s\le t}|\Xi^i(s)|^p\right)\le& C\int_0^t\mathbb{E}\left(\sup_{0\le u\le s}|Y^i(u)-Y^{i,N}(u)|^p\right)ds\\
&+C\int_0^t\left[\mathbb{E}|Y^i(s-\rho)-Y^{i,N}(s-\rho)|^p\right]^{(p-\varepsilon)/p}ds\\
&+C\mathbb{E}\int_0^t\sum_{v=1}^{r}\mathbb{W}_p^p(\mathbb{L}_{Y^i(s+\bar{s}_v)},\mathbb{L}_{Y^N(s+\bar{s}_v)})ds.
\end{aligned}
$$

Therefore, we get from Assumption 2.6 and the technique in the estimation of $J_1^i(t) + J_2^i(t)$ that

$$\begin{aligned}
&\mathbb{E}\left(\sup_{0\le s\le t} |Y^i(s) - Y^{i,N}(s)|^p\right) \\
\le& C\mathbb{E}\left(\sup_{0\le s\le t} |\Xi^i(s)|^p\right) + C\mathbb{E}\left(\sup_{0\le s\le t} |D(Y^i(s-\rho)) - D(Y^{i,N}(s-\rho))|^p\right) \\
\le& C\int_0^t \mathbb{E}\left(\sup_{0\le u\le s} |Y^i(u) - Y^{i,N}(u)|^p\right) ds + C\left[\mathbb{E}\left(\sup_{0\le u\le t} |Y^i(u-\rho) - Y^{i,N}(u-\rho)|^p\right)\right]^{(p-\varepsilon)/p} \\
&+ C\mathbb{E}\int_0^t \sum_{v=1}^r \mathbb{W}_p^p(\mathbb{L}_{Y^i(s+\bar{s}_v)}, \mathbb{L}_{Y^N(s+\bar{s}_v)})ds.
\end{aligned}$$

Using Gronwall's inequality again yields that

$$\begin{aligned}
&\mathbb{E}\left(\sup_{0\le u\le t} |Y^i(u) - Y^{i,N}(u)|^p\right) \\
\le& C\left[\mathbb{E}\left(\sup_{0\le u\le t} |Y^i(u-\rho) - Y^{i,N}(u-\rho)|^p\right)\right]^{(p-\varepsilon)/p} + C\mathbb{E}\int_0^t \sum_{v=1}^r \mathbb{W}_p^p(\mathbb{L}_{Y^i(s+\bar{s}_v)}, \mathbb{L}_{Y^N(s+\bar{s}_v)})ds.
\end{aligned} \tag{3.3}$$

For $u \in [0, \rho]$, we get from (3.3) that

$$\mathbb{E}\left(\sup_{0\le u\le \rho} |Y^i(u) - Y^{i,N}(u)|^p\right) \le C\mathbb{E}\int_0^\rho \sum_{v=1}^r \mathbb{W}_p^p(\mathbb{L}_{Y^i(s+\bar{s}_v)}, \mathbb{L}_{Y^N(s+\bar{s}_v)})ds.$$

To deal with the Wasserstein distance, for $v \in \mathbb{S}_r$ and $t \in [0, T]$, we give the definition $\mathbb{L}_{Y^{*,N}(t+\bar{s}_v)}(\cdot)$ by

$$\mathbb{L}_{Y^{*,N}(t+\bar{s}_v)}(\cdot) = \frac{1}{N}\sum_{j=1}^N \delta_{Y^j(t+\bar{s}_v)}(\cdot).$$

One can observe that, for $v \in \mathbb{S}_r$ and $s \in [0, \rho]$,

$$\begin{aligned}
&\mathbb{W}_p^p(\mathbb{L}_{Y^i(s+\bar{s}_v)}, \mathbb{L}_{Y^N(s+\bar{s}_v)}) \\
&\le C\mathbb{W}_p^p(\mathbb{L}_{Y^i(s+\bar{s}_v)}, \mathbb{L}_{Y^{*,N}(s+\bar{s}_v)}) + \mathbb{W}_p^p(\mathbb{L}_{Y^{*,N}(s+\bar{s}_v)}, \mathbb{L}_{Y^N(s+\bar{s}_v)}) \\
&\le C\mathbb{W}_p^p(\mathbb{L}_{Y^i(s+\bar{s}_v)}, \mathbb{L}_{Y^{*,N}(s+\bar{s}_v)}) + C\frac{1}{N}\sum_{j=1}^N \left|Y^j(s+\bar{s}_v) - Y^{j,N}(s+\bar{s}_v)\right|^p.
\end{aligned}$$

Since all j are identically distributed, we have

$$\mathbb{E}\left(\frac{1}{N}\sum_{j=1}^N \left|Y^j(s+\bar{s}_v) - Y^{j,N}(s+\bar{s}_v)\right|^p\right) = \mathbb{E}\left|Y^i(s+\bar{s}_v) - Y^{i,N}(s+\bar{s}_v)\right|^p.$$

Thus,

$$\begin{aligned}
&\mathbb{E}\left(\sup_{0\le u\le \rho} |Y^i(u) - Y^{i,N}(u)|^p\right) \\
&\le C\mathbb{E}\int_0^\rho \sum_{v=1}^r \left|Y^i(s+\bar{s}_v) - Y^{i,N}(s+\bar{s}_v)\right|^p ds + C\mathbb{E}\int_0^\rho \mathbb{W}_p^p(\mathbb{L}_{Y^i(s+\bar{s}_v)}, \mathbb{L}_{Y^{*,N}(s+\bar{s}_v)})ds \\
&\le C\int_0^\rho \mathbb{E}\left(\sup_{0\le u\le s} |Y^i(u) - Y^{i,N}(u)|^p\right) ds + C\mathbb{E}\int_0^\rho \mathbb{W}_p^p(\mathbb{L}_{Y^i(s+\bar{s}_v)}, \mathbb{L}_{Y^{*,N}(s+\bar{s}_v)})ds.
\end{aligned}$$

Applying Gronwall's inequality and Corollary 2.3 yields that

$$\mathbb{E}\left(\sup_{0\le u\le \rho} |Y^i(u) - Y^{i,N}(u)|^p\right) \le C\begin{cases} N^{-1/2}, & \text{if } p > d/2, \\ N^{-1/2}\log(1+N), & \text{if } p = d/2, \\ N^{-p/d}, & \text{if } 2 \le p < d/2. \end{cases}$$

For $u \in [0, 2\rho]$, using Hölder's inequality and (3.3) gives that

$$
\begin{aligned}
&\mathbb{E}\left(\sup_{0\le u\le 2\rho}|Y^i(u)-Y^{i,N}(u)|^p\right)\\
&\le C\left[\mathbb{E}\left(\sup_{0\le u\le 2\rho}|Y^i(u-\rho)-Y^{i,N}(u-\rho)|^p\right)\right]^{(p-\varepsilon)/p}+C\mathbb{E}\int_0^{2\rho}\sum_{v=1}^{r}\mathbb{W}_p^p(\mathbb{L}_{Y^i(s+\bar{s}_v)},\mathbb{L}_{Y^N(s+\bar{s}_v)})ds\\
&\le C\left[\mathbb{E}\left(\sup_{0\le u\le \rho}|Y^i(u)-Y^{i,N}(u)|^p\right)\right]^{(p-\varepsilon)/p}+C\mathbb{E}\int_0^{2\rho}\sum_{v=1}^{r}\mathbb{W}_p^p(\mathbb{L}_{Y^i(s+\bar{s}_v)},\mathbb{L}_{Y^{*,N}(s+\bar{s}_v)})ds\\
&+C\int_0^{2\rho}\mathbb{E}\left(\sup_{0\le u\le s}|Y^i(u)-Y^{i,N}(u)|^p\right)ds.
\end{aligned}
$$

The Gronwall inequality means that

$$
\begin{aligned}
&\mathbb{E}\left(\sup_{0\le u\le 2\rho}|Y^i(u)-Y^{i,N}(u)|^p\right)\\
&\le C\left[\mathbb{E}\left(\sup_{0\le u\le \rho}|Y^i(u)-Y^{i,N}(u)|^p\right)\right]^{(p-\varepsilon)/p}ds+C\mathbb{E}\int_0^{2\rho}\sum_{v=1}^{r}\mathbb{W}_p^p(\mathbb{L}_{Y^i(s+\bar{s}_v)},\mathbb{L}_{Y^{*,N}(s+\bar{s}_v)})ds\\
&\le C\begin{cases}(N^{-1/2})^{(p-\varepsilon)/p}, & \text{if } p>d/2,\\ [N^{-1/2}\log(1+N)]^{(p-\varepsilon)/p}, & \text{if } p=d/2,\\ (N^{-p/d})^{(p-\varepsilon)/p}, & \text{if } 2\le p<d/2.\end{cases}
\end{aligned}
$$

For $u \in [0, 3\rho]$, we can similarly get that

$$
\begin{aligned}
&\mathbb{E}\left(\sup_{0\le u\le 3\rho}|Y^i(u)-Y^{i,N}(u)|^p\right)\\
&\le C\left[\mathbb{E}\left(\sup_{0\le u\le 2\rho}|Y^i(u)-Y^{i,N}(u)|^p\right)\right]^{(p-\varepsilon)/p}ds+C\mathbb{E}\int_0^{3\rho}\sum_{v=1}^{r}\mathbb{W}_p^p(\mathbb{L}_{Y^i(s+\bar{s}_v)},\mathbb{L}_{Y^{*,N}(s+\bar{s}_v)})ds\\
&\le C\begin{cases}(N^{-1/2})^{(\frac{p-\varepsilon}{p})^2}, & \text{if } p>d/2,\\ [N^{-1/2}\log(1+N)]^{(\frac{p-\varepsilon}{p})^2}, & \text{if } p=d/2,\\ (N^{-p/d})^{(\frac{p-\varepsilon}{p})^2}, & \text{if } 2\le p<d/2.\end{cases}
\end{aligned}
$$

Repeating the same procedures, we obtain that

$$
\mathbb{E}\left(\sup_{0\le u\le[(\lfloor\frac{T}{\rho}\rfloor+1)\rho]}|Y^i(u)-Y^{i,N}(u)|^p\right)\le C\begin{cases}(N^{-1/2})^{\lambda_{T,\rho,p}}, & \text{if } p>d/2,\\ [N^{-1/2}\log(1+N)]^{\lambda_{T,\rho,p}}, & \text{if } p=d/2,\\ (N^{-p/d})^{\lambda_{T,\rho,p}}, & \text{if } 2\le p<d/2,\end{cases}
$$

where $\lambda_{T,\rho,p} = \left(\frac{p-\varepsilon}{p}\right)^{\lfloor\frac{T}{\rho}\rfloor}$ for $\varepsilon \in (0, 1]$. □

Remark 3. From Theorem 3.1, we know that the value of ε influences the rate of convergence of $Y^{i,N}(\cdot)$ to $Y^i(\cdot)$. If the value of ε is close to 0, the convergence rate will become larger but $\bar{p}$ needs to be relatively large to make $(l_U p + \varepsilon)p < \varepsilon\bar{p}$ hold. On the contrary, if the value of ε is close to 1, the requirement for $\bar{p}$ is not strict but the convergence rate will become smaller.

If we impose stronger conditions on delay components, the following theorem reveals the corresponding propagation of chaos.

Theorem 3.2. *Let Assumptions 2.4–2.6 hold with $U_2(x_r, y_r) = 1$, $K_3U_3(x_r, y_r) = K_D \in (0, 1)$. Then there exists a constant C independent of N such that, for any $i \in \mathbb{S}_N$ and $2 \le p < \bar{p}$,*

$$
\mathbb{E}\left(\sup_{0\le t\le T}|Y^i(t)-Y^{i,N}(t)|^p\right)\le C\begin{cases}N^{-1/2}+N^{-(\bar{p}-p)/\bar{p}}, & \text{if } p>d/2 \text{ and } \bar{p}\ne 2p,\\ N^{-1/2}\log(1+N)+N^{-(\bar{p}-p)/\bar{p}}, & \text{if } p=d/2 \text{ and } \bar{p}\ne 2p,\\ N^{-p/d}+N^{-(\bar{p}-p)/\bar{p}}, & \text{if } 2\le p<d/2.\end{cases}
$$

Proof. Let the notations in this proof be the same as these in Theorem 3.1. We only show the main differences of the proof but omit the same procedures. From the proof of Theorem 3.1, we know that

$$\begin{aligned}&\mathbb{E}\left[\sup_{0\le s\le t}(J_1^i(s)+J_2^i(s))\right]\\ \le& C\mathbb{E}\int_0^t\left|\Xi^i(s)\right|^{p-2}\sum_{v=1}^{r}\left[|Y^i(s+\bar{s}_v)-Y^{i,N}(s+\bar{s}_v)|^2+\mathbb{W}_2^2(\mathbb{L}_{Y^i(s+\bar{s}_v)},\mathbb{L}_{Y^N(s+\bar{s}_v)})\right]ds\\ \le& C\int_0^t\mathbb{E}|\Xi^i(s)|^pds+C\int_0^t\mathbb{E}\left(\sup_{0\le u\le s}|Y^i(u)-Y^{i,N}(u)|^p\right)ds+C\mathbb{E}\int_0^t\sum_{v=1}^{r}\mathbb{W}_p^p(\mathbb{L}_{Y^i(s+\bar{s}_v)},\mathbb{L}_{Y^N(s+\bar{s}_v)})ds.\end{aligned}$$

This result with the estimation of $J_3^i(t)$ in the proof of Theorem 3.1 gives that

$$\begin{aligned}&\mathbb{E}\left(\sup_{0\le s\le t}|\Xi^i(s)|^p\right)\\ \le& C\int_0^t\mathbb{E}\left(\sup_{0\le u\le s}|Y^i(u)-Y^{i,N}(u)|^p\right)ds+C\mathbb{E}\int_0^t\sum_{v=1}^{r}\mathbb{W}_p^p(\mathbb{L}_{Y^i(s+\bar{s}_v)},\mathbb{L}_{Y^N(s+\bar{s}_v)})ds,\end{aligned}\tag{3.4}$$

where the Gronwall inequality has been used. Recall the elementary inequality

$$|a+b|^p\le\frac{1}{K_D^{p-1}}|a|^p+\frac{1}{(1-K_D)^{p-1}}|b|^p,\tag{3.5}$$

for $p\ge 2$, $0<K_D<1$ and $a,b\in\mathbb{R}^d$. For any $t\in[0,T]$, using (3.5) and Assumption 2.6 leads to

$$\begin{aligned}&|Y^i(t)-Y^{i,N}(t)|^p\\ =&\left|\Xi^i(s)+D(Y^i(t-\rho))-D(Y^{i,N}(t-\rho))\right|^p\\ \le&\frac{1}{K_D^{p-1}}\left|D(Y^i(t-\rho))-D(Y^{i,N}(t-\rho))\right|^p+\frac{1}{(1-K_D)^{p-1}}|\Xi^i(s)|^p\\ \le& K_D\left|Y^i(t-\rho)-Y^{i,N}(t-\rho)\right|^p+\frac{1}{(1-K_D)^{p-1}}|\Xi^i(s)|^p.\end{aligned}$$

Thus,

$$\begin{aligned}&\sup_{0\le s\le t}|Y^i(s)-Y^{i,N}(s)|^p\\ \le& K_D\sup_{0\le s\le t}\left|Y^i(s-\rho)-Y^{i,N}(s-\rho)\right|^p+\frac{1}{(1-K_D)^{p-1}}\sup_{0\le s\le t}|\Xi^i(s)|^p\\ \le&\frac{1}{(1-K_D)^p}\sup_{0\le s\le t}|\Xi^i(s)|^p.\end{aligned}$$

By the above inequality, (3.4), Gronwall's inequality and the technique in the proof of Theorem 3.1, we have

$$\begin{aligned}&\mathbb{E}\left(\sup_{0\le s\le t}|Y^i(s)-Y^{i,N}(s)|^p\right)\\ &\le C\mathbb{E}\int_0^t\sum_{v=1}^{r}\mathbb{W}_p^p(\mathbb{L}_{Y^i(s+\bar{s}_v)},\mathbb{L}_{Y^N(s+\bar{s}_v)})ds\\ &\le C\int_0^t\mathbb{E}\left(\sup_{0\le u\le s}|Y^i(u)-Y^{i,N}(u)|^p\right)ds+C\mathbb{E}\int_0^t\sum_{v=1}^{r}\mathbb{W}_p^p(\mathbb{L}_{Y^i(s+\bar{s}_v)},\mathbb{L}_{Y^{*,N}(s+\bar{s}_v)})ds.\end{aligned}$$

Applying the Gronwall inequality and Theorem 2.2 leads to

$$\begin{aligned}&\mathbb{E}\left(\sup_{0\le s\le t}|Y^i(s)-Y^{i,N}(s)|^p\right)\\ &\le C\mathbb{E}\int_0^t\sum_{v=1}^{r}\mathbb{W}_p^p(\mathbb{L}_{Y^i(s+\bar{s}_v)},\mathbb{L}_{Y^{*,N}(s+\bar{s}_v)})ds\\ &\le C\begin{cases}N^{-1/2}+N^{-(\bar{p}-p)/\bar{p}}, & \text{if } p>d/2 \text{ and } \bar{p}\ne 2p,\\ N^{-1/2}\log(1+N)+N^{-(\bar{p}-p)/\bar{p}}, & \text{if } p=d/2 \text{ and } \bar{p}\ne 2p,\\ N^{-p/d}+N^{-(\bar{p}-p)/\bar{p}}, & \text{if } 2\le p<d/2.\end{cases}\end{aligned}$$

□

Remark 4. By comparing Theorems 3.1 and 3.2, we can find that: to relax the constraints of the delay term and neutral term, we introduce the functions $U_2(\cdot,\cdot)$ and $U_3(\cdot,\cdot)$, which makes the convergence rate of $Y^{i,N}(\cdot)$ to $Y^i(\cdot)$ less ideal.

Remark 5. The techniques in Theorems 3.1 and 3.2 can be used in the theory about the propagation of chaos in $\mathcal{L}^p$ sense.

4. Numerical scheme for NMSMVEs

In this section, we establish the tamed EM scheme for superlinear NMSMVE. Firstly, the moment boundedness of the numerical solution is analyzed. Then the strong convergence rate is obtained by using propagation of chaos. To investigate the tamed EM scheme for (3.2), for $\Delta \in (0,1)$ and $\gamma \in (0,1/2]$, define

$$\alpha_\Delta(\Gamma(\varphi),\mathbb{L}_{\Gamma(\varphi)}) = \frac{\alpha(\Gamma(\varphi),\mathbb{L}_{\Gamma(\varphi)})}{1+\Delta^\gamma|\alpha(\Gamma(\varphi),\mathbb{L}_{\Gamma(\varphi)})|}, \tag{4.1}$$

where $\Gamma(\varphi) \in (\mathbb{R}^d)^r$, $\mathbb{L}_{\Gamma(\varphi)} \in (\mathcal{P}_2(\mathbb{R}^d))^r$. Assume that there exist two positive integers M and M_T such that $\Delta = \frac{\rho}{M} = \frac{T}{M_T}$ and other positive integers $\bar{k}_2, \bar{k}_3, \dots, \bar{k}_{r-1}$ such that $\frac{-\bar{s}_2}{\bar{k}_2} = \frac{-\bar{s}_3}{\bar{k}_3} = \cdots = \frac{-\bar{s}_{r-1}}{\bar{k}_{r-1}} = \Delta$. Set $t_k = k\Delta$, $k = -M, \dots, 0, 1, \dots, M_T$. Define the tamed EM scheme as:

$$\begin{cases} Z^{i,N}(t_k) = \xi^i(t_k), \quad k = -M, -M+1, \dots, 0, \\ Z^{i,N}(t_{k+1}) - D(Z^{i,N}(t_{k+1-M})) = Z^{i,N}(t_k) - D(Z^{i,N}(t_{k-M})) + \alpha_\Delta(\Gamma(Z^{i,N}_{t_k}), \mathbb{L}_{\Gamma(Z^N_{t_k})})\Delta \\ \qquad\qquad + \beta(\Gamma(Z^{i,N}_{t_k}), \mathbb{L}_{\Gamma(Z^N_{t_k})})\Delta B^i_k, \quad k = 0, 1, \dots, M_T - 1, \end{cases}$$

where

$$\Delta B^i_k = B^i(t_{k+1}) - B^i(t_k),$$

$$\Gamma(Z^{i,N}_{t_k}) = \left(Z^{i,N}(t_k), Z^{i,N}(t_{k-\bar{k}_2}), \dots, Z^{i,N}(t_{k-M})\right),$$

$$\mathbb{L}_{\Gamma(Z^N_{t_k})} = \left(\mathbb{L}_{Z^N(t_k)}, \mathbb{L}_{Z^N(t_{k-\bar{k}_2})}, \dots, \mathbb{L}_{Z^N(t_{k-M})}\right),$$

and

$$\mathbb{L}_{Z^N(t_{k-\bar{k}_v})}(\cdot) = \frac{1}{N}\sum_{j=1}^N \delta_{Z^{j,N}(t_{k-\bar{k}_v})}(\cdot), \quad v \in \mathbb{S}_r.$$

Remark 6. In this discrete-time numerical scheme, we use $\left(Z^{i,N}(t_k), Z^{i,N}(t_{k-\bar{k}_2}), \dots, Z^{i,N}(t_{k-M})\right)$ to approximate $\left(Y^{i,N}(t_k), Y^{i,N}(t_{k-\bar{k}_2}), \dots, Y^{i,N}(t_{k-M})\right)$ for any $i \in \mathbb{S}_N$. This can be achieved in the numerical simulation. The reason is that: for the largest delay constant ρ, we have assumed that there exist a positive integer M such that $\Delta = \frac{\rho}{M}$, which means that there exist $M-1$ time grids to divide the delay interval into M parts. Then other delays may be at these time grids, otherwise, we can make M larger to achieve this goal. Ideally, we have $\frac{-\bar{s}_2}{\bar{k}_2} = \frac{-\bar{s}_3}{\bar{k}_3} = \cdots = \frac{-\bar{s}_{r-1}}{\bar{k}_{r-1}} = \frac{\rho}{M} = \Delta$ for the most appropriate M.

The continuous-time step process numerical solution on $t \in [-\rho, T]$ is defined by

$$Z^{i,N}(t) = \sum_{k=-M}^{M_T} Z^{i,N}(t_k)\mathbb{I}_{[t_k,t_{k+1})}(t).$$

For $t \in [0,T]$, the continuous-sample numerical scheme is defined by

$$\begin{aligned} &\bar{Z}^{i,N}(t) - D(\bar{Z}^{i,N}(t-\rho)) \\ &= \xi^i(0) - D(\xi^i(-\rho)) + \int_0^t \alpha_\Delta(\Gamma(Z^{i,N}_s), \mathbb{L}_{\Gamma(Z^N_s)})ds + \int_0^t \beta(\Gamma(Z^{i,N}_s), \mathbb{L}_{\Gamma(Z^N_s)})dB^i(s), \end{aligned} \tag{4.2}$$

where $\Gamma(Z^{i,N}_t) = \Gamma(Z^{i,N}_{t_k})$ for $t \in [t_k, t_{k+1})$ and $\mathbb{L}_{\Gamma(Z^N_t)}(\cdot) = \frac{1}{N}\sum_{j=1}^N \delta_{\Gamma(Z^{j,N}_t)}(\cdot)$. We can find that $\mathbb{L}_{\Gamma(Z^N_t)} = \mathbb{L}_{\Gamma(Z^N_{t_k})}$ for any $t \in [t_k, t_{k+1})$. Similar to [28,29], we know that $\bar{Z}^{i,N}(t_k) = Z^{i,N}(t_k) = Z^{i,N}(t)$ for any $t \in [t_k, t_{k+1})$. We get from (4.1) that

$$|\alpha_\Delta(\Gamma(\varphi), \mathbb{L}_{\Gamma(\varphi)})| \le \Delta^{-\gamma} \wedge |\alpha(\Gamma(\varphi), \mathbb{L}_{\Gamma(\varphi)})|, \tag{4.3}$$

for any $\Gamma(\varphi) \in (\mathbb{R}^d)^r$, $\mathbb{L}_{\Gamma(\varphi)} \in (\mathcal{P}_2(\mathbb{R}^d))^r$. By (2.5), we derive that

$$(x_1 - D(x_r))^T \alpha_\Delta(x^{(r)}, \mu^{(r)}) \le C\left[1 + \sum_{i=1}^{r-1}|x_i|^2 + U_2^2(x_r, 0)|x_r|^2 + \sum_{i=1}^r \mathcal{W}_2^2(\mu_i)\right], \tag{4.4}$$

for any $x^{(r)} \in (\mathbb{R}^d)^r$ and $\mu^{(r)} \in (\mathcal{P}_2(\mathbb{R}^d))^r$. For simplicity, define $t_\Delta = \lfloor \frac{t}{\Delta} \rfloor \Delta$ for any $t \in [-\rho, T]$.

Lemma 4.1. *Let Assumptions 2.4–2.6 hold. Then for any $\bar{p} > 0$, we have*

$$\max_{i\in\mathbb{S}_N} \sup_{0<\Delta\le 1} \sup_{0\le t\le T} \mathbb{E}|\bar{Z}^{i,N}(t)|^{\bar{p}} \le C, \quad \forall T > 0.$$

Proof. Let $\tilde{p} \geq 4$ first. Applying Itô's formula gives that

$$
\begin{aligned}
&\left|\bar{Z}^{i,N}(t) - D(\bar{Z}^{i,N}(t-\rho))\right|^{\tilde{p}} - \left|\xi^i(0) - D(\xi^i(-\rho))\right|^{\tilde{p}} \\
&\leq \tilde{p}\int_0^t \left|\bar{Z}^{i,N}(s) - D(\bar{Z}^{i,N}(s-\rho))\right|^{\tilde{p}-2}\left(\bar{Z}^{i,N}(s) - D(\bar{Z}^{i,N}(s-\rho))\right)^T \alpha_\Delta\left(\Gamma(Z_s^{i,N}), \mathbb{L}_{\Gamma(Z_s^N)}\right)ds \\
&+ \frac{\tilde{p}(\tilde{p}-1)}{2}\int_0^t \left|\bar{Z}^{i,N}(s) - D(\bar{Z}^{i,N}(s-\rho))\right|^{\tilde{p}-2}\left|\beta\left(\Gamma(Z_s^{i,N}), \mathbb{L}_{\Gamma(Z_s^N)}\right)\right|^2 ds \\
&+ \tilde{p}\int_0^t |\bar{Z}^{i,N}(s) - D(\bar{Z}^{i,N}(s-\rho))|^{\tilde{p}-2}(\bar{Z}^{i,N}(s) - D(\bar{Z}^{i,N}(s-\rho)))^T \beta\left(\Gamma(Z_s^{i,N}), \mathbb{L}_{\Gamma(Z_s^N)}\right)dB^i(s) \\
&=: Q_1^{i,N}(t) + Q_2^{i,N}(t) + Q_3^{i,N}(t).
\end{aligned}
$$

Moreover,

$$
\begin{aligned}
&Q_1^{i,N}(t) \\
&= C\int_0^t \left|\bar{Z}^{i,N}(s) - D(\bar{Z}^{i,N}(s-\rho))\right|^{\tilde{p}-2} \\
&\quad \cdot \left(\bar{Z}^{i,N}(s) - D(\bar{Z}^{i,N}(s-\rho)) - Z^{i,N}(s) + D(Z^{i,N}(s-\rho))\right)^T \alpha_\Delta\left(\Gamma(Z_s^{i,N}), \mathbb{L}_{\Gamma(Z_s^N)}\right)ds \\
&+ C\int_0^t \left|\bar{Z}^{i,N}(s) - D(\bar{Z}^{i,N}(s-\rho))\right|^{\tilde{p}-2}\left(Z^{i,N}(s) - D(Z^{i,N}(s-\rho))\right)^T \alpha_\Delta\left(\Gamma(Z_s^{i,N}), \mathbb{L}_{\Gamma(Z_s^N)}\right)ds. \\
&=: Q_{11}^{i,N}(t) + Q_{12}^{i,N}(t).
\end{aligned}
$$

By (4.2), Young's inequality, Hölder's inequality and Assumptions 2.4, 2.5, we have

$$
\begin{aligned}
&\mathbb{E}Q_{11}^{i,N}(t) \\
&\leq C\mathbb{E}\int_0^t \left|\bar{Z}^{i,N}(s) - D(\bar{Z}^{i,N}(s-\rho))\right|^{\tilde{p}-2}\Bigg|\int_{s_\Delta}^s \alpha_\Delta\left(\Gamma(Z_u^{i,N}), \mathbb{L}_{\Gamma(Z_u^N)}\right)du \\
&\quad + \int_{s_\Delta}^s \beta\left(\Gamma(Z_u^{i,N}), \mathbb{L}_{\Gamma(Z_u^N)}\right)dB^i(u)\Bigg|\left|\alpha_\Delta\left(\Gamma(Z_s^{i,N}), \mathbb{L}_{\Gamma(Z_s^N)}\right)\right|ds. \\
&\leq C\mathbb{E}\int_0^t \left|\bar{Z}^{i,N}(s) - D(\bar{Z}^{i,N}(s-\rho))\right|^{\tilde{p}}ds \\
&\quad + C\Delta^{-\gamma\tilde{p}/2}\mathbb{E}\int_0^t \left(\left|\int_{s_\Delta}^s \alpha_\Delta\left(\Gamma(Z_u^{i,N}), \mathbb{L}_{\Gamma(Z_u^N)}\right)du\right|^{\tilde{p}/2} + \left|\int_{s_\Delta}^s \beta\left(\Gamma(Z_u^{i,N}), \mathbb{L}_{\Gamma(Z_u^N)}\right)dB^i(u)\right|^{\tilde{p}/2}\right)ds \\
&\leq C\mathbb{E}\int_0^t \left|\bar{Z}^{i,N}(s) - D(\bar{Z}^{i,N}(s-\rho))\right|^{\tilde{p}}ds \\
&\quad + C\left(\Delta^{(1/2-\gamma)\tilde{p}} + \Delta^{-\gamma\tilde{p}/2}\mathbb{E}\int_0^t \left|\int_{s_\Delta}^s \beta\left(\Gamma(Z_u^{i,N}), \mathbb{L}_{\Gamma(Z_u^N)}\right)dB^i(u)\right|^{\tilde{p}/2}ds\right) \\
&\leq C\mathbb{E}\int_0^t \left|\bar{Z}^{i,N}(s) - D(\bar{Z}^{i,N}(s-\rho))\right|^{\tilde{p}}ds \\
&\quad + C\Delta^{(1/2-\gamma)\tilde{p}} + C\Delta^{-\gamma\tilde{p}/2}\int_0^t \mathbb{E}\left(\int_{s_\Delta}^s \left|\beta\left(\Gamma(Z_u^{i,N}), \mathbb{L}_{\Gamma(Z_u^N)}\right)\right|^2 du\right)^{\tilde{p}/4}ds \\
&\leq C\mathbb{E}\int_0^t \left|\bar{Z}^{i,N}(s) - D(\bar{Z}^{i,N}(s-\rho))\right|^{\tilde{p}}ds + C\Delta^{(1/2-\gamma)\tilde{p}} \\
&\quad + C\Delta^{(1/2-\gamma)\tilde{p}/2}\int_0^t \mathbb{E}\Bigg[1 + \sum_{v=1}^{r-1}|Z^{i,N}(s+\bar{s}_v)|^{\tilde{p}/2} \\
&\quad + U_2^{\tilde{p}/2}(Z^{i,N}(s-\rho), 0)|Z^{i,N}(s-\rho)|^{\tilde{p}/2} + \sum_{i=1}^r \mathcal{W}_2^{\tilde{p}/2}(\mathbb{L}_{Z^N(s+\bar{s}_v)})\Bigg]ds \\
&\leq C\mathbb{E}\int_0^t \left|\bar{Z}^{i,N}(s) - D(\bar{Z}^{i,N}(s-\rho))\right|^{\tilde{p}}ds \\
&\quad + C\left(1 + \int_0^t \mathbb{E}\left[\sum_{v=1}^{r-1}|Z^{i,N}(s+\bar{s}_v)|^{\tilde{p}} + |Z^{i,N}(s-\rho)|^{(l_U+1)\tilde{p}} + \sum_{i=1}^r \mathcal{W}_{\tilde{p}}^{\tilde{p}}(\mathbb{L}_{Z^N(s+\bar{s}_v)})\right]ds\right) \\
&\leq C\mathbb{E}\int_0^t \left|\bar{Z}^{i,N}(s) - D(\bar{Z}^{i,N}(s-\rho))\right|^{\tilde{p}}ds + C\left(1 + \int_0^t \sup_{0\leq u\leq s}\mathbb{E}\left|\bar{Z}^{i,N}(u)\right|^{\tilde{p}}ds\right) \\
&\quad + C\int_0^t \mathbb{E}|Z^{i,N}(s-\rho)|^{(l_U+1)\tilde{p}}ds.
\end{aligned}
$$

By Hölder's inequality, Young's inequality, (4.4) and Assumptions 2.4–2.6, one can see that

$$
\begin{aligned}
&\mathbb{E}\left(Q_{12}^{i,N}(t)+Q_{2}^{i,N}(t)\right)\\
&\leq C\mathbb{E}\int_0^t \left|\bar{Z}^{i,N}(s)-D(\bar{Z}^{i,N}(s-\rho))\right|^{\tilde{p}-2}\\
&\quad\cdot\left[1+\sum_{v=1}^{r-1}|Z^{i,N}(s+\bar{s}_v)|^2 +U_2^2(Z^{i,N}(s-\rho),0)|Z^{i,N}(s-\rho)|^2+\sum_{i=1}^{r}\mathcal{W}_2^2(\mathbb{L}_{Z^N(s+\bar{s}_v)})\right]ds\\
&\leq C\mathbb{E}\int_0^t \left[1+|Z^{i,N}(s)|^{\tilde{p}}+U_3^{\tilde{p}}(\bar{Z}^{i,N}(s-\rho),0)|\bar{Z}^{i,N}(s-\rho)|^{\tilde{p}}+\sum_{v=1}^{r-1}|Z^{i,N}(s+\bar{s}_v)|^{\tilde{p}}\right.\\
&\quad\left.+U_2^{\tilde{p}}(Z^{i,N}(s-\rho),0)|Z^{i,N}(s-\rho)|^{\tilde{p}}+\sum_{v=1}^{r}\mathcal{W}_{\tilde{p}}^{\tilde{p}}(\mathbb{L}_{Z^N(s+\bar{s}_v)})\right]ds\\
&\leq C\mathbb{E}\int_0^t \left[1+|Z^{i,N}(s)|^{\tilde{p}}+U_3^{2\tilde{p}}(\bar{Z}^{i,N}(s-\rho),0)+|\bar{Z}^{i,N}(s-\rho)|^{2\tilde{p}}\right.\\
&\quad\left.+\sum_{v=1}^{r-1}|Z^{i,N}(s+\bar{s}_v)|^{\tilde{p}}+U_2^{2\tilde{p}}(Z^{i,N}(s-\rho),0)+|Z^{i,N}(s-\rho)|^{2\tilde{p}}+\sum_{v=1}^{r}\mathcal{W}_{\tilde{p}}^{\tilde{p}}(\mathbb{L}_{Z^N(s+\bar{s}_v)})\right]ds\\
&\leq C\int_0^t\left[1+\sup_{0\leq u\leq s}\mathbb{E}\left|\bar{Z}^{i,N}(u)\right|^{\tilde{p}}+|\bar{Z}^{i,N}(s-\rho)|^{2\tilde{p}l_U}+|\bar{Z}^{i,N}(s-\rho)|^{2\tilde{p}}\right.\\
&\quad\left.+|Z^{i,N}(s-\rho)|^{2\tilde{p}l_U}+|Z^{i,N}(s-\rho)|^{2\tilde{p}}\right]ds\\
&\leq C\int_0^t\left[1+\sup_{0\leq u\leq s}\mathbb{E}\left|\bar{Z}^{i,N}(u)\right|^{\tilde{p}}\right]ds+C\int_0^t\mathbb{E}\left(|\bar{Z}^{i,N}(s-\rho)|^{\tilde{p}l_U^*}+Z^{i,N}(s-\rho)|^{\tilde{p}l_U^*}\right)ds.
\end{aligned}
$$

where $l_U^*=2l_U+2$. Combining these inequalities with the Gronwall inequality leads to

$$
\begin{aligned}
&\sup_{0\leq s\leq t}\mathbb{E}\left|\bar{Z}^{i,N}(s)-D(\bar{Z}^{i,N}(s-\rho))\right|^{\tilde{p}}\\
&\leq C\left(1+\int_0^t\sup_{0\leq u\leq s}\mathbb{E}|\bar{Z}^{i,N}(u)|^{\tilde{p}}ds+\int_0^t\mathbb{E}\left(|\bar{Z}^{i,N}(s-\rho)|^{\tilde{p}l_U^*}+Z^{i,N}(s-\rho)|^{\tilde{p}l_U^*}\right)ds\right).
\end{aligned}
$$

Thus, we can immediately get that

$$
\begin{aligned}
\sup_{0\leq s\leq t}\mathbb{E}|\bar{Z}^{i,N}(s)|^{\tilde{p}}&\leq C\left(\sup_{0\leq s\leq t}\mathbb{E}\left|\bar{Z}^{i,N}(s)-D(\bar{Z}^{i,N}(s-\rho))\right|^{\tilde{p}}+\sup_{0\leq s\leq t}\mathbb{E}|D(\bar{Z}^{i,N}(s-\rho))|^{\tilde{p}}\right)\\
&\leq C\left(1+\int_0^t\sup_{0\leq u\leq s}\mathbb{E}|\bar{Z}^{i,N}(u)|^{\tilde{p}}ds+\sup_{0\leq s\leq t}\mathbb{E}\left|\bar{Z}^{i,N}(s-\rho)\right|^{\tilde{p}l_U^*}\right).
\end{aligned}
$$

Thanks to Gronwall's inequality, one can see that

$$
\sup_{0\leq s\leq t}\mathbb{E}|\bar{Z}^{i,N}(s)|^{\tilde{p}}\leq C\left(1+\sup_{0\leq s\leq t}\mathbb{E}\left|\bar{Z}^{i,N}(s-\rho)\right|^{\tilde{p}l_U^*}\right).
$$

By recalling the proof of Theorem 2.8, we can construct a finite sequence $\{\tilde{p}_1,\tilde{p}_2,\dots,\tilde{p}_{M_T+1}\}$ such that $\tilde{p}_{j+1}l_U^*<\tilde{p}_j$ and $\tilde{p}_{M_T+1}=\tilde{p}$ for $j=1,2,\dots,M_T+1$. Then using the same technique means the desired result when $\tilde{p}\geq 4$. The case when $\tilde{p}\in(0,4)$ follows from the Hölder inequality. □

Lemma 4.2. *Let Assumptions 2.4–2.6 hold. Then for any $\bar{p}\in\left[2,\frac{\tilde{p}}{2l_U+2}\right]$, we have*

$$
\max_{i\in\mathbb{N}_N}\sup_{0<\Delta\leq 1}\mathbb{E}\left(\sup_{0\leq t\leq T}|\bar{Z}^{i,N}(t)|^{\bar{p}}\right)\leq C,\quad\forall T>0.
$$

Proof. After analysis, we know that the key difference is to estimate $\mathbb{E}\left(\sup_{0\le s\le t} Q_3^{i,N}(s)\right)$. Using BDG's inequality, Young's inequality and Hölder's inequality gives that

$$
\begin{aligned}
&\mathbb{E}\left(\sup_{0\le s\le t} Q_3^{i,N}(s)\right)\\
&\le C\mathbb{E}\left[\int_0^t \left|\bar{Z}^{i,N}(s)-D(\bar{Z}^{i,N}(s-\rho))\right|^{2\bar{p}-2}\left|\beta\left(\Gamma(Z_s^{i,N}),\mathbb{L}_{\Gamma(Z_s^N)}\right)\right|^2 ds\right]^{1/2}\\
&\le C\mathbb{E}\left[\sup_{0\le s\le t}\left|\bar{Z}^{i,N}(s)-D(\bar{Z}^{i,N}(s-\rho))\right|^{\bar{p}-1}\left(\int_0^t\left|\beta\left(\Gamma(Z_s^{i,N}),\mathbb{L}_{\Gamma(Z_s^N)}\right)\right|^2 ds\right)^{1/2}\right]\\
&\le \frac{1}{2}\mathbb{E}\left(\sup_{0\le s\le t}\left|\bar{Z}^{i,N}(s)-D(\bar{Z}^{i,N}(s-\rho))\right|^{\bar{p}}\right)\\
&+C\mathbb{E}\left[\int_0^t\left[1+\sum_{v=1}^{r-1}|Z^{i,N}(s+\bar{s}_v)|^2+U_2^2(Z^{i,N}(s-\rho),0)|Z^{i,N}(s-\rho)|^2+\sum_{i=1}^{r}\mathcal{W}_2^2(\mathbb{L}_{Z^N(s+\bar{s}_v)})\right]ds\right]^{\bar{p}/2}\\
&\le \frac{1}{2}\mathbb{E}\left(\sup_{0\le s\le t}\left|\bar{Z}^{i,N}(s)-D(\bar{Z}^{i,N}(s-\rho))\right|^{\bar{p}}\right)\\
&+C\int_0^t\left(1+\sup_{0\le u\le s}\mathbb{E}|\bar{Z}^{i,N}(u)|^{\bar{p}}+\sup_{0\le u\le s}\mathbb{E}|\bar{Z}^{i,N}(u)|^{(l_U+1)\bar{p}}\right)ds\\
&\le \frac{1}{2}\mathbb{E}\left(\sup_{0\le s\le t}\left|\bar{Z}^{i,N}(s)-D(\bar{Z}^{i,N}(s-\rho))\right|^{\bar{p}}\right)+C.
\end{aligned}
$$

Then by the results in the proof of Lemma 4.1, we derive that

$$
\begin{aligned}
&\mathbb{E}\left(\sup_{0\le s\le t}\left|\bar{Z}^{i,N}(s)-D(\bar{Z}^{i,N}(s-\rho))\right|^{\bar{p}}\right)\\
&\le C+C\int_0^t\sup_{0\le u\le s}\mathbb{E}|\bar{Z}^{i,N}(u)|^{\bar{p}}ds+C\int_0^t\sup_{0\le u\le s}\mathbb{E}|\bar{Z}^{i,N}(u-\rho)|^{\bar{p}l_U^*}ds\le C,
\end{aligned}
$$

where $\bar{p}l_U^*\le\tilde{p}$ has been used. Thus, we derive that

$$
\begin{aligned}
&\mathbb{E}\left(\sup_{0\le s\le t}|\bar{Z}^{i,N}(s)|^{\bar{p}}\right)\\
&\le C\mathbb{E}\left(\sup_{0\le s\le t}\left|\bar{Z}^{i,N}(s)-D(\bar{Z}^{i,N}(s-\rho))\right|^{\bar{p}}\right)+C\mathbb{E}\left(\sup_{0\le s\le t}|D(\bar{Z}^{i,N}(s-\rho))|^{\bar{p}}\right)\\
&\le C+C\mathbb{E}\left(\sup_{0\le s\le t}\left|\bar{Z}^{i,N}(s-\rho)\right|^{\bar{p}l_U^*}\right).
\end{aligned}
$$

Similar to the last part in the proof of Lemma 4.1, we get the desired result. □

Lemma 4.3. *Let Assumptions 2.4–2.6 hold. Then for any $\hat{p}\in[2,\frac{\bar{p}}{2l_U+2}]$, it holds that*

$$
\max_{i\in\mathbb{N}_N}\mathbb{E}\left(\sup_{0\le k\le M_T}\sup_{t_k\le t\le t_{k+1}}|\bar{Z}^{i,N}(t)-Z^{i,N}(t)|^{\hat{p}}\right)\le C\Delta^{\hat{p}/2}.
$$

Proof. By (4.2), for any $t\in[t_k,t_{k+1})$, we have

$$
\bar{Z}^{i,N}(t)-Z^{i,N}(t)=D(\bar{Z}^{i,N}(t-\rho))-D(Z^{i,N}(t-\rho))+\psi(t),
$$

where $\psi(t):=\int_{t_k}^t\alpha_\Delta\left(\Gamma(Z_s^{i,N}),\mathbb{L}_{\Gamma(Z_s^N)}\right)ds+\int_{t_k}^t\beta\left(\Gamma(Z_s^{i,N}),\mathbb{L}_{\Gamma(Z_s^N)}\right)dB^i(s)$. By Assumption 2.5, Lemma 4.2 and (4.3), we derive that

$$
\begin{aligned}
&\mathbb{E}\left(\sup_{t_k\le t\le t_{k+1}}|\psi(t)|^{\hat{p}}\right)\\
&\le 2^{\hat{p}-1}\mathbb{E}\left(\sup_{t_k\le t\le t_{k+1}}\left|\int_{t_k}^t\alpha_\Delta\left(\Gamma(Z_s^{i,N}),\mathbb{L}_{\Gamma(Z_s^N)}\right)ds\right|^{\hat{p}}\right)\\
&\quad+2^{\hat{p}-1}\mathbb{E}\left(\sup_{t_k\le t\le t_{k+1}}\left|\int_{t_k}^t\beta\left(\Gamma(Z_s^{i,N}),\mathbb{L}_{\Gamma(Z_s^N)}\right)dB^i(s)\right|^{\hat{p}}\right)
\end{aligned}
$$

$$
\begin{aligned}
&\le C\Delta^{(1-\gamma)\hat{p}} + C\mathbb{E}\left[\sup_{t_k\le t\le t_{k+1}}\left|\beta\left(\Gamma(Z^{i,N}_{t_k}),\mathbb{L}_{\Gamma(Z^N_{t_k})}\right)\left(B^i(t)-B^i(t_k)\right)\right|^{\hat{p}}\right]\\
&\le C\Delta^{(1-\gamma)\hat{p}} + C\Delta^{\hat{p}/2}\mathbb{E}\left(1+\sum_{v=1}^{r-1}|Z^{i,N}(t_{k-\bar{k}_v})|^{\hat{p}} + U_2^{\hat{p}}(Z^{i,N}(t_{k-M}),0)|Z^{i,N}(t_{k-M})|^{\hat{p}} + \sum_{v=1}^{r}\mathcal{W}_2^{\hat{p}}(\mathbb{L}_{Z^N(t_{k-\bar{k}_v})})\right)\\
&\le C\Delta^{\hat{p}/2}.
\end{aligned}
$$

Thus, using Assumption 2.6 means that

$$
\begin{aligned}
&\mathbb{E}\left(\sup_{t_k\le t\le t_{k+1}}|\bar{Z}^{i,N}(t)-Z^{i,N}(t)|^{\hat{p}}\right)\\
&\le C\mathbb{E}\left(\sup_{t_k\le t\le t_{k+1}}|D(\bar{Z}^{i,N}(t-\rho))-D(Z^{i,N}(t-\rho))|^{\hat{p}}\right) + C\mathbb{E}\left(\sup_{t_k\le t\le t_{k+1}}|\psi(t)|^{\hat{p}}\right)\\
&\le C\left[\mathbb{E}\left(\sup_{t_k\le t\le t_{k+1}}U_3^{2\hat{p}}(\bar{Z}^{i,N}(t-\rho),Z^{i,N}(t-\rho))\right)\right]^{1/2}\left[\mathbb{E}\left(\sup_{t_k\le t\le t_{k+1}}|\bar{Z}^{i,N}(t-\rho)-Z^{i,N}(t-\rho)|^{2\hat{p}}\right)\right]^{1/2}\\
&\quad + C\Delta^{\hat{p}/2}\\
&\le C\Delta^{\hat{p}/2} + C\left[\mathbb{E}\left(\sup_{t_k\le t\le t_{k+1}}|\bar{Z}^{i,N}(t-\rho)-Z^{i,N}(t-\rho)|^{2\hat{p}}\right)\right]^{1/2}.
\end{aligned}
\tag{4.5}
$$

Define

$$
\hat{p}_j = (M_T+2-j)\hat{p}2^{M_T+1-j},\quad j=1,2,\dots,M_T+1.
$$

One can observe that

$$
2\hat{p}_{j+1} < \hat{p}_j \quad\text{and}\quad \hat{p}_{M_T+1}=\hat{p},\quad j=1,2,\dots,M_T+1.
$$

We only discuss the case when $T>\rho$ (i.e., $M_T>M$), otherwise, the result is immediately obtained. For $0\le k\le M-1$, (4.5) leads to

$$
\mathbb{E}\left(\sup_{t_k\le t\le t_{k+1}}|\bar{Z}^{i,N}(t)-Z^{i,N}(t)|^{\hat{p}_1}\right)\le C\Delta^{\hat{p}_1/2}. \tag{4.6}
$$

For $M\le k\le 2M-1$, combining (4.5) and (4.6) with the Hölder inequality gives that

$$
\begin{aligned}
&\mathbb{E}\left(\sup_{t_k\le t\le t_{k+1}}|\bar{Z}^{i,N}(t)-Z^{i,N}(t)|^{\hat{p}_2}\right)\\
&\le C\Delta^{\hat{p}_2/2} + C\left[\mathbb{E}\left(\sup_{t_k\le t\le t_{k+1}}|\bar{Z}^{i,N}(t-\rho)-Z^{i,N}(t-\rho)|^{\hat{p}_1}\right)\right]^{\hat{p}_2/\hat{p}_1}\\
&\le C\Delta^{\hat{p}_2/2}.
\end{aligned}
$$

The desired result is obtained by induction. □

Theorem 4.4. *Let Assumptions 2.4–2.6 hold. Then, for any* $p\in\left[2,\bar{p}\left(\frac{\hat{p}}{\bar{p}+\hat{p}l_1}\wedge\frac{1}{2(l_U+1)}\right)\right]$ *and* $4l_U+4\le(2l_U+2)\hat{p}\le\bar{p}$, *we have*

$$
\max_{i\in\mathbb{N}_N}\mathbb{E}\left(\sup_{0\le t\le T}|Y^{i,N}(t)-\bar{Z}^{i,N}(t)|^p\right)\le C\Delta^{\gamma p},\quad \forall T>0, \tag{4.7}
$$

and

$$
\max_{i\in\mathbb{N}_N}\mathbb{E}\left(\sup_{0\le t\le T}|Y^{i,N}(t)-Z^{i,N}(t)|^p\right)\le C\Delta^{\gamma p},\quad \forall T>0. \tag{4.8}
$$

Proof. For any $i\in\mathbb{S}_N$ and $t\in[0,T]$, set

$$
\Theta^{i,N}(t) = Y^{i,N}(t)-\bar{Z}^{i,N}(t)-D(Y^{i,N}(t-\rho))+D(\bar{Z}^{i,N}(t-\rho)).
$$

Application of Itô's formula yields that

$$
\begin{aligned}
&|\Theta^{i,N}(t)|^p\\
\le& p\int_0^t |\Theta^{i,N}(s)|^{p-2}(\Theta^{i,N}(s))^T\left[\alpha\left(\Gamma(Y_s^{i,N}),\mathbb{L}_{\Gamma(Y_s^N)}\right)-\alpha_\Delta\left(\Gamma(Z_s^{i,N}),\mathbb{L}_{\Gamma(Z_s^N)}\right)\right]ds\\
&+\frac{p(p-1)}{2}\int_0^t |\Theta^{i,N}(s)|^{p-2}\left|\beta\left(\Gamma(Y_s^{i,N}),\mathbb{L}_{\Gamma(Y_s^N)}\right)-\beta\left(\Gamma(Z_s^{i,N}),\mathbb{L}_{\Gamma(Z_s^N)}\right)\right|^2 ds\\
&+p\int_0^t |\Theta^{i,N}(s)|^{p-2}(\Theta^{i,N}(s))^T\left[\beta\left(\Gamma(Y_s^{i,N}),\mathbb{L}_{\Gamma(Y_s^N)}\right)-\beta\left(\Gamma(Z_s^{i,N}),\mathbb{L}_{\Gamma(Z_s^N)}\right)\right]dB^i(s)\\
\le& p\int_0^t |\Theta^{i,N}(s)|^{p-2}(\Theta^{i,N}(s))^T\left[\alpha\left(\Gamma(Y_s^{i,N}),\mathbb{L}_{\Gamma(Y_s^N)}\right)-\alpha\left(\Gamma(\bar{Z}_s^{i,N}),\mathbb{L}_{\Gamma(\bar{Z}_s^N)}\right)\right]ds\\
&+p\int_0^t |\Theta^{i,N}(s)|^{p-2}(\Theta^{i,N}(s))^T\left[\alpha\left(\Gamma(\bar{Z}_s^{i,N}),\mathbb{L}_{\Gamma(\bar{Z}_s^N)}\right)-\alpha\left(\Gamma(\bar{Z}_s^{i,N}),\mathbb{L}_{\Gamma(Z_s^N)}\right)\right]ds\\
&+p\int_0^t |\Theta^{i,N}(s)|^{p-2}(\Theta^{i,N}(s))^T\left[\alpha\left(\Gamma(\bar{Z}_s^{i,N}),\mathbb{L}_{\Gamma(Z_s^N)}\right)-\alpha\left(\Gamma(Z_s^{i,N}),\mathbb{L}_{\Gamma(Z_s^N)}\right)\right]ds\\
&+p\int_0^t |\Theta^{i,N}(s)|^{p-2}(\Theta^{i,N}(s))^T\cdot\left[\alpha\left(\Gamma(Z_s^{i,N}),\mathbb{L}_{\Gamma(Z_s^N)}\right)-\alpha_\Delta\left(\Gamma(Z_s^{i,N}),\mathbb{L}_{\Gamma(Z_s^N)}\right)\right]ds\\
&+\frac{p(p-1)}{2}\int_0^t |\Theta^{i,N}(s)|^{p-2}\left|\beta\left(\Gamma(Y_s^{i,N}),\mathbb{L}_{\Gamma(Y_s^N)}\right)-\beta\left(\Gamma(Z_s^{i,N}),\mathbb{L}_{\Gamma(Z_s^N)}\right)\right|^2 ds\\
&+p\int_0^t |\Theta^{i,N}(s)|^{p-2}(\Theta^{i,N}(s))^T\left[\beta\left(\Gamma(Y_s^{i,N}),\mathbb{L}_{\Gamma(Y_s^N)}\right)-\beta\left(\Gamma(Z_s^{i,N}),\mathbb{L}_{\Gamma(Z_s^N)}\right)\right]dB^i(s)\\
=:& A_1^{i,N}(t)+A_2^{i,N}(t)+A_3^{i,N}(t)+A_4^{i,N}(t)+A_5^{i,N}(t)+A_6^{i,N}(t).
\end{aligned}
$$

Using Lemma 4.2, Young's inequality and Assumption 2.4, we have

$$
\begin{aligned}
&\mathbb{E}\left(\sup_{0\le s\le t}A_1^{i,N}(s)\right)\\
\le& C\mathbb{E}\int_0^t\left(\left|Y^{i,N}(s)-\bar{Z}^{i,N}(s)\right|^{p-2}+\left|D(Y^{i,N}(s-\rho))-D(\bar{Z}^{i,N}(s-\rho))\right|^{p-2}\right)\\
&\cdot\left[\sum_{v=1}^{r-1}|Y^{i,N}(s+\bar{s}_v)-\bar{Z}^{i,N}(s+\bar{s}_v)|^2\right.\\
&\left.+U_2^2(Y^{i,N}(s-\rho),\bar{Z}^{i,N}(s-\rho))|Y^{i,N}(s-\rho)-\bar{Z}^{i,N}(s-\rho)|^2+\sum_{v=1}^{r}\mathbb{W}_2^2(\mathbb{L}_{Y^N(s+\bar{s}_v)},\mathbb{L}_{\bar{Z}^N(s+\bar{s}_v)})\right]ds\\
\le& C\mathbb{E}\int_0^t\left[\sum_{v=1}^{r-1}|Y^{i,N}(s+\bar{s}_v)-\bar{Z}^{i,N}(s+\bar{s}_v)|^p\right.\\
&+\left[U_2(Y^{i,N}(s-\rho),\bar{Z}^{i,N}(s-\rho))\vee U_3(Y^{i,N}(s-\rho),\bar{Z}^{i,N}(s-\rho))\right]^p\\
&\left.\cdot|Y^{i,N}(s-\rho)-\bar{Z}^{i,N}(s-\rho)|^p+\sum_{v=1}^{r}\mathbb{W}_2^2(\mathbb{L}_{Y^N(s+\bar{s}_v)},\mathbb{L}_{\bar{Z}^N(s+\bar{s}_v)})\right]ds\\
\le& C\int_0^t\mathbb{E}\left(\sup_{0\le u\le s}|Y^{i,N}(u)-\bar{Z}^{i,N}(u)|^p\right)ds+C\int_0^t\left(\mathbb{E}\left|Y^{i,N}(s-\rho)-\bar{Z}^{i,N}(s-\rho)\right|^{2p}\right)^{1/2}ds\\
&+C\mathbb{E}\int_0^t\sum_{v=1}^{r}\frac{1}{N}\sum_{j=1}^{N}\left|Y^{j,N}(s+\bar{s}_v)-\bar{Z}^{j,N}(s+\bar{s}_v)\right|^p ds\\
\le& C\int_0^t\mathbb{E}\left(\sup_{0\le u\le s}|Y^{i,N}(u)-\bar{Z}^{i,N}(u)|^p\right)ds+C\int_0^t\left(\mathbb{E}\left|Y^{i,N}(s-\rho)-\bar{Z}^{i,N}(s-\rho)\right|^{2p}\right)^{1/2}ds.
\end{aligned}
$$

Applying Lemma 4.3 and Assumption 2.4 gives that

$$
\begin{aligned}
&\mathbb{E}\left(\sup_{0\le s\le t}A_2^{i,N}(s)\right)\\
\le& C\mathbb{E}\int_0^t\left|\Theta^{i,N}(s)\right|^{p-1}\sum_{v=1}^{r}\mathbb{W}_2(\mathbb{L}_{\bar{Z}^N(s+\bar{s}_v)},\mathbb{L}_{Z^N(s+\bar{s}_v)})ds
\end{aligned}
$$

$$
\begin{aligned}
&\leq C\mathbb{E}\left[\sup_{0\leq s\leq t}\left|\Theta^{i,N}(s)\right|^{p-1}\int_0^t\sum_{v=1}^r \mathbb{W}_2(\mathbb{L}_{\bar{Z}^N(s+\bar{s}_v)},\mathbb{L}_{Z^N(s+\bar{s}_v)})ds\right]\\
&\leq \frac{1}{5}\mathbb{E}\left(\sup_{0\leq s\leq t}\left|\Theta^{i,N}(s)\right|^p\right)+C\mathbb{E}\int_0^t\sum_{v=1}^r\frac{1}{N}\sum_{j=1}^N\left|\bar{Z}^{j,N}(s+\bar{s}_v)-Z^{j,N}(s+\bar{s}_v)\right|^p ds\\
&\leq \frac{1}{5}\mathbb{E}\left(\sup_{0\leq s\leq t}\left|\Theta^{i,N}(s)\right|^p\right)+C\Delta^{\frac{p}{2}}.
\end{aligned}
$$

By Lemma 4.2, Hölder's inequality, Young's inequality and Assumptions 2.4, 2.7, we derive that

$$
\begin{aligned}
&\mathbb{E}\left(\sup_{0\leq s\leq t}A_3^{i,N}(s)\right)\\
&\leq C\mathbb{E}\int_0^t\left|\Theta^{i,N}(s)\right|^{p-1}\sum_{v=1}^r\left[U_1(\bar{Z}^{i,N}(s+\bar{s}_v),Z^{i,N}(s+\bar{s}_v))|\bar{Z}^{i,N}(s+\bar{s}_v)-Z^{i,N}(s+\bar{s}_v)|\right]ds\\
&\leq C\mathbb{E}\left[\sup_{0\leq s\leq t}\left|\Theta^{i,N}(s)\right|^{p-1}\int_0^t\sum_{v=1}^r\left[U_1(\bar{Z}^{i,N}(s+\bar{s}_v),Z^{i,N}(s+\bar{s}_v))|\bar{Z}^{i,N}(s+\bar{s}_v)-Z^{i,N}(s+\bar{s}_v)|\right]ds\right]\\
&\leq \frac{1}{5}\mathbb{E}\left(\sup_{0\leq s\leq t}\left|\Theta^{i,N}(s)\right|^p\right)\\
&\quad+C\mathbb{E}\left(\int_0^t\sum_{v=1}^r\left[U_1(\bar{Z}^{i,N}(s+\bar{s}_v),Z^{i,N}(s+\bar{s}_v))|\bar{Z}^{i,N}(s+\bar{s}_v)-Z^{i,N}(s+\bar{s}_v)|\right]ds\right)^p\\
&\leq \frac{1}{5}\mathbb{E}\left(\sup_{0\leq s\leq t}\left|\Theta^{i,N}(s)\right|^p\right)+C\int_0^t\sum_{v=1}^r\left[\mathbb{E}\left(1+|\bar{Z}^{i,N}(s+\bar{s}_v)|^{pl_1}+|Z^{i,N}(s+\bar{s}_v)|^{pl_1}\right)^{\frac{\bar{p}}{pl_1}}\right]^{\frac{pl_1}{\bar{p}}}\\
&\qquad\cdot\left[\mathbb{E}\left(|\bar{Z}^{i,N}(s+\bar{s}_v)-Z^{i,N}(s+\bar{s}_v)|^p\right)^{\frac{\bar{p}}{\bar{p}-pl_1}}\right]^{\frac{\bar{p}-pl_1}{\bar{p}}}ds\\
&\leq \frac{1}{5}\mathbb{E}\left(\sup_{0\leq s\leq t}\left|\Theta^{i,N}(s)\right|^p\right)+C\int_0^t\left[\mathbb{E}\left(\sup_{0\leq u\leq s}(|\bar{Z}^{i,N}(u)-Z^{i,N}(u)|^{\frac{p\bar{p}}{\bar{p}-pl_1}})\right)\right]^{\frac{\bar{p}-pl_1}{\bar{p}}}ds\\
&\leq \frac{1}{5}\mathbb{E}\left(\sup_{0\leq s\leq t}\left|\Theta^{i,N}(s)\right|^p\right)+C\Delta^{\frac{p}{2}}.
\end{aligned}
$$

By Assumption 2.4, Hölder's inequality, Young's inequality and (4.1), we get that

$$
\begin{aligned}
&\mathbb{E}\left(\sup_{0\leq s\leq t}A_4^{i,N}(s)\right)\\
&\leq C\mathbb{E}\int_0^t\left|\Theta^{i,N}(s)\right|^{p-1}\left|\alpha\left(\Gamma(Z_s^{i,N}),\mathbb{L}_{\Gamma(Z_s^N)}\right)-\alpha_\Delta\left(\Gamma(Z_s^{i,N}),\mathbb{L}_{\Gamma(Z_s^N)}\right)\right|ds\\
&\leq C\mathbb{E}\left(\sup_{0\leq s\leq t}\left|\Theta^{i,N}(s)\right|^{p-1}\int_0^t\left|\alpha\left(\Gamma(Z_s^{i,N}),\mathbb{L}_{\Gamma(Z_s^N)}\right)-\alpha_\Delta\left(\Gamma(Z_s^{i,N}),\mathbb{L}_{\Gamma(Z_s^N)}\right)\right|ds\right)\\
&\leq \frac{1}{5}\mathbb{E}\left(\sup_{0\leq s\leq t}\left|\Theta^{i,N}(s)\right|^p\right)+C\mathbb{E}\int_0^t\left|\alpha\left(\Gamma(Z_s^{i,N}),\mathbb{L}_{\Gamma(Z_s^N)}\right)-\alpha_\Delta\left(\Gamma(Z_s^{i,N}),\mathbb{L}_{\Gamma(Z_s^N)}\right)\right|^p ds\\
&\leq \frac{1}{5}\mathbb{E}\left(\sup_{0\leq s\leq t}\left|\Theta^{i,N}(s)\right|^p\right)+C\Delta^{p\gamma}\int_0^t\mathbb{E}\left[\frac{\left|\alpha\left(\Gamma(Z_s^{i,N}),\mathbb{L}_{\Gamma(Z_s^N)}\right)\right|^{2p}}{\left(1+\Delta^\gamma|\alpha(\Gamma(Z_s^{i,N}),\mathbb{L}_{\Gamma(Z_s^N)})|\right)^p}\right]ds\\
&\leq \frac{1}{5}\mathbb{E}\left(\sup_{0\leq s\leq t}\left|\Theta^{i,N}(s)\right|^p\right)\\
&\quad+C\Delta^{p\gamma}\int_0^t\mathbb{E}\left(1+\sum_{v=1}^r\left[\mathcal{W}_2(\mathbb{L}_{Z^N(s+\bar{s}_v)})+U_1(Z^{i,N}(s+\bar{s}_v),0)|Z^{i,N}(s+\bar{s}_v)|\right]\right)^{2p}ds\\
&\leq \frac{1}{5}\mathbb{E}\left(\sup_{0\leq s\leq t}\left|\Theta^{i,N}(s)\right|^p\right)+C\Delta^{p\gamma}.
\end{aligned}
$$

Applying Young's inequality, Hölder's inequality, Assumptions 2.5–2.7 and the estimation of $A_1^{i,N}(s)$ gives that

$$
\begin{aligned}
&\mathbb{E}\left(\sup_{0\le s\le t} A_5^{i,N}(s)\right)\\
&\le C\mathbb{E}\int_0^t \left(\left|Y^{i,N}(s)-\bar{Z}^{i,N}(s)\right|^{p-2}+\left|D(Y^{i,N}(s-\rho))-D(\bar{Z}^{i,N}(s-\rho))\right|^{p-2}\right)\\
&\quad\cdot\left[\sum_{v=1}^{r-1}|Y^{i,N}(s+\bar{s}_v)-Z^{i,N}(s+\bar{s}_v)|^2+U_2^2(Y^{i,N}(s-\rho),Z^{i,N}(s-\rho))|Y^{i,N}(s-\rho)-Z^{i,N}(s-\rho)|^2\right.\\
&\quad\left.+\sum_{v=1}^{r}\mathbb{W}_2^2(\mathbb{L}_{Y^N(s+\bar{s}_v)},\mathbb{L}_{Z^N(s+\bar{s}_v)})\right]ds\\
&\le C\mathbb{E}\int_0^t\left[\left|Y^{i,N}(s)-\bar{Z}^{i,N}(s)\right|^p+U_3^p(Y^{i,N}(s-\rho),\bar{Z}^{i,N}(s-\rho))|Y^{i,N}(s-\rho)-\bar{Z}^{i,N}(s-\rho)|^p\right.\\
&\quad+\sum_{v=1}^{r-1}|Y^{i,N}(s+\bar{s}_v)-Z^{i,N}(s+\bar{s}_v)|^p+U_2^p(Y^{i,N}(s-\rho),Z^{i,N}(s-\rho))|Y^{i,N}(s-\rho)-Z^{i,N}(s-\rho)|^p\\
&\quad\left.+\sum_{v=1}^{r}\mathbb{W}_2^p(\mathbb{L}_{Y^N(s+\bar{s}_v)},\mathbb{L}_{Z^N(s+\bar{s}_v)})\right]ds\\
&\le C\int_0^t\mathbb{E}\left(\sup_{0\le u\le s}\left|Y^{i,N}(u)-\bar{Z}^{i,N}(u)\right|^p\right)ds+C\int_0^t\left(\mathbb{E}|Y^{i,N}(s-\rho)-\bar{Z}^{i,N}(s-\rho)|^{2p}\right)^{1/2}ds\\
&\quad+C\int_0^t\mathbb{E}\left(\sum_{v=1}^{r}\left|\bar{Z}^{i,N}(s+\bar{s}_v)-Z^{i,N}(s+\bar{s}_v)\right|^p\right)ds\\
&\quad+C\int_0^t\mathbb{E}\left[\sum_{v=1}^{r}\left(\frac{1}{N}\sum_{j=1}^{N}\left|Y^{j,N}(s+\bar{s}_v)-Z^{j,N}(s+\bar{s}_v)\right|^p\right)\right]ds\\
&\le C\int_0^t\mathbb{E}\left(\sup_{0\le u\le s}\left|Y^{i,N}(u)-\bar{Z}^{i,N}(u)\right|^p\right)ds+C\int_0^t\left(\mathbb{E}|Y^{i,N}(s-\rho)-\bar{Z}^{i,N}(s-\rho)|^{2p}\right)^{1/2}ds+C\Delta^{\frac{p}{2}}.
\end{aligned}
$$

Similar to the estimation of $A_5^{i,N}(t)$, using BDG's inequality, Young's inequality means that

$$
\begin{aligned}
&\mathbb{E}\left(\sup_{0\le s\le t} A_6^{i,N}(s)\right)\\
&\le C\mathbb{E}\left[\int_0^t\left|\Theta^{i,N}(s)\right|^{2p-2}\left|\beta\left(\Gamma(Y_s^{i,N}),\mathbb{L}_{\Gamma(Y_s^N)}\right)-\beta\left(\Gamma(Z_s^{i,N}),\mathbb{L}_{\Gamma(Z_s^N)}\right)\right|^2ds\right]^{1/2}\\
&\le\frac{1}{5}\mathbb{E}\left(\sup_{0\le s\le t}\left|\Theta^{i,N}(s)\right|^p\right)+C\mathbb{E}\left[\int_0^t\left|\beta\left(\Gamma(Y_s^{i,N}),\mathbb{L}_{\Gamma(Y_s^N)}\right)-\beta\left(\Gamma(Z_s^{i,N}),\mathbb{L}_{\Gamma(Z_s^N)}\right)\right|^2ds\right]^{p/2}\\
&\le\frac{1}{5}\mathbb{E}\left(\sup_{0\le s\le t}\left|\Theta^{i,N}(s)\right|^p\right)+C\mathbb{E}\int_0^t\left[\sum_{v=1}^{r-1}|Y^{i,N}(s+\bar{s}_v)-Z^{i,N}(s+\bar{s}_v)|^p\right.\\
&\quad\left.+U_2^p(Y^{i,N}(s-\rho),Z^{i,N}(s-\rho))|Y^{i,N}(s-\rho)-Z^{i,N}(s-\rho)|^p+\sum_{v=1}^{r}\mathbb{W}_2^p(\mathbb{L}_{Y^N(s+\bar{s}_v)},\mathbb{L}_{Z^N(s+\bar{s}_v)})\right]ds\\
&\le\frac{1}{5}\mathbb{E}\left(\sup_{0\le s\le t}\left|\Theta^{i,N}(s)\right|^p\right)+C\Delta^{\frac{p}{2}}.
\end{aligned}
$$

Combining these inequalities gives that

$$
\begin{aligned}
\mathbb{E}\left(\sup_{0\le s\le t}\left|\Theta^{i,N}(s)\right|^p\right)\le{}&\frac{4}{5}\mathbb{E}\left(\sup_{0\le s\le t}\left|\Theta^{i,N}(s)\right|^p\right)+C\int_0^t\mathbb{E}\left(\sup_{0\le u\le s}\left|Y^{i,N}(u)-\bar{Z}^{i,N}(u)\right|^p\right)ds\\
&+C\int_0^t\left(\mathbb{E}|Y^{i,N}(s-\rho)-\bar{Z}^{i,N}(s-\rho)|^{2p}\right)^{1/2}ds+C\Delta^{p\gamma}.
\end{aligned}
$$

Therefore, using Assumption 2.6 yields that

$$
\begin{aligned}
&\mathbb{E}\left(\sup_{0\le s\le t}\left|Y^{i,N}(s)-\bar{Z}^{i,N}(s)\right|^p\right)\\
&\le C\mathbb{E}\left(\sup_{0\le s\le t}\left|\Theta^{i,N}(s)\right|^p\right)+C\mathbb{E}\left(\sup_{0\le s\le t}\left|D(Y^{i,N}(s-\rho))-D(\bar{Z}^{i,N}(s-\rho))\right|^p\right)\\
&\le C\int_0^t\mathbb{E}\left(\sup_{0\le u\le s}\left|Y^{i,N}(u)-\bar{Z}^{i,N}(u)\right|^p\right)ds+C\left[\mathbb{E}\left(\sup_{0\le s\le t}|Y^{i,N}(s-\rho)-\bar{Z}^{i,N}(s-\rho)|^{2p}\right)\right]^{1/2}+C\Delta^{p\gamma}.
\end{aligned}
$$

Thanks to Gronwall's inequality, we have

$$\mathbb{E}\left(\sup_{0\le s\le t}\left|Y^{i,N}(s)-\bar{Z}^{i,N}(s)\right|^{p}\right)\le C\Delta^{p\gamma}+C\left[\mathbb{E}\left(\sup_{0\le s\le t}|Y^{i,N}(s-\rho)-\bar{Z}^{i,N}(s-\rho)|^{2p}\right)\right]^{1/2}.$$

Define $p_j=(M_T+2-j)p2^{M_T+1-j}$, $j=1,2,\dots,M_T+1$. Applying the same technique in Lemma 4.3 gives the desired result (4.7). Then (4.8) is obtained by Lemma 4.3 and (4.7) immediately. □

Remark 7. The optimal convergence rate is $\frac{1}{2}$, which can be obtained in Theorem 4.4 by choosing $\gamma=\frac{1}{2}$.

Theorem 4.5. *Let all conditions in Theorems 3.1 and 4.4 hold. Then,*

$$\mathbb{E}\left(\sup_{0\le t\le T}|Y^{i}(t)-\bar{Z}^{i,N}(t)|^{p}\right)\le C\begin{cases}(N^{-1/2})^{\lambda_{T,\rho,p}}+\Delta^{p\gamma}, & \text{if } p>d/2,\\ [N^{-1/2}\log(1+N)]^{\lambda_{T,\rho,p}}+\Delta^{p\gamma}, & \text{if } p=d/2,\\ (N^{-p/d})^{\lambda_{T,\rho,p}}+\Delta^{p\gamma}, & \text{if } 2\le p<d/2,\end{cases}$$

where $\lambda_{T,\rho,p}=\left(\frac{p-\varepsilon}{p}\right)^{\lfloor\frac{T}{\rho}\rfloor}$.

The above theorem follows directly from Theorems 3.1 and 4.4.

Theorem 4.6. *Let all conditions in Theorems 3.2 and 4.4 hold. Then,*

$$\mathbb{E}\left(\sup_{0\le t\le T}|Y^{i}(t)-\bar{Z}^{i,N}(t)|^{p}\right)\le C\begin{cases}N^{-1/2}+N^{-(\bar{p}-p)/\bar{p}}+\Delta^{p\gamma}, & \text{if } p>d/2 \text{ and } \bar{p}\ne 2p,\\ N^{-1/2}\log(1+N)+N^{-(\bar{p}-p)/\bar{p}}+\Delta^{p\gamma}, & \text{if } p=d/2 \text{ and } \bar{p}\ne 2p,\\ N^{-p/d}+N^{-(\bar{p}-p)/\bar{p}}+\Delta^{p\gamma}, & \text{if } 2\le p<d/2.\end{cases}$$

The above theorem can be achieved easily by using Theorems 3.2 and 4.4.

5. Example

We would like to point out that our theory can cover the numerical example in [28]. Moreover, the delay component in our results can be more general. In the following two examples, we can see that the neutral term is highly nonlinear and the constraint of delay variable is general.

Example 5.1. Consider the scalar NMSMVE

$$\begin{aligned}&d[Y(t)+Y^{3}(t-2)]\\&=[-2Y(t)+Y(t-0.25)-2Y^{5}(t-2)-\int_{\mathbb{R}}|Y(t)-y|\mathbb{L}_{Y(t)}(dy)+\mathbb{E}Y(t-0.5)]dt\\&\quad+[Y(t)+0.25Y(t-0.125)+\mathbb{E}Y(t-0.5)]dB(t),\end{aligned}\tag{5.1}$$

with the initial data $\xi(\theta)=|\sin(B(\theta+2))|$, $\theta\in[-2,0]$. One can observe that $\rho_2=0.125$, $\rho_3=0.25$, $\rho_4=0.5$, $\rho_5=2$. All assumptions are satisfied. We only check (2.3):

$$\begin{aligned}&(x_1+x_5^3-y_1-y_5^3)(-2x_1+x_3-2x_5^5+2y_1-y_3+2y_5^5)\\&=(x_1-y_1)(-2x_1+2y_1+x_3-y_3)+(x_1-y_1)(-2x_5^5+2y_5^5)+(x_5^3-y_5^3)(-2x_1+2y_1+x_3-y_3)\\&\quad+(x_5^3-y_5^3)(-2x_5^5+2y_5^5)\\&\le-2|x_1-y_1|^2+\frac{1}{2}|x_1-y_1|^2+\frac{1}{2}|x_3-y_3|^2+\frac{1}{2}|x_1-y_1|^2+\frac{1}{2}|x_5-y_5|^2|x_5^4+x_5^3y_5+x_5^2y_5^2+x_5y_5^3+y_5^4|^2\\&\quad+\frac{1}{2}|x_5-y_5|^2|x_5^2+x_5y_5+y_5^2|^2+4|x_1-y_1|^2+|x_3-y_3|^2\\&\quad+|x_5-y_5|^2(x_5^2+x_5y_5+y_5^2)(x_5^4+x_5^3y_5+x_5^2y_5^2+x_5y_5^3+y_5^4)\\&\le 3|x_1-y_1|^2+\frac{3}{2}|x_3-y_3|^2+25(1+x_5^8+y_5^8)|x_5-y_5|^2.\end{aligned}$$

Hence, Assumption 2.4 holds with $U_2(x_5,y_5)=1+x_5^4+y_5^4$. Thus, the NMSMVE (5.1) admits a unique strong solution. In the numerical simulation, set $T=4$ and $N=500$. The numerical solution with $\Delta=2^{-12}$ is regarded as the true solution, since the true solution cannot be expressed explicitly. We focus on the error between the numerical solution of tamed EM scheme and the interacting particle system with $N=500$, which is defined by

$$err=\left[\frac{1}{N}\sum_{i=1}^{N}|Z_{u}^{i,N}(T)-Z_{u_l}^{i,N}(T)|^2\right]^{\frac{1}{2}},\tag{5.2}$$

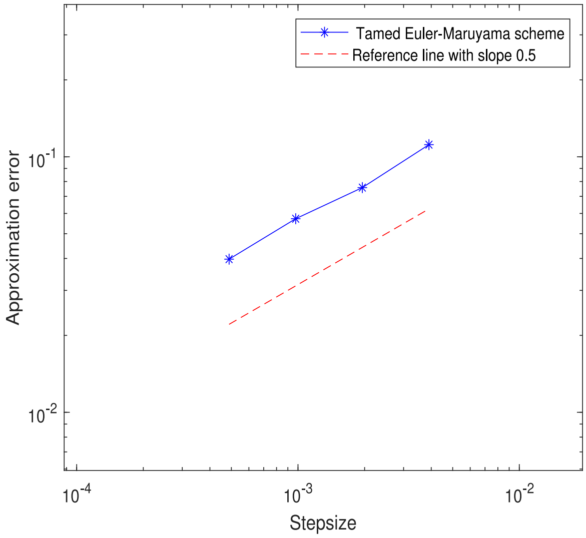

Fig. 1. Convergence rate of tamed EM scheme for (5.1).

where $Z_u^{i,N}(T)$ is the numerical approximation to the tamed EM scheme of $Z^{i,N}$ at time T with $\Delta = 2^{-12}$ and $Z_{u_l}^{i,N}(T)$ is the numerical approximation of $Z^{i,N}$ at time T with the level of the time discretization u_l, and $u_l \in \{u_1, u_2, u_3, u_4\}$ which matches $\Delta \in \{2^{-11}, 2^{-10}, 2^{-9}, 2^{-8}\}$. Let $\gamma = 0.5$ in the definition of α_Δ (4.1), so

$$\alpha_\Delta(\Gamma(Z_{t_k}^{i,N}), \mathbb{L}_{\Gamma(Z_{t_k}^N)}) = \frac{\alpha_{\Delta,k}^*}{1 + \Delta^{0.5}\left|\alpha_{\Delta,k}^*\right|},$$

where

$$\alpha_{\Delta,k}^* = -2Z^{i,N}(t_k) + Z^{i,N}(t_{k-\frac{\rho_3}{\Delta}}) - 2(Z^{i,N}(t_{k-\frac{\rho_5}{\Delta}}))^5 - \frac{1}{N}\sum_{j=1}^{N}|Z^{i,N}(t_k) - Z^{j,N}(t_k)| + \frac{1}{N}\sum_{j=1}^{N} Z^{j,N}(t_{k-\frac{\rho_4}{\Delta}}).$$

After simulation, Fig. 1 gives the approximation error defined by (5.2) as the function of stepsize with the level of the time discretization $u_l \in \{u_1, u_2, u_3, u_4\}$. From Fig. 1, we observe that the rate of convergence is about 0.5, which supports the finding.

Example 5.2. Consider the two-dimensional NMSMVE

$$\begin{aligned} &d\begin{bmatrix} Y_1(t) + 2\sin(Y_1(t-4)) \\ Y_2(t) + 4Y_2^2(t-4) \end{bmatrix} \\ &= \begin{bmatrix} -3Y_2(t) + Y_1(t-0.125) - 4Y_1^3(t-4) + \mathbb{E}Y_2(t-0.125) - 3\mathbb{E}Y_1(t-1) \\ -4(Y_2(t) + 4Y_2^2(t-4))|Y_2(t) + 4Y_2^2(t-4)| + 30Y_2^2(t-4) + 6Y_2(t) - 2Y_2^5(t-4) \end{bmatrix} dt \\ &\quad + \begin{bmatrix} Y_1(t) + \mathbb{E}Y_1(t-1) \\ Y_2(t) + \mathbb{E}Y_2(t-1) \end{bmatrix} dB(t), \end{aligned} \tag{5.3}$$

with the initial data $\xi(\theta) = |\theta|^{\frac{2}{3}} + 1$, $\theta \in [-4, 0]$. We know that $\rho_2 = 0.125$, $\rho_3 = 1$, $\rho_4 = 4$. Let us make an explanation for $x_{ab} \in \mathbb{R}$: it is the bth value of the ath element in the solution $Y(t)$. For example, $Y_1(t) = x_{11}, Y_2(t) = x_{21}, Y_1(t-4) = x_{14}, Y_2(t-4) = x_{24}$. All assumptions are satisfied. We only check (2.3):

$$\begin{aligned} &\left[\begin{pmatrix} x_{11} + 2\sin(x_{14}) \\ x_{21} + 4x_{24}^2 \end{pmatrix} - \begin{pmatrix} y_{11} + 2\sin(y_{14}) \\ y_{21} + 4y_{24}^2 \end{pmatrix}\right]^T \\ &\left[\begin{pmatrix} -3x_{21} + x_{12} - 4x_{14}^3 \\ -4(x_{21} + 4x_{24}^2)|x_{21} + 4x_{24}^2| + 30x_{24}^2 + 6x_{21} - 2x_{24}^5 \end{pmatrix} \right. \\ &\quad \left. - \begin{pmatrix} -3y_{21} + y_{12} - 4y_{14}^3 \\ -4(y_{21} + 4y_{24}^2)|y_{21} + 4y_{24}^2| + 30y_{24}^2 + 6y_{21} - 2y_{24}^5 \end{pmatrix}\right] \end{aligned}$$

$$\begin{aligned}&= [x_{11} + 2\sin(x_{14}) - y_{11} - 2\sin(y_{14})][-3x_{21} + x_{12} - 4x_{14}^3 + 3y_{21} - y_{12} + 4y_{14}^3]\\ &\quad + [x_{21} + 4x_{24}^2 - y_{21} - 4y_{24}^2][-4(x_{21} + 4x_{24}^2)|x_{21} + 4x_{24}^2| + 30x_{24}^2 + 6x_{21} - 2x_{24}^5\\ &\qquad + 4(y_{21} + 4y_{24}^2)|y_{21} + 4y_{24}^2| - 30y_{24}^2 - 6y_{21} + 2y_{24}^5]\\ &=: NU_1 + NU_2.\end{aligned}$$

Using the Young inequality and the inequality $|\sin A - \sin B| \le |A - B|$ for any $A, B \in \mathbb{R}$ gives that

$$\begin{aligned}NU_1 =& -3(x_{11} - y_{11})(x_{21} - y_{21}) + (x_{11} - y_{11})(x_{12} - y_{12})\\ &- 4(x_{11} - y_{11})(x_{14}^3 - y_{14}^3 - 6(\sin(x_{14}) - \sin(y_{14}))(x_{21} - y_{21}))\\ &+ 2(\sin(x_{14}) - \sin(y_{14}))(x_{12} - y_{12}) - 8(\sin(x_{14}) - \sin(y_{14}))(x_{14}^3 - y_{14}^3)\\ \le& \frac{3}{2}|x_{11} - y_{11}|^2 + \frac{3}{2}|x_{21} - y_{21}|^2 + \frac{1}{2}|x_{11} - y_{11}|^2 + \frac{1}{2}|x_{12} - y_{12}|^2\\ &+ 2|x_{11} - y_{11}|^2 + 2|x_{14} - y_{14}|^2|x_{14}^2 + x_{14}y_{14} + y_{14}^2|^2\\ &+ 3|x_{14} - y_{14}|^2 + 3|x_{21} - y_{21}|^2 + |x_{14} - y_{14}|^2 + |x_{12} - y_{12}|^2\\ &+ 4|x_{14} - y_{14}|^2 + 4|x_{14} - y_{14}|^2|x_{14}^2 + x_{14}y_{14} + y_{14}^2|^2.\end{aligned}$$

Note that

$$\begin{aligned}NU_2 =&[(x_{21} + 4x_{24}^2) - (y_{21} + 4y_{24}^2)][-4(x_{21} + 4x_{24}^2)|x_{21} + 4x_{24}^2|\\ &\quad + 4(y_{21} + 4y_{24}^2)|y_{21} + 4y_{24}^2| + 6(x_{21} + 4x_{24}^2) - 6(y_{21} + 4y_{24}^2)]\\ &+ 6[(x_{21} + 4x_{24}^2) - (y_{21} + 4y_{24}^2)](x_{24}^2 - y_{24}^2)\\ &- 2[(x_{21} + 4x_{24}^2) - (y_{21} + 4y_{24}^2)](x_{24}^5 - y_{24}^5).\end{aligned}$$

Applying the inequality $(|A| + |B|)(|A| - |B|)^2 \le (A - B)(A|A| - B|B|)$ for any $A, B \in \mathbb{R}$ leads to

$$\begin{aligned}&[(x_{21} + 4x_{24}^2) - (y_{21} + 4y_{24}^2)][-4(x_{21} + 4x_{24}^2)|x_{21} + 4x_{24}^2|\\ &\quad + 4(y_{21} + 4y_{24}^2)|y_{21} + 4y_{24}^2| + 6(x_{21} + 4x_{24}^2) - 6(y_{21} + 4y_{24}^2)]\\ &\le -4[|x_{21} + 4x_{24}^2| + |y_{21} + 4y_{24}^2|][|x_{21} + 4x_{24}^2| - |y_{21} + 4y_{24}^2|]^2\\ &\quad + 6|(x_{21} + 4x_{24}^2) - (y_{21} + 4y_{24}^2)|^2\\ &\le 12|x_{21} - y_{21}|^2 + 192|x_{24} - y_{24}|^2|x_{24} + y_{24}|^2.\end{aligned}$$

Thus,

$$\begin{aligned}NU_2 \le& 12|x_{21} - y_{21}|^2 + 192|x_{24} - y_{24}|^2|x_{24} + y_{24}|^2\\ &+ 3|x_{21} - y_{21}|^2 + 3|x_{24} - y_{24}|^2|x_{24} + y_{24}|^2 + 24|x_{24} - y_{24}|^2|x_{24} + y_{24}|^2\\ &+ |x_{21} - y_{21}|^2 + |x_{24} - y_{24}|^2|x_{24}^4 + x_{24}^3y_{24} + x_{24}^2y_{24}^2 + x_{24}y_{24}^3 + y_{24}^4|^2\\ &+ 8|x_{24} - y_{24}|^2|x_{24} + y_{24}\,||\,x_{24}^4 + x_{24}^3y_{24} + x_{24}^2y_{24}^2 + x_{24}y_{24}^3 + y_{24}^4|.\end{aligned}$$

Combining these results means that

$$\begin{aligned}NU_1 + NU_2 \le& 21\left|\begin{pmatrix} x_{11}\\ x_{21}\end{pmatrix} - \begin{pmatrix} y_{11}\\ y_{21}\end{pmatrix}\right|^2 + \frac{3}{2}\left|\begin{pmatrix} x_{12}\\ x_{22}\end{pmatrix} - \begin{pmatrix} y_{12}\\ y_{22}\end{pmatrix}\right|^2\\ &+ 590\left(1 + \left|\begin{pmatrix} x_{14}\\ x_{24}\end{pmatrix}\right|^8 + \left|\begin{pmatrix} y_{14}\\ y_{24}\end{pmatrix}\right|^8\right)\left|\begin{pmatrix} x_{14}\\ x_{24}\end{pmatrix} - \begin{pmatrix} y_{14}\\ y_{24}\end{pmatrix}\right|^2.\end{aligned}$$

Hence, Assumption 2.4 is satisfied with

$$U_2\left(\begin{pmatrix} x_{14}\\ x_{24}\end{pmatrix}, \begin{pmatrix} y_{14}\\ y_{24}\end{pmatrix}\right) = 1 + \left|\begin{pmatrix} x_{14}\\ x_{24}\end{pmatrix}\right|^8 + \left|\begin{pmatrix} y_{14}\\ y_{24}\end{pmatrix}\right|^8.$$

Obviously, the NMSMVE (5.3) admits a unique strong solution. In the numerical simulation, set $T = 4$, $N = 500$ and the numerical solution with $\Delta = 2^{-12}$ is seen as the true solution. We also focus on the error between the numerical solution of tamed EM scheme and the interacting particle system, which is defined by (5.2). Moreover,

$$\alpha_\Delta(\Gamma(Z_{t_k}^{i,N}), \mathbb{L}_{\Gamma(Z_{t_k}^N)}) = \frac{\begin{pmatrix} \alpha_{\Delta,k}^{*,(1)}\\ \alpha_{\Delta,k}^{*,(2)}\end{pmatrix}}{1 + \Delta^{0.5}\sqrt{|\alpha_{\Delta,k}^{*,(1)}|^2 + |\alpha_{\Delta,k}^{*,(2)}|^2}},$$

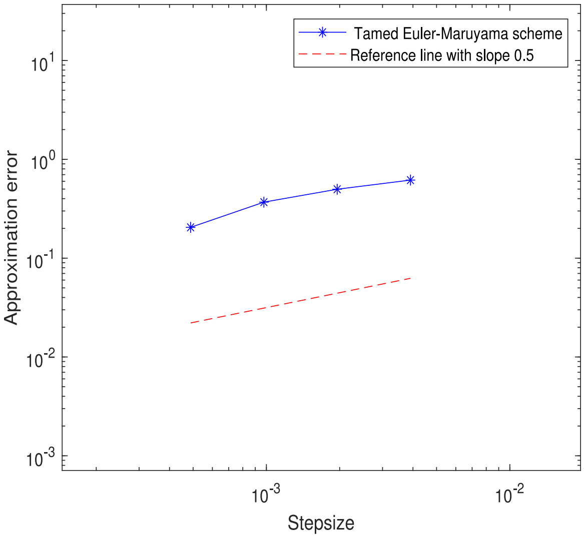

Fig. 2. Convergence rate of tamed EM scheme for (5.3).

where

$$\alpha^{*,(1)}_{\Delta,k} = -3Z^{i,N}_2(t_k) + Z^{i,N}_1(t_{k-\frac{\rho_2}{\Delta}}) - 4(Z^{i,N}_1(t_{k-\frac{\rho_4}{\Delta}}))^3 + \frac{1}{N}\sum_{j=1}^{N} Z^{j,N}_2(t_{k-\frac{\rho_2}{\Delta}}) - \frac{3}{N}\sum_{j=1}^{N} Z^{j,N}_1(t_{k-\frac{\rho_3}{\Delta}}),$$

and

$$\alpha^{*,(2)}_{\Delta,k} = -4(Z^{i,N}_2(t_k) + 4(Z^{i,N}_2(t_{k-\frac{\rho_4}{\Delta}}))^2)|Z^{i,N}_2(t_k) + 4(Z^{i,N}_2(t_{k-\frac{\rho_4}{\Delta}}))^2| + 30(Z^{i,N}_2(t_{k-\frac{\rho_4}{\Delta}}))^2 + 6Z^{i,N}_2(t_k) - 2(Z^{i,N}_2(t_{k-\frac{\rho_4}{\Delta}}))^5.$$

Fig. 2 shows the approximation error for the NMSMVE (5.3). We observe that the convergence rate is approximately 0.5, which is consistent with the theory.

Data availability

Data will be made available on request.

Acknowledgments

The authors thank the associated editor and reviewers for the helpful comments and suggestions.

Funding

This work is supported by the National Natural Science Foundation of China (Grant Nos. 12271368, 11871343 and 62373383) and Shanghai Rising-Star Program, China (Grant No. 22QA1406900).

References

[1] R. Buckdahn, J. Li, J. Ma, A mean-field stochastic control problem with partial observations, Ann. Appl. Probab. 27 (5) (2017) 3201–3245.
[2] R. Carmona, F. Delarue, Probabilistic Theory of Mean Field Games with Applications. I, in: Probability Theory and Stochastic Modelling, vol. 83, Springer, Cham, 2018.
[3] J. Wen, X. Wang, S. Mao, X. Xiao, Maximum likelihood estimation of McKean-Vlasov stochastic differential equation and its application, Appl. Math. Comput. 274 (2016) 237–246.
[4] H.P. McKean, A class of Markov processes associated with nonlinear parabolic equations, Proc. Natl. Acad. Sci. 56 (6) (1966) 1907–1911.
[5] H.P. McKean, Fluctuations in the kinetic theory of gases, Comm. Pure Appl. Math. 28 (4) (1975) 435–455.
[6] H.P. McKean, Propagation of chaos for a class of non-linear parabolic equations, Lect. Ser. Differ. Equ. 2 (1967) 41–57.
[7] M. Kac, Foundations of Kinetic Theory. Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954-1955, III, University of California Press, Berkeley and Los Angeles, 1956, pp. 171–197.

[8] G.D. Reis, W. Salkeld, J. Tugaut, Freidlin–Wentzell LDP in path space for McKean-Vlasov equations and the functional iterated logarithm law, Ann. Appl. Probab. 29 (3) (2019) 1487–1540.
[9] A.S. Sznitman, Topics in Propagation of Chaos, Springer, Berlin, Heidelberg, 1991.
[10] F.Y. Wang, Distribution dependent SDEs for Landau type equations, Stochastic Process. Appl. 128 (2) (2018) 595–621.
[11] X. Fan, X. Huang, Y. Suo, C. Yuan, Distribution dependent SDEs driven by fractional Brownian motions, Stochastic Process. Appl. 151 (2022) 23–67.
[12] X. Huang, P. Ren, F.Y. Wang, Distribution dependent stochastic differential equations, Front. Math. China 16 (2) (2021) 257–301.
[13] X. Huang, F.Y. Wang, Distribution dependent SDEs with singular coefficients, Stochastic Process. Appl. 129 (11) (2019) 4747–4770.
[14] M. Röckner, X. Zhang, Well-posedness of distribution dependent SDEs with singular drifts, Bernoulli 27 (2) (2021) 1131–1158.
[15] J. Shao, D. Wei, Propagation of chaos and conditional McKean-Vlasov SDEs with regime-switching, Front. Math. China (2021) 1–16.
[16] H. Wu, J. Hu, S. Gao, C. Yuan, Stabilization of stochastic McKean-Vlasov equations with feedback control based on discrete-time state observation, SIAM J. Control Optim. 60 (5) (2022) 2884–2901.
[17] M. Hutzenthaler, A. Jentzen, P.E. Kloeden, Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients, Proc. R. Soc. A 467 (2130) (2011) 1563–1576.
[18] J.F. Chassagneux, A. Jacquier, I. Mihaylov, An explicit Euler scheme with strong rate of convergence for financial SDEs with non-Lipschitz coefficients, SIAM J. Financial Math. 7 (1) (2016) 993–1021.
[19] M. Hutzenthaler, A. Jentzen, P.E. Kloeden, Strong convergence of an explicit numerical method for SDEs with nonglobally Lipschitz continuous coefficients, Ann. Appl. Probab. 22 (4) (2012) 1611–1641.
[20] S. Sabanis, A note on tamed Euler approximations, Electron. Commun. Probab. 18 (2013) 1–10.
[21] G.D. Reis, S. Engelhardt, G. Smith, Simulation of McKean-Vlasov SDEs with super-linear growth, IMA J. Numer. Anal. 42 (1) (2022) 874–922.
[22] J. Bao, C. Reisinger, P. Ren, W. Stockinger, First-order convergence of Milstein schemes for McKean-Vlasov equations and interacting particle systems, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 477 (2245) (2021) 20200258.
[23] C. Kumar, Neelima, C. Reisinger, W. Stockinger, Well-posedness and tamed schemes for McKean-Vlasov equations with common noise, Ann. Appl. Probab. 32 (2022) 3283–3330.
[24] J. Bao, C. Reisinger, P. Ren, W. Stockinger, Milstein schemes and antithetic multi-level Monte-Carlo sampling for delay McKean-Vlasov equations and interacting particle systems, 2020, arXiv preprint, arXiv:2005.01165.
[25] X. Huang, X. Wang, Path dependent McKean-Vlasov SDEs with Hölder continuous diffusion, Discrete Contin. Dyn. Syst. Ser. S 16 (5) (2023) 982–998.
[26] Neelima, S. Biswas, C. Kumar, G.D. Reis, C. Reisinger, Well-posedness and tamed Euler schemes for McKean-Vlasov equations driven by Leǘy noise, 2020, arXiv preprint, arXiv:2010.08585.
[27] P. Ren, J.L. Wu, Least squares estimator for path-dependent McKean-Vlasov SDEs via discrete-time observations, Acta Math. Sci. Ser. B 39 (3) (2019) 691–716.
[28] Y. Cui, X. Li, Y. Liu, C. Yuan, Explicit numerical approximations for McKean-Vlasov neutral stochastic differential delay equations, Discrete Contin. Dyn. Syst. Ser. S 16 (5) (2023) 1111–1141.
[29] S. Deng, C. Fei, W. Fei, X. Mao, Tamed EM schemes for neutral stochastic differential delay equations with superlinear diffusion coefficients, J. Comput. Appl. Math. 388 (2021) 113269.
[30] Y. Ji, C. Yuan, Tamed EM scheme of neutral stochastic differential delay equations, J. Comput. Appl. Math. 326 (2017) 337–357.
[31] X. Li, W. Cao, On mean-square stability of two-step Maruyama methods for nonlinear neutral stochastic delay differential equations, Appl. Math. Comput. 261 (2015) 373–381.
[32] L. Tan, Almost sure convergence rate of theta-EM scheme for neutral SDDEs, J. Comput. Appl. Math. 342 (2018) 25–36.
[33] L. Tan, C. Yuan, Strong convergence of a tamed theta scheme for NSDDEs with one-sided Lipschitz drift, Appl. Math. Comput. 338 (2018) 607–623.
[34] J. Bao, C. Yuan, Convergence rate of EM scheme for SDDEs, Proc. Amer. Math. Soc. 141 (9) (2013) 3231–3243.
[35] F. Wu, X. Mao, K. Chen, The Cox-Ingersoll-Ross model with delay and strong convergence of its Euler–Maruyama approximate solutions, Appl. Numer. Math. 59 (10) (2009) 2641–2658.
[36] A. Ahlborn, U. Parlitz, Stabilizing unstable steady states using multiple delay feedback control, Phys. Rev. Lett. 93 (26) (2004) 264101.
[37] C. Fei, W. Fei, X. Mao, M. Shen, L. Yan, Stability analysis of highly nonlinear hybrid multiple-delay stochastic differential equations, J. Appl. Anal. Comput. 9 (3) (2019) 1053–1070.
[38] X. Li, Q. Zhu, D. O'Regan, pTh Moment exponential stability of impulsive stochastic functional differential equations and application to control problems of NNs, J. Franklin Inst. 351 (9) (2014) 4435–4456.
[39] A. Rathinasamy, K. Balachandran, Mean square stability of semi-implicit Euler method for linear stochastic differential equations with multiple delays and Markovian switching, Appl. Math. Comput. 206 (2) (2008) 968–979.
[40] J. Bao, X. Huang, Approximations of McKean-Vlasov stochastic differential equations with irregular coefficients, J. Theoret. Probab. 35 (2) (2022) 1187–1215.
[41] Y. Li, X. Mao, Q. Song, F. Wu, G. Yin, Strong convergence of Euler–Maruyama schemes for McKean-Vlasov stochastic differential equations under local Lipschitz conditions of state variables, IMA J. Numer. Anal. (2022).
[42] N. Fournier, A. Guillin, On the rate of convergence in Wasserstein distance of the empirical measure, Probab. Theory Related Fields 162 (3) (2015) 707–738.
[43] L. Tan, C. Yuan, Convergence rates of theta-method for NSDDEs under non-globally Lipschitz continuous coefficients, Bull. Math. Sci. 9 (03) (2019) 1950006.
[44] M. Bossy, D. Talay, A stochastic particle method for the McKean-Vlasov and the Burgers equation, Math. Comp. 66 (217) (1997) 157–192.
[45] M. Bossy, D. Talay, Convergence rate for the approximation of the limit law of weakly interacting particles: application to the Burgers equation, Ann. Appl. Probab. 6 (3) (1996) 818–861.