Approximations of McKean-Vlasov SDEs with irregular coefficients[†]

Jianhai Bao, Xing Huang*
Center for Applied Mathematics, Tianjin University, Tianjin 300072, China

January 4, 2021

Abstract

The goal of this paper is to approximate two kinds of McKean-Vlasov SDEs with irregular coefficients via weakly interacting particle systems. More precisely, propagation of chaos and convergence rate of Euler-Maruyama scheme associated with the consequent weakly interacting particle systems are investigated for McKean-Vlasov SDEs, where (i) the diffusion terms are Hölder continuous by taking advantage of Yamada-Watanabe's approximation approach and (ii) the drifts are Hölder continuous by freezing distributions followed by invoking Zvonkin's transformation trick.

AMS subject Classification: 65C05, 65C30, 65C35.
Keywords: McKean-Vlasov SDE, Yamada-Watanabe approximation, Zvonkin's transformation, Hölder continuity

1 Introduction and main results

The pioneer work on McKean-Vlasov SDEs whose coefficients depend on state components and their laws is initiated in [28]. In terminology, McKean-Vlasov SDEs are also referred to as distribution-dependent SDEs or mean-field SDEs, which are derived as limits of interacting diffusions. Since McKean's work, McKean-Vlasov SDEs have been applied extensively in stochastic control, queue systems, mathematical finance, multi-factor stochastic volatility and hybrid models, to name a few; see, for example, [4, 7]. So far, McKean-Vlasov SDEs have been investigated considerably on wellposedness [7, 16, 36], ergodicity [18, 33], Feyman-Kac Formulae [5, 14, 33], and Harnack inequalities [21, 38] among others.

In general, McKean-Vlasov SDEs cannot be solved explicitly so it is desirable to devise implementable numerical algorithms so that they can be simulated. With contrast to the standard

[†]Supported in part by NNSFC (11801406, 11771326, 11831014).
*Corresponding author: Xing Huang. Email: xinghuang@tju.edu.cn.

SDEs, the primary challenge to simulate McKean-Vlasov SDEs lies in approximating distributions at each step. At present, there exist a few of results on numerical approximations for McKean-Vlasov SDEs; see e.g. [6, 26]. In particular, [17] is concerned with strong convergence of the tamed Euler-Maruyama (EM for short) scheme for McKean-Vlasov SDEs, where the drift terms involved are of superlinear growth w.r.t. the spatial variables. Subsequently, [17] is extended to the higher-order numerical schemes (e.g., the tamed Milstein scheme and the adaptive EM scheme); see e.g. [8, 12, 24]. It is worthy to emphasize that the strong convergence of numerical algorithms in [8, 12, 17, 24] is analyzed under the smooth condition (e.g., the drift terms are locally Lipschitz and of polynomial growth w.r.t. the state variables). In the meantime, the weak convergence analysis of numerical algorithms concerning McKean-Vlasov SDEs has been treated; see e.g. [1, 11, 37], where the algorithm in [1] is based on new particle representations and [37] is concerned with the (antithetic) multilevel Monte-Carlo algorithms. Finally, we refer to [15] for importance sampling Monte Carlo methods for McKean-Vlasov SDEs with smooth drifts.

The literature [39] initiates the strong convergence analysis of EM numerical scheme for McKean-Vlasov SDEs with irregular coefficients, whereas the distribution involved is not simulated and the convergence rate is not revealed. Compared with the strong convergence analysis of numerical approximations for classical SDEs with irregular coefficients (see e.g. [2, 19, 31]), the counterpart for McKean-Vlasov SDEs with irregular coefficients is rather scarce. Nevertheless, in the present work, for two kinds of McKean-Vlasov SDEs with irregular coefficients (e.g., Hölder continuous drifts or diffusions), we aim to investigate the overall strong convergence rate of EM schemes based on stochastic interacting particle systems.

Next we start with some notation. Let $(\mathbb{R}^d, \langle\cdot,\cdot\rangle, |\cdot|)$ be the d-dimensional Euclidean space and $\mathbb{R}^d\otimes\mathbb{R}^d$ the set of all $d\times d$-matrices. $C^1(\mathbb{R}^d;\mathbb{R}^d\otimes\mathbb{R}^d)$ stands for the family of all Fréchet differentiable functions $f:\mathbb{R}^d\to\mathbb{R}^d\otimes\mathbb{R}^d$. $\mathscr{P}(\mathbb{R}^d)$ means the collection of all probability measures on $\mathbb{R}^d$. For $p>0$, if $\mu\in\mathscr{P}(\mathbb{R}^d)$ has a finite p-th moment, i.e., $\mu(|\cdot|^p):=\int_{\mathbb{R}^d}|x|^p\mu(\mathrm{d}x)<\infty$, we then formulate $\mu\in\mathscr{P}_p(\mathbb{R}^d)$. For $\mu,\nu\in\mathscr{P}_p(\mathbb{R}^d)$, the $\mathbb{W}_p$-Wasserstein distance between μ and ν is defined by

$$\mathbb{W}_p(\mu,\nu)=\inf_{\pi\in\mathcal{C}(\mu,\nu)}\Big(\int_{\mathbb{R}^d\times\mathbb{R}^d}|x-y|^p\pi(\mathrm{d}x,\mathrm{d}y)\Big)^{\frac{1}{1\vee p}},$$

where $\mathcal{C}(\mu,\nu)$ stands for the set of all couplings of μ and ν, that is, $\pi\in\mathcal{C}(\mu,\nu)$ if and only if $\pi(\cdot,\mathbb{R}^d)=\mu(\cdot)$ and $\pi(\mathbb{R}^d,\cdot)=\nu(\cdot)$. Let δ_x be Dirac's delta measure centered at the point $x\in\mathbb{R}^d$. For a random variable ξ, its law is written by $\mathscr{L}_\xi$. For any $t\geq 0$, let $C([0,t];\mathbb{R}^d)$ be the set of all continuous functions $f:[0,t]\to\mathbb{R}^d$ endowed with the uniform norm $\|f\|_{\infty,t}:=\sup_{0\leq s\leq t}|f(s)|$. $\lfloor a\rfloor$ stipulates the integer part of the real number $a\geq 0$.

Let $(\Omega,\mathscr{F},\mathbb{P})$ be a complete probability space endowed with the filtration $(\mathscr{F}_t)_{t\geq0}$, which satisfies the usual conditions (i.e., $\mathscr{F}_0$ contains all $\mathbb{P}$-null sets and $\mathscr{F}_t=\mathscr{F}_{t+}:=\lim_{s\downarrow t}\cap\mathscr{F}_s$, $t\geq 0$), and is rich enough such that, for any $\mu\in\mathscr{P}(\mathbb{R})$, there exists an $\mathbb{R}$-valued random variable ξ on $(\Omega,\mathscr{F}_0,\mathbb{P})$ such that $\mathscr{L}_\xi=\mu$. Let $W=(W_t)_{t\geq0}$ be a standard 1-dimensional Brownian motion on $(\Omega,\mathscr{F},(\mathscr{F}_t)_{t\geq0},\mathbb{P})$, which implies that W is adapted to $(\mathscr{F}_t)_{t\geq0}$ and W_t-W_s is independent of $\mathscr{F}_s$ for any $t>s\geq 0$.

The classical Cox-Ingersoll-Ross (CIR) model (see e.g. [13])

$$\mathrm{d}X_t=(\alpha-\delta X_t)\mathrm{d}t+|X_t|^h\mathrm{d}W_t,\quad \alpha,\delta>0,\quad h\geq\frac{1}{2}$$

deals basically with the development of interest rates. Nowadays, the CIR model has been extended considerably in different manners to characterize stochastic volatility. On the other hand, in applications, the distribution of a stochastic process can be regarded as a macro property. On account of the points above, in this paper we consider the following McKean-Vlasov SDE

$$\mathrm{d}X_t = b(X_t, \mu_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t, \quad t \geq 0, \quad X_0 = \xi, \tag{1.1}$$

where $\mu_t := \mathscr{L}_{X_t}$ is the law of X at time t, $b := b_1 + b_2$, $b_i : \mathbb{R} \times \mathscr{P}(\mathbb{R}) \to \mathbb{R}$, $i = 1, 2$, $\sigma : \mathbb{R} \to \mathbb{R}$. Throughout the paper, we assume that the initial value ξ is $\mathscr{F}_0$-measurable which implies that ξ is independent of W.

Now we introduce the definition of strong solutions to (1.1), which is standard in literature; see e.g. [38, Definition 1.1].

Definition 1.1. A continuous adapted process $(X_t)_{t\geq 0}$ on $\mathbb{R}$ is called a (strong) solution of (1.1), if

$$\int_0^t \mathbb{E}\big\{|b(X_s, \mu_s)| + |\sigma(X_s)|^2\big\}\mathrm{d}s < \infty, \quad t \geq 0,$$

and $\mathbb{P}$-a.s.

$$X_t = X_0 + \int_0^t b(X_s, \mu_s)\mathrm{d}s + \int_0^t \sigma(X_s)\mathrm{d}W_s, \qquad t \geq 0.$$

Remark 1.1. *By Burkhold-Davis-Gundy's inequality, Definition 1.1 yields* $\mathbb{E}\|X\|_{\infty,t} < \infty$ *if* $\mathbb{E}|X_0| < \infty$.

With regard to the drift term b and the diffusion term σ in (1.1), we assume

($\mathbf{H}_1$) For fixed $\mu \in \mathscr{P}(\mathbb{R})$, $\mathbb{R} \ni x \mapsto b_1(x, \mu)$ is continuous and non-increasing, and there exist constants $K_1 > 0$ and $\beta \in (0, 1]$ such that, for $x, y \in \mathbb{R}$ and $\mu, \nu \in \mathscr{P}_1(\mathbb{R})$,

$$|b_1(x, \mu) - b_1(x, \nu)| \leq K_1 \mathbb{W}_1(\mu, \nu), \ |b_1(x, \mu) - b_1(y, \mu)| \leq K_1|x - y|^\beta, \tag{1.2}$$

$$|b_2(x, \mu) - b_2(y, \nu)| \leq K_1\big\{|x - y| + \mathbb{W}_1(\mu, \nu)\big\}. \tag{1.3}$$

($\mathbf{H}_2$) There exist constants $K_2 > 0$ and $\alpha \in [\frac{1}{2}, 1]$ such that $|\sigma(x) - \sigma(y)| \leq K_2|x - y|^\alpha$, $x, y \in \mathbb{R}$.

The theorem below addresses the strong wellposedness of (1.1).

Theorem 1.2. *Assume* ($\mathbf{H}_1$) *and* ($\mathbf{H}_2$). *Then, for any* $X_0^\xi = \xi \in L^p(\Omega \to \mathbb{R}; \mathscr{F}_0, \mathbb{P})$, $p \geq 2$, *(1.1) has a unique strong solution* $(X_t^\xi)_{t\geq 0}$ *with the initial value* $X_0^\xi = \xi$ *such that*

$$\mathbb{E}\|X_t^\xi\|_{\infty,T}^p \leq C_{p,T}(1 + \mathbb{E}|\xi|^p) \tag{1.4}$$

for some constant $C_{p,T} > 0$ *dependent on the parameters* p *and* T.

The strong wellposedness of McKean-Vlasov SDEs with regular coefficients has been investigated considerably; see [7, 36, 38] among others. Meanwhile, the strong wellposedness of McKean-Vlasov SDEs with irregular coefficients has also received much attention. In particular, [3] treats (1.1) with an additive noise and b has the decomposable form $b(x,\mu) = \hat{b}(x,\mu)+\tilde{b}(x,\mu)$, where $\hat{b}$ is merely measurable and bounded and $\tilde{b}$ is Lipschitz continuous w.r.t. the spatial variables; [10] investigates the following non-degenerate McKean-Vlasov SDE

$$\mathrm{d}X_t = b(X_t, \mu_t(\psi_1))\mathrm{d}t + \sigma(X_t, \mu_t(\psi_2))\mathrm{d}W_t, \quad \mu_t(\psi_1) := \int_{\mathbb{R}^d} \psi_1(x)\mu_t(\mathrm{d}x),$$

in which σ is Lipschitz continuous w.r.t. the first component; [34] handles the following mean-filed SDE

$$\mathrm{d}X_t = \int_{\mathbb{R}^d} b(X_t, y)\mu_t(\mathrm{d}y)\mathrm{d}t + \int_{\mathbb{R}^d} \sigma(X_t, y)\mu_t(\mathrm{d}y)\mathrm{d}W_t \tag{1.5}$$

with a singular distribution dependent drift and a constant matrix σ; [21] investigates non-degenerate McKean-Vlasov SDEs under integrability condition, which whereas excludes the setup that the drift is of linear growth. Under the pathwise uniqueness, the unique strong solutions is constructed in [29] by the Euler polygonal approximation rather than using the famous Yamada-Watanabe theorem. As for the weak wellposedness of McKean-Vlasov SDEs, we refer to e.g. [20] under measure-dependent Lyapunov type conditions (see (3.3) therein), [27, 30] for the framework (1.5), and [25] (resp. [3]) for nondegenerate McKean-Vlasov SDEs with bounded (resp. unbounded) measurable drifts. Whereas Theorem 1.2 above shows that the McKean-Vlasov SDE we are interested in is strongly wellposed although the drift term and the diffusion term are Hölder continuous (so that they are irregular in a certain sense).

Since (1.1) is distribution-dependent, we exploit stochastic interacting particle systems to approximate it. Let $N \geq 1$ be an integer and $(X_0^i, W_t^i)_{1\leq i\leq N}$ be i.i.d. copies of (X_0, W_t). Consider the following stochastic non-interacting particle systems

$$\mathrm{d}X_t^i = b(X_t^i, \mu_t^i)\mathrm{d}t + \sigma(X_t^i)\mathrm{d}W_t^i, \quad t \geq 0, \quad i \in \mathcal{S}_N := \{1, \cdots, N\}, \tag{1.6}$$

where $\mu_t^i := \mathscr{L}_{X_t^i}$. By the weak uniqueness due to Theorem 1.2, we have $\mu_t = \mu_t^i, i \in \mathcal{S}_N$. Let $\tilde{\mu}_t^N$ be the empirical distribution associated with $X_t^1, \cdots, X_t^N$, i.e.,

$$\tilde{\mu}_t^N = \frac{1}{N}\sum_{j=1}^N \delta_{X_t^j}. \tag{1.7}$$

Furthermore, we need to consider the so-called stochastic N-interacting particle systems:

$$\mathrm{d}X_t^{i,N} = b(X_t^{i,N}, \hat{\mu}_t^N)\mathrm{d}t + \sigma(X_t^{i,N})\mathrm{d}W_t^i, \ t \geq 0, \ X_0^{i,N} = X_0^i, \ i \in \mathcal{S}_N, \tag{1.8}$$

where $\hat{\mu}_t^N$ means the empirical distribution corresponding to $X_t^{1,N}, \cdots, X_t^{N,N}$, namely,

$$\hat{\mu}_t^N := \frac{1}{N}\sum_{j=1}^N \delta_{X_t^{j,N}}.$$

We remark that particles $(X^i)_{i\in\mathcal{S}_N}$ are mutually independent whereas particles $(X^{i,N})_{i\in\mathcal{S}_N}$ are not independent but identically distributed. Under $(\mathbf{H}_1)$ and $(\mathbf{H}_2)$, the stochastic N-interacting particle systems (1.8) are strongly wellposed; see Lemma 3.1 below for more details.

To discretize (1.8) in time, we introduce the continuous time EM scheme defined as below: for any $\delta\in(0,\mathrm{e}^{-1})$,

$$\mathrm{d}X_t^{\delta,i,N}=b(X_{t_\delta}^{\delta,i,N},\hat\mu_{t_\delta}^{\delta,N})\mathrm{d}t+\sigma(X_{t_\delta}^{\delta,i,N})\mathrm{d}W_t^i,\quad t\ge0,\quad X_0^{\delta,i,N}=X_0^{i,N},\tag{1.9}$$

where $t_\delta:=\lfloor t/\delta\rfloor\delta$ and

$$\hat\mu_{k\delta}^{\delta,N}:=\frac{1}{N}\sum_{j=1}^N\delta_{X_{k\delta}^{\delta,j,N}},\qquad k\ge0.$$

The following result states that the continuous time EM scheme corresponding to stochastic interacting particle systems converges strongly to the non-interacting particle systems whenever the particle number goes to infinity and the stepsize approaches to zero. Most importantly, the corresponding overall convergence rate is provided.

Theorem 1.3. *Assume $(\mathbf{H}_1)$ and $(\mathbf{H}_2)$ and suppose further $X_0\in L^p(\Omega\to\mathbb{R};\mathscr{F}_0,\mathbb{P})$ for some $p>4$. Then, for any $T>0$, there exists a constant $C_T>0$, independent of δ and N, such that*

$$\sup_{i\in\mathcal{S}_N}\mathbb{E}\,\|X^i-X^{\delta,i,N}\|_{\infty,T}\le C_T\begin{cases}N^{-\frac18}+\Big(\frac{1}{\ln\frac1\delta}\Big)^{1/2}, & \alpha=\frac12\\ N^{-\frac{2\alpha-1}{4}}+\delta^{\frac{(2\alpha-1)^2}{2}}+\delta^{\frac{\beta(2\alpha-1)}{2}}, & \alpha\in(\frac12,1]\end{cases}\tag{1.10}$$

and

$$\sup_{i\in\mathcal{S}_N}\mathbb{E}\,\|X^i-X^{\delta,i,N}\|_{\infty,T}^2\le C_T\begin{cases}N^{-\frac14}+\frac{1}{\ln\frac1\delta}, & \alpha=\frac12\\ N^{-\frac14}+\delta^{\frac{2\alpha-1}{2}}+\delta^{\frac{\beta}{2}}, & \alpha\in(\frac12,1)\\ N^{-\frac14}+\delta^{\beta}, & \alpha=1.\end{cases}\tag{1.11}$$

The assumption on $X_0\in L^p(\Omega\to\mathbb{R};\mathscr{F}_0,\mathbb{P})$ for some $p>4$ is set to ensure that the Glivenko-Cantelli convergence under the Wasserstein distance (see e.g. [7, Theorem 5.8]) is available. According to Theorem 1.3, it is preferable to measure the convergence between the non-interacting particle systems and the continuous time EM schemes of the corresponding stochastic interacting particle systems in a lower order moment. Moreover, Theorem 1.3 extends [2, 19] to McKean-Vlasov SDEs with Hölder continuous diffusions.

In the preceding section, we focused mainly on McKean-Vlasov SDEs, where, in particular, the diffusion term is Hölder continuous. We now move forward to consider McKean-Vlasov SDEs, in which the drift coefficients are Hölder continuous w.r.t. the spatial variables, the diffusion terms are assumed to be non-degenerate, and both of them are Lipschitz in law under the $\mathbb{W}_2$-distance (i.e., $\mathbb{W}_2$-Lipschitz). In the sequel, for $d\ge1$, we work on the following McKean-Vlasov SDE

$$\mathrm{d}X_t=b(X_t,\mu_t)\mathrm{d}t+\sigma(X_t,\mu_t)\mathrm{d}W_t,\qquad t\ge0,\quad X_0=\xi,\tag{1.12}$$

where $b:\mathbb{R}^d\times\mathscr{P}(\mathbb{R}^d)\to\mathbb{R}^d$, $\sigma:\mathbb{R}^d\times\mathscr{P}(\mathbb{R}^d)\to\mathbb{R}^d\otimes\mathbb{R}^d$, $(W_t)_{t\geq0}$ is a d-dimensional Brownian motion on some complete filtration probability space $(\Omega,\mathscr{F},(\mathscr{F}_t)_{t\geq0},\mathbb{P})$, and ξ is an $\mathscr{F}_0$-measurable random variable.

Concerning (1.12), we assume

($\mathbf{A}_1$) For each $\mu\in\mathscr{P}_2(\mathbb{R}^d)$, $\sigma(\cdot,\mu)\in C^1(\mathbb{R}^d;\mathbb{R}^d\otimes\mathbb{R}^d)$, $\mathbb{R}^d\ni x\mapsto\sigma(x,\mu)$ is invertible for all $\mu\in\mathscr{P}_2(\mathbb{R}^d)$, and

$$\|b\|_\infty+\|\sigma\|_\infty+\|\nabla\sigma\|_\infty+\|\sigma^{-1}\|_\infty<\infty, \tag{1.13}$$

where ∇ denotes the gradient operator w.r.t. the space variable and $\|\cdot\|_\infty$ means the uniform norm over the space variables and the measure variables.

($\mathbf{A}_2$) There exist constants $K>0,\alpha\in(0,1]$ such that for any $x,y\in\mathbb{R}^d$ and $\mu,\nu\in\mathscr{P}_2(\mathbb{R}^d)$,

$$|b(x,\mu)-b(y,\nu)|\leq K\big\{|x-y|^\alpha+\mathbb{W}_2(\mu,\nu)\big\}, \tag{1.14}$$

$$|\sigma(x,\mu)-\sigma(x,\nu)|\leq K\mathbb{W}_2(\mu,\nu). \tag{1.15}$$

Theorem 1.4. *Assume* ($\mathbf{A}_1$) *and* ($\mathbf{A}_2$)*. Then, for any $X_0\in L^2(\Omega\to\mathbb{R}^d;\mathscr{F}_0,\mathbb{P})$,* (1.12) *has a unique strong solution $(X_t)_{t\geq0}$ such that for some constant $C_T>0$,*

$$\mathbb{E}\|X\|_{T,\infty}^2\leq C_T(1+\mathbb{E}|X_0|^2). \tag{1.16}$$

The proof of Theorem 1.4 is deferred into the Appendix A. One can also refer to [3, Theorems 2.3 & 2.7] or [25, Theorem 3.2] for the weak wellposedness when σ is not dependent on the measure variable.

Consider the stochastic non-interacting particle systems associated with (1.12)

$$\mathrm{d}X_t^i=b(X_t^i,\mu_t^i)\mathrm{d}t+\sigma(X_t^i,\mu_t^i)\mathrm{d}W_t^i,\quad t\geq0,\quad i\in\mathcal{S}_N, \tag{1.17}$$

where $(X_0^i,W^i)_{1\leq i\leq N}$ are i.i.d copies of (X_0,W). The stochastic interacting particle systems corresponding to (1.12) solve

$$\mathrm{d}X_t^{i,N}=b(X_t^{i,N},\hat\mu_t^N)\mathrm{d}t+\sigma(X_t^{i,N},\hat\mu_t^N)\mathrm{d}W_t^i,\quad t\geq0,\quad i\in\mathcal{S}_N. \tag{1.18}$$

To discretize (1.18), we further need to consider the following continuous time EM scheme

$$\mathrm{d}X_t^{\delta,i,N}=b(X_{t_\delta}^{\delta,i,N},\hat\mu_{t_\delta}^{\delta,N})\mathrm{d}t+\sigma(X_{t_\delta}^{\delta,i,N},\hat\mu_{t_\delta}^{\delta,N})\mathrm{d}W_t^i,\quad t\geq0,\quad X_0^{\delta,i,N}=X_0^{i,N}. \tag{1.19}$$

By following the same line of the proof for Lemma 3.1 below, (1.18) has a weak solution. On the other hand, employing Zvonkin's transformation, we infer that (1.18) is pathwise unique. Henceforth, the Yamada-Watanabe theorem (see e.g. [32, Theorem E.1.8]) yields that (1.18) is strongly wellposed under ($\mathbf{A}_1$) and ($\mathbf{A}_2$).

Another contribution in the present paper is concerned with strong convergence analysis between non-interacting particle systems and continuous time EM schemes of stochastic interacting particle systems corresponding to the McKean-Vlasov SDE (1.12), where the drift is Hölder continuous w.r.t. the spatial variables and the diffusion is Lipschitz continuous under the $\mathbb{W}_2$-Wasserstein distance.

Theorem 1.5. *Assume* ($\mathbf{A}_1$) *and* ($\mathbf{A}_2$) *and suppose further* $X_0 \in L^p(\Omega \to \mathbb{R}^d; \mathscr{F}_0, \mathbb{P})$ *for some* $p > 4$. *Then, for any* $T > 0$, *there exists a constant* $C_T > 0$, *independent of* N *and* δ, *such that*

$$\sup_{i\in\mathcal{S}_N} \mathbb{E}\|X^i - X^{\delta,i,N}\|^2_{\infty,T} \le C_T \begin{cases} N^{-\frac{1}{2}} + \delta^\alpha, & d < 4 \\ N^{-\frac{1}{2}}\log N + \delta^\alpha, & d = 4 \\ N^{-\frac{2}{d}} + \delta^\alpha, & d > 4. \end{cases} \tag{1.20}$$

Remark 1.6. *For the standard SDEs, the boundedness on the drifts can be dropped via a stopping time technique (e.g., a localization approach); see e.g. [2]. However, concerning the McKean-Vlasov SDEs, as stated in [17, Remark 3.4], the localization approach does not work any more since applying a stopping time to a single particle does not allow us to fully bound the coefficients and moreover destroys the result of all particles being identically distributed. By a close inspection of the proof of Theorem 1.5, the assertion in Theorem 1.5 can indeed be extended to the case* $b = b_1 + b_2$, *where* b_1 *satisfies* ($\mathbf{A}_1$) *and* ($\mathbf{A}_2$) *and* b_2 *is Lipschitz continuous with respect to both space and measure variables which allows* b_2 *to be unbounded.*

The remainder of this paper is structured as follows: In Section 2, the strong wellposedness of (1.1) is addressed by Yamada-Watanabe's approximation; Section 3 is devoted to completing the proof of Theorem 1.3 via Yamada-Watanabe's approach; Section 4 aims to finish the proof of Theorem 1.5 by employing Zvonkin's transformation; In the Appendix section, we show that (1.12) is strongly wellposed.

2 Proof of Theorem 1.2

To complete the proofs of Theorems 1.2 and 1.3, we shall adopt the Yamada-Watanabe approximation approach (see e.g. [19, 23]), where the essential ingredient is to approximate the function $\mathbb{R} \ni x \mapsto |x|$ in an appropriate manner. For $\gamma > 1$ and $\varepsilon \in (0, 1)$, one trivially has $\int_{\varepsilon/\gamma}^{\varepsilon} \frac{1}{x}\mathrm{d}x = \ln\gamma$ so that there exists a continuous function $\psi_{\gamma,\varepsilon} : \mathbb{R}_+ \to \mathbb{R}_+$ with the support $[\varepsilon/\gamma, \varepsilon]$ such that

$$0 \le \psi_{\gamma,\varepsilon}(x) \le \frac{2}{x\ln\gamma}, \quad x > 0, \quad \int_{\varepsilon/\gamma}^{\varepsilon} \psi_{\gamma,\varepsilon}(r)\mathrm{d}r = 1.$$

By a direct calculation, the following mapping

$$\mathbb{R} \ni x \mapsto V_{\gamma,\varepsilon}(x) := \int_0^{|x|}\int_0^y \psi_{\gamma,\varepsilon}(z)\mathrm{d}z\mathrm{d}y \tag{2.1}$$

is twice differentiable (in this sense, we write $V_{\gamma,\varepsilon} \in C^2$) and satisfies

$$|x| - \varepsilon \le V_{\gamma,\varepsilon}(x) \le |x|, \quad \mathrm{sgn}(x)V'_{\gamma,\varepsilon}(x) \in [0, 1], \quad x \in \mathbb{R}, \tag{2.2}$$

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

$$0 \le V''_{\gamma,\varepsilon}(x) \le \frac{2}{|x|\ln\gamma}\mathbf{1}_{[\varepsilon/\gamma,\varepsilon]}(|x|), \quad x \in \mathbb{R}. \tag{2.3}$$

Herein, $V'_{\gamma,\varepsilon}$ (resp. $V''_{\gamma,\varepsilon}$) denotes the first(resp. second) order derivative of $V_{\gamma,\varepsilon}$.

To obtain existence of solutions to (1.1), for $k \geq 1$, we consider the distribution-iterated SDE

$$\mathrm{d}X_t^{(k)} = b(X_t^{(k)}, \mu_t^{(k-1)})\mathrm{d}t + \sigma(X_t^{(k)})\mathrm{d}W_t, \; X_t^{(0)} = X_0^{(k)} = X_0, \; t \in [0, T], \tag{2.4}$$

where $\mu_t^{(k)} := \mathscr{L}_{X_t^{(k)}}$. For each fixed $k \geq 1$, according to [23, Theorem 3.2, p.168], (2.4) has a unique solution $(X_t^{(k)})_{t \geq 0}$. The lemma below shows that the second order moment is uniformly bounded in a finite time interval.

Lemma 2.1. *Assume* $(\mathbf{H}_1)$ *and* $(\mathbf{H}_2)$*. Then, for each* $k \geq 1$ *and* $X_0 \in L^2(\Omega \to \mathbb{R}^d; \mathscr{F}_0, \mathbb{P})$*, there exists a nondecreasing positive function* $T \mapsto C_T$ *independent of* k *such that*

$$\mathbb{E}\, \|X^{(k)}\|_{\infty,T}^2 \leq C_T\big(1 + \mathbb{E}|X_0|^2\big). \tag{2.5}$$

Proof. The proof of Lemma 2.1 is based on an inductive argument. From $(\mathbf{H}_1)$ and $(\mathbf{H}_2)$, it is easy to see that for any $x \in \mathbb{R}$ and $\mu \in \mathscr{P}_1(\mathbb{R})$,

$$|b(x, \mu)| \leq 2K_1\big(|x| + \mathbb{W}_1(\mu, \delta_0)\big) + c_1, \quad |\sigma(x)|^2 \leq 2K_2^2|x|^2 + c_2, \tag{2.6}$$

where $c_1 := K_1 + |b(0, \delta_0)|$ and $c_2 := 2(K_2 + |\sigma(0)|)^2$. For each integer $N \geq 1$, define the stopping time $\tau_N = \inf\big\{t \geq 0 : |X_t^{(1)}| \geq N\big\}$. Below, we set $t \in [0, T]$. By Hölder's inequality and Burkhold-Davis-Gundy's inequality, it follows from (2.6) that

$$\begin{aligned}
\mathbb{E}\|X^{(1)}\|_{\infty,t\wedge\tau_N}^2 &\leq 3\Big\{\mathbb{E}|X_0|^2 + t\,\mathbb{E}\int_0^{t\wedge\tau_N} |b(X_s^{(1)}, \mu_s^{(0)})|^2\mathrm{d}s + 4\,\mathbb{E}\int_0^{t\wedge\tau_N} |\sigma(X_s^{(1)})|^2\mathrm{d}s\Big\} \\
&\leq C_1(1+t)\Big\{1 + \mathbb{E}|X_0|^2 + \mathbb{E}\int_0^{t\wedge\tau_N} \big\{|X_s^{(1)}|^2 + \mathbb{W}_1(\mu_s^{(0)}, \delta_0)^2\big\}\mathrm{d}s\Big\} \\
&\leq C_2(1+t)\Big\{1 + \mathbb{E}|X_0|^2 + \int_0^t \mathbb{E}|X_{s\wedge\tau_N}^{(1)}|^2 + \int_0^t \mu_s^{(0)}(|\cdot|)^2\mathrm{d}s\Big\} \\
&\leq C_2(1+t)^2\big(1 + \mathbb{E}|X_0|^2\big) + C_2(1+t)\int_0^t \mathbb{E}|X_{s\wedge\tau_N}^{(1)}|^2\mathrm{d}s
\end{aligned}$$

for some constants $C_1, C_2 > 0$. This, together with Gronwall's inequality, implies

$$\mathbb{E}\|X^{(1)}\|_{\infty,t\wedge\tau_N}^2 \leq C_2(1+t)^2\mathrm{e}^{C_2(1+t)t}\big(1 + \mathbb{E}|X_0|^2\big).$$

Thus, (2.5) holds with $k = 1$ by making use of Fatou's lemma. Next, we aim to show that (2.5) still holds true for $k = n+1$ once (2.5) is valid for some $k = n$. Indeed, this can be handled in the same manner by using the triple $(X^{(n+1)}, X^{(n)}, \mu^{(n)})$ in lieu of $(X^{(1)}, X^{(0)}, \mu^{(0)})$. We therefore complete the proof. □

With the approximate function $V_{\gamma,\varepsilon}$, introduced in (2.1), and Lemma 2.1 at hand, we are in a position to complete

Proof of Theorem 1.2. Below, we shall fix the time terminal $T > 0$. For notation brevity, we set $Z_t^{(k)} := X_t^{(k)} - X_t^{(k-1)}$ and $V_\varepsilon := V_{\mathrm{e}^{\frac{1}{\varepsilon}},\varepsilon}$ (i.e., $\gamma = \mathrm{e}^{\frac{1}{\varepsilon}}$ in (2.1)). By Itô's formula, for any $\lambda \geq 0$, it follows that

$$
\begin{aligned}
\mathrm{e}^{-\lambda t}V_\varepsilon(Z_t^{(k+1)}) &= -\lambda\int_0^t \mathrm{e}^{-\lambda s}V_\varepsilon(Z_s^{(k+1)})\mathrm{d}s\\
&\quad+\int_0^t \mathrm{e}^{-\lambda s}V_\varepsilon'(Z_s^{(k+1)})\big\{b(X_s^{(k+1)},\mu_s^{(k)})-b(X_s^{(k)},\mu_s^{(k-1)})\big\}\mathrm{d}s\\
&\quad+\frac{1}{2}\int_0^t \mathrm{e}^{-\lambda s}V_\varepsilon''(Z_s^{(k+1)})\big\{\sigma(X_s^{(k+1)})-\sigma(X_s^{(k)})\big\}^2\mathrm{d}s\\
&\quad+\int_0^t \mathrm{e}^{-\lambda s}V_\varepsilon'(Z_s^{(k+1)})\big\{\sigma(X_s^{(k+1)})-\sigma(X_s^{(k)})\big\}\mathrm{d}W_s\\
&=: I_{1,\varepsilon}^\lambda(t)+I_{2,\varepsilon}^\lambda(t)+I_{3,\varepsilon}^\lambda(t)+I_{4,\varepsilon}^\lambda(t).
\end{aligned}\tag{2.7}
$$

By virtue of (2.2), one obviously has

$$
I_{1,\varepsilon}^\lambda(t)\le \lambda\varepsilon t-\lambda\int_0^t \mathrm{e}^{-\lambda s}|Z_s^{(k+1)}|\mathrm{d}s.\tag{2.8}
$$

By recalling $b=b_1+b_2$, we deduce that

$$
\begin{aligned}
I_{2,\varepsilon}^\lambda(t) &\le \int_0^t \mathrm{e}^{-\lambda s}V_\varepsilon'(Z_s^{(k+1)})\big\{b_1(X_s^{(k+1)},\mu_s^{(k)})-b_1(X_s^{(k)},\mu_s^{(k)})\big\}\mathrm{d}s\\
&\quad+\int_0^t \mathrm{e}^{-\lambda s}\big|V_\varepsilon'(Z_s^{(k+1)})\big|\cdot\big|b_1(X_s^{k},\mu_s^{(k)})-b_1(X_s^{(k)},\mu_s^{(k-1)})\big|\mathrm{d}s\\
&\quad+\int_0^t \mathrm{e}^{-\lambda s}\big|V_\varepsilon'(Z_s^{(k+1)})\big|\cdot\big|b_2(X_s^{(k+1)},\mu_s^{(k)})-b_2(X_s^{(k)},\mu_s^{(k-1)})\big|\mathrm{d}s\\
&\le \int_0^t \mathrm{e}^{-\lambda s}\big|V_\varepsilon'(Z_s^{(k+1)})\big|\cdot\big|b_1(X_s^{k},\mu_s^{(k)})-b_1(X_s^{(k)},\mu_s^{(k-1)})\big|\mathrm{d}s\\
&\quad+\int_0^t \mathrm{e}^{-\lambda s}\big|V_\varepsilon'(Z_s^{(k+1)})\big|\cdot\big|b_2(X_s^{(k+1)},\mu_s^{(k)})-b_2(X_s^{(k)},\mu_s^{(k-1)})\big|\mathrm{d}s\\
&\le 2K_1\int_0^t \mathrm{e}^{-\lambda s}\big\{|Z_s^{(k+1)}|+\mathbb{W}_1(\mu_s^{(k)},\mu_s^{(k-1)})\big\}\mathrm{d}s.
\end{aligned}\tag{2.9}
$$

Herein, the second inequality holds since

$$
\begin{aligned}
&V_\varepsilon'(x-y)\ge 0,\quad b_1(x,\cdot)-b_1(y,\cdot)\le 0,\quad x\ge y;\\
&V_\varepsilon'(x-y)< 0,\quad b_1(x,\cdot)-b_1(y,\cdot)\ge 0,\quad x< y,
\end{aligned}
$$

where we used the fact that $x\mapsto b_1(x,\cdot)$ is non-increasing thanks to $(\mathbf{H}_1)$ and $\mathrm{sgn}(x)V_\varepsilon'(x)\in[0,1]$ due to (2.2), and the third inequality is true by taking advantage of (1.2), (1.3) as well as (2.2). Next, by utilizing $(\mathbf{H}_2)$ and (2.3) with $\gamma=\mathrm{e}^{\frac{1}{\varepsilon}}$ and using $\alpha\in[1/2,1]$, we infer

$$
I_{3,\varepsilon}^\lambda(t)\le\frac{K_2^2\varepsilon}{2}\int_0^t \mathrm{e}^{-\lambda s}|Z_s^{(k+1)}|^{2\alpha-1}\mathbf{1}_{[\frac{\varepsilon}{\mathrm{e}^{1/\varepsilon}},\varepsilon]}(|Z_s^{(k+1)}|)\mathrm{d}s\le\frac{1}{2}c_\lambda K_2^2 t\varepsilon,\tag{2.10}
$$

where $c_\lambda := \big\{1\mathbf{1}_{\{\lambda=0\}} + \frac{1}{\lambda}\mathbf{1}_{\{\lambda>0\}}\big\}$. So, taking advantage of (2.2), (2.9) and (2.10) leads to

$$\begin{aligned}\mathrm{e}^{-\lambda t}|Z_t^{(k+1)}| &\le \big(1+(\lambda+c_\lambda K_2^2/2)t\big)\varepsilon - (\lambda-2K_1)\int_0^t \mathrm{e}^{-\lambda s}|Z_s^{(k+1)}|\mathrm{d}s\\ &\quad + 2K_1\int_0^t \mathrm{e}^{-\lambda s}\mathbb{E}|Z_s^{(k)}|\mathrm{d}s + I_{4,\varepsilon}^\lambda(t).\end{aligned} \tag{2.11}$$

By (2.2) and (2.5), we have $\mathbb{E}I_{4,\varepsilon}^\lambda(t) = 0$. Whence, choosing $\lambda = 0$, approaching $\varepsilon \downarrow 0$, and employing Gronwall's inequality gives

$$\mathbb{E}|Z_t^{(k+1)}| \le 2K_1\mathrm{e}^{2K_1 t}\int_0^t \mathbb{E}|Z_s^{(k)}|\mathrm{d}s. \tag{2.12}$$

For notation simplicity, set

$$|[Z^{(k)}]|_{\lambda,t} := \big\|\mathrm{e}^{-\lambda\cdot}\mathbb{E}|Z_\cdot^{(k)}|\big\|_{\infty,t}, \qquad \||[Z^{(k)}]|\|_{\lambda,t} := \mathbb{E}\big\|\mathrm{e}^{-\lambda\cdot}|Z_\cdot^{(k)}|\big\|_{\infty,t}, \quad t\ge 0.$$

In the sequel, we take $\lambda \ge 2K_1\mathrm{e}^{1+2K_1T}$ and let $t\in[0,T]$. In terms of (2.12), it follows from an inductive trick that

$$\begin{aligned}|[Z^{(k+1)}]|_{\lambda,t} &\le 2K_1\sup_{0\le s\le t}\Big(\mathrm{e}^{-(\lambda-2K_1)s}\int_0^s \mathrm{e}^{\lambda r}\big(\mathrm{e}^{-\lambda r}\mathbb{E}|Z_r^{(k)}|\big)\mathrm{d}r\Big)\\ &\le 2K_1\sup_{0\le s\le t}\Big(\mathrm{e}^{-(\lambda-2K_1)s}|[Z^{(k)}]|_{\lambda,s}\int_0^s \mathrm{e}^{\lambda r}\mathrm{d}r\Big)\\ &\le \frac{2K_1}{\lambda}\mathrm{e}^{2K_1 t}|[Z^{(k)}]|_{\lambda,t} \le \mathrm{e}^{-1}|[Z^{(k)}]|_{\lambda,t} \le \mathrm{e}^{-k}|[Z^{(1)}]|_{\lambda,T},\end{aligned} \tag{2.13}$$

where the third inequality holds since $s\mapsto |[Z^{(k)}]|_{\lambda,s}$ is non-decreasing and the last two inequality is due to $\lambda \ge 2K_1\mathrm{e}^{1+2K_1T}$. Subsequently, by invoking Burkhold-Davis-Gundy's inequality, Jensen's inequality and (2.2) and taking $(\mathbf{H}_2)$ into account followed by approaching $\varepsilon\downarrow 0$, we deduce from (2.11) and $\alpha\in[1/2,1]$ that

$$\begin{aligned}\||[Z^{(k+1)}]|\|_{\lambda,t} &\le 2K_1\int_0^t |[Z^{(k)}]|_{\lambda,s}\mathrm{d}s + 4\sqrt{2}K_2\Big(\int_0^t |[Z^{(k+1)}]|_{\lambda,s}\mathrm{d}s\Big)^{\frac12}\mathbf{1}_{\{\alpha=\frac12\}}\\ &\quad + \Big\{\frac12\||[Z^{(k+1)}]|\|_{\lambda,t} + 16K_2^2\int_0^t |[Z^{(k+1)}]|_{\lambda,s}^{2\alpha-1}\mathrm{d}s\Big\}\mathbf{1}_{\{\alpha\in(\frac12,1]\}}.\end{aligned}$$

This, in addition to (2.13), implies that there exists a constant $C_T>0$ such that

$$\||[Z^{(k+1)}]|\|_{\lambda,t} \le C_T\exp\Big(-\Big(\frac12\mathbf{1}_{\{\alpha=\frac12\}} + (2\alpha-1)\mathbf{1}_{\{\alpha\in(\frac12,1]\}}\Big)k\Big).$$

As a result, there exists an $(\mathscr{F}_t)_{t\in[0,T]}$-adapted continuous stochastic process $(X_t)_{t\in[0,T]}$ with $X_0=\xi$ and $\mu_t = \mathscr{L}_{X_t}$ such that

$$\lim_{k\to\infty}\sup_{t\in[0,T]}\mathbb{W}_1(\mu_t^{(k)},\mu_t) \le \lim_{k\to\infty}\mathbb{E}\|X^{(k)}-X\|_{\infty,t} = 0. \tag{2.14}$$

From $(\mathbf{H}_1)$, we infer that

$$\int_0^t |b(X_s^{(k)},\mu_s^{(k-1)})-b(X_s,\mu_s)|\mathrm{d}s \le \int_0^t \big|b_1(X_s^{(k)},\mu_s)-b_1(X_s,\mu_s)\big|\mathrm{d}s + 2K_1\int_0^t \big\{|X_s^{(k)}-X_s|+\mathbb{W}_1(\mu_s^{(k-1)},\mu_s)\big\}\mathrm{d}s.$$

By (1.2), and the continuity of $b_1(\cdot,\mu)$ for any $\mu\in\mathscr{P}_1(\mathbb{R})$, we can apply (2.14) and the dominated convergence theorem to obtain

$$\lim_{k\to\infty}\int_0^T \mathbb{E}\big|b(X_t^{(k)},\mu_t^{(k-1)})-b(X_t,\mu_t)\big|\mathrm{d}t=0. \tag{2.15}$$

Again, from (2.14) we find

$$\lim_{k\to\infty}\mathbb{E}\Big(\sup_{0\le t\le T}\Big|\int_0^t \big(\sigma(X_s^{(k)})-\sigma(X_s)\big)\mathrm{d}W_s\Big|\Big)=0, \tag{2.16}$$

since, by $(\mathbf{H}_2)$, Burkhold-Davis-Gundy's inequality, Young's inequality and Jensen's inequality, we have

$$\begin{aligned}\mathbb{E}\Big(\sup_{0\le t\le T}\Big|\int_0^t \big(\sigma(X_s^{(k)})-\sigma(X_s)\big)\mathrm{d}W_s\Big|\Big) &\le 4\sqrt{2}K_2\,\mathbb{E}\Big(\int_0^T |X_t^{(k)}-X_t|^{2\alpha}\mathrm{d}t\Big)^{\frac12}\\ &\le 4\sqrt{2}K_2\Big(\int_0^T \mathbb{E}|X_t^{(k)}-X_t|\mathrm{d}t\Big)^{\frac12}\mathbf{1}_{\{\alpha=\frac12\}}\\ &\quad+\Big\{\mathbb{E}\|X^{(k)}-X\|_{\infty,T}+16K_2^2\int_0^T \big(\mathbb{E}|X_t^{(k)}-X_t|\big)^{2\alpha-1}\mathrm{d}t\Big\}\mathbf{1}_{\{\alpha\in(\frac12,1]\}}.\end{aligned}$$

Now with (2.15) and (2.16) at hand, by taking $k\to\infty$ in the following SDE

$$X_t^{(k)}=\xi+\int_0^t b(X_s^{(k)},\mu_s^{(k-1)})\mathrm{d}s+\int_0^t \sigma(X_s^{(k)})\mathrm{d}W_s,\quad k\ge1, t\in[0,T],$$

we derive (by extracting a suitable subsequence) $\mathbb{P}$-a.s.

$$\mathrm{d}X_t=b(X_t,\mu_t)\mathrm{d}t+\sigma(X_t)\mathrm{d}W_t,\quad t\in[0,T],$$

so that the existence of strong solutions to (1.1) is now available.

In the sequel, we prove the uniqueness of (1.1). To this end, we assume that $(X_t^{1,\xi})_{t\ge0}$ and $(X_t^{2,\xi})_{t\ge0}$ are solutions to (1.1) with the same initial value ξ. For $\Gamma_t:=X_t^{1,\xi}-X_t^{2,\xi}$, by following the argument used in deriving (2.12), one has

$$\mathbb{E}|\Gamma_t|\le 2K_1\mathrm{e}^{2K_1t}\int_0^t \mathbb{E}|\Gamma_s|\mathrm{d}s,$$

which, by invoking Gronwall's inequality and Remark 1.1, yields the uniqueness.

Finally, we intend to show that the p-th moment of the solution process is uniformly bounded in a finite time interval. For any $n \geq 1$, define the stopping time $\tau_n = \inf\{t \geq 0 : |X_t| \geq n\}$. Then, applying Hölder's inequality and Burkhold-Davis-Gundy's inequality and utilizing (2.6) yield

$$\begin{aligned}\mathbb{E}\|X\|^p_{\infty,t\wedge\tau_n} &\leq 3^{p-1}\Big\{\mathbb{E}|\xi|^p + t^{p-1}\mathbb{E}\int_0^{t\wedge\tau_n} |b(X_s,\mu_s)|^p \mathrm{d}s + c_p\mathbb{E}\Big(\int_0^{t\wedge\tau_n} |\sigma(X_s)|^2\mathrm{d}s\Big)^{p/2}\Big\}\\ &\leq \frac{1}{2}\mathbb{E}\|X\|^p_{\infty,t\wedge\tau_n} + C_t\int_0^t \mathbb{E}\|X\|^p_{\infty,s\wedge\tau_n}\mathrm{d}s + C_t\big(1+\mathbb{E}|\xi|^p\big) + C_t\int_0^t (\mathbb{E}|X_s|)^p\mathrm{d}s\end{aligned}$$

for some positive increasing function $t \mapsto C_t$. Thus, Gronwall's inequality and Remark 1.1 yield

$$\mathbb{E}\|X\|^p_{\infty,T\wedge\tau_n} \leq C_T\big(1+\mathbb{E}|\xi|^p\big) \tag{2.17}$$

for some constant $C_T > 0$. The estimate above implies

$$C_T\big(1+\mathbb{E}|\xi|^p\big) \geq \mathbb{E}|X_{T\wedge\tau_n}|^p \geq \mathbb{E}\big(|X_{\tau_n}|^p\mathbf{1}_{\{\tau_n\leq T\}}\big) = n^p\,\mathbb{P}(\tau_n \leq T).$$

Whence, one has

$$\mathbb{P}(\tau_n \leq T) \leq \frac{1}{n^p}C_T(1+\mathbb{E}|\xi|^p), \quad p \geq 2,$$

which yields that the series $\sum_{n=1}^\infty \mathbb{P}(\tau_n \leq T)$ is convergent so that, by the Borel-Cantelli lemma, we conclude that $\lim_{n\to\infty}\tau_n =: \tau_\infty > T$ a.s. Due to the arbitrariness of T, we have $\tau_\infty = \infty$, a.s. So, (1.4) holds by taking $n \uparrow \infty$ in (2.17) and using Fatou's lemma. □

3 Proof of Theorem 1.3

In this section, we intend to finish the proof of Theorem 1.3. Before we start, we prepare some auxiliary materials. The lemma below addresses the well-posedness of the stochastic N-interacting particle systems (1.8).

Lemma 3.1. *Assume* $(\mathbf{H}_1)$ *and* $(\mathbf{H}_2)$. *Then, for each* $N \geq 1$ *and any* $X_0^i \in L^2(\Omega \to \mathbb{R}^d; \mathscr{F}_0, \mathbb{P})$, (1.8) *admits a strong solution* $(X_t^{i,N})_{t\geq 0}$ *and there exists a nondecreasing positive function* $T \mapsto C_T$ *independent of* i, *such that*

$$\sup_{i\in\mathcal{S}_N}\mathbb{E}\|X^{i,N}\|^2_{\infty,T} \leq C_T(1+\mathbb{E}|X_0^i|^2). \tag{3.1}$$

Proof. For $x := (x_1,\cdots,x_N)^* \in \mathbb{R}^N$, $x_i \in \mathbb{R}$, set

$$\tilde{\mu}_x^N := \frac{1}{N}\sum_{i=1}^N \delta_{x_i}, \quad \hat{b}(x) := \big(b(x_1,\tilde{\mu}_x^N),\cdots,b(x_N,\tilde{\mu}_x^N)\big)^*,$$
$$\hat{\sigma}(x) := \mathrm{diag}\big(\sigma(x_1),\cdots,\sigma(x_N)\big), \quad \hat{W}_t := \big(W_t^1,\cdots,W_t^N\big)^*.$$

Obviously, $(\hat{W}_t)_{t\geq 0}$ is an N-dimensional Brownian motion. Then, (1.8) can be reformulated as

$$\mathrm{d}X_t = \hat{b}(X_t)\mathrm{d}t + \hat{\sigma}(X_t)\mathrm{d}\hat{W}_t, \quad t \geq 0. \tag{3.2}$$

By the Yamada-Watanabe theorem (see e.g. [32, Theorem E.1.8]), to show that (3.2) has a unique strong solution, it is sufficient to verify that (3.2) possesses a weak solution and that it is pathwise unique. By (1.2), (1.3) and $(\mathbf{H}_2)$, a straightforward calculation yields

$$(3.3)\qquad |\hat{b}(x)|+\|\hat{\sigma}(x)\|_{\mathrm{HS}}\le\left(\sum_{i=1}^N|b(x_i,\tilde{\mu}_x^N)|^2\right)^{\frac12}+\Big(\sum_{i=1}^N\sigma(x_i)^2\Big)^{1/2}\le C_N(1+|x|),\quad x\in\mathbb{R}^N$$

for some constant $C_N>0$, that is, $\hat{b}$ and $\hat{\sigma}$ are at most of linear growth in $\mathbb{R}^N$. Observe that

$$(3.4)\qquad \frac{1}{N}\sum_{j=1}^N(\delta_{x_j}\times\delta_{y_j})\in\mathcal{C}(\tilde{\mu}_x^N,\tilde{\mu}_y^N),\qquad x_j,y_j\in\mathbb{R},$$

so that we have

$$(3.5)\qquad \mathbb{W}_1(\tilde{\mu}_x^N,\tilde{\mu}_y^N)\le\frac{1}{N}\sum_{j=1}^N|x_j-y_j|.$$

This, together with (1.2) and (1.3), besides $(\mathbf{H}_2)$, implies that for any $x,x'\in\mathbb{R}^N$,

$$(3.6)\qquad |\hat{b}(x)-\hat{b}(x')|\le\hat{C}_N\{|x-x'|+|x-x'|^\beta\},\quad \|\hat{\sigma}(x)-\hat{\sigma}(x')\|_{\mathrm{HS}}\le\hat{C}_N|x-x'|^\alpha$$

for some constant $\hat{C}_N>0$ so that $\hat{b}$ and $\hat{\sigma}$ are continuous . Consequently, (3.3) and (3.6) yield that (3.2) has a weak solution; see, for instance, [35, Theorem 175, p.147]. Moreover, by carrying out a similar argument to derive (2.12), we can infer that (1.8) is pathwise unique. As a result, we reach a conclusion that (1.8) has a unique strong solution $(X_t^{i,N})_{t\ge0}$. Finally, with (3.3) at hand, (3.1) can be available via Hölder's inequality, Burkhold-Davis-Gundy's inequality and Gronwall's inequality. □

The following lemma reveals the phenomenon of propagation of chaos and provides the corresponding convergence rate.

Lemma 3.2. *Under the assumptions of Theorem 1.3, for any $T>0$, there exists a constant $C_T>0$, which is independent of N, such that*

$$(3.7)\qquad \sup_{i\in\mathcal{S}_N}\mathbb{E}\big\|X^i-X^{i,N}\big\|_{\infty,T}\le C_T\Big\{N^{-\frac18}\mathbf{1}_{\{\alpha=\frac12\}}+N^{-\frac{2\alpha-1}{4}}\mathbf{1}_{\{\alpha\in(\frac12,1]\}}\Big\},$$

and

$$(3.8)\qquad \sup_{i\in\mathcal{S}_N}\mathbb{E}\big\|X^i-X^{i,N}\big\|_{\infty,T}^2\le C_TN^{-\frac14}.$$

Proof. In what follows, let $i\in\mathcal{S}_N$ and set $Z_t^{i,N}:=X_t^i-X_t^{i,N}$. First of all, we are going to prove that there exists a constant $C_T>0$, independent of N, such that

$$(3.9)\qquad \mathbb{E}|Z_t^{i,N}|\le C_TN^{-\frac14},\quad t\in[0,T].$$

Applying Itô's formula to $V_\varepsilon := V_{e^{\frac{1}{\varepsilon}},\varepsilon}$ and taking $X_0^i = X_0^{i,N}, i \in \mathcal{S}_N$, into consideration yield

$$
\begin{aligned}
V_\varepsilon(Z_t^{i,N}) &= \int_0^t V_\varepsilon'(Z_s^{i,N})\big(b_1(X_s^i,\mu_s^i) - b_1(X_s^{i,N},\mu_s^i)\big)\mathrm{d}s \\
&\quad + \int_0^t V_\varepsilon'(Z_s^{i,N})\big(b_1(X_s^{i,N},\mu_s^i) - b_1(X_s^{i,N},\hat{\mu}_s^N)\big)\mathrm{d}s \\
&\quad + \int_0^t V_\varepsilon'(Z_s^{i,N})\big(b_2(X_s^i,\mu_s^i) - b_2(X_s^{i,N},\hat{\mu}_s^N)\big)\mathrm{d}s \\
&\quad + \frac{1}{2}\int_0^t V_\varepsilon''(Z_s^{i,N})\big(\sigma(X_s^i) - \sigma(X_s^{i,N})\big)^2\mathrm{d}s + M_t^{i,N} \\
&=: \Upsilon_{1,\varepsilon}^i(t) + \Upsilon_{2,\varepsilon}^i(t) + \Upsilon_{3,\varepsilon}^i(t) + \Upsilon_{4,\varepsilon}^i(t) + M_t^{i,N},
\end{aligned}
\tag{3.10}
$$

where

$$
M_t^{i,N} := \int_0^t V_\varepsilon'(Z_s^{i,N})\big(\sigma(X_s^i) - \sigma(X_s^{i,N})\big)\mathrm{d}W_s^i.
$$

Recall that $x \mapsto b_1(x,\cdot)$ is non-increasing. This, together with $\mathrm{sgn}(x)V_\varepsilon'(x) \in [0,1]$ owing to (2.2), leads to

$$
V_\varepsilon'(x-y)\big(b_1(x,\cdot) - b_1(y,\cdot)\big) \le 0, \quad x,y \in \mathbb{R}
\tag{3.11}
$$

so that we infer $\Upsilon_{1,\varepsilon}^i(t) \le 0$. On the other hand, using (1.2), (1.3) and ($\mathbf{H}_2$) and taking advantage of (2.2) and (2.3), we derive that

$$
\begin{aligned}
\sum_{j=2}^4 \Upsilon_{j,\varepsilon}^i(t) &\le 2K_1\int_0^t \big\{|Z_s^{i,N}| + \mathbb{W}_1(\mu_s^i,\hat{\mu}_s^N)\big\}\mathrm{d}s + K_2^2\varepsilon\int_0^t \mathbf{1}_{[\varepsilon/\gamma,\varepsilon]}(|Z_s^{i,N}|)|Z_s^{i,N}|^{2\alpha-1}\mathrm{d}s \\
&\le \big(2K_1 \vee K_2^2\big)\int_0^t \big\{|Z_s^{i,N}| + \mathbb{W}_1(\mu_s^i,\hat{\mu}_s^N) + \varepsilon^{2\alpha}\big\}\mathrm{d}s.
\end{aligned}
$$

We henceforth obtain from (2.2) and (3.10) that

$$
|Z_t^{i,N}| \le \varepsilon + C_1\int_0^t \big\{|Z_s^{i,N}| + \mathbb{W}_1(\mu_s^i,\hat{\mu}_s^N) + \varepsilon^{2\alpha}\big\}\mathrm{d}s + M_t^{i,N}
\tag{3.12}
$$

for some constant $C_1 > 0$. So, by taking expectations on both sides, approaching $\varepsilon \downarrow 0$, and utilizing the triangle inequality for $\mathbb{W}_1$, one obtains

$$
\mathbb{E}|Z_t^{i,N}| \le C_1\int_0^t \big\{\mathbb{E}|Z_s^{i,N}| + \mathbb{E}\mathbb{W}_1(\mu_s^i,\tilde{\mu}_s^N) + \mathbb{E}\mathbb{W}_1(\tilde{\mu}_s^N,\hat{\mu}_s^N)\big\}\mathrm{d}s,
$$

where $\tilde{\mu}^N$ was introduced in (1.7). By the Glivenko-Cantelli theorem (see e.g. [7, Theorem 5.8]), there exists a constant $C_2 > 0$ such that

$$
\mathbb{E}\mathbb{W}_1(\mu_t^i,\tilde{\mu}_t^N) \le C_2 N^{-1/4}.
\tag{3.13}
$$

As a consequence, exploiting (3.5) and (3.13), we derive that

$$\begin{aligned}\mathbb{E}|Z_t^{i,N}| &\le C_1\int_0^t\Big\{\mathbb{E}|Z_s^{i,N}|+\frac{1}{N}\sum_{j=1}^N\mathbb{E}|X_s^j-X_s^{j,N}|+C_2N^{-1/4}\Big\}\mathrm{d}s\\ &\le C_3\int_0^t\big\{\mathbb{E}|Z_s^{i,N}|+N^{-1/4}\big\}\mathrm{d}s\end{aligned}$$

for some constant $C_3>0$, where in the last display we used the fact that $(Z^{j,N})_{1\le j\le N}$ are identically distributed. Subsequently, by employing Gronwall's inequality, (3.9) is available.

Next, by Burkhold-Davis-Gundy's inequality, Young's inequality as well as Jensen's inequality, we derive from (1.2) and (1.3) that there exist constants $C_4,C_5>0$ such that

$$\begin{aligned}\mathbb{E}\Big(\sup_{0\le s\le t}|Z_s^{i,N}|\Big) &\le C_4\int_0^t\big\{\mathbb{E}|Z_s^{i,N}|+N^{-1/4}\big\}\mathrm{d}s+C_4\mathbb{E}\Big(\int_0^t|Z_s^{i,N}|^{2\alpha}\mathrm{d}s\Big)^{1/2}\\ &\le C_4\int_0^t\big\{\mathbb{E}|Z_s^{i,N}|+N^{-1/4}\big\}\mathrm{d}s+C_4\Big(\int_0^t\mathbb{E}|Z_s^{i,N}|\mathrm{d}s\Big)^{1/2}\mathbf{1}_{\{\alpha=1/2\}}\\ &\quad+\Big\{\frac{1}{2}\mathbb{E}\|Z^{i,N}\|_{\infty,t}+C_5\int_0^t\big(\mathbb{E}|Z_s^{i,N}|\big)^{2\alpha-1}\mathrm{d}s\Big\}\mathbf{1}_{\{\alpha\in(1/2,1]\}}.\end{aligned}$$

As a result, (3.7) follows from (3.9).

Again, by applying Hölder's inequality and Burkhold-Davis-Gundy's inequality, it follows from $(\mathbf{H}_2)$ and (3.12) that there exists a constant $C_6>0$ such that

$$\mathbb{E}\Big(\sup_{0\le s\le t}|Z_s^{i,N}|^2\Big)\le C_6t\int_0^t\big\{\mathbb{E}|Z_s^{i,N}|^2+\mathbb{E}\mathbb{W}_1(\mu_s^i,\hat\mu_s^N)^2\big\}\mathrm{d}s+C_6\int_0^t\mathbb{E}|Z_s^{i,N}|^{2\alpha}\mathrm{d}s.$$

Owing to (3.4), we have

$$\mathbb{W}_2\Big(\frac{1}{N}\sum_{j=1}^N\delta_{x_j},\frac{1}{N}\sum_{j=1}^N\delta_{y_j}\Big)^2\le\frac{1}{N}\sum_{j=1}^N|x_j-y_j|^2,\quad x_j,y_j\in\mathbb{R}. \tag{3.14}$$

Whence, it follows that

$$\mathbb{E}\mathbb{W}_2(\tilde\mu_t^N,\hat\mu_t^N)^2\le\frac{1}{N}\sum_{j=1}^N\mathbb{E}|Z_t^{j,N}|^2=\mathbb{E}|Z_t^{i,N}|^2 \tag{3.15}$$

from the fact that $(Z^{j,N})_{1\le j\le N}$ are identically distributed into consideration. Moreover, according to the Glivenko-Cantelli theorem (see e.g. [7, Theorem 5.8]), there exists a constant $C_7>0$ such that

$$\mathbb{E}\mathbb{W}_2(\mu_s^i,\tilde\mu_s^N)^2\le C_7N^{-1/2}. \tag{3.16}$$

Thus, combining (3.15) with (3.16) and employing Young's inequality, we infer that

$$\begin{aligned}\mathbb{E}\|Z^{i,N}\|_{\infty,t}^2 &\le C_8 t\int_0^t \big\{\mathbb{E}|Z_s^{i,N}|^2+\mathbb{E}\mathbb{W}_2(\mu_s^i,\tilde\mu_s^N)^2+\mathbb{E}\mathbb{W}_2(\tilde\mu_s^N,\hat\mu_s^N)^2\big\}\mathrm{d}s+C_6\int_0^t \mathbb{E}|Z_s^{i,N}|^{2\alpha}\mathrm{d}s\\ &\le C_9 t\int_0^t \big\{\mathbb{E}|Z_s^{i,N}|^2+N^{-1/2}\big\}\mathrm{d}s+C_6\int_0^t \mathbb{E}|Z_s^{i,N}|\mathrm{d}s\mathbf{1}_{\{\alpha=\frac12\}}\\ &\quad+C_6\int_0^t\big\{2(1-\alpha)\mathbb{E}|Z_s^{i,N}|+(2\alpha-1)\mathbb{E}|Z_s^{i,N}|^2\big\}\mathrm{d}s\mathbf{1}_{\{\alpha\in(1/2,1]\}}\end{aligned}$$

for some constants $C_8, C_9>0$. Finally, (3.8) holds true from (3.9). □

The lemma below shows that the p-th order moment of the continuous time EM scheme, defined by (1.9), is bounded in a finite time horizontal.

Lemma 3.3. *Assume* $(\mathbf{H}_1)$ *and* $(\mathbf{H}_2)$ *and suppose* $X_0\in L^p(\Omega\to\mathbb{R};\mathscr{F}_0,\mathbb{P})$ *for some* $p>0$. *Then, there exists a constant* $C_{p,T}$, *dependent on* p *and* T *but independent of* N *and* δ, *such that*

$$\mathbb{E}\|X^{\delta,i,N}\|_{\infty,T}^p\le C_{p,T}\big(1+\mathbb{E}|X_0|^p\big). \tag{3.17}$$

Proof. Below, without loss of generality, we set $p\ge2$ since the lower order moment estimate can be achieved by Hölder's inequality. By (2.6), in addition to Burkhold-Davis-Gundy's inequality, it follows that

$$\begin{aligned}\mathbb{E}\|X^{\delta,i,N}\|_{\infty,t}^p &\le 3^{p-1}\Big\{\mathbb{E}|X_0^{i,N}|^p+t^{p-1}\int_0^t\mathbb{E}|b(X_{s_\delta}^{\delta,i,N},\hat\mu_{s_\delta}^{\delta,N})|^p\mathrm{d}s\\ &\quad+\mathbb{E}\Big(\sup_{0\le s\le t}\Big|\int_0^s\sigma(X_{r_\delta}^{\delta,i,N})\mathrm{d}W_r^i\Big|^p\Big)\Big\}\\ &\le C_1\big(1+\mathbb{E}|X_0^{i,N}|^p\big)+C_1\int_0^t\mathbb{E}\big(|X_{s_\delta}^{\delta,i,N}|+\mathbb{W}_1(\hat\mu_{s_\delta}^{\delta,N},\delta_0)\big)^p\mathrm{d}s\\ &\quad+C_1\mathbb{E}\Big(\sup_{0\le s\le t}|X_{s_\delta}^{\delta,i,N}|\int_0^t|X_{s_\delta}^{\delta,i,N})|\mathrm{d}s\Big)^{p/2}\\ &\le C_1\big(1+\mathbb{E}|X_0^{i,N}|^p\big)+C_2\int_0^t\Big(\mathbb{E}|X_{s_\delta}^{\delta,i,N}|^p+\frac1N\sum_{j=1}^N\mathbb{E}|X_{s_\delta}^{\delta,j,N}|^p\Big)\mathrm{d}s\\ &\quad+\frac12\mathbb{E}\Big(\sup_{0\le s\le t}|X_{s_\delta}^{\delta,i,N}|^p\Big)+C_2\int_0^t\mathbb{E}|X_{s_\delta}^{\delta,i,N}|^p\mathrm{d}s,\quad t\in[0,T]\end{aligned}$$

for some constants $C_1=C_1(p,T), C_2=C_2(p,T)>0$, where in the last display we also used the fact that $X_0^{i,N}$ shares the same law with that of X_0. Since

$$\sup_{0\le s\le t}|X_{s_\delta}^{\delta,i,N}|\le\|X^{\delta,i,N}\|_{\infty,t},$$

and $(X_\cdot^{\delta,i,N})_{i\in\mathcal{S}_N}$ are identically distributed, we thus derive that

$$\mathbb{E}\|X^{\delta,i,N}\|_{\infty,t}^p\le C_3(1+\mathbb{E}|X_0|^p)+C_3\int_0^t\mathbb{E}\|X^{\delta,i,N}\|_{\infty,s}^p\mathrm{d}s$$

for some constant $C_3 = C_3(p, T)$. Consequently, Gronwall's inequality yields the desired assertion (3.17). $\square$

The following lemma demonstrates the convergence rate of the continuous time EM scheme associated with (1.8).

Lemma 3.4. *Under the assumptions of Lemma 3.1, for any $T > 0$, there exists a constant $C_T > 0$, independent of N and δ, such that*

$$\sup_{i\in\mathcal{S}_N}\mathbb{E}\|X^{i,N}-X^{\delta,i,N}\|_{\infty,T}\le C_T\Big\{\mathbf{1}_{\{\alpha=\frac12\}}\Big(\frac{1}{\ln\frac1\delta}\Big)^{\frac12}+\mathbf{1}_{\{\alpha\in(\frac12,1]\}}\Big(\delta^{\frac{(2\alpha-1)^2}{2}}+\delta^{\frac{\beta(2\alpha-1)}{2}}\Big)\Big\},\tag{3.18}$$

and

$$\sup_{i\in\mathcal{S}_N}\mathbb{E}\|X^{i,N}-X^{\delta,i,N}\|_{\infty,T}^2\le C_T\Big\{\mathbf{1}_{\{\alpha=\frac12\}}\frac{1}{\ln\frac1\delta}+\mathbf{1}_{\{\alpha\in(\frac12,1)\}}(\delta^{\frac{2\alpha-1}{2}}+\delta^{\frac{\beta}{2}})+\mathbf{1}_{\{\alpha=1\}}\delta^{\beta}\Big\}.\tag{3.19}$$

Proof. For $i \in \mathcal{S}_N$, let $Z_t^{\delta,i,N} = X_t^{i,N} - X_t^{\delta,i,N}$ and $\Lambda_t^{\delta,i,N} = X_t^{\delta,i,N} - X_{t_\delta}^{\delta,i,N}$. Below, we set $t \in [0, T]$. By using Hölder's inequality and Burkhold-Davis-Gundy's inequality, for any $q > 0$, we obtain from (3.17) that there exists $\hat{C}_{T,q} > 0$ such that

$$\mathbb{E}|\Lambda_t^{\delta,i,N}|^q\le\hat{C}_{T,q}\delta^{q/2},\quad t\in[0,T].\tag{3.20}$$

Below, by Itô's formula, it follows that

$$\begin{aligned}\mathrm{d}V_{\gamma,\varepsilon}(Z_t^{\delta,i,N})&=\big\{V'_{\gamma,\varepsilon}(Z_t^{\delta,i,N})\big(b(X_t^{i,N},\hat\mu_t^N)-b(X_{t_\delta}^{\delta,i,N},\hat\mu_{t_\delta}^{\delta,N})\big)\\&\quad+\frac12V''_{\gamma,\varepsilon}(Z_t^{\delta,i,N})\big(\sigma(X_t^{i,N})-\sigma(X_{t_\delta}^{\delta,i,N})\big)^2\big\}\mathrm{d}t+\mathrm{d}\hat M_t^{i,N},\end{aligned}$$

where

$$\mathrm{d}\hat M_t^{i,N}:=V'_{\gamma,\varepsilon}(Z_t^{\delta,i,N})\big(\sigma(X_t^{i,N})-\sigma(X_{t_\delta}^{\delta,i,N})\big)\mathrm{d}W_t^i.$$

Observe from (3.11) that

$$\begin{aligned}&V'_{\gamma,\varepsilon}(Z_t^{\delta,i,N})\big(b_1(X_t^{i,N},\hat\mu_t^N)-b_1(X_{t_\delta}^{\delta,i,N},\hat\mu_{t_\delta}^{\delta,N})\big)\\&=V'_{\gamma,\varepsilon}(Z_t^{\delta,i,N})\big(b_1(X_t^{i,N},\hat\mu_t^N)-b_1(X_t^{\delta,i,N},\hat\mu_t^N)\big)\\&\quad+V'_{\gamma,\varepsilon}(Z_t^{\delta,i,N})\big(b_1(X_t^{\delta,i,N},\hat\mu_t^N)-b_1(X_{t_\delta}^{\delta,i,N},\hat\mu_t^N)\big)\\&\quad+V'_{\gamma,\varepsilon}(Z_t^{\delta,i,N})\big(b_1(X_{t_\delta}^{\delta,i,N}),\hat\mu_t^N)-b_1(X_{t_\delta}^{\delta,i,N},\hat\mu_{t_\delta}^{\delta,N})\big)\\&\le V'_{\gamma,\varepsilon}(Z_t^{\delta,i,N})\big(b_1(X_t^{\delta,i,N},\hat\mu_t^N)-b_1(X_{t_\delta}^{\delta,i,N},\hat\mu_t^N)\big)\\&\quad+V'_{\gamma,\varepsilon}(Z_t^{\delta,i,N})\big(b_1(X_{t_\delta}^{\delta,i,N},\hat\mu_t^N)-b_1(X_{t_\delta}^{\delta,i,N},\hat\mu_{t_\delta}^{\delta,N})\big)\\&\le K_1\big\{\mathbb{W}_1(\hat\mu_t^N,\hat\mu_{t_\delta}^{\delta,N})+|\Lambda_t^{\delta,i,N}|^\beta\big\},\end{aligned}$$

where the last display is due to (1.2) and (2.2). Then, combining this with (1.3) and taking advantage of (2.2), (2.3) as well as (3.20) gives

$$\begin{aligned}\mathbb{E}|Z_t^{\delta,i,N}| &\le \varepsilon + c_1\int_0^t \mathbb{E}\Big\{|Z_s^{\delta,i,N}| + |\Lambda_s^{\delta,i,N}| + |\Lambda_s^{\delta,i,N}|^\beta + \mathbb{W}_1(\hat\mu_t^N, \hat\mu_{t_\delta}^{\delta,N}) \\ &\quad + \frac{1}{|Z_s^{\delta,i,N}|\ln\gamma}\mathbf{1}_{[\varepsilon/\gamma,\varepsilon]}(|Z_s^{\delta,i,N}|)\big(|Z_s^{\delta,i,N}|^{2\alpha} + |\Lambda_s^{\delta,i,N}|^{2\alpha}\big)\Big\}\mathrm{d}s \\ &\le C_{1,T}\Big\{\varepsilon + \frac{\varepsilon^{2\alpha-1}}{\ln\gamma} + \frac{\gamma}{\varepsilon\ln\gamma}\delta^\alpha + \delta^{\frac12} + \delta^{\frac{\beta}{2}} + \int_0^t \mathbb{E}|Z_s^{\delta,i,N}|\mathrm{d}s\Big\}\end{aligned}$$

for some constants $c_1, C_{1,T} > 0$, where we also utilized

$$\mathbb{E}\mathbb{W}_1(\hat\mu_t^N, \hat\mu_{t_\delta}^{\delta,N}) \le \mathbb{E}|\Lambda_t^{\delta,i,N}| + \mathbb{E}|Z_t^{\delta,i,N}|$$

and the fact that $(Z_t^{\delta,i,N})_{1\le i\le N}$ are identically distributed. Thus, Gronwall's inequality yields

$$\mathbb{E}|Z_t^{\delta,i,N}| \le C_{2,T}\Big\{\varepsilon + \frac{\varepsilon^{2\alpha-1}}{\ln\gamma} + \frac{\gamma}{\varepsilon\ln\gamma}\delta^\alpha + \delta^{\frac12} + \delta^{\frac{\beta}{2}}\Big\} \tag{3.21}$$

for some constant $C_{2,T} > 0$. Furthermore, by virtue of Burkhold-Davis-Gundy's inequality and Jensen's inequality, we deduce from $(\mathbf{H}_1)$, $(\mathbf{H}_2)$, (2.2), and (2.3) that

$$\begin{aligned}\mathbb{E}\|Z^{\delta,i,N}\|_{\infty,t} &\le C_{2,T}\Big\{\varepsilon + \frac{\varepsilon^{2\alpha-1}}{\ln\gamma} + \frac{\gamma}{\varepsilon\ln\gamma}\delta^\alpha + \delta^{\frac12} + \delta^{\frac{\beta}{2}}\Big\} \\ &\quad + c_1\Big(\int_0^t (\mathbb{E}|Z_s^{\delta,i,N}| + \mathbb{E}|\Lambda_s^{\delta,i,N}|)\mathrm{d}s\Big)^{1/2}\mathbf{1}_{\{\alpha=1/2\}} \\ &\quad + \Big\{\frac12\mathbb{E}\|Z^{\delta,i,N}\|_{\infty,t} + c_2\int_0^t (\mathbb{E}|Z_s^{\delta,i,N}|)^{2\alpha-1}\mathrm{d}s + c_2\left(\int_0^t \mathbb{E}|\Lambda_s^{\delta,i,N}|^{2\alpha}\mathrm{d}s\right)^{\frac12}\Big\}\mathbf{1}_{\{\alpha\in(1/2,1]\}}\end{aligned} \tag{3.22}$$

and that

$$\begin{aligned}\mathbb{E}\|Z^{\delta,i,N}\|_{t,\infty}^2 &\le C_{3,T}\Big\{\varepsilon + \frac{\varepsilon^{2\alpha-1}}{\ln\gamma} + \frac{\gamma}{\varepsilon\ln\gamma}\delta^\alpha + \delta^{\frac12} + \delta^{\frac{\beta}{2}}\Big\}^2 \\ &\quad + c_3\int_0^t (\mathbb{E}|Z_s^{\delta,i,N}| + \mathbb{E}|\Lambda_s^{\delta,i,N}|)\mathrm{d}s\mathbf{1}_{\{\alpha=1/2\}} \\ &\quad + c_4\Big\{\int_0^t (\mathbb{E}|Z_s^{\delta,i,N}| + \mathbb{E}|Z_s^{\delta,i,N}|^2)\mathrm{d}s + \int_0^t \mathbb{E}|\Lambda_s^{\delta,i,N}|^{2\alpha}\mathrm{d}s\Big\}\mathbf{1}_{\{\alpha\in(1/2,1]\}}\end{aligned} \tag{3.23}$$

for some constants $c_2, c_3, c_4, C_{2,T}, C_{3,T} > 0$. Consequently, the desired assertions (3.18) and (3.19) follows from (3.22) and (3.23) and by taking $\varepsilon = \frac{1}{\ln\frac{1}{\delta}}$ and $\gamma = (1/\delta)^{\frac13}$ for $\alpha = \frac12$ and $\varepsilon = \sqrt{\delta}$ and $\gamma = 2$ for $\alpha \in (1/2, 1]$, respectively. □

Proof of Theorem 1.3. With the help of Lemmas 3.2 and 3.4, we complete directly the proof of Theorem 1.3. □

4 Proof of Theorem 1.5

The proof of Theorem 1.5 is based on two lemmas below, where the first one is concerned with propagation of chaos for non-degenerate McKean-Vlasov SDEs with Hölder continuous drifts and $\mathbb{W}_2$-Lipschitz continuous diffusions.

Lemma 4.1. *Under the assumptions of Theorem 1.5, for any $T > 0$, there exists a constant $C_T > 0$, independent of N, such that*

$$\sup_{i\in\mathcal{S}_N} \mathbb{E}\|X^i - X^{i,N}\|_{\infty,T}^2 \le C_T \begin{cases} N^{-\frac{1}{2}}, & d < 4 \\ N^{-\frac{1}{2}}\log N, & d = 4 \\ N^{-\frac{2}{d}}, & d > 4. \end{cases} \tag{4.1}$$

Proof. Below we set $t \in [0,T]$. For $i \in \mathcal{S}_N$ and $x \in \mathbb{R}^d$, let

$$b_t^{\mu^i}(x) = b(x,\mu_t^i),\ b_t^{\hat\mu^N} = b(x,\hat\mu_t^N),\ \sigma_t^{\mu^i}(x) = \sigma(x,\mu_t^i),\ \sigma_t^{\hat\mu^N}(x) = \sigma(x,\hat\mu_t^N).$$

Then, (1.17) and (1.18) can be rewritten respectively as

$$\mathrm{d}X_t^i = b_t^{\mu^i}(X_t^i)\mathrm{d}t + \sigma_t^{\mu^i}(X_t^i)\mathrm{d}W_t^i, \quad \mathrm{d}X_t^{i,N} = b_t^{\hat\mu^N}(X_t^{i,N})\mathrm{d}t + \sigma_t^{\hat\mu^N}(X_t^{i,N})\mathrm{d}W_t^i.$$

For $\lambda > 0$, consider the following PDE for $u^{\lambda,\mu^i} : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$:

$$\partial_t u_t^{\lambda,\mu^i} + \frac{1}{2}\mathrm{Tr}\big(\sigma_t^{\mu^i}(\sigma_t^{\mu^i})^*\nabla^2 u_t^{\lambda,\mu^i}\big) + \nabla_{b_t^{\mu^i}} u_t^{\lambda,\mu^i} + b_t^{\mu^i} = \lambda u_t^{\lambda,\mu^i}, \quad u_T^{\lambda,\mu^i} = \mathbf{0}, \tag{4.2}$$

where ∇^2 is the second order gradient operator in space. By recurring to [2, Lemma 2.1], for $\lambda > 0$ large enough, (4.2) has a unique solution u^{λ,μ^i} such that

$$\|\nabla u^{\lambda,\mu^i}\|_\infty + \|\nabla^2 u^{\lambda,\mu^i}\|_\infty \le 1/2. \tag{4.3}$$

Using Itô's formula to $\theta_t^{\lambda,\mu^i}(x) := x + u_t^{\lambda,\mu^i}(x)$, $x\in\mathbb{R}^d$ yields

$$\begin{aligned} \mathrm{d}\theta_t^{\lambda,\mu^i}(X_t^i) &= \lambda u_t^{\lambda,\mu^i}(X_t^i)\mathrm{d}t + (\nabla\theta_t^{\lambda,\mu^i}\sigma_t^{\mu^i})(X_t^i)\mathrm{d}W_t^i, \\ \mathrm{d}\theta_t^{\lambda,\mu^i}(X_t^{i,N}) &= \big\{\lambda u_t^{\lambda,\mu^i}(X_t^{i,N}) + \nabla\theta_t^{\lambda,\mu^i}(b_t^{\hat\mu^N} - b_t^{\mu^i})(X_t^{i,N})\big\}\mathrm{d}t + (\nabla\theta_t^{\lambda,\mu^i}\sigma_t^{\hat\mu^N})(X_t^{i,N})\mathrm{d}W_t^i \\ &\quad + \frac{1}{2}\mathrm{Tr}\big[\big(\sigma_t^{\hat\mu^N}(\sigma_t^{\hat\mu^N})^* - \sigma_t^{\mu^i}(\sigma_t^{\mu^i})^*\big)\nabla^2 u_t^{\lambda,\mu^i}\big](X_t^{i,N})\mathrm{d}t. \end{aligned} \tag{4.4}$$

Hence, for $\Lambda_t^{\lambda,i,N} := \theta_t^{\lambda,\mu^i}(X_t^i) - \theta_t^{\lambda,\mu^i}(X_t^{i,N})$, we derive from Hölder's inequality and Burkhold-Davis-Gundy's inequality that

$$\begin{aligned} &\mathbb{E}\|\Lambda^{\lambda,i,N}\|_{\infty,t}^2 \\ &\le C_1\bigg\{\int_0^t \mathbb{E}\big|u_s^{\lambda,\mu^i}(X_s^i) - u_s^{\lambda,\mu^i}(X_s^{i,N})\big|^2\mathrm{d}s \\ &\quad + \int_0^t \mathbb{E}\bigg\{\big|\big(\nabla\theta_s^{\lambda,\mu^i}(b_s^{\hat\mu^N} - b_s^{\mu^i})\big)(X_s^{i,N})\big|^2 + \Big|\mathrm{Tr}[(\sigma_s^{\hat\mu^N}(\sigma_s^{\hat\mu^N})^* - \sigma_s^{\mu^i}(\sigma_s^{\mu^i})^*)\nabla^2 u_s^{\lambda,\mu^i}](X_s^{i,N})\Big|^2\bigg\}\mathrm{d}s\bigg\} \\ &\quad + \int_0^t \mathbb{E}\big\|(\nabla\theta_s^{\lambda,\mu^i}\sigma_s^{\hat\mu^N})(X_s^{i,N}) - (\nabla\theta_s^{\lambda,\mu^i}\sigma_s^{\mu^i})(X_s^i)\big\|_{\mathrm{HS}}^2\mathrm{d}s \\ &=: C_1\{I_{1,i}(t) + I_{2,i}(t)\} + I_{3,i}(t) \end{aligned}$$

for some constant $C_1>0$. Set $Z_t^{i,N}:=X_t^i-X_t^{i,N}$ for convenience. By means of (4.3), one has

$$I_{1,i}(t)\le C_2\int_0^t\mathbb{E}|Z_s^{i,N}|^2\mathrm{d}s. \tag{4.5}$$

for some constant $C_2>0$. Next, via $(\mathbf{A}_1)$-$(\mathbf{A}_2)$ and (4.3), in addition to (3.15), it follows from the triangle inequality that

$$\begin{aligned}I_{2,i}(t)&\le C_3\int_0^t\big\{\mathbb{E}\mathbb{W}_2(\hat\mu_s^N,\tilde\mu_s^N)^2+\mathbb{E}\mathbb{W}_2(\tilde\mu_s^N,\mu_s^i)^2\big\}\mathrm{d}s\\&\le C_3\int_0^t\big\{\mathbb{E}|Z_s^{i,N}|^2+\mathbb{E}\mathbb{W}_2(\tilde\mu_s^N,\mu_s^i)^2\big\}\mathrm{d}s\end{aligned} \tag{4.6}$$

for some constant $C_3>0$. Furthermore, owing to (1.13) and (4.3), we obtain that for some constants $C_4,C_5>0$,

$$\begin{aligned}I_{3,i}(t)&\le 3\int_0^t\mathbb{E}\big\|\nabla\theta_s^{\lambda,\mu^i}(X_s^{i,N})(\sigma_s^{\hat\mu^N}(X_s^{i,N})-\sigma_s^{\hat\mu^N}(X_s^i))\big\|_{\mathrm{HS}}^2\mathrm{d}s\\&\quad+3\int_0^t\mathbb{E}\big\|\nabla\theta_s^{\lambda,\mu^i}(X_s^{i,N})(\sigma_s^{\hat\mu^N}(X_s^i)-\sigma_s^{\mu^i}(X_s^i))\big\|_{\mathrm{HS}}^2\mathrm{d}s\\&\quad+3\int_0^t\mathbb{E}\big\|(\nabla\theta_s^{\lambda,\mu^i}(X_s^{i,N})-\nabla\theta_s^{\lambda,\mu^i}(X_s^i))\sigma_s^{\mu^i}(X_s^i)\big\|_{\mathrm{HS}}^2\mathrm{d}s\\&\le C_4\left(\int_0^t\mathbb{E}|Z_s^{i,N}|^2\mathrm{d}s+\int_0^t\big\{\mathbb{E}\mathbb{W}_2(\hat\mu_s^N,\tilde\mu_s^N)^2+\mathbb{E}\mathbb{W}_2(\tilde\mu_s^N,\mu_s^i)^2\big\}\mathrm{d}s\right)\\&\le C_5\int_0^t\big\{\mathbb{E}|Z_s^{i,N}|^2+\mathbb{E}\mathbb{W}_2(\tilde\mu_s^N,\mu_s^i)^2\big\}\mathrm{d}s.\end{aligned} \tag{4.7}$$

Thus, with the aid of (4.5), (4.6) and (4.7), we find that for some constant $C_6>0$,

$$\mathbb{E}\Big(\sup_{0\le s\le t}|\Lambda_s^{\lambda,i,N}|^2\Big)\le C_6\int_0^t\big\{\mathbb{E}|Z_s^{i,N}|^2+\mathbb{E}\mathbb{W}_2(\tilde\mu_s^N,\mu_s^i)^2\big\}\mathrm{d}s. \tag{4.8}$$

From (4.3), it follows that

$$|Z_t^{i,N}|\le|\Lambda_t^{\lambda,i,N}|+|u_t^{\lambda,\mu^i}(X_t^i)-u_t^{\lambda,\mu^i}(X_t^{i,N})|\le|\Lambda_t^{\lambda,i,N}|+\frac{1}{2}|Z_t^{i,N}|$$

so that $|Z_t^{i,N}|\le 2|\Lambda_t^{\lambda,i,N}|$. This, besides (4.8), gives

$$\mathbb{E}\Big(\sup_{0\le s\le t}|Z_s^{i,N}|^2\Big)\le C_7\int_0^t\big\{\mathbb{E}|Z_s^{i,N}|^2+\mathbb{E}\mathbb{W}_2(\tilde\mu_s^N,\mu_s^i)^2\big\}\mathrm{d}s$$

for some constant $C_7>0$. Hence, the desired assertion (4.1) follows from Gronwall's inequality and the fact that

$$\sup_{0\le t\le T}\mathbb{E}\mathbb{W}_2(\tilde\mu_t^N,\mu_t^i)^2\le C_8\begin{cases}N^{-\frac{1}{2}}, & d<4\\ N^{-\frac{1}{2}}\log N, & d=4\\ N^{-\frac{2}{d}}, & d>4\end{cases} \tag{4.9}$$

for some constant $C_8>0$; see, for instance, [7, Theorem 5.8]. □

Lemma 4.2. *Under the assumptions of Theorem 1.5, for any $T > 0$, there exists a constant $C_T > 0$, independent of N and δ, such that*

$$\sup_{i\in\mathcal{S}_N}\mathbb{E}\|X^{i,N}-X^{\delta,i,N}\|_{\infty,T}^2\le C_T\begin{cases}\delta^\alpha+N^{-\frac{1}{2}}, & d<4\\ \delta^\alpha+N^{-\frac{1}{2}}\log N, & d=4\\ \delta^\alpha+N^{-\frac{2}{d}}, & d>4.\end{cases}\tag{4.10}$$

Proof. In the sequel, let $t\in[0,T]$. For $x\in\mathbb{R}^d$, set

$$b_{k\delta}^{\hat\mu^{\delta,N}}(x):=b(x,\hat\mu_{k\delta}^{\delta,N}),\quad \sigma_{k\delta}^{\hat\mu^{\delta,N}}(x):=\sigma(x,\hat\mu_{k\delta}^{\delta,N})$$

so that (1.19) can be reformulated as

$$\mathrm{d}X_t^{\delta,i,N}=b_{t_\delta}^{\hat\mu^{\delta,N}}(X_{t_\delta}^{\delta,i,N})\mathrm{d}t+\sigma_{t_\delta}^{\hat\mu^{\delta,N}}(X_{t_\delta}^{\delta,i,N})\mathrm{d}W_t^i.$$

Applying Itô's formula to $\theta_t^{\lambda,\mu^i}(x)=x+u_t^{\lambda,\mu^i}(x)$ and taking the fact that u^{λ,μ^i} solves (4.2) into consideration gives that

$$\begin{aligned}\mathrm{d}\theta_t^{\lambda,\mu^i}(X_t^{\delta,i,N})&=\Big\{\lambda u_t^{\lambda,\mu^i}(X_t^{\delta,i,N})+\nabla\theta_t^{\lambda,\mu^i}(X_t^{\delta,i,N})\big(b_{t_\delta}^{\hat\mu^{\delta,N}}(X_{t_\delta}^{\delta,i,N})-b_t^{\mu^i}(X_t^{\delta,i,N})\big)\\&+\frac{1}{2}\sum_{k,j=1}^d\big\langle\big((\sigma_{t_\delta}^{\hat\mu^{\delta,N}}(\sigma_{t_\delta}^{\hat\mu^{\delta,N}})^*)(X_{t_\delta}^{\delta,i,N})-(\sigma_t^{\mu^i}(\sigma_t^{\mu^i})^*)(X_t^{\delta,i,N})\big)e_j,e_l\big\rangle\nabla_{e_j}\nabla_{e_l}u_t^{\lambda,\mu^i}(X_t^{\delta,i,N})\Big\}\mathrm{d}t\\&+\nabla\theta_t^{\lambda,\mu^i}(X_t^{\delta,i,N})\sigma_{t_\delta}^{\hat\mu^{\delta,N}}(X_{t_\delta}^{\delta,i,N})\mathrm{d}W_t^i.\end{aligned}\tag{4.11}$$

Set

$$\Theta_t^{\delta,i,N}:=\theta_t^{\lambda,\mu^i}(X_t^{i,N})-\theta_t^{\lambda,\mu^i}(X_t^{\delta,i,N}),\qquad Z_t^{\delta,i,N}:=X_t^{i,N}-X_t^{\delta,i,N}.$$

Then, from (4.11) and the second SDE in (4.4), we deduce from Hölder's inequality and Burkhold-Davis-Gundy's inequality that

$$\begin{aligned}&\mathbb{E}\|\Theta^{\delta,i,N}\|_{\infty,t}^2\le C\Big\{\int_0^t\mathbb{E}|u_s^{\lambda,\mu^i}(X_s^{i,N})-u_s^{\lambda,\mu^i}(X_s^{\delta,i,N})|^2\mathrm{d}s\\&+\int_0^t\mathbb{E}\big|\nabla\theta_s^{\lambda,\mu^i}(b_s^{\hat\mu^N}-b_s^{\mu^i})(X_s^{i,N})-\nabla\theta_s^{\lambda,\mu^i}(X_s^{\delta,i,N})\big(b_{s_\delta}^{\hat\mu^{\delta,N}}(X_{s_\delta}^{\delta,i,N})-b_s^{\mu^i}(X_s^{\delta,i,N})\big)\big|^2\mathrm{d}s\\&+\int_0^t\mathbb{E}\bigg(\Big|\mathrm{Tr}[(\sigma_s^{\hat\mu^N}(\sigma_s^{\hat\mu^N})^*-\sigma_s^{\mu^i}(\sigma_s^{\mu^i})^*)\nabla^2u_s^{\lambda,\mu^i}](X_s^{i,N})\Big|^2\\&\qquad+\sum_{k,j=1}^d\big|\langle((\sigma_{s_\delta}^{\hat\mu^{\delta,N}}(\sigma_{s_\delta}^{\hat\mu^{\delta,N}})^*)(X_{s_\delta}^{\delta,i,N})-(\sigma_s^{\mu^i}(\sigma_s^{\mu^i})^*)(X_s^{\delta,i,N}))e_j,e_l\rangle\nabla_{e_j}\nabla_{e_l}u_s^{\lambda,\mu^i}(X_s^{\delta,i,N})\big|^2\bigg)\mathrm{d}s\\&+\int_0^t\big\|(\nabla\theta_s^{\lambda,\mu^i}\sigma_s^{\hat\mu^N})(X_s^{i,N})-\nabla\theta_s^{\lambda,\mu^i}(X_s^{\delta,i,N})\sigma_{s_\delta}^{\hat\mu^{\delta,N}}(X_{s_\delta}^{\delta,i,N})\big\|_{\mathrm{HS}}^2\mathrm{d}s\Big\}\\&=C\{J_{1,i}(t)+J_{2,i}(t)+J_{3,i}(t)+J_{4,i}(t)\}\end{aligned}$$

for some constant $C > 0$. In what follows, we intend to estimate $J_{k,i}(t), k = 1, 2, 3, 4$, one-by-one. Owing to (4.3), there exists a constant $c_1 > 0$ such that

$$J_{1,i}(t) \le c_1 \int_0^t \mathbb{E}|Z_s^{\delta,i,N}|^2 \mathrm{d}s. \tag{4.12}$$

Next, thanks to (1.14) and (4.3), it follows from (3.14) that

$$\begin{aligned} J_{2,i}(t) &\le c_2 \int_0^t \big\{\mathbb{E}\mathbb{W}_2(\mu_s^i, \hat{\mu}_s^N)^2 + \mathbb{E}|X_s^{\delta,i,N} - X_{s_\delta}^{\delta,i,N}|^{2\alpha} + \mathbb{E}\mathbb{W}_2(\mu_s^i, \hat{\mu}_{s_\delta}^{\delta,N})^2\big\}\mathrm{d}s \\ &\le c_3 \int_0^t \big\{\delta^\alpha + \mathbb{E}\mathbb{W}_2(\mu_s^i, \tilde{\mu}_s^N)^2 + \mathbb{E}\mathbb{W}_2(\tilde{\mu}_s^N, \hat{\mu}_s^N)^2 + \mathbb{E}\mathbb{W}_2(\tilde{\mu}_s^N, \hat{\mu}_{s_\delta}^{\delta,N})^2\big\}\mathrm{d}s \end{aligned} \tag{4.13}$$

for some constants $c_2, c_3 > 0$, where we have used, for some constant $c_{q,T} > 0$,

$$\mathbb{E}|X_t^{\delta,i,N} - X_{t_\delta}^{\delta,i,N}|^q \le c_{q,T}\delta^{q/2}, \qquad q > 0, \quad t \in [0, T] \tag{4.14}$$

which can be obtained in a standard way under (1.13). Set $Z_t^{i,N} := X_t^i - X_t^{i,N}, i \in \mathcal{S}_N$. On the other hand, by virtue of (3.14) and (4.14), we have for some constant $c_4 > 0$,

$$\begin{aligned} \mathbb{E}\mathbb{W}_2(\tilde{\mu}_t^N, \hat{\mu}_t^N)^2 + \mathbb{E}\mathbb{W}_2(\tilde{\mu}_t^N, \hat{\mu}_{t_\delta}^{\delta,N})^2 &\le \frac{1}{N}\sum_{j=1}^N \big\{\mathbb{E}|Z_t^{j,N}|^2 + \mathbb{E}|X_t^j - X_{t_\delta}^{\delta,j,N}|^2\big\} \\ &\le c_4\delta + \mathbb{E}|Z_t^{i,N}|^2 + 2\mathbb{E}|Z_t^{\delta,i,N}|^2, \end{aligned} \tag{4.15}$$

where in the last display we used the fact that $(Z^{j,N})_{j\in\mathcal{S}_N}$ (resp. $(Z^{\delta,j,N})_{j\in\mathcal{S}_N}$) are identically distributed. Then, plugging (4.15) back into (4.13) gives that

$$J_{2,i}(t) \le c_5 \int_0^t \big\{\delta^\alpha + \mathbb{E}\mathbb{W}_2(\mu_s^i, \tilde{\mu}_s^N)^2 + \mathbb{E}|Z_s^{i,N}|^2 + \mathbb{E}|Z_s^{\delta,i,N}|^2\big\}\mathrm{d}s \tag{4.16}$$

for some constant $c_5 > 0$. Similarly to $J_{2,i}(t)$ and (4.7) for $I_{3,i}(t)$, taking $(\mathbf{A}_1)$-$(\mathbf{A}_2)$, (4.3), and (4.14) into account, we find that there exists a constant $c_6 > 0$ such that

$$J_{3,i}(t) + J_{4,i}(t) \le c_6 \int_0^t \big\{\delta^\alpha + \mathbb{E}\mathbb{W}_2(\mu_s^i, \tilde{\mu}_s^N)^2 + \mathbb{E}|Z_s^{i,N}|^2 + \mathbb{E}|Z_s^{\delta,i,N}|^2\big\}\mathrm{d}s. \tag{4.17}$$

Now, combining (4.12), (4.16) with (4.17), we arrive at

$$\mathbb{E}\|\Theta^{\delta,i,N}\|_{\infty,t}^2 \le c_7 \int_0^t \big\{\delta^\alpha + \mathbb{E}\mathbb{W}_2(\mu_s^i, \tilde{\mu}_s^N)^2 + \mathbb{E}|X_s^i - X_s^{i,N}|^2 + \mathbb{E}|Z_s^{\delta,i,N}|^2\big\}\mathrm{d}s$$

for some constant $c_7 > 0$. This, together with $|Z_t^{\delta,i,N}|^2 \le 4|\Theta_t^{\delta,i,N}|^2$ due to (4.3), yields

$$\mathbb{E}\|Z^{\delta,i,N}\|_{\infty,t}^2 \le c_8 \int_0^t \big\{\delta^\alpha + \mathbb{E}\mathbb{W}_2(\mu_s^i, \tilde{\mu}_s^N)^2 + \mathbb{E}|X_s^i - X_s^{i,N}|^2 + \mathbb{E}|Z_s^{\delta,i,N}|^2\big\}\mathrm{d}s$$

for some constant $c_8 > 0$. Consequently, the desired assertion (4.10) holds true by applying Gronwall's inequality and employing (4.1) and (4.9). □

Proof of Theorem 1.5. On the basis of Lemmas 4.1 and 4.2, the proof of Theorem 1.5 can be complete. □

A Appendix

In the appendix section, we aim to complete

Proof of Theorem 1.4. Below, we follow the line of [22] to complete the proof of Theorem 1.4. For $\mu_\cdot \in C([0,T];\mathscr{P}_2(\mathbb{R}^d))$ and $x\in\mathbb{R}^d$, let $b_t^\mu(x)=b(x,\mu_t)$ and $\sigma_t^\mu(x)=\sigma(x,\mu_t)$. Consider the following time-dependent SDE

$$\mathrm{d}X_t^\mu = b_t^\mu(X_t^\mu)\mathrm{d}t+\sigma_t^\mu(X_t^\mu)\mathrm{d}W_t,\quad t\in[0,T]. \tag{A.1}$$

In terms of [9, Theorem 1.1], (A.1) is strongly wellposed. For any $\mu_\cdot \in C([0,T];\mathscr{P}_2(\mathbb{R}^d))$, let $\Phi_t(\mu)=\mathscr{L}_{X_t^\mu}$, where (X_t^μ) solves (A.1) with $X_0^\mu = X_0$, the initial value of (1.12). For any $\lambda>0$, consider the following PDE for $u^{\lambda,\mu}:[0,T]\times\mathbb{R}^d\to\mathbb{R}^d$:

$$\frac{\partial u_t^{\lambda,\mu}}{\partial t}+\frac{1}{2}\mathrm{Tr}\big(\sigma_t^\mu(\sigma_t^\mu)^*\nabla^2 u_t^{\lambda,\mu}\big)+\nabla_{b_t^\mu}u_t^{\lambda,\mu}+b_t^\mu=\lambda u_t^{\lambda,\mu},\quad u_T^{\lambda,\mu}=\mathbf{0}. \tag{A.2}$$

According to [2, Lemma 2.1], under $(\mathbf{A}_1)$ and $(\mathbf{A}_2)$, for $\lambda>0$ sufficiently large, (A.2) has a unique solution $u^{\lambda,\mu}$ satisfying

$$\|u^{\lambda,\mu}\|_\infty+\|\nabla u^{\lambda,\mu}\|_\infty\le\frac{1}{2},\quad \|\nabla^2 u^{\lambda,\mu}\|_\infty<\infty. \tag{A.3}$$

In the sequel, we set $t\in[0,T]$. For $\Theta_t^{\lambda,\mu}(x):=x+u_t^{\lambda,\mu}(x), x\in\mathbb{R}^d$, (A.2) and the Itô formula yield

$$\begin{aligned}
\mathrm{d}\Theta_t^{\lambda,\mu}(X_t^\mu)&=\lambda u_t^{\lambda,\mu}(X_t^\mu)\mathrm{d}t+(\nabla\Theta_t^{\lambda,\mu}\sigma_t^\mu)(X_t^\mu)\,\mathrm{d}W_t,\\
\mathrm{d}\Theta_t^{\lambda,\mu}(X_t^\nu)&=\lambda u_t^{\lambda,\mu}(X_t^\nu)\mathrm{d}t+(\nabla\Theta_t^{\lambda,\mu}\sigma_t)(X_t^\nu)\,\mathrm{d}W_t\\
&\qquad+\Big\{\big(\nabla\Theta_t^{\lambda,\mu}(b_t^\nu-b_t^\mu)\big)(X_t^\nu)+\frac{1}{2}\mathrm{Tr}\big((\sigma_t^\nu(\sigma_t^\nu)^*-\sigma_t^\mu(\sigma_t^\mu)^*)\nabla^2u_t^{\lambda,\mu}\big)(X_t^\nu)\Big\}\mathrm{d}t.
\end{aligned}$$

Whence, we have

$$\begin{aligned}
\mathrm{d}\big(\Theta_t^{\lambda,\mu}(X_t^\mu)-\Theta_t^{\lambda,\mu}(X_t^\nu)\big)&=\lambda\big(u_t^{\lambda,\mu}(X_t^\mu)-u_t^{\lambda,\mu}(X_t^\nu)\big)\mathrm{d}t\\
&\quad+\big((\nabla\Theta_t^{\lambda,\mu}\sigma_t^\mu)(X_t^\mu)-(\nabla\Theta_t^{\lambda,\mu}\sigma_t^\nu)(X_t^\nu)\big)\mathrm{d}W_t\\
&\quad-\Big\{\big(\nabla\Theta_t^{\lambda,\mu}(b_t^\nu-b_t^\mu)\big)(X_t^\nu)+\frac{1}{2}\mathrm{Tr}[(\sigma_t^\nu(\sigma_t^\nu)^*-\sigma_t^\mu(\sigma_t^\mu)^*)\nabla^2u_t^{\lambda,\mu}](X_t^\nu)\Big\}\mathrm{d}t\\
&=:\Lambda_1(t)\mathrm{d}t+\Lambda_2(t)\mathrm{d}W_t+\Lambda_3(t)\mathrm{d}t.
\end{aligned}$$

By (A.3), there exists a constant $c_1>0$ such that

$$|\Lambda_1(t)|\le c_0|X_t^\mu-X_t^\nu|. \tag{A.4}$$

Moreover, it follows from $(\mathbf{A}_1)$, $(\mathbf{A}_2)$ and (A.3) that for some constant $c_2>0$

$$\begin{aligned}
|\Lambda_3(t)|&\le c_2\mathbb{W}_2(\mu_t,\nu_t),\\
\|\Lambda_2(t)\|_{\mathrm{HS}}^2&\le 3\|(\nabla\Theta_t^{\lambda,\mu}\sigma_t^\mu)(X_t^\mu)-\nabla\Theta_t^{\lambda,\mu}(X_t^\mu)\sigma_t^\mu(X_t^\nu)\|_{\mathrm{HS}}^2\\
&\quad+3\|\nabla\Theta_t^{\lambda,\mu}(X_t^\mu)\sigma_t^\mu(X_t^\nu)-\nabla\Theta_t^{\lambda,\mu}(X_t^\mu)\sigma_t^\nu(X_t^\nu)\|_{\mathrm{HS}}^2\\
&\quad+3\|\nabla\Theta_t^{\lambda,\mu}(X_t^\mu)\sigma_t^\nu(X_t^\nu)-(\nabla\Theta_t^{\lambda,\mu}\sigma_t^\nu)(X_t^\nu)\|_{\mathrm{HS}}^2\\
&\le C_2\big(\mathbb{W}_2(\mu_t,\nu_t)^2+|X_t^\mu-X_t^\nu|^2\big).
\end{aligned} \tag{A.5}$$

Combining (A.4) with (A.5) and

$$|X_t^\mu - X_t^\nu| \le 2|\Theta_t^{\lambda,\mu}(X_t^\nu) - \Theta_t^{\lambda,\mu}(X_t^\nu)|$$

due to (A.3), we find that there exists a constant $c_3 > 0$ such that

$$\mathbb{E}|X_t^\mu - X_t^\nu|^2 \le c_3 \int_0^t \mathbb{E}|X_s^\mu - X_s^\nu|^2 \mathrm{d}t + c_3 \int_0^t \mathbb{W}_2(\mu_s, \nu_s)^2 \mathrm{d}s.$$

Thus, Gronwall's inequality implies

$$\mathbb{E}|X_t^\mu - X_t^\nu|^2 \le c_3 \mathrm{e}^{c_3 t} \int_0^t \mathbb{W}_2(\mu_s, \nu_s)^2 \mathrm{d}s.$$

This further yields

$$\mathbb{W}_2\big(\Phi_t(\mu), \Phi_t(\nu)\big)^2 \le \mathbb{E}|X_t^\mu - X_t^\nu|^2 \le c_3 \mathrm{e}^{c_3 t} \int_0^t \mathbb{W}_2(\mu_s, \nu_s)^2 \mathrm{d}s. \tag{A.6}$$

For $t_0 > 0$ (which indeed is independent of the initial value) such that $c_3 \mathrm{e}^{c_3 t_0} t_0 \le \frac{1}{2}$, set

$$\tilde{E}_{t_0} := \big\{\mu_\cdot \in C([0, t_0]; \mathscr{P}_2(\mathbb{R}^d)) : \mu_0 = \mathscr{L}_{X_0}\big\}$$

equipped with the uniform metric

$$\tilde{\rho}(\mu, \nu) := \sup_{0 \le t \le t_0} \mathbb{W}_2(\mu_t, \nu_t).$$

Hence, from (A.6), we have

$$\tilde{\rho}\big(\Phi(\mu), \Phi(\nu)\big) \le \frac{1}{2} \tilde{\rho}(\mu, \nu)$$

so that Φ is strictly contractive in $\tilde{E}_{t_0}$. Consequently, the Banach fixed point theorem, together with the definition of Φ implies that there exists a unique $\mu \in \tilde{E}_{t_0}$ such that

$$\Phi_t(\mu) = \mu_t = \mathscr{L}_{X_t^\mu}, \quad t \in [0, t_0].$$

Therefore, (1.12) is strongly wellposed in the time interval $[0, t_0]$. Next, by repeating the previous procedure with initial time it_0 and initial value X_{it_0} for $i \ge 1$, in finite many steps we may derive the strong well-posedness up to time T. □

References

[1] Airachid, H., Bossy, M., Ricci, C., Szpruch, L., New particle representations for ergodic McKean-Vlasov SDEs, *ESAIM: Proc. S.*, **65** (2019), 68–83.

[2] Bao, J., Huang, X., Yuan, C., Convergence rate of Euler-Maruyama scheme for SDEs with Hölder-dini Continuous drifts, *J. Theor. Probab.*, **32** (2019), 848–871.

[3] Bauer, M., Meyer-Brandis, T., Proske, F., Strong solutions of mean-field stochastic differential equations with irregular drift, *Electron. J. Probab.*, **23** (2018), 1–35.

[4] Buckdahn, R., Li, J., Ma, J., A mean-field stochastic control problem with partial observations, *Ann. Appl. Probab.*, **27** (2017), 3201–3245.

[5] Buckdahn, R., Li, J., Peng, S., Rainer, C., Mean-field stochastic differential equations and associated PDEs, *Ann. Probab.*, **45** (2017), 824–878.

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

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

[8] Chaman, K., Neelima, Christoph, R., Wolfgang, S., Well-posedness and tamed schemes for McKean-Vlasov equations with common noise, arXiv:2006.00463.

[9] Chaudru de Raynal, P.-E., Strong existence and uniqueness for stochastic differential equation with Hölder drift and degenerate noise, arxiv:1205.6688.

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

[11] Chassagneux, J.-F., Szpruch, L., Tse, A., Weak quantitative propagation of chaos via differential calculus on the space of measures, arXiv:1901.02556.

[12] Christoph, R., Wolfgang, S., An adaptive Euler-Maruyama scheme for McKean SDEs with super-linear growth and application to the mean-field FitzHugh-Nagumo model, arXiv:2005.06034.

[13] Cox, J. C., Ingersoll, J. E., Ross, S. A., An intertemporal general equilibrium model of asset prices, *Econometrica*, **53** (1985), 363–384.

[14] Crisan, D., McMurray, E., Smoothing properties of McKean-Vlasov SDEs, *Probab. Theory Related Fields*, **171** (2018), 97–148.

[15] dos Reis, G., Smith, G., Tankov, P., Importance sampling for McKean-Vlasov SDEs, arXiv:1803.09320.

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

[17] dos Reis, G., Engelhardt, S., Smith, G., Simulation of McKean-Vlasov SDEs with super linear growth, arXiv:1808.05530.

[18] Eberle, A., Guillin, A., Zimmer, R., Quantitative Harris-type theorems for diffusions and McKean-Vlasov processes, *Trans. Amer. Math. Soc.*, **371** (2019), 7135–7173.

[19] Gyöngy, I., Rásonyi, M., A note on Euler approximations for SDEs with Hölder continuous diffusion coefficients, *Stochastic Process. Appl.*, **121** (2011), 2189–2200.

[20] Hammersley, W., Siska, D., Szpruch, L., McKean-Vlasov SDEs under measure dependent Lyapunov conditions, arXiv:1802.03974v2.

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

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

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

[24] Kumar, C., Neelima, On explicit Milstein-type schemes for McKean-Vlasov Stochastic Differential Equations with super-linear drift coefficient, arXiv:2004.01266.

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

[26] Malrieu, F., Convergence to equilibrium for granular media equations and their Euler schemes, *Ann. Appl. Probab.*, **13** (2003), 540–560.

[27] Mehri, S., Stannat, W., Weak solutions to Vlasov-McKean equations under Lyapunov-type conditions, *Stoch. Dyn.*, **19** (2019), 1950042.

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

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

[30] Mishura, Y. S., Veretennikov, A. Yu, Existence and uniqueness theorems for solutions of McKean-Vlasov stochastic equations, arXiv:1603.02212v8.

[31] Ngo, H.-L., Taguchi, D., Strong rate of convergence for the Euler-Maruyama approximation of stochastic differential equations with irregular coefficients, *Math. Comput.*, **85** (2016), 1793–1819.

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

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

[34] Röckner, M., Zhang, X., Well-posedness of distribution dependent SDEs with singular drifts, arXiv:1809.02216.

[35] Situ, R., *Theory of stochastic differential equations with jumps and applications*, Springer, New York, 2005

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

[37] Szpruch, L., Tse, A., Antithetic multilevel particle system sampling method for McKean-Vlasov SDEs, arXiv:1903.07063.

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

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