Well-Posedness for Singular McKean-Vlasov Stochastic Differential Equations *

Xing Huang [a)], **Feng-Yu Wang** [a),b)]

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

xinghuang@tju.edu.cn

b)Department of Mathematics, Swansea University, Singleton Park, SA2 8PP, United Kingdom

wangfy@tju.edu.cn

December 18, 2020

Abstract

By using Zvonkin's transform and the heat kernel parameter expansion with respect to a frozen SDE, the well-posedness is proved for a McKean-Vlasov SDE with distribution dependent noise and singular drift, where the drift may be discontinuous in both weak topology and total variation distance, and is bounded by a linear growth term in distribution multiplying a locally integrable term in time-space. This extends existing results derived in the literature for distribution independent noise or time-space locally integrable drift.

AMS subject Classification: 60H1075, 60G44.
Keywords: Mckean-Vlasov SDEs, Wasserstein distance, weighted total variational distance, heat kernel parameter expansion.

1 Introduction

Let $\mathscr{P}$ be the set of all probability measures on $\mathbb{R}^d$. For $\theta \geq 1$, let

$$\mathscr{P}_\theta = \big\{\gamma \in \mathscr{P} : \|\gamma\|_\theta := \gamma(|\cdot|^\theta)^{\frac{1}{\theta}} < \infty\big\},$$

*Supported in part by NNSFC (11771326, 11831014, 11801406, 11921001) and the DFG through the CRC 1283.

which is a Polish space under the L^θ-Wasserstein distance $\mathbb{W}_\theta$:

$$\mathbb{W}_\theta(\gamma,\tilde\gamma):=\inf_{\pi\in\mathscr{C}(\gamma,\tilde\gamma)}\bigg(\int_{\mathbb{R}^d\times\mathbb{R}^d}|x-y|^\theta\pi(\mathrm{d}x,\mathrm{d}y)\bigg)^{\frac{1}{\theta}},\quad \gamma,\tilde\gamma\in\mathscr{P}_\theta,$$

where $\mathscr{C}(\gamma,\tilde\gamma)$ is the set of all couplings of γ and $\tilde\gamma$. Moreover, $\mathscr{P}_\theta$ is a complete metric space under the weighted variational norm

$$\|\mu-\nu\|_{\theta,TV}:=\sup_{|f|\le 1+|\cdot|^\theta}\big|\mu(f)-\nu(f)\big|,\quad \mu,\nu\in\mathscr{P}_\theta.$$

By [15, Theorem 6.15], there exists a constant $\kappa>0$ such that

$$\|\mu-\nu\|_{TV}+\mathbb{W}_\theta(\mu,\nu)\le\kappa\|\mu-\nu\|_{\theta,TV}, \tag{1.1}$$

where $\|\cdot\|_{TV}:=\|\cdot\|_{0,TV}$ is the total variation norm.

Consider the following distribution dependent SDE on $\mathbb{R}^d$:

$$\mathrm{d}X_t=b_t(X_t,\mathscr{L}_{X_t})\mathrm{d}t+\sigma_t(X_t,\mathscr{L}_{X_t})\mathrm{d}W_t,\quad t\in[0,T] \tag{1.2}$$

for some fixed time $T>0$, where W_t is an m-dimensional Brownian motion on a complete filtration probability space $(\Omega,\{\mathscr{F}_t\}_{t\ge0},\mathbb{P})$, $\mathscr{L}_{X_t}$ is the law of X_t, and

$$b:\mathbb{R}_+\times\mathbb{R}^d\times\mathscr{P}_\theta\to\mathbb{R}^d,\quad \sigma:\mathbb{R}_+\times\mathbb{R}^d\times\mathscr{P}_\theta\to\mathbb{R}^d\otimes\mathbb{R}^m$$

are measurable. This type equations, known as Mckean-Vlasov or mean field SDEs, have been intensively investigated and applied, see for instance the monograph [3] and references therein.

In this paper, we investigate the well-posedness of (1.2) with $b_t(x,\mu)$ singular in x and Lipschitz continuous in μ merely under $\|\cdot\|_{\theta,TV}$. To measure the time-space singularity of $b_t(x,\mu)$, we introduce the following class

$$\mathscr{K}:=\Big\{(p,q):p,q>1,\frac{d}{p}+\frac{2}{q}<1\Big\}.$$

For any $t>s\ge0$, we write $f\in\tilde L_p^q([s,t])$ if $f:[s,t]\times\mathbb{R}^d\to\mathbb{R}$ is measurable with

$$\|f\|_{\tilde L_p^q([s,t])}:=\sup_{z\in\mathbb{R}^d}\bigg\{\int_s^t\bigg(\int_{B(z,1)}|f(u,x)|^p\mathrm{d}x\bigg)^{\frac{q}{p}}\mathrm{d}u\bigg\}^{\frac{1}{q}}<\infty,$$

where $B(z,1):=\{x\in\mathbb{R}^d:|x-z|\le1\}$ is the unit ball at point z. When $s=0$, we simply denote

$$\tilde L_p^q(t)=\tilde L_p^q([0,t]),\quad \|f\|_{\tilde L_p^q(t)}=\|f\|_{\tilde L_p^q([0,t])}.$$

We will adopt the following assumption.

(A) Let $\theta \geq 1$.

(A_1) There exists a constant $K > 0$ such that for any $t \in [0,T], x, y \in \mathbb{R}^d$ and $\mu, \nu \in \mathscr{P}_\theta$,

$$\begin{aligned}
&\|\sigma_t(x,\mu)\|^2 \vee \|(\sigma_t\sigma_t^*)^{-1}(x,\mu)\| \leq K,\\
&\|\sigma_t(x,\mu) - \sigma_t(y,\nu)\| \leq K\big(|x-y| + \mathbb{W}_\theta(\mu,\nu)\big),\\
&\|\{\sigma_t(x,\mu) - \sigma_t(y,\mu)\} - \{\sigma_t(x,\nu) - \sigma_t(y,\nu)\}\| \leq K|x-y|\mathbb{W}_\theta(\mu,\nu).
\end{aligned}$$

(A_2) There exists nonnegative $f \in \tilde{L}_p^q(T)$ for some $(p,q) \in \mathscr{K}$ such that

$$\begin{aligned}
&|b_t(x,\mu)| \leq (1 + \|\mu\|_\theta) f_t(x),\\
&|b_t(x,\mu) - b_t(x,\nu)| \leq f_t(x)\|\mu - \nu\|_{\theta,TV}, \quad t \in [0,T], x \in \mathbb{R}^d, \mu, \nu \in \mathscr{P}_\theta.
\end{aligned}$$

Remark 1.1. (1) *It is easy to see that the third inequality in (A_1) holds if $\sigma_t(x,\mu)$ is differentiable in x with*

$$\|\nabla\sigma_t(\cdot,\mu)(x) - \nabla\sigma_t(\cdot,\nu)(x)\| \leq K\mathbb{W}_\theta(\mu,\nu), \quad \mu,\nu \in \mathscr{P}_\theta, x \in \mathbb{R}^d.$$

Indeed, this implies

$$\begin{aligned}
&\|\{\sigma_t(x,\mu) - \sigma_t(y,\mu)\} - \{\sigma_t(x,\nu) - \sigma_t(y,\nu)\}\|\\
&= \bigg\|\int_0^1 \big\{\nabla_{x-y}\sigma_t(y+s(x-y),\mu) - \nabla_{x-y}\sigma_t(y+s(x-y),\nu)\big\}\mathrm{d}s\bigg\|\\
&\leq \int_0^1 \big\|\nabla_{x-y}\sigma_t(y+s(x-y),\mu) - \nabla_{x-y}\sigma_t(y+s(x-y),\nu)\big\|\mathrm{d}s \leq K|x-y|\mathbb{W}_\theta(\mu,\nu).
\end{aligned}$$

(2) *Let $\sigma\sigma^*$ be uniformly positive definite. When the noise coefficient $\sigma_t(x,\mu) = \sigma_t(x)$ does not depend on μ, the well-posedness of* (1.2) *has been presented in [14] for $b_\cdot(\cdot,\mu) \in \tilde{L}_p^q(T)$ for some $(p,q) \in \mathscr{K}$ and $b_t(x,\cdot)$ being weakly continuous and Lipschitz continuous in $\|\cdot\|_{\theta,TV}$, and in [9] for $b = \bar{b} + \hat{b}$ with $\bar{b}_\cdot(\cdot,\mu) \in L_p^q(T)$ for some $(p,q) \in \mathscr{K}$, $\hat{b}_t(x,\mu)$ having linear growth in x, and $b_t(x,\cdot)$ being Lipschitz continuous in $\|\cdot\|_{TV} + \mathbb{W}_\theta$. In these conditions, the continuity of $b_t(x,\mu)$ in μ is stronger than that presented in (A_2) where $b_t(x,\mu)$ is allowed to be discontinuous in both $\|\cdot\|_{TV}$ and the weak topology. When $\sigma = \sqrt{2}I_d$ (where I_d is the $d\times d$ identity matrix) and $b_t(x,\mu) = \int_{\mathbb{R}^d} K_t(x-y)\mu(\mathrm{d}y)$ with $K \in \tilde{L}_p^q(T)$, the well-posedness of* (1.2) *is proved in [14] for $(p,q) \in \mathscr{K}$, while the weak existence is presented in [19] for some $p,q > 1$ with $\frac{d}{p} + \frac{2}{q} < 2$. When $\sigma_t(x,\mu)$ has linear functional derivative in μ which is Lipschitz continuous in the space variable uniformly in μ and t, the well-posedness is derived in [21] for $b_\cdot(\cdot,\mu) \in \tilde{L}_p^q(T)$ uniformly in μ for some $(p,q) \in \mathscr{K}$, while in [5] for $b_t(x,\mu)$ being bounded and Lipschitz continuous in μ under $\|\cdot\|_{TV}$. See also [1, 2, 4, 6, 8, 7, 12, 13, 16, 20] for earlier results on the well-posedness under different type or stronger conditions. Comparing with conditions in [5, 21], (A_2) allows $b_t(x,\mu)$ to have linear growth in μ and (A_1) does not require $\sigma_t(x,\mu)$ having linear functional derivative in μ. To include drifts with linear growth in the space variable, we hope that the first inequality in (A_2) could be weakened as*

$$|b_t(x,\mu)| \leq (1 + \|\mu\|_\theta)(K|x| + f_t(x))$$

for some constant $K > 0$ and $f \in \tilde{L}_p^q(T)$. But with this condition there is essential difficulty in the proof of Lemma 2.4 below on the heat kernel expansion.

Let $\hat{\mathscr{P}}$ be a subspace of $\mathscr{P}$. We call (1.2) well-posed for initial distributions in $\hat{\mathscr{P}}$, if for any $\mathscr{F}_0$-measurable random variable X_0 with $\mathscr{L}_{X_0} \in \hat{\mathscr{P}}$ and any $\mu_0 \in \hat{\mathscr{P}}$, (1.2) has a unique solution starting at X_0 as well as a unique weak solution starting at μ_0.

Theorem 1.2. *Assume* **(A)**. *Then* (1.2) *is well-posed for initial distributions in* $\mathscr{P}_{\theta+} := \cap_{m>\theta}\mathscr{P}_m$, *and the solution satisfies* $\mathscr{L}_{X_\cdot} \in C([0,T];\mathscr{P}_\theta)$, *the space of continuous maps from* $[0,T]$ *to* $\mathscr{P}_\theta$ *under the metric* $\mathbb{W}_\theta$. *Moreover,*

$$\mathbb{E}\Big[\sup_{t\in[0,T]} |X_t|^\theta\Big] < \infty. \tag{1.3}$$

Note that a Lipschitz function with respect to $\|\cdot\|_{\theta,TV}$ may be discontinuous in the weak topology and the total variation norm, for instance, $F(\mu) := \mu(f)$ with $f := h + |\cdot|^\theta$ for some bounded and discontinuous measurable function h on $\mathbb{R}^d$. So, this result extends existing well-posedness results derived in the above mentioned references.

To prove Theorem 1.2, besides Zvonkin's transform, Krylov's estimate and stochastic Gronwall's inequality used in [14], we will also apply the heat kernel parameter expansion with respect to a frozen SDE. This expansion is useful in the study of heat kernel estimates for distribution dependent SDEs and has been recently used in [10] to estimate the Lion's derivative of the solution to (1.2) with distribution dependent noise.

The remainder of the paper is organized as follows. Section 2 contains necessary preparations including some estimates on the map $\Phi^\gamma_{s,\cdot}$ in (2.2) below induced by (2.1) with a fixed distribution parameter $\mu_\cdot$ replacing $\mathscr{L}_{X_\cdot}$ in the drift term of (1.2). To derive these estimates, the heat kernel parameter expansion with respect to a frozen SDEs is used. With these preparations we prove Theorem 1.2 in Section 3.

2 Preparations

For any $0 \le s < t \le T$, let $C([s,t];\mathscr{P}_\theta)$ be the set of all continuous map from $[s,t]$ to $\mathscr{P}_\theta$ under the metric $\mathbb{W}_\theta$. For $\mu \in C([s,T];\mathscr{P}_\theta)$ and $\gamma \in \mathscr{P}_\theta$, we consider the following SDE with initial distribution $\mathscr{L}_{X^{\gamma,\mu}_{s,s}} = \gamma$ and fixed measure parameter μ_t in the drift:

$$\mathrm{d}X^{\gamma,\mu}_{s,t} = b_t(X^{\gamma,\mu}_{s,t}, \mu_t)\mathrm{d}t + \sigma_t(X^{\gamma,\mu}_{s,t}, \mathscr{L}_{X^{\gamma,\mu}_{s,t}})\mathrm{d}W_t, \quad t \in [s,T]. \tag{2.1}$$

According to Lemma 2.1 below, (A_1) and (A_2) imply the strong and weak well-posedness of (2.1) for initial distributions in $\mathscr{P}_\theta$, and the solution satisfies $\mathscr{L}_{X^{\gamma,\mu}_{s,\cdot}} \in C([s,T];\mathscr{P}_\theta)$. Consider the map

$$\Phi^\gamma_{s,\cdot} : C([s,T];\mathscr{P}_\theta) \to C([s,T];\mathscr{P}_\theta); \quad \Phi^\gamma_{s,t}(\mu) := \mathscr{L}_{X^{\gamma,\mu}_{s,t}}, \quad t \in [s,T], \mu \in C([s,T];\mathscr{P}_\theta). \tag{2.2}$$

It is easy to see that if $\mu_s = \gamma$ and μ is a fixed point of $\Phi^\gamma_{s,\cdot}$ (i.e. $\Phi^\gamma_{s,t}(\mu) = \mu_t, t \in [s,T]$), then $(X^{\gamma,\mu}_{s,t})_{t\in[s,T]}$ is a solution of (1.2) with initial distribution γ at time s.

To prove the existence and uniqueness of the fixed point for $\Phi^{\gamma}_{s,\cdot}$, we investigate the contraction of this map with respect to the complete metric

$$\|\mu-\nu\|_{s,t,\theta,TV} := \sup_{r\in[s,t]} \|\mu_r-\nu_r\|_{\theta,TV}, \quad \mu,\nu\in C([s,t];\mathscr{P}_\theta), \quad 0\le s<t\le T$$

in a subspace of $C([s,t];\mathscr{P}_\theta)$ which contains all distributions of solutions to (2.1) up to time t. To this end, in the following we first study the $\mathbb{W}_{s,t,\theta}$-estimate on $\Phi^{\gamma}_{s,\cdot}$ for

$$\mathbb{W}_{s,t,\theta}(\mu,\nu) \;=\; \sup_{r\in[s,t]} \mathbb{W}_\theta(\mu_r,\nu_r), \quad 0\le s\le t\le T, \mu,\nu\in C([s,t];\mathscr{P}_\theta),$$

then present $\Phi^{\gamma}_{s,\cdot}$-invariant subspaces of $C([s,t];\mathscr{P}_\theta)$, and finally study the $\|\cdot\|_{s,t,\theta,TV}$-contraction of this map in such an invariant subspace which implies the well-posedness of (1.2).

2.1 $\mathbb{W}_{s,t,\theta}$-estimate on $\Phi^{\gamma}_{s,\cdot}$

For any $N>0$ and $0\le s\le t\le T$, let

$$\begin{aligned}&\mathscr{P}_{\theta,N} = \big\{\gamma\in\mathscr{P}_\theta : \|\gamma\|_\theta\le N\big\},\\ &\mathscr{P}^{s,t}_{\theta,N} := \big\{\mu\in C([s,t];\mathscr{P}_\theta) : \|\mu_r\|_\theta\le N, r\in[s,t]\big\}.\end{aligned}$$

Lemma 2.1. *Assume* **(A)**.

(1) *For any $s\in[0,T)$ and $\mu\in C([s,T];\mathscr{P}_\theta)$, (2.1) is well-posed for initial distributions in $\mathscr{P}_\theta$, and the unique solution satisfies $\mathscr{L}_{X^{\gamma,\mu}_{s,\cdot}}\in C([s,T];\mathscr{P}_\theta)$.*

(2) *There exist a constant $\epsilon\in(0,1]$ and a function $K:(0,\infty)\to(0,\infty)$ such that for any $\gamma\in\mathscr{P}_{\theta,N}$ and $0\le s\le t\le T$, the map $\Phi^{\gamma}_{s,\cdot}: C([s,t];\mathscr{P}_\theta)\to C([s,t];\mathscr{P}_\theta)$ satisfies*

$$\mathbb{W}_{s,t,\theta}(\Phi^{\gamma}_{s,\cdot}(\mu),\Phi^{\gamma}_{s,\cdot}(\nu)) \le K_N(t-s)^{\epsilon}\|\mu-\nu\|_{s,t,\theta,TV}, \quad N>0, \mu,\nu\in\mathscr{P}^{s,t}_{\theta,N}.$$

Proof. (1) Simply denote

$$b^{\nu}_t(x) = b_t(x,\nu_t), \quad \sigma^{\nu}_t(x)=\sigma_t(x,\nu_t), \quad \nu\in C([s,T];\mathscr{P}_\theta), t\in[s,T].$$

Let $X^{\gamma}_{s,s}$ be $\mathscr{F}_s$-measurable with $\mathscr{L}_{X^{\gamma}_{s,s}}=\gamma$, and let $\mu,\nu\in\mathscr{P}^{s,T}_{\theta,N}$ for some $N>0$. For $\bar\nu,\bar\mu\in C([s,T];\mathscr{P}_\theta)$, consider the SDEs

$$\mathrm{d}X^{\gamma,\mu,\bar\mu}_{s,t} = b^{\mu}_t(X^{\gamma,\mu,\bar\mu}_{s,t})\mathrm{d}t + \sigma^{\bar\mu}_t(X^{\gamma,\mu,\bar\mu}_{s,t})\mathrm{d}W_t, \quad X^{\gamma,\mu,\bar\mu}_{s,s} = X^{\gamma}_{s,s}, \; t\in[s,T], \tag{2.3}$$

$$\mathrm{d}X^{\gamma,\nu,\bar\nu}_{s,t} = b^{\nu}_t(X^{\gamma,\nu,\bar\nu}_{s,t})\mathrm{d}t + \sigma^{\bar\nu}_t(X^{\gamma,\nu,\bar\nu}_{s,t})\mathrm{d}W_t, \quad X^{\gamma,\nu,\bar\nu}_{s,s} = X^{\gamma}_{s,s}, \; t\in[s,T]. \tag{2.4}$$

According to [19], both SDEs are well-posed under assumption **(A)**. For any $\lambda\ge0$, consider the following PDE for $u:[0,T]\times\mathbb{R}^d\to\mathbb{R}^d$:

$$\frac{\partial u_t}{\partial t} + \frac12\mathrm{Tr}(\sigma^{\bar\mu}_t(\sigma^{\bar\mu}_t)^*\nabla^2u_t) + \{\nabla u_t\}b^{\mu}_t + b^{\mu}_t = \lambda u_t, \quad t\in[s,T], u_T=0. \tag{2.5}$$

According to [17, Theorem 3.1], for large enough $\lambda > 0$ depending on N via $\mu \in \mathscr{P}^{s,T}_{\theta,N}$, (A_1) and (A_2) imply that (2.5) has a unique solution $\mathbf{u}^{\lambda,\mu,\bar\mu}$ satisfying

$$\sup_{\mu\in\mathscr{P}^{s,T}_{\theta,N}} \|\nabla^2\mathbf{u}^{\lambda,\mu,\bar\mu}\|_{\tilde L^p_q(T)} \le L \tag{2.6}$$

for some constant $L > 0$ depending on λ and N, and

$$\sup_{\mu\in\mathscr{P}^{s,T}_{\theta,N}} (\|\mathbf{u}^{\lambda,\mu,\bar\mu}\|_\infty + \|\nabla\mathbf{u}^{\lambda,\mu,\bar\mu}\|_\infty) \le \frac{1}{5}. \tag{2.7}$$

Let $\Theta^{\lambda,\mu,\bar\mu}_t(x) = x + \mathbf{u}^{\lambda,\mu,\bar\mu}_t(x), (t,x) \in [s,T]\times\mathbb{R}^d$. By [17, Lemma 4.1 (iii)], we have

$$\begin{aligned}
&\mathrm{d}\Theta^{\lambda,\mu,\bar\mu}_t(X^{\gamma,\mu,\bar\mu}_{s,t}) = \lambda\mathbf{u}^{\lambda,\mu,\bar\mu}_t(X^{\gamma,\mu,\bar\mu}_{s,t})\mathrm{d}t + (\{\nabla\Theta^{\lambda,\mu,\bar\mu}_t\}\sigma^{\bar\mu}_t)(X^{\gamma,\mu,\bar\mu}_{s,t})\,\mathrm{d}W_t,\\
&\mathrm{d}\Theta^{\lambda,\mu,\bar\mu}_t(X^{\gamma,\nu,\bar\nu}_{s,t}) = \big[\lambda\mathbf{u}^{\lambda,\mu,\bar\mu}_t + \{\nabla\Theta^{\lambda,\mu,\bar\mu}_t\}(b^\nu_t - b^\mu_t)\big](X^{\gamma,\nu,\bar\nu}_{s,t})\mathrm{d}t\\
&\qquad + \frac{1}{2}\big[\mathrm{Tr}\{\sigma^{\bar\nu}_t(\sigma^{\bar\nu}_t)^* - \sigma^{\bar\mu}_t(\sigma^{\bar\mu}_t)^*\}\nabla^2\mathbf{u}^{\lambda,\mu,\bar\mu}_t\big](X^{\gamma,\nu,\bar\nu}_{s,t}) + (\{\nabla\Theta^{\lambda,\mu,\bar\mu}_t\}\sigma^{\bar\nu}_t)(X^{\gamma,\nu,\bar\nu}_{s,t})\,\mathrm{d}W_t.
\end{aligned}$$

where by (2.7) and $\gamma \in \mathscr{P}_\theta$ the first equation implies $\mathbb{E}\big[\sup_{t\in[s,T]}|X^{\gamma,\mu,\bar\mu}_{s,t}|^\theta\big] < \infty$, so that

$$\mathscr{L}_{X^{\gamma,\mu,\bar\mu}_{s,\cdot}} \in C([s,T];\mathscr{P}_\theta),\quad \bar\mu \in C([s,T];\mathscr{P}_\theta). \tag{2.8}$$

Moreover, combining these two equations with **(A)**, we find a constant $c_1 > 1$ depending on N such that $\eta_{s,t} := |X^{\gamma,\mu,\bar\mu}_{s,t} - X^{\gamma,\nu,\bar\nu}_{s,t}|$ satisfies

$$\begin{aligned}
c_1^{-1}\eta_{s,t} &\le \big|\Theta^{\lambda,\mu,\bar\mu}_t(X^{\gamma,\mu,\bar\mu}_{s,t}) - \Theta^{\lambda,\mu,\bar\mu}_t(X^{\gamma,\nu,\bar\nu}_{s,t})\big|\\
&\le c_1\int_s^t \big\{\eta_{s,r} + \|\mu_r - \nu_r\|_{\theta,TV}(1 + f_r(X^{\gamma,\nu,\bar\nu}_{s,r}))\\
&\qquad + \mathbb{W}_\theta(\bar\mu_r,\bar\nu_r)\|\nabla^2\mathbf{u}^{\lambda,\mu,\bar\mu}_r(X^{\gamma,\nu,\bar\nu}_{s,r})\|\big\}\mathrm{d}r + \bigg|\int_s^t \Xi_r\mathrm{d}W_r\bigg|,
\end{aligned} \tag{2.9}$$

where $\Xi_r := (\{\nabla\Theta^{\lambda,\mu,\bar\mu}_r\}\sigma^{\bar\mu}_r)(X^{\gamma,\mu,\bar\mu}_{s,r}) - (\{\nabla\Theta^{\lambda,\mu,\bar\mu}_r\}\sigma^{\bar\nu}_r)(X^{\gamma,\nu,\bar\nu}_{s,r})$ satisfies

$$\|\Xi_r\| \le c_1\eta_{s,r} + c_1\mathbb{W}_\theta(\bar\mu_r,\bar\nu_r) + c_1\|\nabla\mathbf{u}^{\lambda,\mu,\bar\mu}_r(X^{\gamma,\mu,\bar\mu}_{s,r}) - \nabla\mathbf{u}^{\lambda,\mu,\bar\mu}_r(X^{\gamma,\nu,\bar\nu}_{s,r})\|. \tag{2.10}$$

Since $\eta_{s,s} = 0$, by (2.9), (2.10) and $(A1)$, for $2m > \theta$, we find a constant $c_2 > 0$ depending on N and a local martingale $(M_t)_{t\in[s,T]}$ such that

$$\begin{aligned}
\eta^{2m}_{s,t} &\le c_2\int_s^t \eta^{2m}_{s,r}\mathrm{d}A_r + c_2\int_s^t \mathbb{W}_\theta(\bar\mu_r,\bar\nu_r)^{2m}\mathrm{d}r\\
&\quad + c_2\|\mu-\nu\|^{2m}_{s,t,\theta,TV}\bigg|\int_s^t (1 + f_r(X^{\gamma,\nu,\bar\nu}_{s,r}))\mathrm{d}r\bigg|^{2m} + M_t,\quad t\in[s,T]
\end{aligned} \tag{2.11}$$

holds for

$$A_t := \int_s^t \Big\{1 + K^2 + \|\nabla^2\mathbf{u}^{\lambda,\mu,\bar\mu}_r\|(X^{\gamma,\nu,\bar\nu}_{s,r})$$

$$
\begin{aligned}
&+\Big[\big(\mathscr{M}|\nabla^2\mathbf{u}_r^{\lambda,\mu,\bar\mu}|\big)(X_{s,r}^{\gamma,\mu,\bar\mu})+\big(\mathscr{M}|\nabla^2\mathbf{u}_r^{\lambda,\mu,\bar\mu}|\big)(X_{s,r}^{\gamma,\nu,\bar\nu})\Big]^2\Big\}\mathrm{d}r,\\
&\mathscr{M}g(x):=\sup_{r\in[0,1]}\frac{1}{|B(x,r)|}\int_{B(x,r)}g(y)\mathrm{d}y,\quad g\in L^1_{loc}(\mathbb{R}^d),\\
&B(x,r):=\{y\in\mathbb{R}^d:|x-y|\le r\},\quad x\in\mathbb{R}^d.
\end{aligned}
$$

By Krylov's and Khasminskii's estimates [17, (4.1),(4.2)] and (2.6), and applying the stochastic Gronwall inequality [18, Lemma 3.8], we find constants $c_3>0$ depending on N such that (2.11) yields

$$
\begin{aligned}
\text{(2.12)}\qquad &\{\mathbb{W}_\theta(\mathscr{L}_{X_{s,t}^{\gamma,\mu,\bar\mu}},\mathscr{L}_{X_{s,t}^{\gamma,\nu,\bar\nu}})\}^{2m}\le(\mathbb{E}\eta_{s,t}^\theta)^{\frac{2m}{\theta}}\\
&\le c_3\left(\int_s^t\mathbb{W}_\theta(\bar\mu_r,\bar\nu_r)^{2m}\mathrm{d}r+\|\mu-\nu\|_{s,t,\theta,TV}^{2m}(t-s)^{2m\epsilon}\right),\quad t\in[s,T]
\end{aligned}
$$

for some $\epsilon\in(0,1)$. Taking $\mu=\nu$ gives

$$
\mathbb{W}_{s,t,\theta}(\mathscr{L}_{X_{s,t}^{\gamma,\mu,\bar\mu}},\mathscr{L}_{X_{s,t}^{\gamma,\mu,\bar\nu}})\le\{c_3(t-s)\}^{\frac{1}{2m}}\mathbb{W}_{s,t,\theta}(\bar\mu,\bar\nu),\quad t\in[s,T].
$$

Letting $t_0=\frac{1}{2c_4}$, we conclude from this and (2.8) that the map

$$
\bar\mu\mapsto\mathscr{L}_{X_{s,\cdot}^{\gamma,\mu,\bar\mu}}
$$

is contractive in $C([s,(s+t_0)\wedge T];\mathscr{P}_\theta)$ under the complete metric $\mathbb{W}_{s,(s+t_0)\wedge T,\theta}$. Therefore, it has a unique fixed point $\bar\mu=\mathscr{L}_{X_{s,\cdot}^{\gamma,\mu,\bar\mu}}\in C([s,(s+t_0)\wedge T];\mathscr{P}_\theta)$, so that $X_{s,\cdot}^{\gamma,\mu,\bar\mu}$ is the unique solution of (2.1) up to time $(s+t_0)\wedge T$. Due to this and the well-posedness of (2.3), tthe modified Yamada-Watanabe principle [9, Lemma 2.1] also implies the well-posedness of (2.1) up to time $(s+t_0)\wedge T$ for initial distributions in $\mathscr{P}_\theta$. So, if $s+t_0\ge T$ then we have proved the first assertion. Otherwise, by the same argument we may consider (2.1) from time $s+t_0$ to conclude that it is well-posed up to time $(s+2t_0)\wedge T$. Repeating finite many times we prove the well-posedness of (2.1) up to time T.

(2) By taking $\bar\mu=\Phi_{s,\cdot}^\gamma(\mu)$, $\bar\nu=\Phi_{s,\cdot}^\gamma(\nu)$ in (2.12), and applying Gronwall's inequality, we find a constant $C>0$ depending on N such that

$$
\mathbb{W}_{s,t,\theta}(\Phi_{s,\cdot}^\gamma(\mu),\Phi_{s,\cdot}^\gamma(\nu))^{2m}\le C(t-s)^{2m\epsilon}\|\mu-\nu\|_{s,t,\theta,TV}^{2m},\quad t\in[s,T],\mu,\nu\in\mathscr{P}_{\theta,N}^{s,t}.
$$

This finishes the proof. □

2.2 Invariant subspaces of $\Phi_{s,\cdot}^\gamma$

Lemma 2.2. *Assume* **(A)**. *There exists a constant $N_0>0$ such that for any $N\ge N_0$ and $0\le s<t\le T$, $\gamma\in\mathscr{P}_\theta$, the class*

$$
\mathscr{P}_{\theta,N}^{s,t,\gamma}:=\Big\{\mu\in C([s,t];\mathscr{P}_\theta):\mu_s=\gamma,\ \sup_{r\in[s,t]}(1+\|\mu_r\|_\theta)\mathrm{e}^{-N(r-s)}\le2(1+\|\gamma\|_\theta)\Big\}
$$

is $\Phi_{s,\cdot}^\gamma$-invariant, i.e. $\mu\in\mathscr{P}_{\theta,N}^{s,t,\gamma}$ implies $\Phi_{s,\cdot}^\gamma(\mu)\in\mathscr{P}_{\theta,N}^{s,t,\gamma}$.

Proof. Simply denote $\xi_r = X^{\gamma,\mu}_{s,r}, M_r = \int_s^r \sigma_u(\xi_u, \mathscr{L}_{\xi_u})\mathrm{d}W_u, r \in [s,T]$. By **(A)**, (2.1), and $\mu \in \mathscr{P}^{s,t,\gamma}_{\theta,N}$, we have

$$\begin{aligned}|\xi_r|\mathrm{e}^{-N(r-s)} &\le |\xi_0|\mathrm{e}^{-N(r-s)} + \mathrm{e}^{-N(r-s)}\int_s^r (1+\|\mu_u\|_\theta) f_u(\xi_u)\mathrm{d}u + \mathrm{e}^{-N(r-s)}\big|M_r\big| \\ &\le |\xi_0|\mathrm{e}^{-N(r-s)} + 2(1+\|\gamma\|_\theta)\int_s^r \mathrm{e}^{-N(r-u)} f_u(\xi_u)\mathrm{d}u + \mathrm{e}^{-N(r-s)}\big|M_r\big|.\end{aligned}$$

Let $q' \in (1,q)$ such that $(p,q') \in \mathscr{K}$. Combining this with Krylov's estimate [17, (4.1),(4.2)], the BDG inequality, and $\|\sigma\sigma^*\| \le K$, we find a constant $c_1 > 0$ such that

$$\begin{aligned}&\mathrm{e}^{-N(r-s)}\|\Phi^\gamma_{s,r}(\mu)\|_\theta = \mathrm{e}^{-N(r-s)}(\mathbb{E}|\xi_r|^\theta)^{\frac{1}{\theta}} \\ &\le \mathrm{e}^{-N(r-s)}\|\gamma\|_\theta + 2(1+\|\gamma\|_\theta)\bigg(\mathbb{E}\bigg|\int_s^r \mathrm{e}^{-N(r-u)} f_u(\xi_u)\mathrm{d}u\bigg|^\theta\bigg)^{\frac{1}{\theta}} + \mathrm{e}^{-N(r-s)}\big(\mathbb{E}|M_r|^\theta\big)^{\frac{1}{\theta}} \\ &\le \mathrm{e}^{-N(r-s)}\|\gamma\|_\theta + c_1(1+\|\gamma\|_\theta)\big(\|\mathrm{e}^{-N(r-\cdot)}f\|_{\tilde{L}^{q'}_p([s,r])} + \mathrm{e}^{-N(r-s)}\sqrt{r-s}\big), \quad r \in [s,t].\end{aligned} \tag{2.13}$$

Noting that Hölder's inequality yields

$$\begin{aligned}\sup_{s\in[0,T),r\in[s,T]} \|\mathrm{e}^{-N(r-\cdot)}f\|_{\tilde{L}^{q'}_p([s,r])} &\le \sup_{s\in[0,T),r\in[s,T]} \bigg(\int_s^r \mathrm{e}^{-N(r-u)\frac{qq'}{q-q'}}\mathrm{d}u\bigg)^{\frac{q-q'}{qq'}} \|f\|_{\tilde{L}^q_p([T])} \\ &\le \bigg(N\frac{qq'}{q-q'}\bigg)^{-\frac{q-q'}{qq'}} \|f\|_{\tilde{L}^q_p(T)},\end{aligned}$$

we obtain

$$\lim_{N\to\infty} \sup_{s\in[0,T),r\in[s,T]} \big(\|\mathrm{e}^{-N(r-\cdot)}f\|_{\tilde{L}^{q'}_p([s,r])} + \mathrm{e}^{-N(r-s)}\sqrt{r-s}\big) = 0.$$

Combining this with (2.13), we find a constant $N_0 > 0$ such that

$$\sup_{r\in[s,t]} (1+\|\Phi^\gamma_{s,r}(\mu)\|_\theta)\mathrm{e}^{-N(r-s)} \le 2(1+\|\gamma\|_\theta), \quad N \ge N_0, \mu \in \mathscr{P}^{s,t,\gamma}_{\theta,N}.$$

That is, $\Phi^\gamma_{s,\cdot}(\mu) \in \mathscr{P}^{s,t,\gamma}_{\theta,N}$ for $N \ge N_0$ and $\mu \in \mathscr{P}^{s,t,\gamma}_{\theta,N}$. □

2.3 $\|\cdot\|_{s,t,\theta,TV}$-contraction of $\Phi^\gamma_{s,\cdot}$

To prove the $\|\cdot\|_{s,t,\theta,TV}$-contraction of $\Phi^\gamma_{s,\cdot}$, for any $\mu \in C([s,T];\mathscr{P}_\theta)$, we make use of the parameter expansion of $p^{\gamma,\mu}_{s,t}$ with respect to the heat kernel of a frozen SDE whose solution is a Gaussian Markov process, where $p^{\gamma,\mu}_{s,t}(x,\cdot)$ is the distribution density function of the unique solution to the SDE

$$\mathrm{d}X^{x,\gamma,\mu}_{s,t} = b_t(X^{x,\gamma,\mu}_{s,t},\mu_t)\mathrm{d}t + \sigma_t(X^{x,\gamma,\mu}_{s,t},\Phi^\gamma_{s,t}(\mu))\mathrm{d}W_t, \quad t \in [s,T], \quad X^{x,\gamma,\mu}_{s,s} = x. \tag{2.14}$$

According to [19], **(A)** is enough to ensure the well-posedness of this SDE. By the standard Markov property of solutions to (2.14), the solution to (2.1) satisfies

$$\mathbb{E} f(X_{s,t}^{\gamma,\mu}) = \int_{\mathbb{R}^d} \gamma(\mathrm{d}x) \int_{\mathbb{R}^d} f(y) p_{s,t}^{\gamma,\mu}(x,y)\mathrm{d}y, \quad t > s, f \in \mathscr{B}_b(\mathbb{R}^d), \gamma \in \mathscr{P}_\theta, \tag{2.15}$$

where $\mathscr{B}_b(\mathbb{R}^d)$ is the class of bounded measurable functions on $\mathbb{R}^d$.

For any $z \in \mathbb{R}^d, t \in [s,T]$ and $\mu \in C([s,t];\mathscr{P}_\theta)$, let $p_{s,r}^{\gamma,\mu,z}(x,\cdot)$ be the distribution density function of the random variable

$$X_{s,r}^{x,\gamma,\mu,z} := x + \int_s^r \sigma_u(z, \Phi_{s,u}^\gamma(\mu))\mathrm{d}W_u, \quad r \in [s,t], x \in \mathbb{R}^d.$$

Let

$$a_{s,r}^{\gamma,\mu,z} := \int_s^r (\sigma_u \sigma_u^*)(z, \Phi_{s,u}^\gamma(\mu))\mathrm{d}u, \quad r \in [s,t]. \tag{2.16}$$

Then

$$p_{s,r}^{\gamma,\mu,z}(x,y) = \frac{\exp[-\frac{1}{2}\langle (a_{s,r}^{\gamma,\mu,z})^{-1}(y-x), y-x\rangle]}{(2\pi)^{\frac{d}{2}}(\det\{a_{s,r}^{\gamma,\mu,z}\})^{\frac{1}{2}}}, \quad x,y \in \mathbb{R}^d, r \in (s,t]. \tag{2.17}$$

Obviously, (A_1) implies

$$\begin{aligned} &\|a_{s,r}^{\gamma,\mu,z} - a_{s,r}^{\gamma,\nu,z}\| \le K(r-s)\mathbb{W}_{s,r,\theta}\big(\Phi_{s,\cdot}^\gamma(\mu), \Phi_{s,\cdot}^\gamma(\nu)\big),\\ &\frac{1}{K(r-s)} \le \|(a_{s,r}^{\gamma,\mu,z})^{-1}\| \le \frac{K}{r-s}, \quad r \in [s,t]. \end{aligned} \tag{2.18}$$

For any $r \in [s,t)$ and $y,z \in \mathbb{R}^d$, let

$$\begin{aligned} H_{r,t}^{\gamma,\mu}(y,z) :=& \big\langle -b_r(y,\mu_r), \nabla p_{r,t}^{\gamma,\mu,z}(\cdot,z)(y)\big\rangle\\ &+ \frac{1}{2}\mathrm{tr}\left[\big\{(\sigma_r\sigma_r^*)(z,\Phi_{s,r}^\gamma(\mu)) - (\sigma_r\sigma_r^*)(y,\Phi_{s,r}^\gamma(\mu))\big\}\nabla^2 p_{r,t}^{\gamma,\mu,z}(\cdot,z)(y)\right]. \end{aligned} \tag{2.19}$$

By **(A)**, we have the parameter expansion formula

$$p_{s,t}^{\gamma,\mu}(x,z) = p_{s,t}^{\gamma,\mu,z}(x,z) + \sum_{m=1}^\infty \int_s^t \mathrm{d}r \int_{\mathbb{R}^d} H_{r,t}^{\gamma,\mu,m}(y,z) p_{s,r}^{\gamma,\mu,z}(x,y)\mathrm{d}y, \tag{2.20}$$

where $H_{r,t}^{\gamma,\mu,m}$ for $m \in \mathbb{N}$ are defined by

$$\begin{aligned} &H_{r,t}^{\gamma,\mu,1} := H_{r,t}^{\gamma,\mu},\\ &H_{r,t}^{\gamma,\mu,m}(y,z) := \int_r^t \mathrm{d}u \int_{\mathbb{R}^d} H_{u,t}^{\gamma,\mu,m-1}(z',z) H_{r,u}^{\gamma,\mu}(y,z')\mathrm{d}z', \ m \ge 2. \end{aligned} \tag{2.21}$$

Note that (2.20) follows from the parabolic equations for the heat kernels $p_{s,t}^{\gamma,\mu}$ and $p_{s,t}^{\gamma,\mu,z}$, see for instance the paragraph after [11, Lemma 3.1] for an explanation.

Let

$$\tilde{p}^K_{s,r}(x,y)=\frac{\exp[-\frac{1}{4K(r-s)}|y-x|^2]}{(4K\pi(r-s))^{\frac{d}{2}}},\quad x,y\in\mathbb{R}^d, r\in(s,t]. \tag{2.22}$$

By multiplying the time parameter with T^{-1} to make it stay in $[0,1]$, we deduce from [21, (2.3), (2.4)] with $\beta=\beta'=1$ and $\lambda=\frac{1}{8KT}$ that

$$\begin{aligned}&\int_s^t\int_{\mathbb{R}^d}\tilde{p}^K_{s,r}(x,y')(r-s)^{-\frac{1}{2}}g_r(y')(t-r)^{-\frac{1}{2}}\tilde{p}^{2K}_{r,t}(y',y)\mathrm{d}y'\\&\le c(t-s)^{-\frac{1}{2}+\frac{1}{2}(1-\frac{d}{p}-\frac{2}{q})}\tilde{p}^{2K}_{s,t}(x,y)\|g\|_{\tilde{L}^q_p([s,t])},\quad 0\le s\le t\le T, g\in\tilde{L}^q_p([s,t])\end{aligned} \tag{2.23}$$

holds for some constant $c>0$ depending on T,d,p,q and K. By (A_1) and (2.17), there exists a constant $c_1>0$ such that

$$\begin{aligned}&p^{\gamma,\mu,z}_{s,t}(x,y)\bigg(1+\frac{|x-y|^4}{(t-s)^2}\bigg)\\&\le c_1\tilde{p}^K_{s,t}(x,y),\quad x,y,z\in\mathbb{R}^d, 0\le s\le t\le T,\gamma\in\mathscr{P}_\theta,\mu\in C([s,t];\mathscr{P}_\theta).\end{aligned} \tag{2.24}$$

Lemma 2.3. *Assume (A_1). There exists a constant $c>0$ such that for any $0\le s<t\le T, x,y,z\in\mathbb{R}^d,\gamma\in\mathscr{P}_\theta$, and $\mu,\nu\in C([s,t];\mathscr{P}_\theta)$,*

$$\bigg(1+\frac{|x-y|^2}{t-s}\bigg)|p^{\gamma,\mu,z}_{s,t}(x,y)-p^{\gamma,\nu,z}_{s,t}(x,y)|\le c\tilde{p}^K_{s,t}(x,y)\mathbb{W}_{s,t,\theta}(\Phi^\gamma_{s,\cdot}(\mu),\Phi^\gamma_{s,\cdot}(\nu)), \tag{2.25}$$

$$\sqrt{t-s}|\nabla p^{\gamma,\mu,z}_{s,t}(\cdot,y)(x)|+(t-s)\|\nabla^2p^{\gamma,\mu,z}_{s,t}(\cdot,y)(x)\|\le c\tilde{p}^K_{s,t}(x,y), \tag{2.26}$$

$$\begin{aligned}&\sqrt{t-s}|\nabla p^{\gamma,\mu,z}_{s,t}(\cdot,y)(x)-\nabla p^{\gamma,\nu,z}_{s,t}(\cdot,y)(x)|\\&+(t-s)\|\nabla^2p^{\gamma,\mu,z}_{s,t}(\cdot,y)(x)-\nabla^2p^{\gamma,\nu,z}_{s,t}(\cdot,y)(x)\|\\&\le c\tilde{p}^K_{s,t}(x,y)\mathbb{W}_{s,t,\theta}(\Phi^\gamma_{s,\cdot}(\mu),\Phi^\gamma_{s,\cdot}(\nu)).\end{aligned} \tag{2.27}$$

Proof. (1) Let $F(a,\mu)=\langle(a^{\mu,z}_{s,t})^{-1}(y-x),y-x\rangle$ and $F(a,\nu)$ be defined with ν in place of μ. It is easy to see that

$$\begin{aligned}&|p^{\gamma,\mu,z}_{s,t}(x,y)-p^{\gamma,\nu,z}_{s,t}(x,y)|\\&=\left|\frac{\exp[-\frac{1}{2}F(a,\mu)]}{(2\pi)^{\frac{d}{2}}(\det\{a^{\mu,z}_{s,t}\})^{\frac{1}{2}}}-\frac{\exp[-\frac{1}{2}F(a,\nu)]}{(2\pi)^{\frac{d}{2}}(\det\{a^{\nu,z}_{s,t}\})^{\frac{1}{2}}}\right|\\&\le\frac{\left|\exp[-\frac{1}{2}F(a,\mu)]-\exp[-\frac{1}{2}F(a,\nu)]\right|}{(2\pi)^{\frac{d}{2}}(\det\{a^{\mu,z}_{s,t}\})^{\frac{1}{2}}}\\&\quad+\frac{\exp[-\frac{1}{2}F(a,\nu)]}{(2\pi)^{\frac{d}{2}}}\left|(\det\{a^{\mu,z}_{s,t}\})^{-\frac{1}{2}}-(\det\{a^{\nu,z}_{s,t}\})^{-\frac{1}{2}}\right|\\&=:I_1+I_2,\quad y\in\mathbb{R}^d,t>s.\end{aligned} \tag{2.28}$$

Combining this with (A_1) which implies (2.18), we find a constant $c_1 > 0$ such that

$$|F(a,\mu) - F(a,\nu)| = \big|\langle\{(a^{\mu,z}_{s,t})^{-1} - (a^{\nu,z}_{s,t})^{-1}\}(y-x), y-x\rangle\big|$$
$$\le c_1\frac{|y-x|^2}{t-s}\mathbb{W}_{s,t,\theta}(\Phi^\gamma_{s,\cdot}(\mu), \Phi^\gamma_{s,\cdot}(\nu)),$$

which together with (2.24) yields that for some constant $c_2 > 0$,

$$\left(1+\frac{|x-y|^2}{t-s}\right)I_1 \le c_2\tilde{p}^K_{s,t}(x,y)\mathbb{W}_{s,t,\theta}(\Phi^\gamma_{s,\cdot}(\mu), \Phi^\gamma_{s,\cdot}(\nu)).$$

Next, by (2.18) and (2.24), we find a constant $c_3 > 0$ such that

$$\left(1+\frac{|x-y|^2}{t-s}\right)I_2 \le c_3\tilde{p}^K_{s,t}(x,y)\mathbb{W}_{s,t,\theta}(\Phi^\gamma_{s,\cdot}(\mu), \Phi^\gamma_{s,\cdot}(\nu)).$$

Combining these with (2.28), we arrive at

$$\left(1+\frac{|x-y|^2}{t-s}\right)|p^{\gamma,\mu,z}_{s,t}(x,y) - p^{\gamma,\nu,z}_{s,t}(x,y)| \le (c_2+c_3)\tilde{p}^K_{s,t}(x,y)\mathbb{W}_{s,t,\theta}(\Phi^\gamma_{s,\cdot}(\mu), \Phi^\gamma_{s,\cdot}(\nu)).$$

(2) By (2.17) we have

$$\nabla p^{\gamma,\mu,z}_{s,t}(\cdot,y)(x) = (a^{\mu,z}_{s,t})^{-1}(y-x)p^{\gamma,\mu,z}_{s,t}(x,y), \tag{2.29}$$

$$\nabla^2 p^{\gamma,\mu,z}_{s,t}(\cdot,y)(x) = p^{\gamma,\mu,z}_{s,t}(x,y)\Big(\big\{(a^{\mu,z}_{s,t})^{-1}(y-x)\big\}\otimes\big\{(a^{\mu,z}_{s,t})^{-1}(y-x)\big\} - (a^{\mu,z}_{s,t})^{-1}\Big) \tag{2.30}$$

So, by (2.18) and (2.24) we find a constant $c > 0$ such that (2.26) holds. Moreover, (2.29) implies

$$|\nabla p^{\gamma,\mu,z}_{s,t}(\cdot,y)(x) - \nabla p^{\gamma,\nu,z}_{s,t}(\cdot,y)(x)|$$
$$\le \big|\{(a^{\mu,z}_{s,t})^{-1} - (a^{\nu,z}_{s,t})^{-1}\}(y-x)\big|\, p^{\gamma,\mu,z}_{s,t}(x,y) + \big|\{p^{\gamma,\mu,z}_{s,t}(x,y) - p^{\gamma,\nu,z}_{s,t}(x,y)\}(a^{\nu,z}_{s,t})^{-1}(y-x)\big|.$$

Combining this with (2.18), (2.24) and (2.25), we find a constant $c > 0$ such that

$$|\nabla p^{\gamma,\mu,z}_{s,t}(\cdot,y)(x) - \nabla p^{\gamma,\nu,z}_{s,t}(\cdot,y)(x)| \le \frac{c\mathbb{W}_{s,t,\theta}(\Phi^\gamma_{s,\cdot}(\mu), \Phi^\gamma_{s,\cdot}(\nu))}{\sqrt{t-s}}\tilde{p}^K_{s,t}(x,y).$$

Similarly, combining (2.30) with (2.18), (2.24) and (2.25), we find a constant $c > 0$ such that

$$\|\nabla^2 p^{\gamma,\mu,z}_{s,t}(\cdot,y)(x) - \nabla^2 p^{\gamma,\nu,z}_{s,t}(\cdot,y)(x)\| \le \frac{c\mathbb{W}_{s,t,\theta}(\Phi^\gamma_{s,\cdot}(\mu), \Phi^\gamma_{s,\cdot}(\nu))}{t-s}\tilde{p}^K_{s,t}(x,y).$$

Therefore, (2.27) holds for some constant $c > 0$. □

For $0 \le s \le t \le T, \gamma \in \mathscr{P}_\theta$ and $\mu,\nu \in C([s,t];\mathscr{P}_\theta)$, let

$$\Lambda_{s,t,\gamma}(\mu,\nu) = \mathbb{W}_{s,t,\theta}(\Phi^\gamma_{s,\cdot}(\mu), \Phi^\gamma_{s,\cdot}(\nu)) + \|\mu-\nu\|_{s,t,\theta,TV}. \tag{2.31}$$

Lemma 2.4. *Assume* (**A**). *Let*

$$\delta = \frac{1}{2}\left(1 - \frac{d}{p} - \frac{2}{q}\right) > 0,$$
$$S^{\mu}_{s,t} := \sup_{r\in[s,t]} (1 + \|\mu_r\|_\theta),$$
$$S^{\mu,\nu}_{s,t} := S^{\mu}_{s,t} \vee S^{\nu}_{s,t}, \quad 0 \le s \le t \le T, \mu \in C([0,T]; \mathscr{P}_\theta).$$

Then there exists a constant $C \ge 1$ *such that for any* $0 \le s \le t \le T$, $y, z \in \mathbb{R}^d$, $\mu, \nu \in C([0,T]; \mathscr{P}_\theta)$, *and* $m \ge 1$,

$$|H^{\gamma,\mu,m}_{s,t}(y,z)| \le f_s(y)(CS^{\mu}_{s,t})^m (t-s)^{-\frac{1}{2}+\delta(m-1)} \tilde{p}^{2K}_{s,t}(x,y), \tag{2.32}$$

$$\begin{aligned} &\|H^{\gamma,\mu,m}_{s,t}(y,z) - H^{\gamma,\nu,m}_{s,t}(y,z)\| \\ &\le m f_s(y)(CS^{\mu,\nu}_{s,t})^m (t-s)^{-\frac{1}{2}+\delta(m-1)} \tilde{p}^{2K}_{s,t}(x,y)\Lambda_{s,t,\gamma}(\mu,\nu). \end{aligned} \tag{2.33}$$

Proof. (1) By (2.19), (2.26) and (*A*1)-(*A*2), we find a constant $c_1 > 0$ such that for any $0 \le s < t \le T, \mu \in C([0,T]; \mathscr{P}_\theta)$ and $y, z \in \mathbb{R}^d$,

$$|H^{\gamma,\mu}_{s,t}(y,z)| \le c_1(t-s)^{-\frac{1}{2}}\{(1 + \|\mu_s\|_\theta) f_s(y)\} \tilde{p}^K_{s,t}(y,z). \tag{2.34}$$

So, (2.32) holds for $m = 1$ and $C = c_1$. Thanks to [21, (2.3), (2.4)] with $\beta = \beta' = 1$, $\lambda = \frac{1}{8K}$, we have

$$\begin{aligned} I_k &:= \int_s^t \int_{\mathbb{R}^d} (t-u)^{-\frac{1}{2}} (t-u)^{\delta(k-1)} \tilde{p}^{2K}_{u,t}(y,z) f_u(y)(u-s)^{-\frac{1}{2}} \tilde{p}^K_{s,u}(x,y) \mathrm{d}y \mathrm{d}u \\ &\le c_2(t-s)^{-\frac{1}{2}} \tilde{p}^{2K}_{s,t}(x,z)(t-s)^{\frac{1}{2}(1-\frac{d}{p}-\frac{2}{q})} \|f\|_{\tilde{L}^q_p([s,t])}(t-s)^{\delta(k-1)} \\ &= c_3(t-s)^{-\frac{1}{2}} \tilde{p}^{2K}_{s,t}(x,z)(t-s)^{\delta k}. \quad 0 \le s < t \le T, k \ge 1 \end{aligned} \tag{2.35}$$

where $c_3 := c_2\|f\|_{\tilde{L}^q_p([s,t])}$. Let $C := 1 \vee c_1^2 \vee (4c_3^2)$. If for some $k \ge 1$ we have

$$|H^{\gamma,\mu,k}_{s,t}(y,z)| \le (CS^{\mu}_{s,t})^k f_s(y) \tilde{p}^{2K}_{s,t}(y,z)(t-s)^{-\frac{1}{2}+\delta(k-1)}$$

for all $y, z \in \mathbb{R}^d$ and $0 \le s \le t \le T$, then by combining with (2.34) and (2.35), we arrive at

$$\begin{aligned} |H^{\gamma,\mu,k+1}_{s,t}(y,z)| &\le \int_s^t \mathrm{d}u \int_{\mathbb{R}^d} |H^{\gamma,\mu,k}_{u,t}(z',z) H^{\gamma,\mu}_{s,u}(y,z')| \mathrm{d}z' \\ &\le C^k \sqrt{C} (S^{\mu}_{s,t})^{k+1} f_s(y) I_k \\ &\le C^{k+1} (S^{\mu}_{s,t})^{k+1} f_s(y)(t-s)^{-\frac{1}{2}+\delta k} \tilde{p}^{2K}_{s,t}(y,z). \end{aligned}$$

Therefore, (2.32) holds for all $m \ge 1$.

(2) By (2.26), (2.27), (2.18) and (*A*1)-(*A*2), we find a constant $c > 0$ such that for any $0 \le s < t \le T, \mu, \nu \in C([0,T]; \mathscr{P}_\theta)$ and $y, z \in \mathbb{R}^d$,

$$\begin{aligned} &|H^{\gamma,\mu}_{s,t}(y,z) - H^{\gamma,\nu}_{s,t}(y,z)| \\ &\le c(t-s)^{-\frac{1}{2}} \tilde{p}^K_{s,t}(y,z) S^{\mu,\nu}_{s,t} f_s(y) \left(\|\mu - \nu\|_{s,t,\theta,TV} + \mathbb{W}_{s,t,\theta}(\Phi^{\gamma}_{s,\cdot}(\mu), \Phi^{\gamma}_{s,\cdot}(\nu))\right) \end{aligned} \tag{2.36}$$

Let, for instance, $L = 1 + 4C^2 + 4c^2$, where C is in (2.32). If for some $k \geq 1$ we have

$$\|H^{\gamma,\mu,k}_{s,t}(z',z) - H^{\gamma,\nu,k}_{s,t}(z',z)\| \leq k(LS^{\mu,\nu}_{s,t})^k f_s(z')\tilde{p}^{2K}_{s,t}(z',z)(t-s)^{-\frac{1}{2}+\delta(k-1)}\Lambda_{s,t,\gamma}(\mu,\nu),$$

for any $0 \leq s < t \leq T$ and $z, z' \in \mathbb{R}^d$, then (2.32), (2.35) and (2.36) imply

$$\begin{aligned}
&\|H^{\gamma,\mu,k+1}_{s,t}(y,z) - H^{\gamma,\nu,k+1}_{s,t}(y,z)\| \\
&\leq \int_s^t \mathrm{d}r \int_{\mathbb{R}^d} \Big\{ \|H^{\gamma,\mu,k}_{r,t}(z',z) - H^{\gamma,\nu,k}_{r,t}(z',z)\| \cdot |H^{\gamma,\mu}_{s,r}(y,z')| \\
&\qquad\qquad + |H^{\gamma,\nu,k}_{r,t}(z',z)| \cdot \|H^{\gamma,\mu}_{s,r}(y,z') - H^{\gamma,\nu}_{s,r}(y,z')\| \Big\} \mathrm{d}z' \\
&\leq (k+1)(LS^{\mu,\nu}_{s,t})^{k+1} f_s(y)\tilde{p}^{2K}_{s,t}(y,z)(t-s)^{-\frac{1}{2}+\delta k}\Lambda_{s,t,\gamma}(\mu,\nu).
\end{aligned}$$

Therefore, (2.33) holds for some constant $C > 0$. □

We are now ready to prove the following main result in this part, which ensures the $\|\cdot\|_{s,t,\theta,TV}$-contraction of $\Phi^\gamma_{s,\cdot}$ for small $t-s$.

Lemma 2.5. *Assume* **(A)**. *There exist constants* $\varepsilon_0, \varepsilon \in (0,1)$ *and a function* $\phi : (0,\infty) \to [0,\infty)$ *such that*

$$\|\Phi^\gamma_{s,\cdot}(\mu) - \Phi^\gamma_{s,\cdot}(\nu)\|_{s,t,\theta,TV} \leq \phi(N)(t-s)^\varepsilon \|\mu - \nu\|_{s,t,\theta,TV}\gamma\big(1 + |\cdot|^\theta\big)$$

holds for any $N > 0, s \in [0,T), t \in [s, (s + \varepsilon_0 N^{-1/\delta}) \wedge T], \mu, \nu \in \mathscr{P}^{s,t}_{\theta,N}$ *and* $\gamma \in \mathscr{P}_{\theta,N}$.

Proof. By Lemma 2.1, for any $N > 0$, we find a constant $\phi_1(N) > 0$ such that

$$\begin{aligned}
&\mathbb{W}_{s,t,\theta}(\Phi^\gamma_{s,\cdot}(\mu), \Phi^\gamma_{s,\cdot}(\nu)) \\
&\leq \phi_1(N)(t-s)^\epsilon \|\mu-\nu\|_{s,t,\theta,TV}, \quad 0 \leq s \leq t \leq T, \mu,\nu \in \mathscr{P}^{s,t,\gamma}_{\theta,N}, \gamma \in \mathscr{P}_{\theta,N}.
\end{aligned}$$

Combining this with (2.24), Lemma 2.3-Lemma 2.4, (2.20), (2.23) and (A_2), we find $\phi_2(N) > 0$ such that for $\varepsilon := \epsilon \wedge \delta$ and $\varepsilon_0 := (2C)^{-1/\delta}$,

$$\begin{aligned}
&|p^{\gamma,\mu}_{s,t}(x,z) - p^{\gamma,\nu}_{s,t}(x,z)| \\
&\leq c\tilde{p}^K_{s,t}(x,z)\mathbb{W}_{s,t,\theta}(\Phi^\gamma_{s,\cdot}(\mu), \Phi^\gamma_{s,\cdot}(\nu)) + \sum_{m=1}^\infty \int_s^t \mathrm{d}r \int_{\mathbb{R}^d} |H^{\gamma,\mu,m}_{r,t}(y,z) - H^{\gamma,\nu,m}_{r,t}(y,z)| p^{\gamma,\nu,z}_{s,r}(x,y)\mathrm{d}y \\
&\qquad + \sum_{m=1}^\infty \int_s^t \mathrm{d}r \int_{\mathbb{R}^d} |H^{\gamma,\mu,m}_{r,t}(y,z)||p^{\gamma,\mu,z}_{s,r}(x,y) - p^{\gamma,\nu,z}_{s,r}(x,y)|\mathrm{d}y \\
&\leq c\tilde{p}^K_{s,t}(x,z)\mathbb{W}_{s,t,\theta}(\Phi^\gamma_{s,\cdot}(\mu), \Phi^\gamma_{s,\cdot}(\nu)) \\
&+ \sum_{m=1}^\infty (m+1)(CN)^m \Lambda_{s,t,\gamma}(\mu,\nu)(t-s)^{\frac{1}{2}+\delta(m-1)} \\
&\qquad \times \int_s^t \int_{\mathbb{R}^d} (t-r)^{-\frac{1}{2}}\tilde{p}^{2K}_{r,t}(y,z) f_r(y)(r-s)^{-\frac{1}{2}}\tilde{p}^K_{s,r}(x,y)\mathrm{d}y\mathrm{d}r
\end{aligned}$$

$$
\begin{aligned}
&\le c\tilde{p}^K_{s,t}(x,z)\mathbb{W}_{s,t,\theta}(\Phi^\gamma_{s,\cdot}(\mu),\Phi^\gamma_{s,\cdot}(\nu))\\
&\qquad + (t-s)^\delta \Lambda_{s,t,\gamma}(\mu,\nu)\tilde{p}^{2K}_{s,t}(x,z)\sum_{m=1}^\infty (m+1)(CN)^m(t-s)^{\delta(m-1)}\\
&\le \phi_2(N)(t-s)^\varepsilon \|\mu-\nu\|_{s,t,\theta,TV}\tilde{p}^{2K}_{s,t}(x,z),\quad x,z\in\mathbb{R}^d
\end{aligned}
$$

holds for any $N>0, 0\le s\le t\le (s+\varepsilon_0 N^{-1/\delta})\wedge T,\ \mu,\nu\in\mathscr{P}^{s,t,\gamma}_{\theta,N}$, and $\gamma\in\mathscr{P}_{\theta,N}$. So, by (2.22) and the definitions of $\Phi^\gamma_{s,t}$ and $\|\cdot\|_{\theta,TV}$, we find a constant $\phi(N)>0$ such that

$$
\begin{aligned}
&\|\Phi^\gamma_{s,t}(\mu)-\Phi^\gamma_{s,t}(\nu)\|_{\theta,TV}\\
&= \sup_{|g|\le 1+|\cdot|^\theta}\left|\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} g(z)p^{\gamma,\mu}_{s,t}(x,z)\mathrm{d}z\gamma(\mathrm{d}x) - \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} g(z)p^{\gamma,\nu}_{s,t}(x,z)\mathrm{d}z\gamma(\mathrm{d}x)\right|\\
&\le \phi_2(N)(t-s)^\varepsilon\|\mu-\nu\|_{s,t,\theta,TV}\int_{\mathbb{R}^d\times\mathbb{R}^d}(1+|z|^\theta)\tilde{p}^{2K}_{s,t}(x,z)\mathrm{d}z\gamma(\mathrm{d}x)\\
&\le \phi(N)\gamma\big(1+|\cdot|^\theta\big)(t-s)^\varepsilon\|\mu-\nu\|_{s,t,\theta,TV},\quad t\in[s,(s+\varepsilon_0N^{-1/\delta})\wedge T].
\end{aligned}
$$

Then the proof is finished. □

3 Proof of Theorem 1.2

To prove Theorem 1.2 using the contraction result Lemma 2.5, we need the following priori-estimates on the solution of (1.2).

Lemma 3.1. *Assume* (**A**) *and let* $m>\theta$. *Then there exists a constant* $N_1>0$ *such that for any* $s\in[0,T)$, $X_{s,s}\in L^m(\Omega\to\mathbb{R}^d,\mathscr{F}_s,\mathbb{P})$, *and* $T'\in(0,T]$, *a solution* $(X_{s,t})_{t\in[s,T']}$ *of* (1.2) *with initial value at* $X_{s,s}$ *from* s *up to* T' *satisfies*

$$
\mathbb{E}\Big[\sup_{t\in[s,T']}(1+|X_{s,t}|^2)^{\frac{\theta}{2}}\Big]\le \mathrm{e}^{N_1T'}\left(\mathbb{E}(1+|X_{s,s}|^2)^{\frac{m}{2}}\right)^{\frac{\theta}{m}},\quad t\in[s,T'].
$$

Proof. Without loss of generality, we may and do assume that $s=0$ and denote $X_t=X_{0,t}$. By Itô's formula, (A_2) and the boundedness of σ, there exists a constant $c_2>0$ such that

$$
\begin{aligned}
&\mathrm{d}(1+|X_t|^2)^{\frac{m}{2}}\\
&\le \frac{m}{2}(1+|X_t|^2)^{\frac{m}{2}-1}\left(2\langle X_t, b_t(X_t,\mathscr{L}_{X_t})\rangle\mathrm{d}t + \|\sigma_t(X_t,\mathscr{L}_{X_t})\|^2_{HS}\mathrm{d}t\right)\\
&+\frac{m}{4}(\frac{m}{2}-1)(1+|X_t|^2)^{\frac{m}{2}-2}|\sigma_t(X_t,\mathscr{L}_{X_t})^*X_t|^2\mathrm{d}t\\
&+m(1+|X_t|^2)^{\frac{m}{2}-1}\langle X_t,\sigma_t(X_t,\mathscr{L}_{X_t})\mathrm{d}W_t\rangle\\
&\le mK(1+|X_t|^2)^{\frac{m}{2}}f_t(X_t)\mathrm{d}t + (1+\|\mathscr{L}_{X_t}\|_\theta)^m f_t(X_t)\mathrm{d}t + c_2\mathrm{d}t\\
&+m(1+|X_t|^2)^{\frac{m}{2}-1}\langle X_t,\sigma_t(X_t,\mathscr{L}_{X_t})\mathrm{d}W_t\rangle.
\end{aligned}
$$

Thanks to Krylov's estimate and Khasminskii's estimate [17, (4.1),(4.2)], $(A1)$-$(A2)$ and the stochastic Gronwall inequality [18, Lemma 3.8] yield

$$\left(\mathbb{E}\sup_{t\in[0,T']}(1+|X_t|^2)^{\frac{\theta}{2}}\right)^{\frac{m}{\theta}}\le \mathrm{e}^{N_1T'}\mathbb{E}(1+|X_0|^2)^{\frac{m}{2}}$$

for some constant $N_1>0$ independent of X_0 and $T'\in[0,T]$. □

Proof of Theorem 1.2. Since (1.3) is implied by Lemma 3.1, we only prove the well-posedness of (1.2) for initial distributions in $\mathscr{P}_m$ for some $m>\theta$. Since according to [17] the assumption **(A)** implies the well-posedness of the SDE

$$\mathrm{d}X_t^\mu=b_t^\mu(X_t^\mu)+\sigma_t^\mu(X_t^\mu)\mathrm{d}W_t,\quad X_0^\mu=X_0$$

for $\mu\in C([0,T];\mathscr{P}_\theta)$ and $\mathbb{E}|X_0|^\theta<\infty$, by the modified Yamada-Watanabe principle [9, Lemma 2.1] we only need to prove the strong well-posedness of (1.2) with an initial value X_0 such that $E|X_0|^m<\infty$.

(1) Let N_0 and N_1 be in Lemmas 2.2 and 3.1 respectively. Take

$$N=\max\left\{N_0,\mathrm{e}^{N_1(1+T)}\mathbb{E}(1+|X_0|^2)^{\frac{\theta}{2}}\right\}. \tag{3.1}$$

Since $N\ge N_0$, Lemma 2.2 implies

$$\Phi_{s,\cdot}^\gamma:\ \mathscr{P}_{\theta,N}^{s,t,\gamma}\to\mathscr{P}_{\theta,N}^{s,t,\gamma},\quad 0\le s\le t\le T.$$

Moreover, by Lemma 2.5, there exists constant $C>0$ depending on N such that

$$\begin{aligned}&\|\Phi_{s,\cdot}^\gamma(\mu)-\Phi_{s,\cdot}^\gamma(\nu)\|_{s,t,\theta,TV}\\&\le C(t-s)^\varepsilon\|\mu-\nu\|_{s,t,\theta,TV},\quad \mu,\nu\in\mathscr{P}_{\theta,N}^{s,t,\gamma},\gamma\in\mathscr{P}_{\theta,N},t\in[s,(s+\varepsilon_0N^{-\frac{1}{\delta}})\wedge T].\end{aligned}$$

Taking $t_0\in(0,\varepsilon_0N^{-\frac{1}{\delta}})$ such that $Ct_0<1$, we conclude that $\Phi_{s,\cdot}^\gamma$ is contractive in $\mathscr{P}_{\theta,N}^{s,(s+t_0)\wedge T,\gamma}$ for any $s\in[0,T)$ and $\gamma\in\mathscr{P}_{\theta,N}$. Below we prove that this implies the existence and uniqueness of solution of (1.2).

(2) Let $s=0$ and $\gamma=\mathscr{L}_{X_0}$. By (1) and the fixed point theorem, there exists a unique $\mu\in\mathscr{P}_{\theta,N}^{0,t_0\wedge T,\gamma}$ such that $\mu_t=\Phi_{s,t}^\gamma(\mu)$ for $t\in[0,t_0\wedge T]$. Combining this with the definition of $\Phi_{s,t}^\gamma(\mu)$, we conclude that $X_{s,t}^{\gamma,\mu}$ is a solution of (1.2) up to time $t_0\wedge T$. Moreover, it is easy to see that the distribution of a solution to (1.2) is a fixed point of the map $\Phi_{0,\cdot}^\gamma$, and by Lemma 3.1 and (3.1) a solution of (1.2) up to time $t_0\wedge T$ must in the space $\mathscr{P}_{\theta,N}^{0,t_0\wedge T,\gamma}$. Therefore, (1.2) has a unique solution up to time $t_0\wedge T$.

(3) If $t_0\ge T$ then the proof is done. Assume that for some integer $k\ge1$ the equation (1.2) has a unique solution $(X_t)_{t\in[0,kt_0]}$ up to time $kt_0\le T$, we take $s=kt_0$ and $\gamma=\mathscr{L}_{X_{kt_0}}$. By Lemma 3.1 and (3.1), we have $\gamma\in\mathscr{P}_{\theta,N}$, so that $\Phi_{kt_0,\{(k+1)t_0\}\wedge T}^\gamma$ is contractive in $\mathscr{P}_{\theta,N}^{kt_0,\{(k+1)t_0\}\wedge T,\gamma}$. Hence, as explained in (2) that the SDE (1.2) has a unique solution from time $s=kt_0$ up to $\{(k+1)t_0\}\wedge T$. This together with the assumption we conclude that (1.2) has a unique solution up to time $\{(k+1)t_0\}\wedge T$. In conclusion, we have proved the existence and uniqueness of solution to (1.2).

□

References

[1] M. Bauer, T. M-Brandis, *Existence and Regularity of Solutions to Multi-Dimensional Mean-Field Stochastic Differential Equations with Irregular Drift,* arXiv:1912.05932.

[2] M. Bauer, T. M-Brandis, F. Proske, *Strong Solutions of Mean-Field Stochastic Differential Equations with irregular drift,* arXiv:1806.11451.

[3] P. Cardaliaguet, F. Delarue, J.-M. Lasry, P.-L. Lions, *The Master Equation and the Convergence Problem in Mean Field Games,* Princeton University Press, 2019.

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

[5] P. E. Chaudru de Raynal, N. Frikha *Well-posedness for some non-linear diffusion processes and related pde on the wasserstein space,* arXiv:1811.06904.

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

[7] X. Huang, M. Röckner, F.-Y. Wang, *Nonlinear Fokker–Planck equations for probability measures on path space and path-distribution dependent SDEs,* Disc. Cont. Dyn. Syst. Ser. A 39(2019), 3017-3035.

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

[9] X. Huang, F.-Y. Wang, *McKean-Vlasov SDEs with drifts discontinuous under Wasserstein distance,* Disc. Cont. Dyn. Syst. Ser. A. doi: 10.3934/dcds.2020336.

[10] X. Huang, F.-Y. Wang, *Derivative Estimates on Distributions of McKean-Vlasov S-DEs,* arXiv:2006.16731.

[11] V. Konakov, E. Mammen, *Local limit theorems for transition densities of Markov chains converging to diffusions,* Probab. Theory Relat. Fields 117(2000), 551-587.

[12] D. Lacker, *On a strong form of propagation of chaos for McKean-Vlasov equations,* arXiv:1805.04476.

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

[14] M. Röckner, X. Zhang, *Well-posedness of distribution dependent SDEs with singular drifts,* arXiv:1809.02216. To appear in Bernoulli.

[15] C. Villani *Optimal Transport, Old and New,* Springer-Verlag, 2009.

[16] F.-Y. Wang, *Distribution-dependent SDEs for Landau type equations,* Stoch. Proc. Appl. 128(2018), 595-621.

[17] P. Xia, L. Xie, X. Zhang, G. Zhao, *$L^q(L^p)$-theory of stochastic differential equations,* arXiv:1908.01255.

[18] L. Xie, X. Zhang, *Ergodicity of stochastic differential equations with jumps and singular coefficients,* Ann. Inst. Henri Poincaré Probab. Stat. 56(2020), 175-229.

[19] X. Zhang, *Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients,* Electron. J. Probab. 16(2011), 1096-1116.

[20] X. Zhang, *Weak solutions of McKean-Vlasov SDEs with supercritical drifts,* arXiv:2010.15330.

[21] G. Zhao, *On Distribution depend SDEs with singular drifts,* arXiv:2003.04829.