WELL-POSEDNESS OF DISTRIBUTION DEPENDENT SDES WITH SINGULAR DRIFTS

Michael Röckner and Xicheng Zhang

Abstract. In this paper we consider the following distribution dependent SDE:

$$\mathrm{d}X_t = \sigma_t(X_t, \mu_{X_t})\mathrm{d}W_t + b_t(X_t, \mu_{X_t})\mathrm{d}t,$$

where μ_{X_t} stands for the distribution of X_t. We show the strong well-posedness of the above SDE under some integrability assumptions in the spatial variable and Lipschitz continuity in μ about b and σ. In particular, we extend the results of Krylov-Röckner to the distribution dependent case.

Keywords: Distribution dependent SDEs, McKean-Vlasov system, Zvonkin's transformation, Singular drifts, Superposition principle

AMS 2010 Mathematics Subject Classification: Primary: 60H10, 35K55.

1. Introduction

Let $\mathcal{P}(\mathbb{R}^d)$ be the space of all probability measures over $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, which is endowed with the weak convergence topology. Consider the following distribution dependent SDEs (abbreviated as DDSDEs):

$$\mathrm{d}X_t = b_t(X_t, \mu_{X_t})\mathrm{d}t + \sigma_t(X_t, \mu_{X_t})\mathrm{d}W_t, \tag{1.1}$$

where $b : \mathbb{R}_+ \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$ and $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d \otimes \mathbb{R}^d$ are two Borel measurable functions, $\mu_{X_t} := \mathbb{P} \circ X_t^{-1}$ is the probability distribution measure of X_t, W is a d-dimensional standard Brownian motion. By Itô's formula, it is easy to see that μ_{X_t} satisfies the following non-linear Fokker-Planck equation (abbreviated as FPE) in the distributional sense:

$$\partial_t \mu_{X_t} = (\mathscr{L}_t^{\sigma^X})^* \mu_{X_t} + \mathrm{div}(b_t^X \mu_{X_t}), \tag{1.2}$$

where $\sigma_t^X(x) := \sigma_t(x, \mu_{X_t})$, $b_t^X(x) := b_t(x, \mu_{X_t})$, and $(\mathscr{L}_t^{\sigma^X})^*$ is the adjoint operator of the following second order partial differential operator

$$\mathscr{L}_t^{\sigma^X} f(x) := \frac{1}{2} \sum_{i,j,k=1}^d \sigma_t^{ik}(x, \mu_{X_t}) \sigma_t^{jk}(x, \mu_{X_t}) \partial_i \partial_j f(x). \tag{1.3}$$

Notice that if

$$\sigma_t^X(x) = \int_{\mathbb{R}^d} \sigma_t(x, y) \mu_{X_t}(\mathrm{d}y), \quad b_t^X(x) = \int_{\mathbb{R}^d} b_t(x, y) \mu_{X_t}(\mathrm{d}y),$$

then DDSDE (1.1) is also called mean-field SDE or McKean-Vlasov SDE in the literatures, which naturally appears in the studies of interacting particle systems and mean-field games (see [12, 18, 22, 3, 4] and references therein).

Up to now, there are numerous papers devoted to the study of this type of nonlinear FPEs and DDSDE (1.1). In [10], Funaki showed the existence of martingale solutions for (1.1) under broad conditions of Lyapunov's type and also the uniqueness under global Lipschitz assumptions. His method is based on a suitable time discretization. Thus, the well-posedness of FPE

This work is supported by NNSFC grant of China (No. 11731009) and the DFG through the CRC 1283 "Taming uncertainty and profiting from randomness and low regularity in analysis, stochastics and their applications".

(1.2) is also obtained. More recently, under some one-side Lipschitz assumptions, Wang [26] showed the strong well-posedness and some functional inequalities to DDSDE (1.1). In [7], Hammersley, Sitsa and Szpruch proved the existence of weak solutions to SDE (1.1) on a domain $D \subset \mathbb{R}^d$ with continuous and unbounded coefficients under Lyapunov-type conditions. Moreover, uniqueness is also obtained under some functional Lyapunov conditions. Notice that all the above results require the continuity of coefficients. In [5], Chiang obtained the existence of weak solutions for time-independent SDE (1.1) with drifts that have some discontinuities. When the diffusion matrix is uniformly non-degenerate and b, σ are only measurable and of at most linear growth, by using Krylov's estimate, Mishura and Veretennikov [19] showed the existence of weak and strong solutions. The uniqueness is also proved when σ does not depend on μ and is Lipschitz continuous in x and b is Lipschitz continuous with respect to μ with Lipschitz constant linearly depending on x. It should be noted that by Schauder's fixed point theorem and Girsanov's theorem, Li and Min [15] also obtained the existence and uniqueness of weak solutions when b is bounded measurable and σ is nondegenerate and Lipschitz continuous. On the other hand, by a purely analytic argument, Manita and Shaposhnikov [17] and Manita, Romanov and Shaposhnikov [16] showed the existence and uniqueness of solutions to the nonlinear FPE (1.2) under quite general assumptions. By a result of Trevisan [23] (see Theorem 5.1 below), one in fact can obtain the well-posedness of DDSDE (1.1) from [17] and [16]. In [1], a technique is developed to prove weak existence of solutions to (1.1) by first solving (1.2) which works also for coefficients whose dependence on μ_{X_t} is of "Nemytskii-type", i.e., are not continuous in μ_{X_t} in the weak topology.

In this work we are interested in extending Krylov-Röckner's result [13] to the singular distribution dependent case, that is not covered by all of the above results. More precisely, we want to show the well-posedness of the following DDSDE:

$$\mathrm{d}X_t = \left(\int_{\mathbb{R}^d} b_t(x-y)\mu_{X_t}(\mathrm{d}y)\right)\mathrm{d}t + \sqrt{2}\mathrm{d}W_t, \tag{1.4}$$

where $b \in L^q_{loc}(\mathbb{R}_+; L^p(\mathbb{R}^d))$ for some $p, q \in (2, \infty)$ with $\frac{d}{p} + \frac{2}{q} < 1$. Notice that the above equation is not covered by Huang and Wang's results [8] since $\mu \mapsto \int_{\mathbb{R}^d} b_t(x-y)\mu(\mathrm{d}y)$ is not weakly continuous. In fact, if we let

$$B_t(x, \mu) := \int_{\mathbb{R}^d} b_t(x-y)\mu(\mathrm{d}y), \ \mu \in \mathcal{P}(\mathbb{R}^d), \tag{1.5}$$

then we only have

$$\|B_t(\cdot, \mu) - B_t(\cdot, \mu')\|_p \leqslant \|b_t\|_p \|\mu - \mu'\|_{TV},$$

where $\|\cdot\|_{TV}$ is the total variation distance.

One of the main results of this paper is stated as follows:

Theorem 1.1. *Let $b : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ be a measurable vector field and $m > 2$. We assume that for some $p, q \in (2, \infty)$ with $\frac{d}{p} + \frac{2}{q} < 1$,*

$$b \in L^q_{loc}(\mathbb{R}_+; L^p(\mathbb{R}^d)) + L^\infty(\mathbb{R}_+ \times \mathbb{R}^d).$$

Then for any initial random variable X_0 with m-order finite moment, there is a unique strong solution to SDE (1.4). *Moreover, the following assertions hold:*

(i) The law μ_t of X_t uniquely solves the following nonlinear FPE in the distributional sense:

$$\partial_t \mu_t = \Delta\mu_t + \mathrm{div}\left(\mu_t(b_t(x-\cdot))\mu_t\right), \ \lim_{t\downarrow 0}\mu_t(\mathrm{d}y) = \mathbb{P}\circ X_0^{-1}(\mathrm{d}y) =: \mu_0(\mathrm{d}y) \tag{1.6}$$

in the class that $t \mapsto \mu_t$ *is continuous and*

$$\int_0^T \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |b_t(x-y)| \mu_t(\mathrm{d}y) \mu_t(\mathrm{d}x) \mathrm{d}t < \infty, \quad \forall T > 0,$$

(ii) $\mu_t(\mathrm{d}y) = \rho_t^X(y)\mathrm{d}y$ *and* $(t, y) \mapsto \rho_t^X(y)$ *is continuous on* $(0, \infty) \times \mathbb{R}^d$ *and satisfies the following two-sided estimate: for any* $T > 0$*, there are constants* $\gamma_0, c_0 \geqslant 1$ *such that for all* $t \in [0, T]$*,*

$$c_0^{-1} P_{t/\gamma_0} \mu_0(y) \leqslant \rho_t^X(y) \leqslant c_0 P_{\gamma_0 t} \mu_0(y),$$

where $P_t \mu_0(y) := (2\pi t)^{-d/2} \int_{\mathbb{R}^d} \mathrm{e}^{-|x-y|^2/(2t)} \mu_0(\mathrm{d}x)$ *is the Gaussian heat semigroup.*

(iii) If $\mathrm{div} b = 0$*, then* $\rho_t^X(y) \in C^1(\mathbb{R}^d)$ *and we have the following gradient estimate: for any* $T > 0$*, there are constants* $\gamma_1, c_1 \geqslant 1$ *such that for all* $t \in [0, T]$*,*

$$|\nabla \rho_t^X(y)| \leqslant c_1 t^{-1/2} P_{\gamma_1 t} \mu_0(y).$$

Example 1.2. *Let* $b(x) := x/|x|^\alpha$ *for some* $\alpha \in (1, 2)$*. Then it is easy to see that* $b \in (L^p + L^\infty)(\mathbb{R}^d)$ *for some* $p > d$*.*

This paper is organized as follows: In Section 2, we prepare some well-known results and tools for later use. In Section 3, we show the existence of weak and strong solutions to DDSDE (1.1) when the drift is the sum of a singular part and a dissipative part, and the diffusion coefficient is uniformly nondegenerate and bounded Hölder continuous. In Section 4, we prove the uniqueness of weak and strong solutions to (1.1) in two cases: the coefficients b and σ are Lipschitz continuous in the third variable with respect to the Wasserstein metric; drift b is Lipschitz continuous in the third variable with respect to the total variation distance and the diffusion coefficient does not depend on the distribution. In Section 5, we present some applications to nonlinear FPE (1.2) and prove Theorem 1.1.

Finally we collect some frequently used notations and conventions for later use.

- For $\theta > 0$, $\mathcal{P}_\theta(\mathbb{R}^d) := \left\{\mu \in \mathcal{P}(\mathbb{R}^d) : \int_{\mathbb{R}^d} |x|^\theta \mu(\mathrm{d}x) < \infty\right\}$.
- For $R > 0$, set $B_R := \{x \in \mathbb{R}^d : |x| < R\}$.
- For a function $f : \mathbb{R}^d \to \mathbb{R}$, $\mathcal{M}f(x) := \sup_{R>0} \fint_{B_R} |f|(x+y)\mathrm{d}y$.
- Let $\mathbf{S}_{\mathrm{toch}}$ be the set of all measurable stochastic processes that are stochastically continuous.
- Let $b : \mathbb{R}_+ \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$ be a measurable vector field. For $X \in \mathbf{S}_{\mathrm{toch}}$, define
$$b_t^X(x) := b_t(x, \mu_{X_t}), \quad \mu_{X_t} := \mathbb{P} \circ X_t^{-1}. \tag{1.7}$$
If b has a subscript, then we shall write the above function as $b_1^X(t, x)$.
- For a signed measure μ, we denote by $\|\mu\|_{TV} := \sup_{\|f\|_\infty \leqslant 1} |\mu(f)|$ the total variation of μ.
- We use $A \lesssim B$ (resp. $\asymp$) to denote $A \leqslant CB$ (resp. $CB^{-1} \leqslant A \leqslant CB$) for some unimportant constant $C \geqslant 1$, whose dependence on the parameter can be traced from the context.

2. Preliminaries

In this section we recall some well-known results. We first introduce the following spaces and notations for later use. For $p, q \in [1, \infty]$ and $T > S \geqslant 0$, let $\mathbb{L}_p^q(S, T)$ be the space of all Borel functions on $[S, T] \times \mathbb{R}^d$ with norm

$$\|f\|_{\mathbb{L}_p^q(S,T)} := \left(\int_S^T \left(\int_{\mathbb{R}^d} |f(t, x)|^p \mathrm{d}x \right)^{q/p} \mathrm{d}t \right)^{1/q} < \infty.$$

For $p = \infty$ or $q = \infty$, the above norm is understood as the usual L^∞-norm. We shall simply write

$$\mathbb{L}_p^q(T) := \mathbb{L}_p^q(0, T), \quad \mathbb{L}^p(T) := \mathbb{L}_p^p(T).$$

For $(p,\alpha) \in [1,\infty] \times (0,2] \setminus \{\infty\} \times \{1,2\}$, let $H^\alpha_p := (\mathbb{I}-\Delta)^{-\alpha/2}(L^p(\mathbb{R}^d))$ be the usual Bessel potential space with norm

$$\|f\|_{\alpha,p} := \|(\mathbb{I}-\Delta)^{\alpha/2} f\|_p,$$

where $\|\cdot\|_p$ is the usual L^p-norm in $\mathbb{R}^d$. For $p=\infty$ and $j=1,2$, we define H^j_∞ as the space of functions with finite norm

$$\|f\|_{j,\infty} := \|f\|_\infty + \|\nabla^j f\|_\infty < \infty.$$

In the following, given $T>0$, $\alpha \in (0,2]$ and $q,p \in [1,\infty]$, we write

$$\mathbb{H}^{\alpha,q}_p(T) := L^q([0,T]; H^\alpha_p).$$

Let $\sigma_t(x,\mu) = \sigma_t(x) : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ be a Borel measurable function, which satisfies:

($\mathbf{H}^\sigma_0$) There are constants $c_0 \geqslant 1$ and $\beta \in (0,1)$ such that for all $t>0$ and $x,y,\xi \in \mathbb{R}^d$,

$$c_0^{-1}|\xi|^2 \leqslant |\sigma^*_t(x)\xi|^2 \leqslant c_0|\xi|^2, \quad \|\sigma_t(x)-\sigma_t(y)\|_{HS} \leqslant c_0|x-y|^\beta,$$

where $\|\cdot\|_{HS}$ stands for the Hilbert-Schmidt norm of a matrix.

For $\lambda, T>0$, consider the following backward second order parabolic equation:

$$\partial_t u + (\mathscr{L}^\sigma_t - \lambda)u + b\cdot\nabla u = f, \quad u(T,x) = 0. \tag{2.1}$$

We have the following result, which is taken from [27].

Theorem 2.1. *Let $p \in (d/2 \vee 1, \infty), q \in (1,\infty)$ and $T>0$. Assume that* ($\mathbf{H}^\sigma_0$) *holds and for some $p_1 \in [p,\infty]$ and $q_1 \in [q,\infty]$ with $\frac{d}{p_1}+\frac{2}{q_1}<1$,*

$$\|b\|_{\mathbb{L}^{q_1}_{p_1}(T)} \leqslant \kappa_0 < \infty.$$

Let $\Theta := (\beta, c_0, d, p, q, p_1, q_1, \kappa_0)$ be the parameter set. Then there is a $\lambda_0 = \lambda_0(\Theta) \geqslant 1$ such that for all $\lambda \geqslant \lambda_0$ and $f \in \mathbb{L}^q_p(T)$, there exists a unique solution $u \in \mathbb{H}^{2,q}_p(T)$ to equation (2.1). Moreover, we have the following conclusions:

(i) There exists a constant $c_1 = c_1(\Theta,T) > 0$ such that for all $\lambda \geqslant \lambda_0$,

$$\|\nabla^2 u\|_{\mathbb{L}^q_p(T)} \leqslant c_1\|f\|_{\mathbb{L}^q_p(T)}. \tag{2.2}$$

(ii) For any $\vartheta \in [0,2)$ and $p' \in [p,\infty]$, $q' \in [q,\infty]$ satisfying

$$\frac{d}{p}+\frac{2}{q} < 2-\vartheta+\frac{d}{p'}+\frac{2}{q'}, \tag{2.3}$$

there exists a constant $c_2 = c_2(\Theta,T,\vartheta,q',p') > 0$ such that for all $\lambda \geqslant \lambda_0$,

$$\lambda^{\frac{1}{2}(2-\vartheta+\frac{d}{p'}+\frac{2}{q'}-\frac{d}{p}-\frac{2}{q})}\|u\|_{\mathbb{H}^{\vartheta,q'}_{p'}(T)} \leqslant c_2\|f\|_{\mathbb{L}^q_p(T)}. \tag{2.4}$$

(iii) Let (σ',b',f') be another group of coefficients satisfying the same assumptions as (σ,b,f). Let u' be the solution of (2.1) *corresponding to (σ',b',f'). Then for the same index in* (2.3)*, there exists a constant $c_3 = c_3(\Theta,T,\vartheta,q',p') > 0$ such that for all $\lambda \geqslant \lambda_0$,*

$$\begin{aligned}&\lambda^{\frac{1}{2}(2-\vartheta+\frac{d}{p'}+\frac{2}{q'}-\frac{d}{p}-\frac{2}{q})}\|u-u'\|_{\mathbb{H}^{\vartheta,q'}_{p'}(T)} \\ &\leqslant c_3\|f\|_{\mathbb{L}^q_p(T)}\Big(\|\sigma-\sigma'\|_{\mathbb{L}^\infty(T)} + \|b-b'\|_{\mathbb{L}^{q_1}_{p_1}(T)}\Big) + c_3\|f-f'\|_{\mathbb{L}^q_p(T)}.\end{aligned} \tag{2.5}$$

Proof. The existence and uniqueness of $u \in \mathbb{H}^{2,q}_p(T)$ as well as the first two conclusions are proved in [27, Theorem 4.3]. We only show (iii). Let $w = u'-u$. Then

$$\partial_t w + (\mathscr{L}^{\sigma'}_t - \lambda)w + b'\cdot\nabla w = (\mathscr{L}^\sigma_t - \mathscr{L}^{\sigma'}_t)u + (b-b')\cdot\nabla u + f' - f.$$

By (2.4) and Hölder's inequality we have

$$\lambda^{\frac{1}{2}(2-\vartheta+\frac{d}{p'}+\frac{2}{q'}-\frac{d}{p}-\frac{2}{q})}\|w\|_{\mathbb{H}^{\vartheta,q'}_{p'}(T)} \lesssim \|(\mathscr{L}^{\sigma}_t-\mathscr{L}^{\sigma'}_t)u+(b-b')\cdot\nabla u+f'-f\|_{\mathbb{L}^q_p(T)}$$
$$\lesssim \|\sigma'-\sigma\|_{\mathbb{L}^\infty(T)}\|\nabla^2 u\|_{\mathbb{L}^q_p(T)}+\|b'-b\|_{\mathbb{L}^{q_1}_{p_1}(T)}\cdot\|\nabla u\|_{\mathbb{L}^{q_2}_{p_2}(T)}+\|f'-f\|_{\mathbb{L}^q_p(T)},$$

where $\frac{1}{q_2}+\frac{1}{q_1}=1$ and $\frac{1}{p_2}+\frac{1}{p_1}=1$. The desired estimate now follows by (2.2) and (2.4). □

The following stochastic Gronwall's inequality for continuous martingales is proved by Scheutzow [20]. For general discontinuous martingales, it is due to [27].

Lemma 2.2 (Stochastic Gronwall's inequality)**.** *Let $\xi(t)$ and $\eta(t)$ be two nonnegative càdlàg $\mathscr{F}_t$-adapted processes, A_t a continuous nondecreasing $\mathscr{F}_t$-adapted process with $A_0=0$, M_t a local martingale with $M_0=0$. Suppose that*

$$\xi(t)\leqslant\eta(t)+\int_0^t\xi(s)\mathrm{d}A_s+M_t,\ \forall t\geqslant 0. \tag{2.6}$$

Then for any $0<q<p<1$ and $\tau>0$, we have

$$\left[\mathbb{E}(\xi(\tau)^*)^q\right]^{1/q}\leqslant\Big(\tfrac{p}{p-q}\Big)^{1/q}\Big(\mathbb{E}\mathrm{e}^{pA_\tau/(1-p)}\Big)^{(1-p)/p}\mathbb{E}(\eta(\tau)^*), \tag{2.7}$$

where $\xi(t)^*:=\sup_{s\in[0,t]}\xi(s)$.

We also recall the following result about maximal functions (for example, see [27, Lemma 8.1]). Let f be a locally integrable function on $\mathbb{R}^d$. The Hardy-Littlewood maximal function of f is defined by

$$\mathcal{M}f(x):=\sup_{r>0}\fint_{B_r}|f(x+y)|\mathrm{d}y,$$

where $\fint_{B_r}:=\frac{1}{|B_r|}\int_{B_r}$ and $|B_r|$ denotes the Lebesgue measure of the ball $B_r:=\{x:|x|<r\}$. We have

Lemma 2.3. *(i) Let f be a locally integrable function with $\nabla f\in L^1_{loc}(\mathbb{R}^d)$. Then there is a Lebesgue zero set E such that for all $x,y\notin E$,*

$$|f(x)-f(y)|\leqslant C_d|x-y|\Big(\mathcal{M}|\nabla f|(x)+\mathcal{M}|\nabla f|(y)\Big). \tag{2.8}$$

(ii) For $p\in(1,\infty]$, there is a constant $C_{d,p}>0$ such that for all $f\in L^p(\mathbb{R}^d)$,

$$\|\mathcal{M}f\|_p\leqslant C_{d,p}\|f\|_p. \tag{2.9}$$

Finally we recall the following Krylov estimate proved in [27, Lemma 5.5].

Lemma 2.4. *(Krylov's estimate) Let X be an Itô's process of the form*

$$X_t=X_0+\int_0^t\sigma_s(X_s)\mathrm{d}W_s+\int_0^t\xi(s)\mathrm{d}s, \tag{2.10}$$

where X_0 is an $\mathscr{F}_0$-measurable random variable, $\xi(t)$ is a measurable $\mathscr{F}_t$-adapted process. Let $T>0$. Under $(\mathbf{H}^\sigma_0)$, for any $p,q\in(1,\infty)$ with $\frac{d}{p}+\frac{2}{q}<1$ and each $\delta>0$, there is a constant $C_\delta=C_\delta(T,c_0,\beta,d,p,q)>0$ such that for any $0\leqslant t_0\leqslant t_1\leqslant T$ and $f\in\mathbb{L}^q_p(t_0,t_1)$,

$$\mathbb{E}\left(\int_{t_0}^{t_1}f(s,X_s)\mathrm{d}s\Big|\mathscr{F}_{t_0}\right)\leqslant\|f\|_{\mathbb{L}^q_p(t_0,t_1)}\left[C_\delta+\delta\mathbb{E}\left(\int_{t_0}^{t_1}|\xi(s)|\mathrm{d}s\Big|\mathscr{F}_{t_0}\right)\right]. \tag{2.11}$$

Moreover, if $\xi\equiv 0$, then we can relax p,q to satisfy $\frac{d}{p}+\frac{2}{q}<2$.

3. Existence of weak and strong solutions

In this section we show the weak existence and strong existence of DDSDEs with singular drifts of at most polynomial growth. First of all we recall the notions of martingale solutions and weak solutions for (1.1). Let $\mathbb{C}$ be the space of all continuous functions from $\mathbb{R}_+$ to $\mathbb{R}^d$, which is endowed with the usual Borel σ-field $\mathcal{B}(\mathbb{C})$. All the probability measures over $(\mathbb{C}, \mathcal{B}(\mathbb{C}))$ is denoted by $\mathcal{P}(\mathbb{C})$. Let w_t be the coordinate process over $\mathbb{C}$, that is,

$$w_t(\omega) = \omega_t, \ \ \omega \in \mathbb{C}.$$

For $t \geqslant 0$, let $\mathcal{B}_t(\mathbb{C}) = \sigma\{w_s : s \leqslant t\}$ be the natural filtration. For a probability measure $\mathbb{P} \in \mathcal{P}(\mathbb{C})$, the expectation with respect to $\mathbb{P}$ will be denoted by $\mathbb{E}$ if there is no confusion.

Definition 3.1. *(Martingale solutions) We call a probability measure $\mathbb{P} \in \mathcal{P}(\mathbb{C})$ a martingale solution of DDSDE* (1.1) *with initial distribution $\nu \in \mathcal{P}(\mathbb{R}^d)$ if $\mathbb{P} \circ w_0^{-1} = \nu$ and for any $f \in C^\infty$,*

$$\int_0^t |\mathscr{L}_s^{\sigma^{\mathbb{P}}} f|(w_s)\mathrm{d}s + \int_0^t |b_s^{\mathbb{P}} \cdot \nabla f|(w_s)\mathrm{d}s < \infty, \ \ \mathbb{P} - a.s, \forall t > 0,$$

where $\sigma_t^{\mathbb{P}}(x) := \sigma_t(x, \mu_t^{\mathbb{P}})$ and $b_t^{\mathbb{P}}(x) := b_t(x, \mu_t^{\mathbb{P}})$, $\mu_t^{\mathbb{P}} := \mathbb{P} \circ w_t^{-1}$, and

$$M_t^f := f(w_t) - f(w_0) - \int_0^t (\mathscr{L}_s^{\sigma^{\mathbb{P}}} f)(w_s)\mathrm{d}s - \int_0^t (b_s^{\mathbb{P}} \cdot \nabla f)(w_s)\mathrm{d}s, \tag{3.1}$$

is a continuous local $\mathcal{B}_t(\mathbb{C})$-martingale under $\mathbb{P}$. All the martingale solutions of DDSDE (1.1) *with coefficients σ, b and initial distribution ν is denoted by $\mathscr{M}_\nu^{\sigma,b}$.*

Definition 3.2 (Weak solutions)**.** *Let (X, W) be two $\mathbb{R}^d$-valued continuous adapted processes on some filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathbf{P})$. We call $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathbf{P}; X, W)$ a weak solution of DDSDE* (1.1) *with initial distribution $\nu \in \mathcal{P}(\mathbb{R}^d)$ if*

(i) $\mathbf{P} \circ X_0^{-1} = \nu$ and W is an $\mathscr{F}_t$-Brownian motion.

(ii) For all $t > 0$,

$$\int_0^t |b_s|(X_s, \mu_{X_s})\mathrm{d}s + \int_0^t \|\sigma_s\sigma_s^*\|_{HS}(X_s, \mu_{X_s})\mathrm{d}s < \infty, \ \ \mathbf{P} - a.s.$$

and

$$X_t = X_0 + \int_0^t b_s(X_s, \mu_{X_s})\mathrm{d}s + \int_0^t \sigma_s(X_s, \mu_{X_s})\mathrm{d}W_s, \ \ \mathbf{P} - a.s. \tag{3.2}$$

It is well known that weak solutions and martingale solutions are equivalent (cf. [21]), which means that for any $\mathbb{P} \in \mathscr{M}_\nu^{\sigma,b}$, there is a weak solution $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathbf{P}; X, W)$ such that

$$\mathbb{P} = \mathbf{P} \circ X^{-1}.$$

We now prove the following convergence lemma, which has independent interest.

Lemma 3.3. *Let $X^n, Y^n, X, Y \in \mathbf{S}_{\text{toch}}$ be such that for each $t \geqslant 0$, X_t^n converges to X_t almost surely and Y_t^n converges to Y_t in distribution. Let $p, q > 1$ and $m > 0$. Suppose that for any $T > 0$, there are constants $C_1, C_2 > 0$ such that for all $f \in \mathbb{L}_p^q(T)$,*

$$\sup_n \sup_{t\in[0,T]} \mathbb{E}|X_t^n|^m \leqslant C_1, \ \ \mathbb{E}\left(\int_0^T f(s, X_s^n)\mathrm{d}s\right) \leqslant C_2\|f\|_{\mathbb{L}_p^q(T)}. \tag{3.3}$$

Moreover, we assume that $b = b_1 + b_2$ satisfies the following assumptions:

(i) For each (t, x), $\mu \mapsto b(t, x, \mu)$ is continuous with respect to the weak convergence topology.

(ii) For some $\gamma > 1$ and $\vartheta \in [0, m/\gamma]$, there is a constant $\kappa_0 > 0$ such that for all $Z \in \mathbf{S}_{\rm toch}$,

$$\|b_1^Z\|_{\mathbb{L}^{\gamma q}_{\gamma p}(T)} + \sup_{(s,x)\in\mathbb{R}_+\times\mathbb{R}^d} \frac{|b_2^Z(s,x)|}{1+|x|^\vartheta} \leqslant \kappa_0, \tag{3.4}$$

where b_i^Z is defined by (1.7).

Then for each $T > 0$, we have

$$\lim_{n\to\infty} \mathbb{E}\left(\int_0^T |b^{Y_n}(s, X_s^n) - b^Y(s, X_s)|\mathrm{d}s\right) = 0. \tag{3.5}$$

Proof. To prove (3.5), it suffices to show the following:

$$\lim_{n\to\infty} \mathbb{E}\left(\int_0^T \left|b^{Y^n}(s, X_s^n) - b^Y(s, X_s^n)\right| \mathrm{d}s\right) = 0, \tag{3.6}$$

$$\lim_{n\to\infty} \mathbb{E}\left(\int_0^T \left|b^{Y}(s, X_s^n) - b^Y(s, X_s)\right| \mathrm{d}s\right) = 0. \tag{3.7}$$

We first look at (3.6). Since $\mu_{Y_s^n}$ weakly converges to μ_{Y_s} for each $s \geqslant 0$, by assumption (i), we have

$$b^{Y_n}(s,x) \overset{n\to\infty}{\to} b^Y(s,x), \quad \forall (s,x) \in \mathbb{R}_+ \times \mathbb{R}^d. \tag{3.8}$$

Moreover, for fixed $R > 1$, by the assumption (ii), one sees that

$$\sup_n \left\| |b^{Y_n}|^\gamma 1_{B_R}\right\|_{\mathbb{L}^q_p(T)} = \sup_n \left\|b^{Y_n} 1_{B_R}\right\|^\gamma_{\mathbb{L}^{\gamma q}_{\gamma p}(T)} \leqslant \left(\kappa_0 + \kappa_0(1+R^\vartheta)\right)^\gamma < \infty.$$

Hence, $(b^{Y_n} 1_{B_R})_{n\in\mathbb{N}}$ is uniformly integrable. Thus by (3.3) and (3.8), we have

$$\lim_{n\to\infty} \mathbb{E}\left(\int_0^T 1_{|X_s^n|\leqslant R} \left|b^{Y_n}(s, X_s^n) - b^Y(s, X_s^n)\right| \mathrm{d}s\right) \leqslant C_2 \lim_{n\to\infty} \left\|(b^{Y_n} - b^Y)1_{B_R}\right\|_{\mathbb{L}^q_p(T)} = 0. \tag{3.9}$$

On the other hand, by Hölder's inequality and (3.3), we have

$$\begin{aligned}
&\mathbb{E}\left(\int_0^T 1_{|X_s^n|>R} \left|b^{Y_n}(s, X_s^n) - b^Y(s, X_s^n)\right| \mathrm{d}s\right) \\
&\leqslant \int_0^T \mathbb{P}(|X_s^n| > R)^{\frac{\gamma-1}{\gamma}} \left(\mathbb{E}\left|b^{Y_n}(s, X_s^n) - b^Y(s, X_s^n)\right|^\gamma\right)^{\frac{1}{\gamma}} \mathrm{d}s \\
&\leqslant \sup_{s\in[0,T]} \mathbb{P}(|X_s^n| > R)^{\frac{\gamma-1}{\gamma}} \left(\int_0^T \mathbb{E}\left|b^{Y_n}(s, X_s^n) - b^Y(s, X_s^n)\right|^\gamma \mathrm{d}s\right)^{\frac{1}{\gamma}} \\
&\leqslant \sup_{s\in[0,T]} \left(\frac{\mathbb{E}|X_s^n|^m}{R^m}\right)^{\frac{\gamma-1}{\gamma}} \left(C_2^\gamma \left\|b_1^{Y_n} - b_1^Y\right\|^\gamma_{\mathbb{L}^{\gamma q}_{\gamma p}(T)} + T\kappa_0^\gamma \sup_{s\in[0,T]} \mathbb{E}\left(1 + |X_s^n|^\vartheta\right)^\gamma\right)^{\frac{1}{\gamma}},
\end{aligned}$$

which implies by the assumptions that

$$\lim_{R\to\infty} \sup_n \mathbb{E}\left(\int_0^T 1_{|X_s^n|>R} \left|b^{Y_n}(s, X_s^n) - b^Y(s, X_s^n)\right| \mathrm{d}s\right) = 0.$$

Combining this with (3.9), we obtain (3.6).

Next we show (3.7). Let $b^Y_\varepsilon(s,x) := b^Y(s,\cdot) * \varrho_\varepsilon(x)$ be the mollifying approximation of b^Y. Using the same argument as above, one can show

$$\lim_{\varepsilon\to 0} \sup_{n\in\mathbb{N}\cup\{\infty\}} \mathbb{E}\left(\int_0^T \left|b^Y_\varepsilon(s, X_s^n) - b^Y(s, X_s^n)\right| \mathrm{d}s\right) = 0, \tag{3.10}$$

where we have used the convention $X_s^\infty = X_s$. In fact, for any $R > 0$, we have

$$\lim_{\varepsilon\to 0}\sup_{n\in\mathbb{N}\cup\{\infty\}}\mathbb{E}\left(\int_0^T 1_{|X_s^n|\leqslant R}\left|b_\varepsilon^Y(s,X_s^n)-b^Y(s,X_s^n)\right|\mathrm{d}s\right)\leqslant C_2\lim_{\varepsilon\to 0}\|1_{B_R}(b_\varepsilon^Y-b^Y)\|_{\mathbb{L}_p^q(T)}=0,$$

where the last equality is due to (3.4) and the dominated convergence theorem, and

$$\sup_{\varepsilon}\sup_{n\in\mathbb{N}\cup\{\infty\}}\mathbb{E}\left(\int_0^T 1_{|X_s^n|>R}\left|b_\varepsilon^Y(s,X_s^n)-b^Y(s,X_s^n)\right|\mathrm{d}s\right)\leqslant\frac{C}{R^{m(\gamma-1)/\gamma}}\to 0 \text{ as } R\to\infty.$$

On the other hand, for fixed $\varepsilon > 0$, by (3.3) and (3.4), we have

$$\mathbb{E}\left(\int_0^T |b_\varepsilon^Y(s,X_s^n)|^\gamma\mathrm{d}s\right)\leqslant C_2\|b_1^Y\|^\gamma_{\mathbb{L}_{\gamma p}^{\gamma q}(T)}+\kappa_0\mathbb{E}\left(\int_0^T(1+|X_s^n|^\vartheta)^\gamma\mathrm{d}s\right)\leqslant C,$$

where C does not depend on n. Thus $(b_\varepsilon^Y(s,X_s^n))_{n\in\mathbb{N}}$ is uniformly integrable as random variables of (s,ω). Therefore, for fixed $\varepsilon > 0$,

$$\lim_{n\to\infty}\mathbb{E}\left(\int_0^T\left|b_\varepsilon^Y(s,X_s^n)-b_\varepsilon^Y(s,X_s)\right|\mathrm{d}s\right)=0,$$

which together with (3.10) yields (3.7). The proof is complete. □

In the above lemma, condition (i) sometimes may be not satisfied. For example, consider the following interesting example:

$$b(t,x,\mu)=\int_{\mathbb{R}^d}\tilde{b}(t,x,y)\mu(\mathrm{d}y), \tag{3.11}$$

where $\tilde{b}:\mathbb{R}_+\times\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}$ is a bounded measurable function. Obviously the weak continuity of $\mu\mapsto b(t,x,\mu)$ does not hold. However, in this case we still have the following limiting result.

Lemma 3.4. *Let $X^n, Y^n, X, Y \in \mathbf{S}_{\rm toch}$ be such that for each $t \geqslant 0$, the random variables (X_t^n, Y_t^n) almost surely converge to (X_t, Y_t). Let $p, q > 1$ and $m > 0$. Suppose that for any $T > 0$, there are constants $C_1, C_2 > 0$ such that for all $f \in \mathbb{L}_p^q(T)$,*

$$\sup_n\sup_{t\in[0,T]}\mathbb{E}\left(|X_t^n|^m+|Y_t^n|^m\right)\leqslant C_1,\quad\mathbb{E}\left(\int_0^T f(s,X_s^n)\mathrm{d}s\right)\leqslant C_2\|f\|_{\mathbb{L}_p^q(T)}. \tag{3.12}$$

Moreover, we assume b taken form (3.11) with $\tilde{b} = \tilde{b}_1 + \tilde{b}_2$, where $\tilde{b}_1, \tilde{b}_2$ satisfy that for some $\gamma > 1$ and $\vartheta \in [0, m/\gamma]$, there is a constant $\kappa_0 > 0$ such that for all $Z \in \mathbf{S}_{\rm toch}$,

$$\|b_1^Z\|_{\mathbb{L}_{\gamma p}^{\gamma q}(T)}+\sup_{(s,x,y)\in\mathbb{R}_+\times\mathbb{R}^d\times\mathbb{R}^d}\frac{|\tilde{b}_2(s,x,y)|}{1+|x|^\vartheta+|y|^\vartheta}\leqslant\kappa_0,$$

where $b_1^Z(s,x) := \mathbb{E}|\tilde{b}_1|(s,x,Z_s)$. Then for each $T > 0$, we have

$$\lim_{n\to\infty}\mathbb{E}\left(\int_0^T|b(s,X_s^n,\mu_{Y_s^n})-b(s,X_s,\mu_{Y_s})|\mathrm{d}s\right)=0. \tag{3.13}$$

Proof. Since b only depends on the distribution of Y^n, without loss of generality we may assume that $(X^n)_{n\in\mathbb{N}}$ and $(Y^n)_{n\in\mathbb{N}}$ are independent. Thus, to show (3.13), it suffices to show

$$\lim_{n\to\infty}\mathbb{E}\int_0^T|\tilde{b}(s,X_s^n,Y_s^n)-\tilde{b}(s,X_s,Y_s)|\mathrm{d}s=0. \tag{3.14}$$

By the independence of X^n and Y^n and (3.12), we have

$$\mathbb{E}\left(\int_0^T|\tilde{b}_1(s,X_s^n,Y_s^n)|^\gamma\mathrm{d}s\right)=\mathbb{E}\left(\int_0^T|\tilde{b}_1^{Y^n}(s,X_s^n)|^\gamma\mathrm{d}s\right)\leqslant C_2\|\tilde{b}_1^{Y^n}\|_{\mathbb{L}_{\gamma p}^{\gamma q}(T)}\leqslant C_2\kappa_0,$$

and

$$\mathbb{E}\left(\int_0^T |\tilde{b}_2(s, X_s^n, Y_s^n)|^\gamma \mathrm{d}s\right) \leqslant \kappa_0 \mathbb{E}\left(\int_0^T \left(1 + |X_s^n|^\vartheta + |Y_s^n|^\vartheta\right)^\gamma \mathrm{d}s\right) \leqslant \kappa_0 T(1 + C_1).$$

So, $(\tilde{b}(s, X_s^n, Y_s^n))_{n\in\mathbb{N}}$ are uniformly integrable as random variables of (s, ω). Thus (3.14) holds, and so does (3.13). □

Now we make the following assumptions about σ and b:

$(\mathbf{H}_1^\sigma)$ There are constants $c_0 \geqslant 1$ and $\beta \in (0, 1)$ such that for all $t > 0$, $x, y, \xi \in \mathbb{R}^d$ and $\mu \in \mathcal{P}(\mathbb{R}^d)$,

$$c_0^{-1}|\xi|^2 \leqslant |\sigma_t^*(x, \mu)\xi|^2 \leqslant c_0|\xi|^2, \quad \|\sigma_t(x, \mu) - \sigma_t(y, \mu)\|_{HS} \leqslant c_0|x - y|^\beta.$$

$(\mathbf{H}_1^b)$ $b = b_1 + b_2$, where b_1 is the singular part satisfying that for some $\frac{d}{p} + \frac{2}{q} < 1$,

$$\sup_{Z\in\mathbf{S}_{\rm toch}} \|b_1^Z\|_{\mathbb{L}_p^q(T)} \leqslant \kappa_0 < \infty,$$

and b_2 is the dissipative part which satisfies for some $\kappa_1, \kappa_2, \kappa_3 > 0$, $\vartheta \geqslant 0$ and any $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$ and $\mu \in \mathcal{P}(\mathbb{R}^d)$,

$$\langle x, b_2(t, x, \mu)\rangle \leqslant -\kappa_1|x|^{1+\vartheta} + \kappa_2(1 + |x|^2) \quad \text{and} \quad |b_2(t, x, \mu)| \leqslant \kappa_3(1 + |x|^\vartheta). \tag{3.15}$$

Moreover, $\mu \mapsto b_t(x, \mu)$ is weakly continuous for each t, x.

$(\mathbf{H}_2^b)$ b has the form (3.11) with $\tilde{b} = \tilde{b}_1 + \tilde{b}_2$, where $\tilde{b}_1$ satisfies that for some $\frac{d}{p} + \frac{2}{q} < 1$,

$$\sup_{Z\in\mathbf{S}_{\rm toch}} \|b_1^Z\|_{\mathbb{L}_p^q(T)} \leqslant \kappa_0 < \infty, \quad b_1^Z(t, x) := \mathbb{E}|\tilde{b}_1|(t, x, Z_t),$$

and $\tilde{b}_2$ is the dissipative part which satisfies for some $\kappa_1, \kappa_2, \kappa_3 > 0$, $\vartheta \geqslant 0$, and any $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$ and $\mu \in \mathcal{P}(\mathbb{R}^d)$,

$$\langle x, \tilde{b}_2(t, x, y)\rangle \leqslant -\kappa_1|x|^{1+\vartheta} + \kappa_2(1 + |x|^2) \quad \text{and} \quad |\tilde{b}_2(t, x, y)| \leqslant \kappa_3(1 + |x|^\vartheta). \tag{3.16}$$

Notice that if $\vartheta = 1$, then (3.15) equivalently says that b_2 is linear growth in x uniformly in t, μ.

To show the existence of weak solutions, we first establish the following apriori estimates.

Lemma 3.5. *Let $m > 2$ and $Z \in \mathbf{S}_{\rm toch}$. Under $(\mathbf{H}_1^\sigma)$ and $(\mathbf{H}_1^b)$ or $(\mathbf{H}_2^b)$, for any initial distribution $\nu \in \mathcal{P}_{m(\vartheta\vee 1)}(\mathbb{R}^d)$, let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathbf{P}; X, W)$ be a solution of the following SDE:*

$$\mathrm{d}X_t = b_t^Z(X_t)\mathrm{d}t + \sigma_t^Z(X_t)\mathrm{d}W_t, \quad \mathbf{P} \circ X_0^{-1} = \nu.$$

Let $\Theta = (d, p, q, c_0, \beta, \kappa_0, \kappa_1, \kappa_2, \kappa_3, \vartheta)$ be the parameter set in the assumptions. We have

(i) For any $T > 0$, there is a constant $C_1 = C_1(\Theta, T, m) > 0$ such that

$$\sup_{t\in[0,T]} \mathbf{E}|X_t|^m \leqslant C_1\left(\mathbf{E}|X_0|^m + 1\right), \quad \sup_{t\neq t'\in[0,T]} \frac{\mathbf{E}|X_t - X_{t'}|^m}{|t - t'|^{m/2}} \leqslant C_1\left(\mathbf{E}|X_0|^{m\vartheta} + 1\right).$$

(ii) Let $p', q' \in (2, \infty)$ with $\frac{d}{p'} + \frac{2}{q'} < 1$. For any $T > 0$, there is a constant $C_2 > 0$ depending only on $p', q', \Theta, T, \nu, m$ such that for all $f \in \mathbb{L}_{p'}^{q'}(T)$,

$$\mathbf{E}\left(\int_0^T f(s, X_s)\mathrm{d}s\right) \leqslant C_2\|f\|_{\mathbb{L}_{p'}^{q'}(T)}. \tag{3.17}$$

(iii) If $\vartheta = 0$ in (3.15) or (3.16), then for any $p', q' \in (2, \infty)$ with $\frac{d}{p'} + \frac{2}{q'} < 2$ and $T > 0$, there is a constant $C_3 = C_3(p', q', \Theta, T) > 0$ such that for all $S \in [0, T]$ and $f \in \mathbb{L}_{p'}^{q'}(S, T)$,

$$\mathbf{E}\left(\int_S^T f(s, X_s)\mathrm{d}s \Big| \mathscr{F}_S\right) \leqslant C_3\|f\|_{\mathbb{L}_{p'}^{q'}(S,T)}. \tag{3.18}$$

Proof. We use the Zvonkin transformation to kill the singular part. For $\lambda, T > 0$, consider the following backward PDE:

$$\partial_t u + (\mathscr{L}_t^{\sigma^Z} - \lambda)u + b_1^Z \cdot \nabla u + b_1^Z = 0, \ u(T, x) = 0.$$

By Theorem 2.1, for λ_0 large enough and all $\lambda \geqslant \lambda_0$, there is a unique solution $u \in \mathbb{H}_p^{2,q}(T)$ solving the above PDE, and there is a constant $c_1 > 0$ such that for all $\lambda \geqslant \lambda_0$,

$$\lambda^{\frac{1}{2}(1-\frac{d}{p}-\frac{2}{q})}\|u\|_{\mathbb{H}_\infty^{1,\infty}(T)} + \|\nabla^2 u\|_{\mathbb{L}_p^q(T)} \leqslant c_1\|b_1\|_{\mathbb{L}_p^q(T)}. \tag{3.19}$$

In particular, if we choose λ large enough, then

$$\|u\|_{\mathbb{L}^\infty(T)} + \|\nabla u\|_{\mathbb{L}^\infty(T)} \leqslant 1/2.$$

Now if we define

$$\Phi(t, x) := x + u(t, x),$$

then it is easy to see that

$$|x - y|/2 \leqslant |\Phi(t, x) - \Phi(t, y)| \leqslant 2|x - y| \tag{3.20}$$

and

$$\partial_t \Phi + \mathscr{L}_t^{\sigma^Z}\Phi + b^Z \cdot \nabla\Phi = \lambda u + b_2^Z \cdot \nabla\Phi. \tag{3.21}$$

By Itô's formula, we have

$$Y_t := \Phi(t, X_t) = \Phi(0, X_0) + \int_0^t \tilde{b}_2(s, X_s)\mathrm{d}s + \int_0^t \tilde{\sigma}(s, X_s)\mathrm{d}W_s, \tag{3.22}$$

where

$$\tilde{\sigma} := \sigma^Z \cdot \nabla\Phi, \ \tilde{b}_2 := \lambda u + b_2^Z \cdot \nabla\Phi.$$

By (3.20), one sees that

$$1 + |Y_t| \asymp 1 + |X_t|. \tag{3.23}$$

Moreover, it is easy to see that

$$\|\tilde{\sigma}\|_{HS} \leqslant 2\|\sigma^Z\|_{HS}, \tag{3.24}$$

and for λ large enough and for some $\tilde{\kappa}_i > 0, i = 1, 2,$

$$\langle\Phi(t, x), \tilde{b}_2(t, x)\rangle \leqslant -\tilde{\kappa}_1|x|^{1+\vartheta} + \tilde{\kappa}_2(1 + |x|^2) \quad \text{and} \quad |\tilde{b}_2(t, x)| \leqslant \tilde{\kappa}_3(1 + |x|^\vartheta). \tag{3.25}$$

In fact, by the definition we have

$$\begin{aligned}\langle\Phi(t, x), \tilde{b}_2(t, x)\rangle &= \langle x + u(t, x), b_2^Z(t, x) \cdot (\mathbb{I} + \nabla u(t, x))\rangle \\ &\leqslant \langle x, b_2^Z(t, x)\rangle + (|x| \cdot \|\nabla u\|_\infty + \|u\|_\infty\|\nabla u\|_\infty)|b_2^Z(t, x)| \\ &\leqslant -\kappa_1|x|^{1+\vartheta} + \kappa_2(1 + |x|^2) + \kappa_3\|\nabla u\|_\infty(|x| + 1)(1 + |x|^\vartheta),\end{aligned}$$

which in turn gives the first inequality in (3.25) by (3.19) with λ large enough so that

$$\kappa_3\|\nabla u\|_\infty \leqslant \kappa_1/2.$$

(i) By equation (3.22) and Itô's formula, we have

$$\begin{aligned}|Y_t|^m = |Y_0|^m &+ m\int_0^t |Y_s|^{m-2}\langle Y_s, \tilde{b}_2(s, X_s)\rangle\mathrm{d}s + m\int_0^t |Y_s|^{m-2}\langle\tilde{\sigma}(s, X_s)^*Y_s, \mathrm{d}W_s\rangle \\ &+ m\Big(\tfrac{m}{2} - 1\Big)\int_0^t |Y_s|^{m-4}|\tilde{\sigma}(s, X_s)^*Y_s|^2\mathrm{d}s + \frac{m}{2}\int_0^t |Y_s|^{m-2}\|\tilde{\sigma}(s, X_s)\|_{HS}^2\mathrm{d}s.\end{aligned}$$

If necessary, by a stopping time technique, by (3.23), (3.24) and (3.25) we obtain

$$\mathbf{E}|Y_t|^m \leqslant \mathbf{E}|Y_0|^m + c\int_0^t \mathbf{E}|Y_s|^m \mathrm{d}s + ct,$$

which yields by Gronwall's inequality that

$$\mathbf{E}|Y_t|^m \leqslant C_t(\mathbf{E}|Y_0|^m + 1) \overset{(3.23)}{\leqslant} C_t(\mathbf{E}|X_0|^m + 1). \tag{3.26}$$

On the other hand, by (3.22) and BDG's inequality, for all $0 \leqslant t' < t \leqslant T$, we have

$$\begin{aligned}\mathbf{E}|Y_t - Y_{t'}|^m &\lesssim \mathbf{E}\left|\int_{t'}^t \tilde{b}_2(s, X_s)\mathrm{d}s\right|^m + \mathbf{E}\left|\int_{t'}^t \tilde{\sigma}(s, X_s)\mathrm{d}W_s\right|^m \\ &\lesssim \mathbf{E}\left|\int_{t'}^t (1 + |Y_s|^\vartheta)\mathrm{d}s\right|^m + |t-t'|^{m/2} \\ &\lesssim |t-t'|^m \sup_{s\in[0,T]} \mathbf{E}|Y_s|^{m\vartheta} + |t-t'|^{m/2},\end{aligned}$$

which together with (3.20), (3.26) yields (i).

(ii) By Lemma 2.4, for any $T, \delta > 0$, there exists a constant $C_\delta > 0$ such that for any $f \in \mathbb{L}^{q'}_{p'}(T)$,

$$\mathbf{E}\left(\int_0^T f(s, Y_s)\mathrm{d}s\right) \leqslant \left(C_\delta + \delta\mathbf{E}\left(\int_0^T |b_1 + b_2|(s, Y_s)\mathrm{d}s\right)\right)\|f\|_{\mathbb{L}^{q'}_{p'}(T)}. \tag{3.27}$$

Since $b_1 \in \mathbb{L}^q_p(T)$ with $\frac{d}{p} + \frac{2}{q} < 1$, we can take $f = |b_1|$ to get

$$\mathbf{E}\left(\int_0^T |b_1|(s, Y_s)\mathrm{d}s\right) \leqslant C_\delta + \delta\|b_1\|_{\mathbb{L}^q_p(T)}\left[\mathbf{E}\left(\int_0^T |b_1|(s, Y_s)\mathrm{d}s\right) + \mathbf{E}\left(\int_0^T |b_2|(s, Y_s)\mathrm{d}s\right)\right].$$

Now, choosing δ small enough in the above inequality such that

$$\delta\|b_1\|_{\mathbb{L}^q_p(T)} \leqslant \delta\kappa_0 \leqslant 1/2,$$

and by (3.25), we obtain

$$\mathbf{E}\left(\int_0^T |b_1|(s, Y_s)\mathrm{d}s\right) \leqslant 2C_\delta + \mathbf{E}\left(\int_0^T |b_2|(s, Y_s)\mathrm{d}s\right). \tag{3.28}$$

Substituting this into (3.27) and by (3.25) and (3.26), we obtain

$$\mathbf{E}\left(\int_0^T f(s, Y_s)\mathrm{d}s\right) \leqslant C\|f\|_{\mathbb{L}^{q'}_{p'}(T)},$$

which yields by the change of variable and (3.20) that

$$\mathbf{E}\left(\int_0^T f(s, X_s)\mathrm{d}s\right) \leqslant C\|f \circ \Phi^{-1}\|_{\mathbb{L}^{q'}_{p'}(T)} \leqslant C\|f\|_{\mathbb{L}^{q'}_{p'}(T)}.$$

(iii) If $\vartheta = 0$, then $\tilde{b}_2$ is bounded by $\lambda + 2\kappa_3$. It was proved in [28, Theorem 2.1] (see also [27, Theorem 5.7]) that for $p', q' \in (2, \infty)$ with $\frac{d}{p'} + \frac{2}{q'} < 2$, there is a constant $C_3 = C_3(p', q', \Theta, T) > 0$ such that for all $S \in [0, T]$ and $f \in \mathbb{L}^{q'}_{p'}(S, T)$,

$$\mathbf{E}\left(\int_S^T f(s, Y_s)\mathrm{d}s\middle|\mathscr{F}_S\right) \leqslant C_3\|f\|_{\mathbb{L}^{q'}_{p'}(S,T)}.$$

By the change of variable and (3.20) again, we obtain (iii). □

Remark 3.6. *(iii) above can be derived from Lemma 2.4 with $\xi = 0$ by Girsanov's theorem. An important conclusion of (iii) is the following Khasminskii's type estimate (see* [27, Lemma 3.5]*): For any $\lambda, T > 0$ and $f \in \mathbb{L}^{q'}_{p'}(T)$,*

$$\mathbf{E}\exp\left(\lambda \int_0^T |f(s, X_s)|\mathrm{d}s\right) \leqslant C_4, \tag{3.29}$$

where C_4 only depends on $\lambda, \Theta, q', p', T$ and $\|f\|_{\mathbb{L}^{q'}_{p'}(T)}$.

Now we can show the following weak existence result.

Theorem 3.7. *Under* $(\mathbf{H}^\sigma_1)$ *and* $(\mathbf{H}^b_1)$ *or* $(\mathbf{H}^b_2)$*, for any initial distribution $\nu \in \mathcal{P}_{m(\vartheta\vee 1)}(\mathbb{R}^d)$, where $m > 2$, there exists a weak solution $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathbf{P}; X, W)$ to DDSDE* (1.1) *with $\mathbf{P}\circ X_0^{-1} = \nu$.*

Proof. Let $X^0_t \equiv X_0$. For $n \in \mathbb{N}$, consider the following approximating SDE:

$$X^n_t = X_0 + \int_0^t b^n(s, X^n_s)\mathrm{d}s + \int_0^t \sigma^n(s, X^n_s)\mathrm{d}W_s, \tag{3.30}$$

where

$$b^n(s, x) := b(s, x, \mu_{X^{n-1}_s}),\ \sigma^n(s, x) := \sigma(s, x, \mu_{X^{n-1}_s}).$$

By the assumptions, one sees that

$$c_0^{-1}|\xi|^2 \leqslant |\sigma^n(t, x)\xi|^2 \leqslant c_0|\xi|^2,\ \forall t \geqslant 0, x, \xi \in \mathbb{R}^d,$$

and $b^n = b^n_1 + b^n_2$ with

$$\sup_n \|b^n_1\|_{\mathbb{L}^q_p(T)} \leqslant \kappa_0 < \infty, \tag{3.31}$$

and b^n_2 satisfying (3.15) or (3.16) with the same constants $\kappa_1, \kappa_2, \kappa_3$. By induction and Lemma 3.5, we can show that there is a weak solution $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathbf{P}; X^n, W)$ to (3.30) with $\mathbf{P}\circ (X^n_0)^{-1} = \nu$ so that the following uniform estimates hold (see [29]):

(i) For any $T > 0$, there is a constant $C_T > 0$ such that for all $n \in \mathbb{N}$,

$$\sup_{t\in[0,T]} \mathbf{E}|X^n_t|^m \leqslant C_T\left(\mathbf{E}|X_0|^m + 1\right), \quad \sup_{t\neq t'\in[0,T]} \frac{\mathbf{E}|X^n_t - X^n_{t'}|^m}{|t - t'|^{m/2}} \leqslant C_T\left(\mathbf{E}|X_0|^{m\vartheta} + 1\right).$$

(ii) Let $p', q' \in (2, \infty)$ with $\frac{d}{p'} + \frac{2}{q'} < 1$. For any $T > 0$, there is a constant $C > 0$ such that for all $f \in \mathbb{L}^{q'}_{p'}(T)$,

$$\sup_n \mathbf{E}\left(\int_0^T f(s, X^n_s)\mathrm{d}s\right) \leqslant C\|f\|_{\mathbb{L}^{q'}_{p'}(T)}.$$

Now by (i), the laws $\mathbb{Q}_n$ of (X^n, W) in $\mathbb{C}\times\mathbb{C}$ are tight. Let $\mathbb{Q}$ be any accumulation point of $\mathbb{Q}_n$. Without loss of generality, we assume that $\mathbb{Q}_n$ weakly converges to some probability measure $\mathbb{Q}$. By Skorokhod's representation theorem, there are a probability space $(\tilde\Omega, \tilde{\mathscr{F}}, \tilde{\mathbf{P}})$ and random variables $(\tilde X^n, \tilde W^n)$ and $(\tilde X, \tilde W)$ defined on it such that

$$(\tilde X^n, \tilde W^n) \to (\tilde X, \tilde W),\ \ \tilde{\mathbf{P}} - a.s. \tag{3.32}$$

and

$$\tilde{\mathbf{P}}\circ(\tilde X^n, \tilde W^n)^{-1} = \mathbb{Q}^n, \quad \tilde{\mathbf{P}}\circ(\tilde X, \tilde W)^{-1} = \mathbb{Q}. \tag{3.33}$$

Define $\tilde{\mathscr{F}}^n_t := \sigma(\tilde W^n_s; s \leqslant t)$. Notice that

$$\mathbf{P}^n(W^n_t - W^n_s \in \cdot|\mathscr{F}^n_s) = \mathbf{P}^n(W^n_t - W^n_s \in \cdot) \Rightarrow \tilde{\mathbf{P}}(\tilde W^n_t - \tilde W^n_s \in \cdot|\tilde{\mathscr{F}}^n_s) = \tilde{\mathbf{P}}(\tilde W^n_t - \tilde W^n_s \in \cdot).$$

In other words, $\tilde W^n$ is an $\tilde{\mathscr{F}}^n_t$-Brownian motion. Thus, by (3.33) we have

$$\tilde X^n_t = \tilde X^n_0 + \int_0^t b_s(\tilde X^n_s, \mu_{\tilde X^{n-1}_s})\mathrm{d}s + \int_0^t \sigma_s(\tilde X^n_s, \mu_{\tilde X^{n-1}_s})\mathrm{d}\tilde W^n_s.$$

By (ii), (3.32), Lemmas 3.3, 3.4 and [11, Theorem 6.22, p383], one can take limits $n \to \infty$ to obtain

$$\tilde X_t = \tilde X_0 + \int_0^t b_s(\tilde X_s, \mu_{\tilde X_s})\mathrm{d}s + \int_0^t \sigma_s(\tilde X_s, \mu_{\tilde X_s})\mathrm{d}\tilde W_s.$$

The proof is complete. □

To obtain the existence of strong solutions, we need a stronger assumption about σ:

($\mathbf{H}^\sigma_2$) In addition to **($\mathbf{H}^\sigma_1$)**, we also assume that for some $p_1, q_1 \in (2, \infty]$ with $\frac{d}{p_1} + \frac{2}{q_1} < 1$,

$$\sup_{Z\in\mathbf{S}_{\rm toch}} \|\nabla \sigma^Z_t\|_{\mathbb{L}^{q_1}_{p_1}(T)} < \infty.$$

Corollary 3.8. *Under* **($\mathbf{H}^\sigma_2$)** *and* **($\mathbf{H}^b_1$)** *or* **($\mathbf{H}^b_2$)***, then for any initial random variable X_0 with finite $m(\vartheta \vee 1)$-order moment, where $m > 2$, there exists a strong solution to DDSDE* (1.1).

Proof. Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathbf{P}; X, W)$ be a weak solution of DDSDE (1.1). Define

$$b^X_t(x) := b_t(x, \mu_{X_t}),\ \ \sigma^X_t(x) := \sigma_t(x, \mu_{X_t}),\ \ \mu_{X_t} := \mathbf{P}\circ X_t^{-1}.$$

Consider the following SDE:

$$\mathrm{d}Z_t = b^X_t(Z_t)\mathrm{d}t + \sigma^X_t(Z_t)\mathrm{d}W_t.$$

Under the assumption of the theorem, it has been shown in [27] that there is a unique strong solution to this equation. Since X also satisfies the above equation, we obtain that $X = Z$ is a strong solution. □

Remark 3.9. *Although we have shown the existence of strong or weak solutions, the uniqueness of strong solutions or weak solutions is a more difficult problem.*

4. Uniqueness of strong and weak solutions

In this section we study the uniqueness of strong and weak solutions. We introduce the following assumptions about the third variable μ:

($\mathbf{A}_\theta$) Let $p, q \in (2, \infty)$ with $\frac{d}{p} + \frac{2}{q} < 1$ and $\theta > 2$. It holds that $\sup_{Z\in\mathbf{S}_{\rm toch}} \|b^Z\|_{\mathbb{L}^q_p(T)} < \infty$ and for any $R > 0$, there are $\ell^R \in L^q_{loc}(\mathbb{R}_+)$ and a constant $c_0 \geqslant 1$ such that for any two random variables X, Y with $\|X\|_\theta \vee \|Y\|_\theta < R$,

$$\|b_t(\cdot, \mu_X) - b_t(\cdot, \mu_Y)\|_p \leqslant \ell^R_t \|X - Y\|_\theta,\ \ \|\sigma_t(\cdot, \mu_X) - \sigma_t(\cdot, \mu_Y)\|_\infty \leqslant c_0 \|X - Y\|_\theta. \tag{4.1}$$

Notice that (4.1) is equivalent to that for all $\mu, \mu' \in \mathcal{P}_\theta(\mathbb{R}^d)$ with $\mu(|\cdot|^\theta) \vee \mu'(|\cdot|^\theta) < R^\theta$,

$$\|b_t(\cdot, \mu) - b_t(\cdot, \mu')\|_p \leqslant \ell^R_t \mathcal{W}_\theta(\mu, \mu'),\ \ \|\sigma_t(\cdot, \mu) - \sigma_t(\cdot, \mu')\|_\infty \leqslant c_0 \mathcal{W}_\theta(\mu, \mu'),$$

where $\mathcal{W}_\theta$ is the usual Wasserstein metric of θ-order. For convenience, we would like to use (4.1) rather than introducing the Wasserstein metric.

Remark 4.1. *We note that in* [8], (4.1) *is assumed to hold for $p = \infty$ and $R = \infty$.*

We first show the following strong uniqueness result.

Theorem 4.2. *Let $\theta > 2$. Under* **($\mathbf{H}^\sigma_2$)** *and* **($\mathbf{A}_\theta$)***, for any initial random variable X_0 with finite θ-order moment, there is a unique strong solution to DDSDE* (1.1) *in the class that for any $T > 0$*

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

Proof. Below we fix $p, q \in (2,\infty)$ satisfying $\frac{d}{p}+\frac{2}{q} < 1$ and always assume $(\mathbf{H}_2^\sigma)$ and $(\mathbf{A})$. Without loss of generality, we consider the time interval $[0,1]$ and assume that for some $\gamma > 1$,

$$\|\ell^R\|_{L^{\gamma q}(0,1)} + \sup_{Z\in\mathbf{S}_{\rm toch}} \|b^Z\|_{\mathbb{L}_p^{\gamma q}(1)} < \infty. \tag{4.2}$$

Otherwise, we may choose $q' < q$ so that $\frac{2}{q'} + \frac{d}{p} < 1$ holds and replace q with q'. The existence of strong solutions has been shown in Corollary 3.8. We only need to prove the pathwise uniqueness. Let X, Y be two strong solutions defined on the same probability space with

$$\sup_{t\in[0,1]} \mathbb{E}|X_t|^\theta \vee \sup_{t\in[0,1]} \mathbb{E}|Y_t|^\theta = R < \infty. \tag{4.3}$$

We divide the proof into three steps.

(i) Let $T \in (0,1)$ and $\lambda > 0$. We consider the following backward PDE:

$$\partial_t u^X + (\mathscr{L}_t^{\sigma^X} - \lambda)u + b^X \cdot \nabla u^X + b^X = 0,\ u^X(T,x) = 0. \tag{4.4}$$

By Theorem 2.1, for λ_0 large enough and all $\lambda \geqslant \lambda_0$, there is a unique solution $u^X \in \mathbb{H}_p^{2;q}(T)$ solving the above PDE, and there is a constant $c_1 > 0$ such that for all $\lambda \geqslant \lambda_0$ and $T \in (0,1)$,

$$\lambda^{\frac{1}{2}(1-\frac{d}{p}-\frac{2}{q})}\|u^X\|_{\mathbb{H}_\infty^{1,\infty}(T)} + \|\nabla^2 u^X\|_{\mathbb{L}_p^q(T)} \leqslant c_1\|b^X\|_{\mathbb{L}_p^q(T)}. \tag{4.5}$$

In particular, we can find $\lambda \geqslant \lambda_0$ large enough so that for all $T \in (0,1)$,

$$\|u^X\|_{\mathbb{L}^\infty(T)} + \|\nabla u^X\|_{\mathbb{L}^\infty(T)} \leqslant 1/2. \tag{4.6}$$

Below we shall fix such a λ and define

$$\Phi^X(t,x) := x + u^X(t,x).$$

It is easy to see that

$$\partial_t\Phi^X + \mathscr{L}_t^X\Phi^X + b^X\cdot\nabla\Phi^X = \lambda u^X.$$

(ii) By Itô's formula, we have

$$\tilde{X}_t := \Phi^X(t,X_t) = \Phi^X(0,X_0) + \lambda\int_0^t u^X(s,X_s)\mathrm{d}s + \int_0^t \tilde{\sigma}^X(s,X_s)\mathrm{d}W_s, \tag{4.7}$$

where

$$\tilde{\sigma}^X := \sigma^X\cdot\nabla\Phi^X.$$

For simplicity we write

$$\xi_t := X_t - Y_t,\ \ \tilde{\xi}_t := \tilde{X}_t - \tilde{Y}_t.$$

Noticing that by (4.6),

$$|x-y| \leqslant 2|\Phi^X(t,x) - \Phi^X(t,y)| \leqslant 2|\Phi^X(t,x) - \Phi^Y(t,y)| + 2\|u^X - u^Y\|_{\mathbb{L}^\infty(T)}$$

and

$$|\Phi^X(t,x) - \Phi^Y(t,y)| \leqslant 2|x-y| + \|u^X - u^Y\|_{\mathbb{L}^\infty(T)},$$

we have

$$|\xi_t| \leqslant 2|\tilde{\xi}_t| + 2\|u^X - u^Y\|_{\mathbb{L}^\infty(T)},\ \ |\tilde{\xi}_t| \leqslant 2|\xi_t| + \|u^X - u^Y\|_{\mathbb{L}^\infty(T)}. \tag{4.8}$$

By (4.7) and Itô's formula again, we have for any $m \geqslant 1$,

$$\begin{aligned}|\tilde{\xi}_t|^m = |\tilde{\xi}_0|^m &+ m\lambda\int_0^t |\tilde{\xi}_s|^{m-2}\langle\tilde{\xi}_s, u^X(s,X_s) - u^Y(s,Y_s)\rangle\mathrm{d}s\\ &+ m\int_0^t |\tilde{\xi}_s|^{m-2}\langle(\tilde{\sigma}^X(s,X_s) - \tilde{\sigma}^Y(s,Y_s))^*\tilde{\xi}_s, \mathrm{d}W_s\rangle\\ &+ m\Big(\tfrac{m}{2}-1\Big)\int_0^t |\tilde{\xi}_s|^{m-4}|(\tilde{\sigma}^X(s,X_s) - \tilde{\sigma}^Y(s,Y_s))^*\tilde{\xi}_s|^2\mathrm{d}s\end{aligned}$$

$$+\frac{m}{2}\int_0^t|\tilde\xi_s|^{m-2}\|\tilde\sigma^X(s,X_s)-\tilde\sigma^Y(s,Y_s)\|_{HS}^2\mathrm{d}s$$
$$:=I_1+I_2+I_3+I_4+I_5.$$

Noticing that by (4.6),

$$|u^X(t,x)-u^Y(t,y)|\leqslant|x-y|+\|u^X-u^Y\|_{\mathbb{L}^\infty(T)},$$

by Young's inequality, we obtain

$$\begin{aligned}I_2&\lesssim\int_0^t|\tilde\xi_s|^m\mathrm{d}s+\lambda\int_0^t|u^X(s,X_s)-u^Y(s,Y_s)|^m\mathrm{d}s\\&\lesssim\int_0^t(|\tilde\xi_s|^m+\lambda|\xi_s|^m)\mathrm{d}s+\lambda^mT\|u^X-u^Y\|^m_{\mathbb{L}^\infty(T)}.\end{aligned}$$

Let $g_s^X(x):=|\nabla^2u_s^X(x)|+|\nabla\sigma_s^X(x)|$. By the definition of $\tilde\sigma^X$, we also have

$$\begin{aligned}|\tilde\sigma^X(s,x)-\tilde\sigma^Y(s,y)|&\leqslant\|\sigma^Y\|_{\mathbb{L}^\infty(T)}|\nabla\Phi^X(s,x)-\nabla\Phi^Y(s,y)|+|\sigma_s^X(x)-\sigma_s^Y(y)|\cdot\|\nabla\Phi^X\|_{\mathbb{L}^\infty(T)}\\&\leqslant\|\sigma^Y\|_{\mathbb{L}^\infty(T)}\Big(|\nabla u^X(s,x)-\nabla u^X(s,y)|+|\nabla u^X(s,y)-\nabla u^Y(s,y)|\Big)\\&\quad+\Big(|\sigma_s^X(x)-\sigma_s^X(y)|+|\sigma_s^X(y)-\sigma_s^Y(y)|\Big)\cdot\|\nabla\Phi^X\|_{\mathbb{L}^\infty(T)}\\&\overset{(2.8)}{\lesssim}|x-y|\Big(\mathcal{M}g_s^X(x)+\mathcal{M}g_s^X(y)\Big)+\|\nabla u^X-\nabla u^Y\|_{\mathbb{L}^\infty(T)}+\|\sigma_s^X-\sigma_s^Y\|_\infty.\end{aligned}$$

Hence,

$$\begin{aligned}I_4+I_5\lesssim&\int_0^t\Big(|\xi_s|^m+|\tilde\xi_s|^m\Big)\Big(\mathcal{M}g_s^X(X_s)+\mathcal{M}g_s^X(Y_s)\Big)^2\mathrm{d}s\\&+T\|\nabla u^X-\nabla u^Y\|^m_{\mathbb{L}^\infty(T)}+\int_0^t\|\sigma_s^X-\sigma_s^Y\|_\infty^m\mathrm{d}s.\end{aligned}$$

Combining the above calculations and by the assumption, we obtain

$$\begin{aligned}|\tilde\xi_t|^m\lesssim|\tilde\xi_0|^m+\|u^X-u^Y\|^m_{\mathbb{H}^{1,\infty}_\infty(T)}+\int_0^t\Big(|\tilde\xi_s|^m+|\xi_s|^m+\|\xi_s\|_\theta^m\Big)\mathrm{d}s\\+\int_0^t\Big(|\xi_s|^m+|\tilde\xi_s|^m\Big)\Big(\mathcal{M}g_s^X(X_s)+\mathcal{M}g_s^X(Y_s)\Big)^2\mathrm{d}s.\end{aligned}\tag{4.9}$$

(iii) Now we define

$$A_t:=t+\int_0^t\Big(\mathcal{M}g_s^X(X_s)+\mathcal{M}g_s^X(Y_s)\Big)^2\mathrm{d}s.$$

By (4.9) and (4.8), we obtain that for all $t\in[0,T]$,

$$|\xi_t|^m\lesssim\|u^X-u^Y\|^m_{\mathbb{H}^{1,\infty}_\infty(T)}A_T+\int_0^T\|\xi_s\|_\theta^m\mathrm{d}s+\int_0^t|\xi_s|^m\mathrm{d}A_s+M_t,$$

where M_t is a continuous local martingale. Note that by $(\mathbf{H}_2^\sigma)$ and (4.5),

$$(s,x)\mapsto|\mathcal{M}g_s^X(x)|^2\in\mathbb{L}^{q_1/2}_{p_1/2}(T)+\mathbb{L}^{q/2}_{p/2}(T).$$

By Khasminskii's estimate (3.29), we have

$$\mathbb{E}\exp\gamma A_T<\infty,\quad\forall\gamma,T>0.$$

Thus we can use the stochastic Gronwall's inequality (2.7) to derive that

$$\sup_{s\in[0,T]}\|\xi_s\|_\theta^m=\Big(\sup_{s\in[0,T]}\mathbb{E}|\xi_s|^\theta\Big)^{m/\theta}\lesssim\|u^X-u^Y\|^m_{\mathbb{H}^{1,\infty}_\infty(T)}+\int_0^T\|\xi_s\|_\theta^m\mathrm{d}s.\tag{4.10}$$

Noticing that by (4.3) and (4.1),

$$\|b^X - b^Y\|_{\mathbb{L}^q_p(T)} \leqslant \left(\int_0^T (\ell^R_t)^q \|X_t - Y_t\|^q_\theta \mathrm{d}t\right)^{1/q} \leqslant \|\ell^R\|_{L^q(0,T)} \sup_{t\in[0,T]} \|\xi_t\|_\theta,$$

and

$$\|\sigma^X - \sigma^Y\|_{\mathbb{L}^\infty(T)} \leqslant c_0 \sup_{t\in[0,T]} \|X_t - Y_t\|_\theta = c_0 \sup_{t\in[0,T]} \|\xi_t\|_\theta,$$

we have by (2.5),

$$\begin{aligned}\|u^X - u^Y\|_{\mathbb{H}^{1,\infty}_\infty(T)} &\lesssim \|b^X - b^Y\|_{\mathbb{L}^q_p(T)} + \|b^X\|_{\mathbb{L}^q_p(T)} \left(\|\sigma^X - \sigma^Y\|_{\mathbb{L}^\infty(T)} + \|b^X - b^Y\|_{\mathbb{L}^q_p(T)}\right) \\ &\lesssim \left(\|\ell^R\|_{L^q(0,T)} + \|b^X\|_{\mathbb{L}^q_p(T)}\right) \sup_{t\in[0,T]} \|\xi_t\|_\theta \overset{(4.2)}{\lesssim} T^{\frac{\gamma-1}{\gamma q}} \sup_{t\in[0,T]} \|\xi_t\|_\theta.\end{aligned}$$

Substituting this into (4.10), we obtain

$$\sup_{s\in[0,T]} \|\xi_s\|^m_\theta \leqslant C T^{\frac{m(\gamma-1)}{\gamma q}} \sup_{t\in[0,T]} \|\xi_t\|^m_\theta, \quad T \in (0,1),$$

where C does not depend on $T \in (0,1)$. By choosing T small enough, we get $\|\xi_t\|^m_\theta = 0$ for all $t \in [0,T]$. By shifting the time T, we obtain the uniqueness. □

It is obvious that b defined in (3.11) does not satisfy (4.1). Below we shall relax it to the weighted total variation norm by Girsanov's transformation. The price we have to pay is that we need to assume that the diffusion coefficient does not depend on the distribution of X. For $\theta \geqslant 1$, let

$$\phi_\theta(x) := 1 + |x|^\theta.$$

We assume

($\mathbf{A}'_\theta$) ($\mathbf{H}^b_1$) or ($\mathbf{H}^b_2$) with $\vartheta = 0$ in (3.15) and (3.16), and for some $\theta \geqslant 1$, there is an $\ell \in L^q_{loc}(\mathbb{R}_+)$ such that for all $\mu, \mu' \in \mathcal{P}(\mathbb{R}^d)$ and $t \geqslant 0$,

$$\|b_1(t,\cdot,\mu) - b_1(t,\cdot,\mu')\|_p + \|b_2(t,\cdot,\mu) - b_2(t,\cdot,\mu')\|_\infty \leqslant \ell_t \|\phi_\theta \cdot (\mu - \mu')\|_{TV}. \tag{4.11}$$

It should be noted that [25, Theorem 6.15],

$$\mathcal{W}_\theta(\mu,\mu') \leqslant c\|\phi_\theta \cdot (\mu - \mu')\|^{1/\theta}_{TV}.$$

Theorem 4.3. *Let $\theta \geqslant 1$. Assume $\sigma(t,x,\mu) = \sigma(t,x)$ satisfies ($\mathbf{H}^\sigma_2$) and ($\mathbf{A}'_\theta$) holds. Then for any initial random variable X_0 with finite m-order moment, where $m > 2\theta$, there is a unique weak (strong) solution to DDSDE* (1.1) *in the class that for any $T > 0$,*

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

Proof. We use the Girsanov transform as used in [19] to show the weak uniqueness, and so also the strong uniqueness. Since under the assumptions of the theorem, weak solutions are also strong solutions (see Corollary 3.8), without loss of generality, let $X^{(i)}, i = 1,2$ be two solutions of SDE (1.1) defined on the same probability space $(\Omega, \mathscr{F}, \mathbf{P})$ and with the same Brownian motion and starting point ξ. That is,

$$\mathrm{d}X^{(i)}_t = \sigma_t(X^{(i)}_t)\mathrm{d}W_t + b_t(X^{(i)}_t, \mu^{(i)}_t)\mathrm{d}t, \quad X^{(i)}_0 = \xi,$$

where $\mu^{(i)}_t = \mathbf{P} \circ (X^{(i)}_t)^{-1}$. We want to show $\mu^{(1)}_t = \mu^{(2)}_t$.

Under ($\mathbf{H}^\sigma_2$), by Theorem 4.2, there is a unique strong solution to SDE

$$\mathrm{d}Z_t = \sigma_t(Z_t)\mathrm{d}W_t, \quad Z_0 = \xi.$$

Let $m > 2\theta$. It is easy to see that

$$\sup_{t\in[0,T]} \mathbb{E}|Z_t|^m \leqslant C\Big(\mathbb{E}|\xi|^m + 1\Big). \tag{4.12}$$

Define

$$\tilde{b}^{(i)}_s(x) := \sigma^{-1}_s(x)\cdot b_s(x,\mu^{(i)}_s),\ \ \tilde{W}^{(i)}_t := W_t - \int_0^t \tilde{b}^{(i)}_s(Z_s)\mathrm{d}s,$$

and

$$\mathscr{E}^{(i)}_T := \exp\left\{\int_0^T \tilde{b}^{(i)}_s(Z_s)\mathrm{d}W_s - \frac{1}{2}\int_0^T |\tilde{b}^{(i)}_s(Z_s)|^2\mathrm{d}s\right\}.$$

Since $\|\tilde{b}^{(i)}\|_{\mathbb{L}^q_p(T)} < \infty$, by Khasminskii's estimate (3.29), we have

$$\mathbf{E}\exp\left\{\gamma\int_0^T |\tilde{b}^{(i)}_s(Z_s)|^2\mathrm{d}s\right\} \leqslant C_{T,\gamma},\ \ \forall\gamma > 0, \tag{4.13}$$

and for any $\gamma\in\mathbb{R}$,

$$\mathbb{E}(\mathscr{E}^{(i)}_T)^\gamma \leqslant C_{T,\gamma} < \infty.$$

Hence, for each $i = 1, 2$, $\mathbf{E}\mathscr{E}^{(i)}_T = 1$, and $\tilde{W}^{(i)}$ is still a Brownian motion under $\mathscr{E}^{(i)}_T\cdot\mathbf{P}$, and

$$\mathrm{d}Z_t = \sigma_t(Z_t)\mathrm{d}\tilde{W}^{(i)}_t + b_t(Z_t,\mu^{(i)}_t)\mathrm{d}t,\ \ Z_0 = \xi.$$

Since the above SDE admits a unique strong solution, we have

$$(\mathscr{E}^{(i)}_T\mathbf{P})\circ Z^{-1}_T = \mathbf{P}\circ(X^{(i)}_T)^{-1} = \mu^{(i)}_T,\ \ i = 1, 2.$$

Therefore, for $\delta = \frac{m}{m-\theta} < 2$, by Hölder's inequality, we get

$$\begin{aligned}\|\phi_\theta\cdot(\mu^{(1)}_T - \mu^{(2)}_T)\|_{TV} &= \|\phi_\theta\cdot((\mathscr{E}^{(1)}_T\mathbf{P})\circ Z^{-1}_T - (\mathscr{E}^{(2)}_T\mathbf{P})\circ Z^{-1}_T)\|_{TV}\\ &\leqslant \mathbf{E}\Big(\phi_\theta(Z_T)|\mathscr{E}^{(1)}_T - \mathscr{E}^{(2)}_T|\Big) \leqslant \|\phi_\theta(Z_T)\|_{\frac{\delta}{\delta-1}}\|\mathscr{E}^{(1)}_T - \mathscr{E}^{(2)}_T\|_\delta.\end{aligned} \tag{4.14}$$

Noticing that

$$\mathrm{d}\mathscr{E}^{(i)}_t = \mathscr{E}^{(i)}_t\tilde{b}^{(i)}_t(Z_t)\mathrm{d}W_t,$$

we have

$$\mathrm{d}(\mathscr{E}^{(1)}_t - \mathscr{E}^{(2)}_t) = (\mathscr{E}^{(1)}_t\tilde{b}^{(1)}_t(Z_t) - \mathscr{E}^{(2)}_t\tilde{b}^{(2)}_t(Z_t))\mathrm{d}W_t.$$

By Itô's formula, we have

$$\begin{aligned}\mathrm{d}|\mathscr{E}^{(1)}_t - \mathscr{E}^{(2)}_t|^2 &= |\mathscr{E}^{(1)}_t\tilde{b}^{(1)}_t(Z_t) - \mathscr{E}^{(2)}_t\tilde{b}^{(2)}_t(Z_t)|^2\mathrm{d}t + M_t,\\ &\leqslant 2|\mathscr{E}^{(1)}_t - \mathscr{E}^{(2)}_t|^2|\tilde{b}^{(1)}_t(Z_t)|^2\mathrm{d}t + 2|\mathscr{E}^{(2)}_t|^2|\tilde{b}^{(1)}_t(Z_t) - \tilde{b}^{(2)}_t(Z_t)|^2\mathrm{d}t + M_t,\end{aligned}$$

where M is a continuous local martingale. Since $\delta < 2$, by stochastic Gronwall's inequality and (4.13), we obtain

$$\Big(\mathbf{E}|\mathscr{E}^{(1)}_T - \mathscr{E}^{(2)}_T|^\delta\Big)^{2/\delta} \lesssim \int_0^T \mathbf{E}|\mathscr{E}^{(2)}_t(\tilde{b}^{(1)}_t(Z_t) - \tilde{b}^{(2)}_t(Z_t))|^2\mathrm{d}t.$$

Let $\gamma\in(1, 1/(d/p+2/q))$. By Hölder's inequality and Krylov's estimate (3.18), we further have

$$\begin{aligned}\Big(\mathbf{E}|\mathscr{E}^{(1)}_T - \mathscr{E}^{(2)}_T|^\delta\Big)^{2/\delta} &\lesssim \left(\int_0^T \mathbf{E}|\tilde{b}^{(1)}_t(Z_t) - \tilde{b}^{(2)}_t(Z_t)|^{2\gamma}\mathrm{d}t\right)^{1/\gamma}\\ &\lesssim \Big\||\tilde{b}^{(1)}_1 - \tilde{b}^{(2)}_1|^{2\gamma}\Big\|^{1/\gamma}_{\mathbb{L}^{q/(2\gamma)}_{p/(2\gamma)}(T)} + \Big\|\tilde{b}^{(1)}_2 - \tilde{b}^{(2)}_2\Big\|^2_{\mathbb{L}^{2\gamma}_\infty(T)}\\ &\lesssim \Big\|\tilde{b}^{(1)}_1 - \tilde{b}^{(2)}_1\Big\|^2_{\mathbb{L}^q_p(T)} + \Big\|\tilde{b}^{(1)}_2 - \tilde{b}^{(2)}_2\Big\|^2_{\mathbb{L}^q_\infty(T)}\\ &\overset{(4.11)}{\lesssim} \left(\int_0^T \ell^q_s\|\phi_\theta(\mu^{(1)}_s - \mu^{(2)}_s)\|^q_{TV}\mathrm{d}s\right)^{2/q},\end{aligned}$$

which together with (4.14) and (4.12) yields

$$\|\phi_\theta(\mu_T^{(1)} - \mu_T^{(2)})\|_{TV}^q \leqslant C \int_0^T \ell_s^q \|\phi_\theta(\mu_s^{(1)} - \mu_s^{(2)})\|_{TV}^q \mathrm{d}s.$$

By Gronwall's inequality, we obtain

$$\|\phi_\theta(\mu_T^{(1)} - \mu_T^{(2)})\|_{TV}^q = 0.$$

The proof is thus complete. □

5. Application to nonlinear Fokker-Planck equations

In this section we present some applications to nonlinear Fokker-Planck equations. First of all we recall the following superposition principle: one-to-one correspondence between DDSDE (1.1) and nonlinear Fokker-Planck equation (1.2), which is originally due to Figalli [9] and Trevisan [23].

Theorem 5.1 (Superposition principle)**.** *Let $\mu_t : \mathbb{R}_+ \to \mathcal{P}(\mathbb{R}^d)$ be a continuous curve such that for each $T > 0$,*

$$\int_0^T \int_{\mathbb{R}^d} \Big(|(\sigma_t^{ik}\sigma_t^{jk})(x, \mu_t)| + |b_t(x, \mu_t)| \Big) \mu_t(\mathrm{d}x)\mathrm{d}t < \infty. \tag{5.1}$$

Then μ_t solves the nonlinear Fokker-Planck equation (1.2) *in the distributional sense if and only if there exists a martingale solution $\mathbb{P} \in \mathscr{M}_\nu^{\sigma,b}$ to DDSDE* (1.1) *so that for each $t > 0$,*

$$\mu_t = \mathbb{P} \circ w_t^{-1}.$$

In particular, if there is at most one element in $\mathcal{M}_\nu^{\sigma,b}$ with time martingale $\mu_t := \mu_{X_t}, t \geqslant 0$, satisfying (5.1)*, then there is at most one solution to* (1.2) *satisfying* (5.1).

Proof. If $\mathbb{P} \in \mathscr{M}_\nu^{\sigma,b}$ and $\mu_t = \mathbb{P} \circ w_t^{-1}$, then by (5.1) and (3.1), it is easy to see that μ_t solves (1.2). Now we assume μ_t solves (1.2). Consider the following linear Fokker-Planck equation:

$$\partial_t \tilde{\mu}_t = (\mathscr{L}_t^{\sigma^\mu})^* \tilde{\mu}_t + \mathrm{div}(b_t^\mu \cdot \tilde{\mu}_t),$$

where $b_t^\mu(x) := b_t(x, \mu_t)$ and $\sigma_t^\mu(x) := \sigma_t(x, \mu_t)$. Since μ_t is a solution of the above linear Fokker-Planck equation, by [23, Theorem 2.5], there is a martingale solution $\mathbb{P} \in \mathscr{M}_\nu^{\sigma^\mu, b^\mu}$ so that

$$\mu_t = \mathbb{P} \circ X_t^{-1}.$$

In particular, $\mathbb{P} \in \mathscr{M}_\nu^{\sigma,b}$. The last assertion is then obvious and thus the proof is complete. □

We have the following useful consequence.

Corollary 5.2. *For $\nu \in \mathcal{P}(\mathbb{R}^d)$, let $\mathbb{M}_\nu$ be the set of all continuous curves $\mu_t : \mathbb{R}_+ \to \mathcal{P}(\mathbb{R}^d)$ with property* (5.1) *and $\mu_0 = \nu$. Then there is exactly one solution for* (1.2) *in the class Θ_ν for all $\nu \in \mathcal{P}(\mathbb{R}^d)$ if and only if there is exactly one $\mathbb{P} \in \mathscr{M}_\nu^{\sigma,b}$ so that $\mathbb{P} \circ w^{-1} \in \mathbb{M}_\nu$ for all $\nu \in \mathcal{P}(\mathbb{R}^d)$.*

Proof. The backward direction "($\Leftarrow$)" follows from the last assertion in Theorem 5.1. The other direction follows from a well-known fact in the theory of martingale problems (see [21]). □

Remark 5.3. *By* [17] *or* [2, Theorem 6.7.8]*,* [16, Theorem 3.1] *and the above corollary, one sees that for a large class of functions σ and b, there exists a unique martingale solution to DDSDE* (1.1)*. For example, if b is bounded measurable and σ satisfies* ($\mathbf{H}_0^\sigma$)*, then there is a unique weak solution for DDSDE* (1.1)*. However, the results in* [16] *does not apply to* (1.5) *with $b \in \mathbb{L}_p^q(T)$ since in this case the Lyapunov condition is not satisfied.*

From the above superposition principle and our well-posedness results, we can obtain the following wellposedness result about the nonlinear Fokker-Planck equations.

Theorem 5.4. *In the situations of Theorems 4.2 and 4.3, there is a unique continuous curve μ_t solving the nonlinear Fokker-Planck equation* (1.2).

Now we turn to the proof of Theorem 1.1.

Proof. The existence and uniqueness of solutions to the nonlinear FPE (1.6) are consequences of Theorem 4.3 and Theorem 5.1. We aim to show the existence and smoothness of the density $\rho^X_t(y)$. Let μ_t be the solution of the Fokker-Planck equation (1.6). We consider the following SDE:

$$\mathrm{d}X_t = b^\mu_t(X_t)\mathrm{d}t + \sqrt{2}\mathrm{d}W_t,\ X_0 = \xi, \tag{5.2}$$

where $b^\mu_t(x) := \int_{\mathbb{R}^d} b_t(x-y)\mu_t(\mathrm{d}y)$. Since $b^\mu \in \cap_{T>0}(\mathbb{L}^q_p(T) + \mathbb{L}^\infty(T))$, where $\frac{d}{p} + \frac{2}{q} < 1$, it is well known that the operator $\Delta + b^\mu \cdot \nabla$ admits a heat kernel $\rho_{b^\mu}(s, x; t, y)$ (see [6, Theorems 1.1 and 1.3]), which is continuous in $(s, x; t, y)$ on $\{(s, x; t, y) : 0 \leqslant s < t < \infty, x, y \in \mathbb{R}^d\}$ and satisfies the following two-sided estimate: For any $T > 0$, there are constants $c_0, \gamma_0 > 1$ such that for all $0 \leqslant s < t \leqslant T$ and $x, y \in \mathbb{R}^d$

$$c_0^{-1}(t-s)^{-d/2}\mathrm{e}^{-\gamma_0|x-y|^2/(t-s)} \leqslant \rho_{b^\mu}(s, x; t, y) \leqslant c_0(t-s)^{-d/2}\mathrm{e}^{-|x-y|^2/(\gamma_0(t-s))},$$

and the gradient estimate: for some $c_1, \gamma_1 > 1$,

$$|\nabla_x\rho_{b^\mu}(s, x; t, y)| \leqslant c_1(t-s)^{-(d+1)/2}\mathrm{e}^{-|x-y|^2/(\gamma_1(t-s))}.$$

If $\mathrm{div} b \equiv 0$, then $\rho_{b^\mu}(s, x; t, y) = \rho_{-b^\mu}(s, y; t, x)$, and so in this case,

$$|\nabla_y\rho_{b^\mu}(s, x; t, y)| \leqslant c_1(t-s)^{-(d+1)/2}\mathrm{e}^{-|x-y|^2/(\gamma_1(t-s))}.$$

In particular, the density of the law of X_t is just given by

$$\rho^X_t(y) = \int_{\mathbb{R}^d} \rho(0, x; t, y)(\mathbb{P} \circ X_0^{-1})(\mathrm{d}x).$$

Strong uniqueness of SDE (5.2) ensures that $\rho^X_t(y)\mathrm{d}y = \mu_t(\mathrm{d}y)$. The desired estimates now follow from the above estimates. □

References

[1] Barbu V. and Röckner M.: Probabilistic Representation for Solutions to Nonlinear FokkerPlanck Equations. *SIAM J. Math. Anal.* **50** (2018), no. 4, 4246-4260.

[2] Bogachev V.I., Krylov N.V. Röckner M. and Shaposhinikov S.V.: *Fokker-Planck-Kolmogorov equations.* AMS, 2015.

[3] Carmona R. and Delarue F.: Probabilistic Analysis of Mean-Field Games. *SIAM J. Control and Optimization*, 51(4):2705-2734, 2013.

[4] Carrillo J.A., Gvalani R.S., Pavliotis G.A. and Schlichting A.: Long-time behavior and phase transitions for the McKean-Vlasov equation on the torus. arXiv: 1806.01719v2.

[5] Chiang T.: McKean-Vlasov equations with discontinuous coefficients. *Soochow J. Math.*, 20(4):507-526, 1994.

[6] Chen Z.Q., Hu E., Xie L. and Zhang X.: Heat kernels for non-symmetric diffusion operators with jumps. *J. Differential Equations* 263 (2017) 6576-6634.

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

[8] Huang X. and Wang F.Y.: Distribution dependent SDEs with singular coefficients. arXiv:1805.01682v1.

[9] Figalli A.: Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients, *J. Funct. Anal.* 254 (2008), no. 1, 109-153.

[10] Funaki T.: A certain class of diffusion processes associalted with nonlinear parabolic equations. *Prob. Theory and Relat. Fields*, 67(3):331-348,1984.

[11] Jacod J. and Shiryaev A. N.: *Limit theorems for stochastic processes*. Springer-Verlag, 2002.

[12] Kac M.: Foundations of Kinetic Theory. In Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume 3: Contributions to Astronomy and Physics, pages 171197, Berkeley, Calif., 1956. University of California Press.
[13] Krylov N.V. and Röckner M.: Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Relat. Fields*, **131** (2005), 154-196.
[14] Kurtz T.G.: Weak and strong solutions of general stochastic models. *Electron. Common. Probab.*, 19(58):1-16, 2014.
[15] Li J. and Min H.: Weak solutions of mean-field stochastic differential equations and application to zero-sum stochastic differential games. *SIAM Journal on Control and Optimization*, 54(3):1826-1858, 2016.
[16] Manita O.A., Romanov M.S., Shaposhnikov S.V.: On uniqueness of solutions to nonlinear Fokker-Planck-Kolmogorov equations. *Nonlinear Anal.* 128 (2015), 199-226.
[17] Manita, O. A.; Shaposhnikov, S. V.: Nonlinear parabolic equations for measures. (Russian) Algebra i Analiz 25 (2013), no. 1, 64–93; translation in St. Petersburg Math. J. 25 (2014), no. 1, 4362 .
[18] McKean H. P.: A class of Markov processes associated with nonlinear parabolic equations. *Proc Nat. Acad Sci USA*, 56(6):1907-1911, 1966.
[19] Mishura Y.S. and Veretennikov A.Y.: Existence and uniqueness theorems for solutions of McKean-Vlasov stochastic equations, arXiv:1603.02212v4.
[20] Scheutzow M.: A stochastic Gronwall's lemma. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, **16**, No. 2 (2013) 1350019 (4 pages).
[21] Stroock D.W. and Varadhan S.S.: *Multidimensional diffusion processes*. Springer-Verlag, Berlin, 1979.
[22] Sznitman A.S.: *Topics in propagation of chaos*. In École d'Été de Prob. de Saint-Flour XIX-1989, Vol. 1464, *Lect. Notes in Math.*, pages 165-251. Springer-Verlag, 1991.
[23] Trevisan D.: Well-posedness of multidimensional diffusion processes with weakly differentiable coefficients. *Electron. J. Probab.***21**, (2016), Paper No. 22, 41 pp.
[24] Vlasov A. A.: The vibrational properties of an electron gas. *Soviet Physics Uspekhi*, 10(6):721, 1968.
[25] Villani C.: *Optimal transport, old and new*. Springer-Verlag, 2009.
[26] Wang F.Y.: Distribution dependent SDEs for Landau type equations. *Stoch. Procc. Appl.*, 128, no. 2, 595-621(2018).
[27] Xie L. and Zhang X.: Ergodicity of stochastic differential equations with jumps and singular coefficients. *arXiv:1705.07402*.
[28] Zhang X.: Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients. *Electron. J. Probab.* **16** (2011), 1096-1116.
[29] Zhang X. and Zhao G.: Heat kernel and ergodicity of SDEs with distributional drifts. arXiv.1710.10537.
[30] Zvonkin, A.K.: A transformation of the phase space of a diffusion process that removes the drift. *Mat. Sbornik*. **93** (135) (1974), 129-149.

Michael Röckner: Fakultät für Mathematik, Universität Bielefeld, 33615, Bielefeld, Germany, Email: roeckner@math.uni-bielefeld.de

Xicheng Zhang: School of Mathematics and Statistics, Wuhan University, Wuhan, Hubei 430072, P.R.China, Email: XichengZhang@gmail.com