Singular McKean-Vlasov (Reflecting) SDEs with Distribution Dependent Noise*

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, Bay Campus, SA1 8EN, United Kingdom

wangfy@tju.edu.cn

May 20, 2022

Abstract

By using Zvonkin's transformation and a two-step fixed point argument in distributions, the well-posedness and regularity estimates are derived for singular McKean-Vlasov SDEs with distribution dependent noise, where the drift contains a term growing linearly in space and distribution and a locally integrable term independent of distribution, while the noise coefficient is weakly differentiable in space and Lipschitz continuous in distribution with respect to the sum of Wasserstein and weighted variation distances. The main results extend existing ones derived for noise coefficients either independent of distribution, or having nice linear functional derivatives in distribution. Singular reflecting SDEs with distribution dependent noise are also studied.

AMS subject Classification: 60H1075, 60G44.
Keywords: McKean-Vlasov SDEs, Wasserstein distance, two-step fixed point argument, weighted variation distance.

1 Introduction

As a crucial stochastic model characterizing nonlinear Fokker-Planck equations and mean field particle systems, the following McKean-Vlasov (i.e. distribution dependent) SDE has been intensively investigated:

$$\mathrm{d}X_t = b_t(X_t, \mathscr{L}_{X_t})\mathrm{d}t + \sigma_t(X_t, \mathscr{L}_{X_t})\mathrm{d}W_t, \quad t \in [0, T], \tag{1.1}$$

*Supported in part by NNSFC (11831014, 11801406, 11921001).

where $T > 0$ is a fixed constant, $(W_t)_{t\in[0,T]}$ is an m-dimensional Brownian motion on a complete filtration probability space $(\Omega, \{\mathscr{F}_t\}_{t\in[0,T]}, \mathbb{P})$, $\mathscr{L}_{X_t}$ is the law of X_t, and for the space $\mathscr{P}$ of probability measures on $\mathbb{R}^d$ equipped with the weak topology,

$$b : [0,T] \times \mathbb{R}^d \times \mathscr{P} \to \mathbb{R}^d, \quad \sigma : [0,T] \times \mathbb{R}^d \times \mathscr{P} \to \mathbb{R}^d \otimes \mathbb{R}^m$$

are measurable. Among many other references, see for instance [1, 2, 5, 6, 9, 10, 13, 14, 16, 18, 27].

When the noise coefficient $\sigma_t(x,\mu) = \sigma_t(x)$, by using Zvonkin's transformation, the well-posedness, regularity estimates and exponential ergodicity have been studied in [15, 19, 20] for the drift $b_t(x,\mu)$ containing a time-spatial locally integrable term in $\tilde{L}_p^q(T)$ for some $(p,q) \in \mathscr{K}$ introduced in [22], see (1.3) and (1.4) below.

Concerning singular McKean-Vlasov SDEs, the well-posedness is derived in [6, 27] when the noise coefficient $\sigma_t(x,\mu)$ has a nice linear functional derivative in μ besides other conditions, where in [6] the drift $b_t(x,\mu)$ is bounded and uniformly Lipschitz continuous in μ with respect to the total variation distance, and in [27] the drift $b_t(x,\mu)$ is Lipschitz continuous in μ with respect to a weighted variation distance uniformly in (t,x), and $\|b_\cdot(\cdot,\mu)\|_{\tilde{L}_p^q(T)} < \infty$ uniformly in μ for some $(p,q) \in \mathscr{K}$.

Comparing with [6, 27], this paper studies (1.1) for $\sigma_t(x,\cdot)$ not necessarily having linear functional derivatives, and for $b_t(x,\mu)$ unbounded in μ and containing a singular distribution independent term. For instance, let $\sigma_t(x,\mu) = \sigma(\mu) := f(\mu) I_{d\times d}$, where $k \geq 1$, $I_{d\times d}$ is the identity matrix, and $f(\mu) := 1 + \mu(|\cdot|^k) \wedge 1$. Then σ is Lipschitz continuous in the k^{th}-Wasserstein distance and hence satisfies assumption (A_1) introduced below, but it does not have bounded linear functional derivative required in [6, 27] , according to (2.3) in [6] and the fact that f is not Lipschitz continuous in the total variation norm.

Instead of the usual fixed point method developed for the well-posedness of distribution dependent SDEs, we will adopt a two-step fixed point argument by freezing the distribution variables in b and σ respectively.

Let $k \in [1,\infty)$. Then

$$\mathscr{P}_k = \left\{ \mu \in \mathscr{P} : \|\mu\|_k := \mu(|\cdot|^k)^{\frac{1}{k}} := \left(\int_{\mathbb{R}^d} |x|^k \mu(\mathrm{d}x) \right)^{\frac{1}{k}} < \infty \right\}$$

is a Polish space under the k^{th}-Wasserstein distance $\mathbb{W}_k$:

$$\mathbb{W}_k(\mu,\nu) := \inf_{\pi\in\mathscr{C}(\mu,\nu)} \left(\int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^k \pi(\mathrm{d}x,\mathrm{d}y) \right)^{\frac{1}{k}}, \quad \mu,\nu \in \mathscr{P}_k,$$

where $\mathscr{C}(\mu,\nu)$ is the set of all couplings of μ and ν. Moreover, $\mathscr{P}_k$ is a complete metric space under the weighted variation norm

$$\|\mu-\nu\|_{k,var} := \sup_{|f|\leq 1+|\cdot|^k} \big|\mu(f) - \nu(f)\big|, \quad \mu,\nu \in \mathscr{P}_k.$$

By [17, Theorem 6.15], there exists a constant $\kappa > 0$ such that

$$\|\mu-\nu\|_{var} + \mathbb{W}_k(\mu,\nu)^k \leq \kappa \|\mu-\nu\|_{k,var}, \tag{1.2}$$

where $\|\cdot\|_{var}$ is the total variation norm. On the other hand, when $k>1$ there is no any constant $c>0$ such that $\|\mu-\nu\|_{k,var}\ge c\mathbb{W}_k(\mu,\nu)$ holds for all $\mu,\nu\in\mathscr{P}_k$.

We call equation (1.1) strongly (weakly) well-posed for distributions in $\mathscr{P}_k$, if for any $\mathscr{F}_0$-measurable initial value X_0 with $\mathscr{L}_{X_0}\in\mathscr{P}_k$ (respectively any initial distribution $\mu\in\mathscr{P}_k$), it has a unique strong solution (respectively weak solution) such that $\mathscr{L}_{X_\cdot}\in C([0,T];\mathscr{P}_k)$, the space of continuous maps from $[0,T]$ to the Polish space $(\mathscr{P}_k,\mathbb{W}_k)$. Moreover, we call (1.1) well-posed for distributions in $\mathscr{P}_k$ if it is strongly and weakly well-posed for distributions in $\mathscr{P}_k$. In this case, we denote

$$P_t^*\mu=\mathscr{L}_{X_t}\ \text{ for the solution with } \mathscr{L}_{X_0}=\mu\in\mathscr{P}_k.$$

To measure the singularity of $b_t(x,\mu)$ in (t,x), we recall locally integrable functional spaces introduced in [22]. For any $t>s\ge 0$ and $p,q\in(1,\infty)$, we write $f\in\tilde{L}_p^q([s,t])$ if $f:[s,t]\times\mathbb{R}^d\to\mathbb{R}$ is measurable with

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

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

$$\tilde{L}_p^q(t)=\tilde{L}_p^q([0,t]),\quad \|f\|_{\tilde{L}_p^q(t)}=\|f\|_{\tilde{L}_p^q([0,t])}. \tag{1.3}$$

We will take (p,q) from the space

$$\mathscr{K}:=\Big\{(p,q):p,q>2,\frac{d}{p}+\frac{2}{q}<1\Big\}. \tag{1.4}$$

For any $\mu\in C([0,T];\mathscr{P}_k)$, let

$$\sigma_t^\mu(x):=\sigma_t(x,\mu_t),\quad b_t^\mu(x):=b_t(x,\mu_t),\quad (t,x)\in[0,T]\times\mathbb{R}^d.$$

We make the following assumption.

(A_0) There exist constants $K>K_0\ge 0$, $l\in\mathbb{N}$, $\{(p_i,q_i):0\le i\le l\}\subset\mathscr{K}$ and $1\le f_i\in\tilde{L}_{p_i}^{q_i}(T)$ for $0\le i\le l$ such that $\sigma_t^\mu(x)$ and $b_t^\mu(x):=b_t^{(1)}(x)+b_t^{\mu,0}(x)$ satisfy the following conditions for all $\mu\in C([0,T];\mathscr{P}_k)$.

(1) $a^\mu:=\sigma^\mu(\sigma^\mu)^*$ is invertible with $\|a^\mu\|_\infty+\|(a^\mu)^{-1}\|_\infty\le K$ and

$$\lim_{\varepsilon\downarrow 0}\sup_{\mu\in C([0,T];\mathscr{P}_k)}\sup_{t\in[0,T],|x-y|\le\varepsilon}\|a_t^\mu(x)-a_t^\mu(y)\|=0.$$

(2) $b^{(1)}$ is locally bounded on $[0,T]\times\mathbb{R}^d$, σ_t^μ is weakly differentiable such that

$$|b_t^{\mu,0}(x)|\le f_0(t,x)+K_0\|\mu_t\|_k,\quad \|\nabla\sigma_t^\mu(x)\|\le\sum_{i=1}^l f_i(t,x),\quad (t,x)\in[0,T]\times\mathbb{R}^d,$$

$$|b_t^{(1)}(x)-b_t^{(1)}(y)|\le K|x-y|,\quad t\in[0,T],x,y\in\mathbb{R}^d.$$

This assumption implies the well-posedness of the SDE with drift $b_t^\mu(x)$ and noise coefficient $\sigma_t^\nu(x)$ for all $\mu,\nu\in C([0,T];\mathscr{P}_k)$, see [15, Theorem 2.1]. To prove the well-posedness of (1.1), we need the following conditions on the distribution dependence.

(A_1) For any $t\in[0,T], x\in\mathbb{R}^d$ and $\mu,\nu\in\mathscr{P}_k$,

$$\|\sigma_t(x,\mu)-\sigma_t(x,\nu)\|+|b_t(x,\mu)-b_t(x,\nu)|\le \mathbb{W}_k(\mu,\nu)\sum_{i=0}^l f_i(t,x).$$

Our first result is the following.

Theorem 1.1. *Assume (A_0) and (A_1). Then the following assertions hold.*

(1) (1.1) *is well-posed for distributions in $\mathscr{P}_k$. Moreover, for any $j\ge k$ there exists a constant $c(j)>0$ such that the solution satisfies*

$$\mathbb{E}\Big[\sup_{t\in[0,T]}|X_t|^j\Big|\mathscr{F}_0\Big]\le c(j)\big\{1+|X_0|^j+(\mathbb{E}[|X_0|^k])^{\frac{j}{k}}\big\}. \tag{1.5}$$

(2) *For any $N>0$ and $j\ge k$, there exists a constant $C_{j,N}>0$ such that for any two solutions X_t^i of* (1.1) *with $\mathbb{E}[|X_0^i|^k]\le N, i=1,2$,*

$$\mathbb{E}\Big(\sup_{t\in[0,T]}|X_t^1-X_t^2|^j\Big|\mathscr{F}_0\Big)\le C_{j,N}\big\{|X_0^1-X_0^2|^j+(\mathbb{E}[|X_0^1-X_0^2|^k])^{\frac{j}{k}}\big\}. \tag{1.6}$$

Consequently,

$$\sup_{t\in[0,T]}\mathbb{W}_k(P_t^*\mu^1,P_t^*\mu^2)\le 2C_{k,N}\mathbb{W}_k(\mu^1,\mu^2),\ \ \mu^1,\mu^2\in\mathscr{P}_k,\ \mu^1(|\cdot|^k),\mu^2(|\cdot|^k)\le N. \tag{1.7}$$

When $K_0=0$, this estimate holds for some constant $C_j>0$ replacing $C_{j,N}$ for any two solutions for distributions in $\mathscr{P}_k$.

Comparing with (A_1), the following assumption allows weaker distribution dependence for $b_t(x,\cdot)$ but needs $b^{(1)}=0$ and stronger conditions on σ.

(A_2) $b^{(1)}=0$, and there exists a constant $\kappa\ge0$ such that the following conditions hold for all $t\in[0,T], x,y\in\mathbb{R}^d$ and $\mu,\nu\in\mathscr{P}_k$.

$$\begin{aligned}&|b_t(x,\mu)-b_t(x,\nu)|\le\big\{\kappa\|\mu-\nu\|_{k,var}+\mathbb{W}_k(\mu,\nu)\big\}\sum_{i=0}^l f_i(t,x),\\&\|\sigma_t(x,\mu)\|^2\vee\|(\sigma_t\sigma_t^*)^{-1}(x,\mu)\|\le K,\\&\|\sigma_t(x,\mu)-\sigma_t(y,\nu)\|\le K\big(|x-y|+\mathbb{W}_k(\mu,\nu)\big),\\&\|\{\sigma_t(x,\mu)-\sigma_t(y,\mu)\}-\{\sigma_t(x,\nu)-\sigma_t(y,\nu)\}\|\le K|x-y|\mathbb{W}_k(\mu,\nu).\end{aligned}$$

Remark 1.1. It is easy to see that the fourth inequality in (A_2) holds if $\sigma_t(x,\mu)$ is differentiable in x with

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

Indeed, this implies

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

Theorem 1.2. *Assume (A_0) and (A_2). Then Theorem 1.1(1) holds. If $\kappa=0$, then for any $N\ge1$, there exists a constant $C(N)>0$, such that*

$$\|P_t^*\mu-P_t^*\nu\|_{var}\le\frac{C(N)}{\sqrt{t}}\mathbb{W}_k(\mu,\nu),\quad t>0,\|\mu\|_k\vee\|\nu\|_k\le N. \tag{1.8}$$

If moreover $K_0=0$, then the constant $C(N)$ can be independent of N.

The above two theorems are proved in Sections 2 and 3 respectively, and Theorem 1.1 will be extended in Section 4 to reflecting SDEs.

2 Proof of Theorem 1.1

Let us explain the main idea of the two-step fixed point argument.

Let X_0 be $\mathscr{F}_0$-measurable with $\gamma:=\mathscr{L}_{X_0}\in\mathscr{P}_k$. Let

$$\mathscr{C}_k^\gamma:=\big\{\mu\in C([0,T];\mathscr{P}_k):\ \mu_0=\gamma\big\}.$$

We solve (1.1) with a fixed distribution parameter $\mu\in\mathscr{C}_k^\gamma$ in the drift:

$$\mathrm{d}X_t^\mu=b_t(X_t^\mu,\mu_t)\mathrm{d}t+\sigma_t(X_t^\mu,\mathscr{L}_{X_t^\mu})\mathrm{d}W_t,\quad t\in[0,T],X_0^\mu=X_0, \tag{2.1}$$

such that the well-posedness of this SDE for distributions in $\mathscr{P}_k$ provides a map

$$\mathscr{C}_k^\gamma\ni\mu\mapsto\Phi_\cdot^\gamma\mu:=\mathscr{L}_{X_\cdot^\mu}\in\mathscr{C}_k^\gamma.$$

Then the well-posedness of (1.1) follows if the map Φ^γ has a unique fixed point in $\mathscr{C}_k^\gamma$.

To solve (2.1), we further fix the distribution parameter $\nu\in\mathscr{C}_k^\gamma$ in σ such that the SDE becomes

$$\mathrm{d}X_t^{\mu,\nu}=b_t(X_t^{\mu,\nu},\mu_t)\mathrm{d}t+\sigma_t(X_t^{\mu,\nu},\nu_t)\mathrm{d}W_t,\quad t\in[0,T],X_0^{\mu,\nu}=X_0,$$

which is well-posed under (A_0) according to [15, Theorem 2.1]. This gives a map

$$\mathscr{C}_k^\gamma \ni \nu \mapsto \Phi_\cdot^{\gamma,\mu}\nu := \mathscr{L}_{X_\cdot^{\mu,\nu}} \in \mathscr{C}_k^\gamma. \tag{2.2}$$

So, we first prove that this map has a unique fixed point such that (2.1) is well-posed, then apply the fixed point theorem to Φ^γ to derive the well-posedness of the original SDE (1.1).

For any $\kappa \geq 0$, let

$$\mathbb{W}_{k,\kappa var}(\mu^1,\mu^2) := \mathbb{W}_k(\mu^1,\mu^2) + \kappa\|\mu^1-\mu^2\|_{k,var}, \quad \mu^1,\mu^2 \in \mathscr{P}_k.$$

To apply the fixed point theorem, we will use the following complete metrics on $\mathscr{C}_k^\gamma$ for $\theta > 0$ and $\kappa \geq 0$:

$$\begin{aligned} &\mathbb{W}_{k,\kappa var,\theta}(\mu,\nu) := \sup_{t\in[0,T]} \mathrm{e}^{-\theta t}\mathbb{W}_{k,\kappa var}(\mu_t,\nu_t), \\ &\mathbb{W}_{k,\theta}(\mu,\nu) := \sup_{t\in[0,T]} \mathrm{e}^{-\theta t}\mathbb{W}_k(\mu_t,\nu_t), \quad \mu,\nu \in \mathscr{C}_k^\gamma. \end{aligned} \tag{2.3}$$

To prove that Φ^γ has a unique fixed point in $\mathscr{C}_k^\gamma$, we need to restrict the map to the following bounded subspaces of $\mathscr{C}_k^\gamma$:

$$\mathscr{C}_k^{\gamma,N} := \Big\{\mu \in \mathscr{C}_k^\gamma : \sup_{t\in[0,T]} \mathrm{e}^{-Nt}(1+\mu_t(|\cdot|^k)) \leq N\Big\}, \quad N > 0, \tag{2.4}$$

and to prove that these spaces are Φ^γ-invariant for large N. This enables us to verify the contraction of Φ^γ in $\mathscr{C}_k^{\gamma,N}$ under a suitable complete metric.

For this purpose, we present the following lemmas. The first one ensures the well-posedness of (2.1).

Lemma 2.1. *Assume (A_0) and that for some constant $\kappa \geq 0$,*

$$\begin{aligned} &|b_t(x,\nu_1) - b_t(x,\nu_2)| \leq \mathbb{W}_{k,\kappa var}(\nu_1,\nu_2)\sum_{i=0}^l f_i(t,x), \\ &\|\sigma_t(x,\nu_1) - \sigma_t(x,\nu_2)\| \leq \mathbb{W}_k(\nu_1,\nu_2)\sum_{i=0}^l f_i(t,x) \end{aligned} \tag{2.5}$$

holds for any $\nu_1,\nu_2 \in \mathscr{P}_k, t \in [0,T]$ and $x \in \mathbb{R}^d$. Then (2.1) is well-posed for distributions in $\mathscr{P}_k$. Moreover, there exist $\theta_0 > 0$ and decreasing function $\beta : [\theta_0,\infty) \to (0,\infty)$ with $\beta(\theta) \downarrow 0$ as $\theta \uparrow \infty$ such that

$$\mathbb{W}_{k,\theta}(\Phi^\gamma\mu,\Phi^\gamma\nu) \leq \beta(\theta)\mathbb{W}_{k,\kappa var,\theta}(\mu,\nu), \quad \mu,\nu \in \mathscr{C}_k^{\gamma,N}. \tag{2.6}$$

Proof. (a) For the well-posedness, it suffices to prove that $\Phi^{\gamma,\mu}$ defined in (2.2) has a unique fixed point in $\mathscr{C}_k^\gamma$.

In general, let $\mu^i \in \mathscr{C}_k^{\gamma^i,N}$ for some $N > 0, \gamma^i \in \mathscr{P}^k, i = 1,2$. For $\nu^i \in \mathscr{C}_k^{\gamma^i}$ and initial value X_0^i with $\mathscr{L}_{X_0^i} = \gamma^i, i = 1,2$, consider the SDEs

$$\mathrm{d}X_t^i = b_t^{\mu^i}(X_t^i)\mathrm{d}t + \sigma_t^{\nu^i}(X_t^i)\mathrm{d}W_t, \quad t \in [0,T], i = 1,2. \tag{2.7}$$

According to [15, Theorem 2.1], under (A_0) these SDEs are well-posed, and by [24, Theorem 2.1], there exist constants $c_0, \lambda_0 \geq 0$ depending on N via $\mu^1 \in \mathscr{C}_k^{\gamma,N}$ due to

$$|b_t^{\mu^1,0}(x)| \leq f_0(t,x) + K_0\|\mu_t^1\|_k,$$

such that for any $\lambda \geq \lambda_0$, the PDE

$$\Big(\partial_t + \frac{1}{2}\mathrm{tr}\{a_t^{\nu^1}\nabla^2\}\Big)u_t + (\nabla u_t)b_t^{\mu^1} = \lambda u_t - b_t^{\mu^1,0}, \quad t \in [0,T], u_T = 0 \tag{2.8}$$

has a unique solution such that

$$\|\nabla^2 u\|_{\tilde{L}_{p_0}^{q_0}(T)} \leq c_0, \quad \|u\|_\infty + \|\nabla u\|_\infty \leq \frac{1}{2}. \tag{2.9}$$

Let $Y_t^i := \Theta_t(X_t^i), i = 1,2, \Theta_t := id + u_t$. By Itô's formula we obtain

$$\begin{aligned}
\mathrm{d}Y_t^1 &= \big\{b_t^{(1)} + \lambda u_t\big\}(X_t^1)\mathrm{d}t + (\{\nabla\Theta_t\}\sigma_t^{\nu^1})(X_t^1)\,\mathrm{d}W_t,\\
\mathrm{d}Y_t^2 &= \big\{\big\{b_t^{(1)} + \lambda u_t + (\nabla\Theta_t)(b_t^{\mu^2} - b_t^{\mu^1})\big\}(X_t^2)\\
&+ \frac{1}{2}\big[\mathrm{tr}\{(a_t^{\nu^2} - a_t^{\nu^1})\nabla^2 u_t\}\big](X_t^2)\big\}\mathrm{d}t + (\{\nabla\Theta_t\}\sigma_t^{\nu^2})(X_t^2)\,\mathrm{d}W_t.
\end{aligned}$$

Let $\eta_t := |X_t^1 - X_t^2|$ and

$$\begin{aligned}
g_r &:= \sum_{i=0}^{l} f_i(r, X_r^2), \quad \tilde{g}_r := g_r\|\nabla^2 u_r(X_r^2)\|,\\
\bar{g}_r &:= \sum_{i=1}^{2}\|\nabla^2 u_r\|(X_r^i) + \sum_{j=1}^{2}\sum_{i=0}^{l} f_i(r, X_r^j), \quad r \in [0,T].
\end{aligned}$$

Since $b_t^{(1)} + \lambda u_t$ is Lipschitz continuous uniformly in $t \in [0,T]$, by (A_0), (2.5) and the maximal functional inequality in [22, Lemma 2.1], there exists a constant $c_1 > 0$ depending on N such that

$$\begin{aligned}
&\big|\big\{b_r^{(1)} + \lambda u_r\big\}(X_r^1) - \big\{b_r^{(1)} + \lambda u_r\big\}(X_r^2)\big| \leq c_1\eta_r,\\
&\big|\big\{(\nabla\Theta_r)(b_r^{\mu^2} - b_r^{\mu^1})\big\}(X_r^2)\big| \leq c_1 g_r \mathbb{W}_{k,\kappa var}(\mu_r^1, \mu_r^2),\\
&\big|\big[\mathrm{tr}\{(a_r^{\nu^2} - a_r^{\nu^1})\nabla^2 u_r\}\big](X_r^2)\big| \leq c_1\tilde{g}_r\mathbb{W}_k(\nu_r^1, \nu_r^2),\\
&\big\|\big\{(\nabla\Theta_r)\sigma_r^{\nu^1}\big\}(X_r^1) - \big\{(\nabla\Theta_r)\sigma_r^{\nu^2}\big\}(X_r^2)\big\|\\
&\leq c_1\bar{g}_r\eta_r + c_1 g_r\mathbb{W}_k(\nu_r^1, \nu_r^2), \quad r \in [0,T].
\end{aligned}$$

So, by Itô's formula, for any $j \geq k$ we find a constant $c_2 > 1$ depending on N such that

$$\mathrm{d}|Y_t^1 - Y_t^2|^{2j} \leq c_2 \eta_t^{2j} \mathrm{d}A_t + c_2(g_t^2 + \tilde{g}_t)\big\{\mathbb{W}_{k,\kappa var}(\mu_t^1, \mu_t^2)^{2j} + \mathbb{W}_k(\nu_t^1, \nu_t^2)^{2j}\big\}\mathrm{d}t + \mathrm{d}M_t \tag{2.10}$$

holds for some martingale M_t with $M_0 = 0$ and

$$A_t := \int_0^t \big\{1 + g_s^2 + \tilde{g}_s + \bar{g}_s^2\big\}\mathrm{d}s.$$

Since $\|\nabla u\|_\infty \leq \frac{1}{2}$ implies $|Y_t^1 - Y_t^2| \geq \frac{1}{2}\eta_t$, this implies

$$\begin{aligned} \eta_t^{2j} &\leq 2^{2j} M_t + 2^{2j}\eta_0^{2j} + 2^{2j} c_2 \int_0^t \eta_r^{2j} \mathrm{d}A_r \\ &+ 2^{2j} c_2 \int_0^t (g_s^2 + \tilde{g}_s)\big\{\mathbb{W}_{k,\kappa var}(\mu_s^1, \mu_s^2)^{2j} + \mathbb{W}_k(\nu_s^1, \nu_s^2)^{2j}\big\}\mathrm{d}s \end{aligned} \tag{2.11}$$

for some constant $c_2 > 0$ and all $t \in [0, T]$. By (2.9), $f_i \in \tilde{L}_{p_i}^{q_i}(T)$ for $(p_i, q_i) \in \mathscr{K}$, Krylov's and Khasminskii's estimates (see [24]), we find an increasing function $\alpha : (0,\infty) \to (0,\infty)$ and a decreasing function $\varepsilon : (0,\infty) \to (0,\infty)$ with $\varepsilon_\theta \to 0$ as $\theta \to \infty$, such that

$$\mathbb{E}[\mathrm{e}^{rA_T}|\mathscr{F}_0] \leq \alpha(r), \quad r > 0,$$

$$\sup_{t\in[0,T]} \mathbb{E}\bigg(\int_0^t \mathrm{e}^{-2k\theta(t-r)}(g_r^2 + \tilde{g}_r)\mathrm{d}r\bigg|\mathscr{F}_0\bigg) \leq \varepsilon_\theta, \quad \theta > 0.$$

By the stochastic Gronwall inequality and the maximal inequality (see [22]), we find a constant $c_3 > 0$ depending on N such that (2.11) yields

$$\begin{aligned} &\bigg\{\mathbb{E}\bigg(\sup_{s\in[0,t]} \eta_s^j\bigg|\mathscr{F}_0\bigg)\bigg\}^2 \\ &\leq c_3 \mathbb{E}\bigg(\eta_0^{2j} + \int_0^t (g_s^2 + \tilde{g}_s)\big\{\mathbb{W}_{k,\kappa var}(\mu_s^1, \mu_s^2)^{2j} + \mathbb{W}_k(\nu_s^1, \nu_s^2)^{2j}\big\}\mathrm{d}s\bigg|\mathscr{F}_0\bigg) \\ &\leq c_3\eta_0^{2j} + c_3 \mathrm{e}^{2k\theta t}\varepsilon_\theta\big\{\mathbb{W}_{k,\kappa var,\theta}(\mu^1, \mu^2)^{2j} + \mathbb{W}_{k,\theta}(\nu^1, \nu^2)^{2j}\big\}. \end{aligned} \tag{2.12}$$

Noting that

$$\mathbb{W}_k(\mathscr{L}_{X_t^1}, \mathscr{L}_{X_t^2})^k \leq \mathbb{E}[|X_t^1 - X_t^2|^k] = \mathbb{E}[\eta_t^k],$$

by taking $j = k$ we obtain

$$\mathbb{W}_{k,\theta}(\mathscr{L}_{X^1}, \mathscr{L}_{X^2})^k \leq \sqrt{c_3}\mathbb{E}[\eta_0^k] + \sqrt{c_3\varepsilon_\theta}\big\{\mathbb{W}_{k,\kappa var,\theta}(\mu^1, \mu^2)^k + \mathbb{W}_{k,\theta}(\nu^1, \nu^2)^k\big\}. \tag{2.13}$$

By taking $X_0^1 = X_0^2 = X_0$ and $\mu^1 = \mu^2 = \mu \in \mathscr{C}_k^{\gamma,N}$, when $\theta > 0$ is large enough such that $\sqrt{c_3\varepsilon_\theta} \leq \frac{1}{2}$, $\Phi^{\gamma,\mu}\nu^i = \mathscr{L}_{X^i}$ satisfies

$$\mathbb{W}_{k,\theta}(\Phi^{\gamma,\mu}\nu^1, \Phi^{\gamma,\mu}\nu^2) \leq \frac{1}{2}\mathbb{W}_{k,\theta}(\nu^1, \nu^2), \quad \nu_1, \nu_2 \in \mathscr{C}_k^\gamma.$$

Thus, $\Phi^{\gamma,\mu}$ has a unique fixed point in $\mathscr{C}_k^\gamma$, so that (2.1) is well-posed for distributions in $\mathscr{P}_k$.

(b) Taking $\nu^i = \Phi^\gamma \mu^i$, we have $\mathscr{L}_{X^i} = \Phi^\gamma \mu^i$, so that (2.13) becomes

$$\mathbb{W}_{k,\theta}(\Phi^\gamma\mu^1, \Phi^\gamma\mu^2) \le (c_3\varepsilon_\theta)^{\frac{1}{2k}}\big\{\mathbb{W}_{k,\kappa var,\theta}(\mu^1,\mu^2) + \mathbb{W}_{k,\theta}(\Phi^\gamma\mu^1, \Phi^\gamma\mu^2)\big\}.$$

Taking $\theta_0 > 0$ large enough such that $c_3\varepsilon_{\theta_0} < 1$ we prove (2.6) for

$$\beta(\theta) := \frac{(c_3\varepsilon_\theta)^{\frac{1}{2k}}}{1-(c_3\varepsilon_\theta)^{\frac{1}{2k}}}, \quad \theta \ge \theta_0.$$

□

Lemma 2.2. *Assume* (A_0).

(1) *There exists a constant* $N_0 > 0$ *such that for any* $N \ge N_0$ *we have* $\Phi^\gamma \mathscr{C}_k^{\gamma,N} \subset \mathscr{C}_k^{\gamma,N}$.

(2) *Solutions to* (1.1) *for distributions in* $\mathscr{P}_k$ *satisfy* (1.5) *for any* $j \ge k$ *and some constant* $c(j) > 0$.

Proof. (1) Simply denote $M_t = \int_0^t \sigma_s(X_s^\mu, \mathscr{L}_{X_s^\mu})\mathrm{d}W_s$. Since $\|\sigma\|_\infty < \infty$ due to (A_0), we have

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

Combining this with Lemma 2.3 below, we find some constants $c_0, c_1 > 0$ such that

$$\begin{aligned}
&\mathbb{E}(1+|X_t^\mu|^k)\\
&\le \mathbb{E}(1+|X_0|^k) + c_0\mathbb{E}\bigg|\int_0^t (K_0\|\mu_s\|_k + f_0(s,X_s^\mu) + |X_s^\mu| + 1)\mathrm{d}s\bigg|^k + \mathbb{E}\big|M_t\big|^k\\
&\le c_1 + c_1\bigg|\int_0^t \|\mu_s\|_k^2\mathrm{d}s\bigg|^{k/2} + c_1\int_0^t \mathbb{E}(1+|X_s^\mu|^k)\mathrm{d}s, \quad t\in[0,T].
\end{aligned}$$

By Gronwall's inequality, we find $c_2, c_3 > 0$ such that

$$\begin{aligned}
\mathbb{E}(1+|X_t^\mu|^k) &\le c_2 + c_2\bigg|\int_0^t \mathrm{e}^{-\frac{2N}{k}s}\|\mu_s\|_k^2 \mathrm{e}^{\frac{2N}{k}s}\mathrm{d}s\bigg|^{k/2}\\
&\le c_3 + c_3 N^{1-k/2}\mathrm{e}^{Nt}, \quad \mu\in\mathscr{C}_k^{\gamma,N}, t\in[0,T].
\end{aligned}$$

Therefore, we find a constant $N_0 > 0$ such that

$$\sup_{t\in[0,T]}(1+\|\Phi_t^\gamma\mu\|_k^k)\mathrm{e}^{-Nt} \le c_3 + c_3N^{1-k/2} \le N, \quad N\ge N_0, \mu\in\mathscr{C}_k^{\gamma,N}.$$

That is, $\Phi^\gamma\mathscr{C}_k^{\gamma,N} \subset \mathscr{C}_k^{\gamma,N}$ for $N \ge N_0$.

(2) Let X_t solve (1.1) with $\gamma := \mathscr{L}_{X_0} \in \mathscr{P}_k$, and denote $\mu_t := \mathscr{L}_{X_t}$. Then $X_t = X_t^\mu$. By (A_0) and Itô's formula, for any $j \geq 1$ we find a constant $c_1 > 0$ such that

$$(2.14)\qquad |X_t|^{2j} - |X_0|^{2j} \leq c_1 \int_0^t \big\{1 + |X_s|^{2j} + |X_s|^{2j-1} f_0(s, X_s) + \|\mu_s\|_k^{2j}\big\}\mathrm{d}s + M_t$$

holds for some martingale M_t with $\mathrm{d}\langle M\rangle_t \leq c_1^2 |X_t|^{2(2j-1)}\mathrm{d}t$. Noting that

$$\begin{aligned} c_1 \int_0^t |X_s|^{2j-1} f_0(s, X_s)\mathrm{d}s &\leq c_1 \Big(\sup_{s\in[0,t]} |X_s|^{2j-1}\Big) \int_0^t f_0(s, X_s)\mathrm{d}s \\ &\leq \frac{1}{2}\sup_{s\in[0,t]} |X_s|^{2j} + c_2 \bigg(\int_0^t f_0(s, X_s)\mathrm{d}s\bigg)^{2j} \end{aligned}$$

holds for some constant $c_2 > 0$, we see that $\eta_t := \sup_{s\in[0,t]} |X_s|^{2j}$ satisfies

$$(2.15)\qquad \eta_t \leq 2|X_0|^{2j} + 2c_1 \int_0^t \big\{1 + \eta_s + \|\mu_s\|_k^{2j}\big\}\mathrm{d}s + 2c_2 \bigg(\int_0^t f_0(s, X_s)\mathrm{d}s\bigg)^{2j} + 2 \sup_{s\in[0,t]} M_s.$$

By $\mathrm{d}\langle M\rangle_t \leq c_1^2 |X_t|^{2(2j-1)}\mathrm{d}t$ and BDG's inequality, we find constants $c_3, c_4 > 0$ such that

$$\begin{aligned} \mathbb{E}\Big(\sup_{s\in[0,t]} M_s\Big|\mathscr{F}_0\Big) &\leq c_3 \mathbb{E}\bigg[\bigg(\int_0^t |X_s|^{2(2j-1)}\mathrm{d}s\bigg)^{\frac{1}{2}}\bigg|\mathscr{F}_0\bigg] \\ &\leq \frac{1}{4}\mathbb{E}\big(\eta_t\big|\mathscr{F}_0\big) + c_4 \int_0^t \big\{1 + \mathbb{E}(\eta_s|\mathscr{F}_0)\big\}\mathrm{d}s. \end{aligned}$$

Combining this with (2.15) and (2.19) below, we find a constant $c_5 > 0$ such that

$$(2.16)\qquad \mathbb{E}\big(\eta_t\big|\mathscr{F}_0\big) \leq c_5 + c_5|X_0|^{2j} + c_5 \int_0^t \big\{\mathbb{E}(\eta_s|\mathscr{F}_0) + \|\mu_s\|_k^{2j}\big\}\mathrm{d}s, \quad t \in [0, T].$$

By Gronwall's inequality, there exists a constant $c_6 > 0$ such that

$$(2.17)\qquad \mathbb{E}\big(\eta_t\big|\mathscr{F}_0\big) \leq c_6 + c_6|X_0|^{2j} + c_6 \int_0^t \|\mu_s\|_k^{2j}\mathrm{d}s, \quad t \in [0, T].$$

In particular, choosing $j = k$ and applying Jensen's inequality, we derive

$$\begin{aligned} \mathbb{E}\Big[\sup_{s\in[0,t]} |X_s|^k\Big|\mathscr{F}_0\Big] &\leq \Big\{\mathbb{E}\big(\eta_t\big|\mathscr{F}_0\big)\Big\}^{\frac{1}{2}} \\ &\leq \sqrt{c_6}\big(1 + |X_0|^k\big) + \frac{c_6}{2}\int_0^t \|\mu_s\|_k^k\mathrm{d}s + \frac{1}{2}\sup_{s\in[0,t]} \|\mu_s\|_k^k. \end{aligned}$$

Noting that $\|\mu_s\|_k^k = \mathbb{E}[|X_s|^k]$, by taking expectation we obtain

$$\|\mu_t\|_k^k \leq \mathbb{E}\Big[\sup_{s\in[0,t]} |X_s|^k\Big] \leq 2\sqrt{c_6}\big(1 + \mathbb{E}[|X_0|^k]\big) + c_6 \int_0^t \|\mu_s\|_k^k\mathrm{d}s, \quad t \in [0, T].$$

By Gronwall's inequality, we find a constant $c > 0$ such that

$$\|\mu_t\|_k^k \le c(1 + \mathbb{E}[|X_0|^k]), \quad t \in [0,T].$$

Substituting into (2.17) we prove (1.5). $\square$

Lemma 2.3. *Assume* (A_0). *For any* $(p,q) \in \mathscr{K}$, *there exist a constant* $c_0 \ge 1$ *and a function* $c : [1,\infty) \to (0,\infty)$ *such that for any* $j \ge 1$ *and* $\mu \in \mathscr{C}_k^\gamma$, *the solution to* (2.1) *satisfies*

$$\mathbb{E}\Big[\mathrm{e}^{\int_0^t |f_s(X_s^\mu)|^2 \mathrm{d}s}\Big|\mathscr{F}_0\Big] \le \mathrm{e}^{c_0 + c_0\int_0^t \|\mu_s\|_k^2 \mathrm{d}s + c_0\|f\|_{\tilde{L}_p^q(t)}^{c_0}}, \tag{2.18}$$

$$\mathbb{E}\left[\left(\int_0^t |f_s(X_s^\mu)|^2 \mathrm{d}s\right)^j \middle| \mathscr{F}_0\right] \le c(j)\left(1 + \int_0^t \|\mu_s\|_k^2 \mathrm{d}s\right)^j \|f\|_{\tilde{L}_p^q(t)}^{2j} \tag{2.19}$$

for any $t \in [0,T]$ *and* $f \in \tilde{L}_p^q(t), t \in [0,T]$.

Proof. Consider the SDE

$$\mathrm{d}\bar{X}_t = b_t^{(1)}(\bar{X}_t)\mathrm{d}t + \sigma_t(\bar{X}_t, \Phi_t^\gamma \mu)\mathrm{d}W_t, \quad \bar{X}_0 = X_0, t \in [0,T].$$

By Khasminskii's estimate (see [24]), there exists a constant $c_1 > 1$ such that

$$\mathbb{E}\Big[\mathrm{e}^{\int_0^t |f_s(\bar{X}_s^\mu)|^2 \mathrm{d}s}\Big|\mathscr{F}_0\Big] \le \mathrm{e}^{c_1 + c_1\|f\|_{\tilde{L}_p^q(t)}^{c_1}}, \quad f \in \tilde{L}_q^p(t), t \in [0,T]. \tag{2.20}$$

By (A_0),

$$\xi_t := \sigma_t(\bar{X}_t, \Phi_t^\gamma \mu)^* \{\sigma_t(\bar{X}_t, \Phi_t^\gamma \mu)\sigma_t(\bar{X}_t, \Phi_t^\gamma \mu)^*\}^{-1} b_t^{\mu,0}(\bar{X}_t)$$

satisfies

$$|\xi_t| \le c_2 f_0(t, \bar{X}_t) + c_2\|\mu_t\|_k, \quad t \in [0,T]$$

for some constant $c_2 > 0$. Combining this with (2.20), we conclude that

$$R_t := \mathrm{e}^{\int_0^t \langle \xi_s, \mathrm{d}W_s\rangle - \frac{1}{2}\int_0^t |\xi_s|^2 \mathrm{d}s}, \quad t \in [0,T]$$

is a martingale satisfying

$$\mathbb{E}[R_t^2|\mathscr{F}_0] \le \mathrm{e}^{c_3 + c_3\int_0^t \|\mu_s\|_k^2 \mathrm{d}s}, \quad t \in [0,T] \tag{2.21}$$

for some constant $c_3 > 0$. By Girsanov's theorem

$$\tilde{W}_t := W_t - \int_0^t \xi_s \mathrm{d}s, \quad t \in [0,T]$$

is m-dimensional Brownian motion under the probability measure $\mathbb{Q}_T := R_T\mathbb{P}$. Since $b^\mu = b^{(1)} + b^{\mu,0}$, we may reformulate the SDE for $\bar{X}_t$ as

$$\mathrm{d}\bar{X}_t = b_t^\mu(\bar{X}_t)\mathrm{d}t + \sigma_t(\bar{X}_t, \Phi_t^\gamma \mu)\mathrm{d}\tilde{W}_t, \quad \bar{X}_0 = X_0, t \in [0,T],$$

so that the weak uniqueness of (2.1) yields $\mathscr{L}_{\bar X|\mathbb{Q}_T} = \mathscr{L}_{X^\mu}$. Combining this with (2.20) and (2.21), we obtain

$$\begin{aligned}&\mathbb{E}\big[\mathrm{e}^{\int_0^t f(s,X_s^\mu)^2\mathrm{d}s}\big|\mathscr{F}_0\big] = \mathbb{E}\big[R_t\mathrm{e}^{\int_0^t f(s,\bar X_s)^2\mathrm{d}s}\big|\mathscr{F}_0\big]\\&\le \big(\mathbb{E}[|R_t|^2|\mathscr{F}_0]\big)^{\frac12}\big(\mathbb{E}[\mathrm{e}^{\int_0^t 2f(s,\bar X_s)^2\mathrm{d}s}|\mathscr{F}_0]\big)^{\frac12} \le \mathrm{e}^{c_4+c_4\int_0^t\|\mu_s\|_k^2\mathrm{d}s+c_4\|f\|^{c_1}_{\tilde L_p^q(t)}}\end{aligned}$$

for some constant $c_4>0$. This implies (2.18) for some constant $c_0>1$.

By choosing large enough constant $C_j>0$ such that $h(r):=\{\log(C_j+r)\}^j$ is concave for $r\ge 0$, using Jensen's inequality and (2.18) we find a constant $\tilde C_j>1$ increasing in $j\ge 1$ such that

$$\begin{aligned}&\mathbb{E}\bigg[\bigg(\int_0^t|f_s(X_s^\mu)|^2\mathrm{d}s\bigg)^j\bigg|\mathscr{F}_0\bigg]\le\mathbb{E}\bigg(\Big[\log\big(C_j+\mathrm{e}^{\int_0^t f_s(X_s^\mu)^2\mathrm{d}s}\big)\Big]^j\bigg|\mathscr{F}_0\bigg)\\&\le\Big[\log\big(C_j+\mathbb{E}[\mathrm{e}^{\int_0^t f_s(X_s^\mu)^2\mathrm{d}s}]\big|\mathscr{F}_0\big)\Big]^j\le\tilde C_j\bigg(1+\int_0^t\|\mu_s\|_k^2\mathrm{d}s+\|f\|^{c_1}_{\tilde L_p^q(t)}\bigg)^j.\end{aligned}$$

Using $\frac{f}{\|f\|_{\tilde L_p^q(t)}}$ replacing f, we derive

$$\mathbb{E}\bigg[\bigg(\int_0^t|f_s(X_s^\mu)|^2\mathrm{d}s\bigg)^j\bigg|\mathscr{F}_0\bigg]\le\|f\|^{2j}_{\tilde L_p^q(t)}\tilde C_j\bigg(1+\int_0^t\|\mu_s\|_k^2\mathrm{d}s+1\bigg)^j$$

which implies (2.19). □

We are now ready to prove Theorem 1.1.

Proof of Theorem 1.1. (1) Since (1.5) is included in Lemma 2.2, it remains to prove that Φ^γ has a unique fixed point in $\mathscr{C}_k^{\gamma,N}$ for $N>N_0$.

Under (A_1), (2.5) holds for $\kappa=0$, so that (2.6) becomes

$$\mathbb{W}_{k,\theta}(\Phi^\gamma\mu^1,\Phi^\gamma\mu^2)\le\beta(\theta)\mathbb{W}_{k,\theta}(\mu^1,\mu^2),\quad \theta\ge\theta_0.$$

Taking large enough θ such that $\beta(\theta)<1$ we prove the contraction of Φ^γ on the complete metric space $(\mathscr{C}_k^{\gamma,N},\mathbb{W}_{k,\theta})$, so that Φ^γ has a unique fixed point in $\mathscr{C}_k^{\gamma,N}$.

(2) Let $\kappa=0$ and $N>0$. For any two solutions X_t^i of (1.1) with $\mathbb{E}[|X_0^i|^k]\le N$, they solve (2.1) for $\mu_t^i=\nu_t^i=\mathscr{L}_{X_t^i}, i=1,2$. By (1.5), there exists a constant $K_N>0$ depending on N such that $\mu,\nu\in\mathscr{C}_k^{\gamma,K_N}$. Since $\kappa=0$ and (2.13) for large θ such that $\sqrt{c_3\varepsilon_\theta}\le\frac14$, where θ and c_3 depend on N, we obtain

$$\mathbb{W}_{k,\theta}(\mu_t^1,\mu_t^2)^k\le 2\sqrt{c_3}\mathbb{E}[|X_0^1-X_0^2|^k].$$

Substituting into (2.12) for $\kappa=0$ yields the estimate (1.6) for some constant $C_{j,N}>0$. When $K_0=0$ we have $|b^{\mu,0}|\le f_0$ for any $\mu\in C([0,T];\mathscr{P}_k)$, so that all the above constants are uniformly bounded in N, hence (1.6) holds for some constant $C_{j,N}=C_j$ independent of N.

Finally, by taking $j=k$ and X_0^1,X_0^2 such that

$$\mathscr{L}_{X_0^1}=\mu^1,\quad \mathscr{L}_{X_0^2}=\mu^2,\quad \mathbb{E}[|X_0^1-X_0^2|^k]=\mathbb{W}_k(\mu^1,\mu^2)^k,$$

we deduce (1.7) from (1.6). □

3 Proof of Theorem 1.2

By Lemma 2.1, (2.1) is well-posed so that the map Φ^γ is well-defined on $\mathscr{C}_k^\gamma$. Moreover, Lemma 2.2 ensures that $\mathscr{C}_k^{\gamma,N}$ is Φ^γ-invariant for $N \geq N_0$. So, for the well-posedness of (1.1), it suffices to prove the contraction of Φ^γ in $\mathscr{C}_k^{\gamma,N}$ for $N > N_0$ under the metric $\mathbb{W}_{k,\kappa var,\theta}$ for large $\theta > 0$. To this end, we will make use of the parametrix expansion for transition densities.

3.1 Parametrix expansion

For any $\mu \in \mathscr{C}_k^\gamma$, and a measurable map Γ on $\mathscr{C}_k^\gamma$, consider the following SDE:

$$\mathrm{d}X_t^{x,\mu} = b_t(X_t^{x,\mu}, \mu_t)\mathrm{d}t + \sigma_t(X_t^{x,\mu}, \Gamma_t\mu)\mathrm{d}W_t, \quad t \in [0,T], \quad X_0^{x,\mu} = x. \tag{3.1}$$

Again by [15, Theorem 2.1], (A_0) implies the well-posedness of this SDE. Moreover, by Theorem 6.2.7(ii)-(iii) in [3], $\mathscr{L}_{X_t^{x,\mu}}$ has a density function $p_t^\mu(x,\cdot)$ (called transition density) with respect to the Lebesgue measure. By the standard Markov property of solutions to (3.1), the solution to (2.1) satisfies

$$\mathbb{E}f(X_t^\mu) = \int_{\mathbb{R}^d} \gamma(\mathrm{d}x) \int_{\mathbb{R}^d} f(y) p_t^\mu(x,y)\mathrm{d}y, \quad t \in (0,T], f \in \mathscr{B}_b(\mathbb{R}^d), \tag{3.2}$$

where $\mathscr{B}_b(\mathbb{R}^d)$ is the class of bounded measurable functions on $\mathbb{R}^d$. So, to estimate $\|\Phi_t^\gamma\mu - \Phi_t^\gamma\nu\|_{k,var}$, it suffices to calculate $|p_t^\mu(x,y) - p_t^\nu(x,y)|$, for which we make use of the parametrix expansion formula.

For any $x, z \in \mathbb{R}^d, 0 \leq s < t \leq T$ and $\mu \in \mathscr{C}_k^\gamma$, let $p_{s,t}^{\mu,z}(x,\cdot)$ be the distribution density function of the random variable

$$X_{s,t}^{x,\mu,z} := x + \int_s^t \sigma_r(z, \Gamma_r\mu)\mathrm{d}W_r.$$

Let

$$a_{s,t}^{\mu,z} := \int_s^t (\sigma_r\sigma_r^*)(z, \Gamma_r\mu)\mathrm{d}r, \quad 0 \leq s < t \leq T. \tag{3.3}$$

We have

$$p_{s,t}^{\mu,z}(x,y) = \frac{\exp[-\frac{1}{2}\langle (a_{s,t}^{\mu,z})^{-1}(y-x), y-x\rangle]}{(2\pi)^{\frac{d}{2}}(\det\{a_{s,t}^{\mu,z}\})^{\frac{1}{2}}}, \quad x, y \in \mathbb{R}^d. \tag{3.4}$$

Obviously, (A_0) and (A_2) imply

$$\begin{aligned} &\|a_{s,t}^{\mu,z} - a_{s,t}^{\nu,z}\| \leq K \int_s^t \mathbb{W}_k\big(\Gamma_r\mu, \Gamma_r\nu\big)\mathrm{d}r, \\ &\frac{1}{K(t-s)} \leq \|(a_{s,t}^{\mu,z})^{-1}\| \leq \frac{K}{t-s}, \quad 0 \leq s < t \leq T, \ \mu, \nu \in \mathscr{C}_k^\gamma. \end{aligned} \tag{3.5}$$

Next, for $\mu \in \mathscr{C}_k^\gamma$, $y, z \in \mathbb{R}^d$ and $0 \le s < t \le T$, let

$$
\begin{aligned}
H^{\mu,1}_{s,t}(y,z) = H^{\mu}_{s,t}(y,z) &:= \big\langle -b_s(y,\mu_s), \nabla p^{\mu,z}_{s,t}(\cdot,z)(y)\big\rangle \\
&\quad + \frac{1}{2}\mathrm{tr}\left[\{(\sigma_s\sigma_s^*)(z,\Gamma_s\mu) - (\sigma_s\sigma_s^*)(y,\Gamma_s\mu)\}\nabla^2 p^{\mu,z}_{s,t}(\cdot,z)(y)\right], \\
H^{\mu,j}_{s,t}(y,z) &:= \int_s^t \mathrm{d}r \int_{\mathbb{R}^d} H^{\mu,j-1}_{r,t}(z',z) H^{\mu}_{s,r}(y,z')\mathrm{d}z',\ \ j \ge 2.
\end{aligned} \tag{3.6}
$$

By the parabolic equations for the transition densities $p^\mu_{s,t}$ and $p^{\mu,z}_{s,t}$, see for instance the paragraph after Lemma 3.1 in [12], we have the parametrix expansion formula

$$
p^\mu_t(x,z) = p^{\mu,z}_{0,t}(x,z) + \sum_{j=1}^\infty \int_0^t \mathrm{d}s \int_{\mathbb{R}^d} H^{\mu,j}_{s,t}(y,z) p^{\mu,z}_{0,s}(x,y)\mathrm{d}y. \tag{3.7}
$$

Let

$$
\tilde{p}^K_{s,t}(x,y) = \frac{\exp[-\frac{1}{4K(t-s)}|y-x|^2]}{(4K\pi(t-s))^{\frac{d}{2}}},\quad x,y \in \mathbb{R}^d, 0 \le s < t \le T. \tag{3.8}
$$

By multiplying the time parameter with T^{-1} to make it stay in $[0,1]$, we deduce from [27, (2.3), (2.4)] with $\beta = \beta' = 1$ and $\lambda = \frac{1}{8KT}$ that

$$
\begin{aligned}
&\int_s^t \int_{\mathbb{R}^d} \tilde{p}^K_{s,r}(x,y')(r-s)^{-\frac{1}{2}} g_r(y')(t-r)^{-\frac{1}{2}} \tilde{p}^{2K}_{r,t}(y',y)\mathrm{d}y' \\
&\le c(t-s)^{-\frac{1}{2}+\frac{1}{2}(1-\frac{d}{p}-\frac{2}{q})} \tilde{p}^{2K}_{s,t}(x,y)\|g\|_{\tilde{L}^q_p([s,t])},\quad 0 \le s < t \le T, g \in \tilde{L}^q_p([s,t])
\end{aligned} \tag{3.9}
$$

holds for some constant $c > 0$ depending on T, d, p, q and K. By the condition on a included in (A_0), we find a constant $c_1 > 0$ such that (3.4) implies

$$
\begin{aligned}
&p^{\mu,z}_{s,t}(x,y)\left(1 + \frac{|x-y|^4}{(t-s)^2}\right) \\
&\le c_1 \tilde{p}^K_{s,t}(x,y),\quad x,y,z \in \mathbb{R}^d, 0 \le s < t \le T, \gamma \in \mathscr{P}_k, \mu \in C([s,t];\mathscr{P}_k).
\end{aligned} \tag{3.10}
$$

Lemma 3.1. *Assume (A_0) and (A_2). Let $p^{\mu,z}_{s,t}(x,y)$ be defined by (3.3) and (3.4) for some map $\Gamma : \mathscr{C}_k^\gamma \to \mathscr{C}_k^\gamma$. There exists a constant $c > 0$ independent of Γ, such that for any $0 \le s < t \le T, x, y, z \in \mathbb{R}^d, \gamma \in \mathscr{P}_k$, and $\mu, \nu \in C([s,t];\mathscr{P}_k)$,*

$$
\left(1 + \frac{|x-y|^2}{t-s}\right)|p^{\mu,z}_{s,t}(x,y) - p^{\nu,z}_{s,t}(x,y)| \le \frac{c\tilde{p}^K_{s,t}(x,y)}{t-s}\int_s^t \mathbb{W}_k(\Gamma_r\mu, \Gamma_r\nu)\mathrm{d}r, \tag{3.11}
$$

$$
\sqrt{t-s}|\nabla p^{\mu,z}_{s,t}(\cdot,y)(x)| + (t-s)\|\nabla^2 p^{\mu,z}_{s,t}(\cdot,y)(x)\| \le c\tilde{p}^K_{s,t}(x,y), \tag{3.12}
$$

$$
\begin{aligned}
&\sqrt{t-s}|\nabla p^{\mu,z}_{s,t}(\cdot,y)(x) - \nabla p^{\nu,z}_{s,t}(\cdot,y)(x)| \\
&+ (t-s)\|\nabla^2 p^{\mu,z}_{s,t}(\cdot,y)(x) - \nabla^2 p^{\nu,z}_{s,t}(\cdot,y)(x)\| \\
&\le \frac{c\tilde{p}^K_{s,t}(x,y)}{t-s}\int_s^t \mathbb{W}_k(\Gamma_r\mu, \Gamma_r\nu)\mathrm{d}r.
\end{aligned} \tag{3.13}
$$

Proof. (1) For fixed $x, y \in \mathbb{R}^d$ and $0 \le s < t \le T$, let

$$F(\mu) := \langle (a_{s,t}^{\mu,z})^{-1}(y-x), y-x\rangle, \quad \mu \in C([s,t];\mathscr{P}_k).$$

It is easy to see that

$$\begin{aligned} &|p_{s,t}^{\mu,z}(x,y) - p_{s,t}^{\nu,z}(x,y)| \\ &= \left| \frac{\exp[-\frac{1}{2}F(\mu)]}{(2\pi)^{\frac{d}{2}}(\det\{a_{s,t}^{\mu,z}\})^{\frac{1}{2}}} - \frac{\exp[-\frac{1}{2}F(\nu)]}{(2\pi)^{\frac{d}{2}}(\det\{a_{s,t}^{\nu,z}\})^{\frac{1}{2}}} \right| \le I_1 + I_2, \end{aligned} \tag{3.14}$$

where

$$\begin{aligned} I_1 &:= \frac{\big|\exp[-\frac{1}{2}F(\mu)] - \exp[-\frac{1}{2}F(\nu)]\big|}{(2\pi)^{\frac{d}{2}}(\det\{a_{s,t}^{\mu,z}\})^{\frac{1}{2}}} \\ I_2 &:= \frac{\exp[-\frac{1}{2}F(\nu)]}{(2\pi)^{\frac{d}{2}}} \Big|(\det\{a_{s,t}^{\mu,z}\})^{-\frac{1}{2}} - (\det\{a_{s,t}^{\nu,z}\})^{-\frac{1}{2}}\Big|. \end{aligned}$$

Combining this with (A_0) and (A_2) which imply (3.5), we find a constant $c_1 > 0$ such that

$$\begin{aligned} &|F(\mu) - F(\nu)| = \big|\langle\{(a_{s,t}^{\mu,z})^{-1} - (a_{s,t}^{\nu,z})^{-1}\}(y-x), y-x\rangle\big| \\ &\le c_1 \frac{|y-x|^2}{(t-s)^2} \int_s^t \mathbb{W}_k(\Gamma_r\mu, \Gamma_r\nu)\mathrm{d}r, \end{aligned}$$

which together with (3.10) and $\frac{|x-y|^2}{t-s} \le \frac{1}{2}(1 + \frac{|x-y|^4}{(t-s)^2})$ yields that for some constant $c_2 > 0$,

$$\left(1 + \frac{|x-y|^2}{t-s}\right) I_1 \le \frac{c_2 \tilde{p}_{s,t}^K(x,y)}{t-s} \int_s^t \mathbb{W}_k(\Gamma_r\mu, \Gamma_r\nu)\mathrm{d}r.$$

Again by (3.5), (3.10) and $\frac{|x-y|^2}{t-s} \le \frac{1}{2}(1 + \frac{|x-y|^4}{(t-s)^2})$, we find a constant $c_3 > 0$ such that

$$\left(1 + \frac{|x-y|^2}{t-s}\right) I_2 \le \frac{c_3 \tilde{p}_{s,t}^K(x,y)}{t-s} \int_s^t \mathbb{W}_k(\Gamma_r\mu, \Gamma_r\nu)\mathrm{d}r.$$

Combining these with (3.14), we arrive at

$$\left(1 + \frac{|x-y|^2}{t-s}\right) |p_{s,t}^{\mu,z}(x,y) - p_{s,t}^{\nu,z}(x,y)| \le \frac{(c_2 + c_3)\tilde{p}_{s,t}^K(x,y)}{t-s} \int_s^t \mathbb{W}_k(\Gamma_r\mu, \Gamma_r\nu)\mathrm{d}r.$$

(2) By (3.4) we have

$$\nabla p_{s,t}^{\mu,z}(\cdot,y)(x) = (a_{s,t}^{\mu,z})^{-1}(y-x) p_{s,t}^{\mu,z}(x,y), \tag{3.15}$$

$$\nabla^2 p_{s,t}^{\mu,z}(\cdot,y)(x) = p_{s,t}^{\mu,z}(x,y)\Big(\big\{(a_{s,t}^{\mu,z})^{-1}(y-x)\big\} \otimes \big\{(a_{s,t}^{\mu,z})^{-1}(y-x)\big\} - (a_{s,t}^{\mu,z})^{-1}\Big). \tag{3.16}$$

So, by (3.5) and (3.10) we find a constant $c > 0$ such that (3.12) holds. Moreover, (3.15) implies

$$\begin{aligned}&|\nabla p^{\mu,z}_{s,t}(\cdot,y)(x) - \nabla p^{\nu,z}_{s,t}(\cdot,y)(x)|\\&\le \left|\{(a^{\mu,z}_{s,t})^{-1} - (a^{\nu,z}_{s,t})^{-1}\}(y-x)\right| p^{\mu,z}_{s,t}(x,y)\\&\quad + \left|p^{\mu,z}_{s,t}(x,y) - p^{\nu,z}_{s,t}(x,y)\right| \cdot \left|(a^{\nu,z}_{s,t})^{-1}(y-x)\right|.\end{aligned}$$

Combining this with (3.5), (3.10) and (3.11), we find a constant $c > 0$ such that

$$|\nabla p^{\mu,z}_{s,t}(\cdot,y)(x) - \nabla p^{\nu,z}_{s,t}(\cdot,y)(x)| \le \frac{c\tilde{p}^K_{s,t}(x,y)}{(t-s)^{\frac{3}{2}}} \int_s^t \mathbb{W}_k(\Gamma_r\mu, \Gamma_r\nu)\mathrm{d}r.$$

Similarly, combining (3.16) with (3.5), (3.10) and (3.11), we find a constant $c > 0$ such that

$$\|\nabla^2 p^{\mu,z}_{s,t}(\cdot,y)(x) - \nabla^2 p^{\nu,z}_{s,t}(\cdot,y)(x)\| \le \frac{c\tilde{p}^K_{s,t}(x,y)}{(t-s)^2} \int_s^t \mathbb{W}_k(\Gamma_r\mu, \Gamma_r\nu)\mathrm{d}r.$$

Therefore, (3.13) holds for some constant $c > 0$. □

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

$$\Lambda_{s,t}(\mu,\nu) := \sup_{r\in[s,t]} \left\{\mathbb{W}_k(\Gamma_r\mu, \Gamma_r\nu) + \mathbb{W}_{k,\kappa var}(\mu_r, \nu_r)\right\}. \tag{3.17}$$

Lemma 3.2. *Assume (A_0) and (A_2). Let $\delta := \frac{1}{2}\left(1 - \frac{d}{p_0} - \frac{2}{q_0}\right) > 0$ and denote*

$$S_\mu := \sup_{t\in\in[0,T]} (1 + \|\mu_t\|_k), \quad S_{\mu,\nu} := S_\mu \vee S_\nu, \quad \nu,\mu \in \mathscr{C}^\gamma_k.$$

Then there exists a constant $C \ge 1$ such that for any $0 \le s < t \le T$, $y, z \in \mathbb{R}^d$, $\mu, \nu \in \mathscr{C}^\gamma_k$, and $j \ge 1$,

$$|H^{\mu,j}_{s,t}(y,z)| \le f_0(s,y)(CS_\mu)^j (t-s)^{-\frac{1}{2}+\delta(j-1)} \tilde{p}^{2K}_{s,t}(x,y), \tag{3.18}$$

$$\begin{aligned}&|H^{\mu,j}_{s,t}(y,z) - H^{\nu,j}_{s,t}(y,z)|\\&\le j f_0(s,y)(CS_{\mu,\nu})^j (t-s)^{-\frac{1}{2}+\delta(j-1)} \tilde{p}^{2K}_{s,t}(x,y)\Lambda_{s,t}(\mu,\nu).\end{aligned} \tag{3.19}$$

Proof. (1) By (3.6), (3.12), (A_0) and (A_2), we find a constant $c_1 > 0$ such that for any $0 \le s < t \le T, \mu \in C([0,T]; \mathscr{P}_k)$ and $y, z \in \mathbb{R}^d$,

$$|H^\mu_{s,t}(y,z)| \le c_1 (t-s)^{-\frac{1}{2}} \{(1 + \|\mu_s\|_k) f_0(s,y)\} \tilde{p}^K_{s,t}(y,z). \tag{3.20}$$

So, (3.18) holds for $j = 1$ and $C = c_1$. Thanks to [27, (2.3), (2.4)] with $\beta = \beta' = 1$, $\lambda = \frac{1}{8K}$, we have

$$I_j := \int_s^t \int_{\mathbb{R}^d} (t-u)^{-\frac{1}{2}} (t-u)^{\delta(j-1)} \tilde{p}^{2K}_{u,t}(y,z) f_0(u,y)(u-s)^{-\frac{1}{2}} \tilde{p}^K_{s,u}(x,y)\mathrm{d}y\mathrm{d}u$$

$$(3.21)\qquad \le c_2(t-s)^{-\frac{1}{2}}\tilde p^{2K}_{s,t}(x,z)(t-s)^{\frac{1}{2}(1-\frac{d}{p_0}-\frac{2}{q_0})}\|f_0\|_{\tilde L^{q_0}_{p_0}([s,t])}(t-s)^{\delta(j-1)}$$
$$= c_3(t-s)^{-\frac{1}{2}}\tilde p^{2K}_{s,t}(x,z)(t-s)^{\delta j}.\ \ 0\le s<t\le T, j\ge 1$$

where $c_3 := c_2\|f_0\|_{\tilde L^q_p([s,t])}$. Let $C := 1\vee c_1^2\vee(4c_3^2)$. If for some $j\ge 1$ we have

$$|H^{\mu,j}_{s,t}(y,z)|\le (CS_\mu)^j f_0(s,y)\tilde p^{2K}_{s,t}(y,z)(t-s)^{-\frac{1}{2}+\delta(j-1)}$$

for all $y,z\in\mathbb{R}^d$ and $0\le s<t\le T$, then by combining with (3.20) and (3.21), we arrive at

$$\begin{aligned}|H^{\mu,j+1}_{s,t}(y,z)| &\le \int_s^t \mathrm{d}u\int_{\mathbb{R}^d}|H^{\mu,j}_{u,t}(z',z)H^{\mu}_{s,u}(y,z')|\mathrm{d}z'\\ &\le C^j\sqrt{C}(S_\mu)^{j+1}f_0(s,y)I_k\\ &\le C^{j+1}(S_\mu)^{j+1}f_0(s,y)(t-s)^{-\frac{1}{2}+\delta j}\tilde p^{2K}_{s,t}(y,z).\end{aligned}$$

Therefore, (3.18) holds for all $j\ge 1$.

(2) By (3.12), (3.13), (3.5), (A_0) and (A_2), we find a constant $c>0$ such that for any $0\le s<t\le T,\mu,\nu\in C([0,T];\mathscr{P}_k)$ and $y,z\in\mathbb{R}^d$,

$$(3.22)\qquad |H^{\mu}_{s,t}(y,z)-H^{\nu}_{s,t}(y,z)|\le c(t-s)^{-\frac{1}{2}}\tilde p^{K}_{s,t}(y,z)S_{\mu,\nu}f_0(s,y)\Lambda_{s,t}(\mu,\nu).$$

Let, for instance, $L=1+4C^2+4c^2$, where C is in (3.18). If for some $j\ge 1$ we have

$$|H^{\mu,j}_{s,t}(z',z)-H^{\nu,j}_{s,t}(z',z)|\le j(LS_{\mu,\nu})^j f_0(s,z')\tilde p^{2K}_{s,t}(z',z)(t-s)^{-\frac{1}{2}+\delta(j-1)}\Lambda_{s,t}(\mu,\nu),$$

for any $0\le s<t\le T$ and $z,z'\in\mathbb{R}^d$, then (3.18), (3.21) and (3.22) imply

$$\begin{aligned}&|H^{\mu,j+1}_{s,t}(y,z)-H^{\nu,j+1}_{s,t}(y,z)|\\ &\le \int_s^t\mathrm{d}r\int_{\mathbb{R}^d}\Big\{|H^{\mu,j}_{r,t}(z',z)-H^{\nu,j}_{r,t}(z',z)|\cdot|H^{\mu}_{s,r}(y,z')|\\ &\qquad\qquad +|H^{\nu,j}_{r,t}(z',z)|\cdot|H^{\mu}_{s,r}(y,z')-H^{\nu}_{s,r}(y,z')|\Big\}\mathrm{d}z'\\ &\le (j+1)(LS_{\mu,\nu})^{j+1}f_0(s,y)\tilde p^{2K}_{s,t}(y,z)(t-s)^{-\frac{1}{2}+\delta j}\Lambda_{s,t}(\mu,\nu).\end{aligned}$$

Therefore, (3.19) holds for some constant $C>0$. □

We are now ready to prove the following main result in this part, which will be used to prove the contraction of Φ^γ on the path space over a small time interval. For $t_0\in(0,T]$, let

$$\mathscr{C}^{\gamma,N}_{k,t_0} := \big\{\mu\in C([0,t_0];\mathscr{P}_k):\ \mu_{\cdot\wedge t_0}\in\mathscr{C}^{\gamma,N}_k\big\},\quad N\ge N_0.$$

Lemma 3.3. *Assume (A_0) and (A_2). For any $N\ge N_0$, there exist $\theta_N>0, t_N\in(0,T]$ such that*

$$\mathbb{W}_{k,\kappa var,\theta_N}(\Phi^\gamma_{\cdot\wedge t_N}\mu,\Phi^\gamma_{\cdot\wedge t_N}\nu)\le\frac{1}{2}\mathbb{W}_{k,\kappa var,\theta_N}(\mu_{\cdot\wedge t_N},\nu_{\cdot\wedge t_N}),\quad \mu,\nu\in\mathscr{C}^{\gamma,N}_{k,t_N}.$$

Proof. By (3.10), Lemma 3.1, Lemma 3.2, (3.7), (3.9) and (A_2), we find constants $c_1, c_2, c_3 > 0$ such that for any $\theta > 0$ and $t_N \in (0, T \wedge (2CN)^{-\frac{1}{\delta}}]$,

$$\begin{aligned}
|p_t^\mu(x,z) - p_t^\nu(x,z)| &\le \frac{c_1 \tilde p_{0,t}^K(x,z)}{t} \int_0^t \mathbb{W}_k(\Gamma_s\mu, \Gamma_s\nu)\mathrm{d}s \\
&\quad + \sum_{n=1}^\infty \int_0^t \mathrm{d}s \int_{\mathbb{R}^d} \big\{|H_{s,t}^{\mu,n} - H_{s,t}^{\nu,n}|(y,z) p_{0,s}^{\nu,z}(x,y) + |H_{s,t}^{\mu,n}(y,z)||p_{0,s}^{\mu,z} - p_{0,s}^{\nu,z}|(x,y)\big\}\mathrm{d}y \\
&\le c_1 \mathrm{e}^{\theta t} \mathbb{W}_{k,\theta}(\Gamma_{\cdot\wedge t}\mu, \Gamma_{\cdot\wedge t}\nu) \tilde p_{0,t}^K(x,z) \\
&+ \sum_{n=1}^\infty (n+1)(CN)^n \Lambda_{0,t}(\mu,\nu) t^{\frac{1}{2}+\delta(n-1)} \\
&\quad \times \int_0^t \int_{\mathbb{R}^d} (t-r)^{-\frac{1}{2}} \tilde p_{r,t}^{2K}(y,z) f_0(r,y) r^{-\frac{1}{2}} \tilde p_{0,r}^K(x,y)\mathrm{d}y\mathrm{d}r \\
&\le c_1 \mathrm{e}^{t\theta} \mathbb{W}_{k,\theta}(\Gamma_{\cdot\wedge t}\mu, \Gamma_{\cdot\wedge t}\nu) \tilde p_{0,t}^K(x,z) + c_2 t^\delta \Lambda_{0,t}(\mu,\nu) \tilde p_{0,t}^{2K}(x,z) \sum_{n=1}^\infty (n+1)(CN)^n t^{\delta(n-1)} \\
&\le c_1 \mathrm{e}^{\theta t} \mathbb{W}_{k,\theta}(\Gamma_{\cdot\wedge t}\mu, \Gamma_{\cdot\wedge t}\nu) \tilde p_{0,t}^K(x,z) + c_3 t^\delta \Lambda_{0,t}(\mu,\nu) \tilde p_{0,t}^{2K}(x,z)
\end{aligned}$$

holds for any $x, z \in \mathbb{R}^d, t \in (0, t_N], \mu, \nu \in \mathscr{C}_k^{\gamma,N}$. Combining this with (3.8), we find a constant $c_4 > 0$ such that

$$\begin{aligned}
&\sup_{|g|\le 1+|\cdot|^k} \left| \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} g(z)(p_t^\mu - p_t^\nu)(x,z)\mathrm{d}z\gamma(\mathrm{d}x) \right| \\
&\le c_1 \mathrm{e}^{\theta t} \mathbb{W}_{k,\theta}(\Gamma_{\cdot\wedge t}\mu, \Gamma_{\cdot\wedge t}\nu) \int_{\mathbb{R}^d\times\mathbb{R}^d} (1+|z|^k) \tilde p_{0,t}^K(x,z)\mathrm{d}z\gamma(\mathrm{d}x) \\
&\quad + c_3 t^\delta \Lambda_{0,t}(\mu,\nu) \int_{\mathbb{R}^d\times\mathbb{R}^d} (1+|z|^k) \tilde p_{0,t}^{2K}(x,z)\mathrm{d}z\gamma(\mathrm{d}x) \\
&\le c_4 \mathrm{e}^{\theta t} \mathbb{W}_{k,\theta}(\Gamma_{\cdot\wedge t}\mu, \Gamma_{\cdot\wedge t}\nu) + c_4 t^\delta \Lambda_{0,t}(\mu,\nu), \quad t \in (0, t_N],\ \mu,\nu \in \mathscr{C}_k^{\gamma,N}.
\end{aligned} \tag{3.23}$$

Taking $\Gamma = \Phi^\gamma$, by the definition of Φ_t^γ, (3.23) and (3.17), we find a constant $c_5 > 0$ such that

$$\begin{aligned}
&\mathbb{W}_{k,\kappa var,\theta}(\Phi^\gamma_{\cdot\wedge t_N}\mu, \Phi^\gamma_{\cdot\wedge t_N}\nu) \\
&\le c_5 \mathbb{W}_{k,\theta}(\Phi^\gamma_{\cdot\wedge t_N}\mu, \Phi^\gamma_{\cdot\wedge t_N}\nu) + c_5 t_N^\delta \mathbb{W}_{k,\kappa var,\theta}(\mu_{\cdot\wedge t_N}, \nu_{\cdot\wedge t_N}), \quad \mu,\nu \in \mathscr{C}_k^{\gamma,N}.
\end{aligned}$$

By (2.6) with $\beta(\theta) \to 0$ as $\theta \to \infty$, we find large $\theta_N > 0$ and small $t_N \in (0, T]$ depending on N such that

$$\begin{aligned}
\mathbb{W}_{k,\kappa var,\theta_N}(\Phi^\gamma_{\cdot\wedge t_N}\mu, \Phi^\gamma_{\cdot\wedge t_N}\nu) &\le c_5(\beta(\theta_N) + t_N^\delta) \mathbb{W}_{k,\kappa var,\theta_N}(\mu_{\cdot\wedge t_N}, \nu_{\cdot\wedge t_N}) \\
&\le \frac{1}{2} \mathbb{W}_{k,\kappa var,\theta_N}(\mu_{\cdot\wedge t_N}, \nu_{\cdot\wedge t_N}).
\end{aligned}$$

□

3.2 Proof of Theorem 1.2

Estimate (1.5) is included in Lemma 2.2(2). It suffices to prove the well-posedness of (1.1) and estimate (1.8) for $\kappa = 0$, where $C(N)$ is bounded in N when $K_0 = 0$.

(a) Well-posedness. By the priori estimate (1.5), there exists a constant $C > 0$ such that for any solution of (1.1) on $[0, T]$ with $\mathscr{L}_{X_0} = \gamma$,

$$\sup_{t\in[0,T]} \mathscr{L}_{X_t}(|\cdot|^k) \le C. \tag{3.24}$$

So, we may fix $N_0 > 0$ depending only on C such that any solution of (1.1) with initial distribution γ satisfies $\mathscr{L}_{X_\cdot} \in \mathscr{C}_k^{\gamma,N_0}$. By Lemma 3.3, there exists $\theta > 0$ and $t_0 \in (0, T]$ depending only on N_0 such that the map $\Phi^\gamma_{\cdot\wedge t_0}$ is contractive in $\mathscr{C}^{\gamma,N_0}_{k,t_0}$ under the metric $\mathbb{W}_{k,\kappa var,\theta}$, and hence (1.1) for $t \in [0, t_0]$ is well-posed for distributions in $\mathscr{P}_k$ and (3.24) holds. Using (t_0, X_{t_0}) replacing $(0, X_0)$, the same argument implies the well-posedness of (1.1) for $t \in [t_0, (2t_0) \wedge T]$ and that (3.24) holds for $(2t_0) \wedge T$ replacing t_0. By repeating the procedure finitely many times, we prove the well-posedness of (1.1) for distributions in $\mathscr{P}_k$.

(b) Estimate (1.8). For any $\mu_0^i \in \mathscr{P}_k$ with $\mu_0^i(|\cdot|^k) \le N, i = 1, 2$, let

$$\mu_t^i = P_t^* \mu_0^i, \quad i = 1, 2, t \in [0, T].$$

By (1.5), there exists a constant $C_N > 0$ such that

$$\sup_{t\in[0,T]} (\mu_t^1 + \mu_t^2)(|\cdot|^k) \le C_N. \tag{3.25}$$

So, there exists a constant $\bar{N}$ depending on C_N such that

$$\mu_\cdot^i \in \mathscr{C}_k^{\gamma,\bar{N}}, \quad i = 1, 2.$$

Consider the SDEs

$$\mathrm{d}X_t^{x,i} = b_t(X_t^{x,i}, \mu_t^i) + \sigma_t(X_t^{x,i}, \mu_t^i)\mathrm{d}W_t, \quad X_0^{x,i} = x \in \mathbb{R}^d, t \in [0, T], i = 1, 2. \tag{3.26}$$

We have

$$\mu_t^i := P_t^* \mu_0^i = \int_{\mathbb{R}^d} \mathscr{L}_{X_t^{x,i}} \mu_0^i(\mathrm{d}x), \quad t \in [0, T], i = 1, 2, \tag{3.27}$$

According to [20, Theorem 2.1(2)], (3.25) and (A_0) imply

$$\|\mathscr{L}_{X_t^{x,i}} - \mathscr{L}_{X_t^{y,i}}\|_{var} \le \frac{c_1}{\sqrt{t}}|x - y|, \quad x, y \in \mathbb{R}^d, t \in (0, T], i = 1, 2$$

for some constant $c_1 > 0$ depending on N. Combining this with (3.27) gives

$$\left\|P_t^*\mu_0^1 - \int_{\mathbb{R}^d} \mathscr{L}_{X_t^{y,1}} \mu_0^2(\mathrm{d}y)\right\|_{var} = \left\|\int_{\mathbb{R}^d} \mathscr{L}_{X_t^{x,1}} \mu_0^1(\mathrm{d}x) - \int_{\mathbb{R}^d} \mathscr{L}_{X_t^{y,1}} \mu_0^2(\mathrm{d}y)\right\|_{var}$$

$$\text{(3.28)} \qquad \le \inf_{\pi\in\mathscr{C}(\mu_0^1,\mu_0^2)} \int_{\mathbb{R}^d\times\mathbb{R}^d} \|\mathscr{L}_{X_t^{x,1}} - \mathscr{L}_{X_t^{y,1}}\|_{var}\pi(\mathrm{d}x,\mathrm{d}y)$$
$$\le \frac{c_1}{\sqrt{t}}\mathbb{W}_1(\mu_0^1,\mu_0^2) \le \frac{c_1}{\sqrt{t}}\mathbb{W}_k(\mu_0^1,\mu_0^2), \quad t\in(0,T].$$

On the other hand, by (3.27) and (3.23) for $\mu=\mu^1$, $\nu=\mu^2$, $\kappa=0$ and $\Gamma=id$, we find constants $c_2>0$ and $t_N\in(0,T]$ depending on N such that

$$\left\| P_t^*\mu_0^2 - \int_{\mathbb{R}^d} \mathscr{L}_{X_t^{y,1}}\mu_0^2(\mathrm{d}y)\right\|_{var} \le c_2 \sup_{t\in[0,T]} \mathbb{W}_k(\mu_t^1,\mu_t^2), \quad t\in[0,t_N].$$

For any $t\in[t_N,T]$, repeating the above argument for the time interval $[t-t_N,t]$ replacing $[0,t_N]$ we prove

$$\left\| P_t^*\mu_0^2 - \int_{\mathbb{R}^d} \mathscr{L}_{X_t^{y,1}}\mu_0^2(\mathrm{d}y)\right\|_{var} \le c \sup_{t\in[0,T]} \mathbb{W}_k(\mu_t^1,\mu_t^2)$$

for some constant $c>0$ depending on N. Combining this with (3.28) and (1.7) which holds since (A_2) with $\kappa=0$ implies (A_1), we prove (1.8) for some constant $C(N)>0$.

Finally, noting that the dependence on N comes from Krylov's and Khasminskii's estimates for the solutions, and when $K_0=0$ we have $|b^{\mu,0}|\le f_0$ for all $\mu\in\mathscr{C}_k$, these estimates are uniform in μ. Thus, in this case (1.8) holds for all $\mu,\nu\in\mathscr{P}_k$ and a constant $C>0$ independent of N.

4 Extension of Theorem 1.1 to reflecting SDEs

Let $D\subset\mathbb{R}^d$ be a connected open domain with $\partial D\in C_b^{2,L}$ in the following sense: there exists a constant $r_0>0$ such that the polar coordinate map

$$\Psi:\partial D\times[-r_0,r_0]\ni(z,r)\mapsto z+r\mathbf{n}(z)\in\partial_{\pm r_0}D := \big\{x\in\mathbb{R}^d : \rho_\partial(x) := \mathrm{dist}(x,\partial D)\le r_0\big\}$$

is a C^2-diffeomorphism, such that $\Psi^{-1}(x)$ have bounded and continuous first and second order derivatives in $x\in\partial_{\pm r_0}D$, and $\nabla^2\rho_\partial$ is Lipschitz continuous on $\partial_{\pm r_0}D$.

Consider the following distribution dependent reflecting SDE on the closure $\bar D$ of D:

$$\text{(4.1)} \qquad \mathrm{d}X_t = b_t(X_t,\mathscr{L}_{X_t})\mathrm{d}t + \sigma_t(X_t,\mathscr{L}_{X_t})\mathrm{d}W_t + \mathbf{n}(X_t)\mathrm{d}l_t, \quad t\in[0,T],$$

where $\mathbf{n}$ is the unit inward normal vector field on the boundary ∂D and l_t is a continuous adapted increasing process with $\mathrm{d}l_t$ supported on $\{t: X_t\in\partial D\}$. Let $\tilde L_p^q(T)$ and $\mathscr{P}_k$ be defined as before for $\bar D$ replacing $\mathbb{R}^d$. When $\sigma_t(x,\mu)=\sigma_t(x)$ does not depend on μ, the well-posedness of (4.1) has been proved in [21] under the following assumption, where $a_t^\mu := (\sigma_t\sigma_t^*)(\cdot,\mu_t)$.

- (B) Assumptions (A_0) and (A_1) hold for $\bar D$ replacing $\mathbb{R}^d$. Moreover, there exists a constant $c>0$ such that for any $\mu\in C([0,T];\mathscr{P}_k)$, the Neumann semigroup $\{P_{s,t}^\mu\}_{0\le s\le t\le T}$ generated by the operator $L_t^\mu := \frac{1}{2}\mathrm{tr}\{a_t^\mu\nabla^2\} + b_t^{(1)}\cdot\nabla$ on $\bar D$ satisfies

$$\text{(4.2)} \qquad \|\nabla^i P_{s,t}^\mu\phi\|_\infty \le c(t-s)^{-\frac{i}{2}}\|\phi\|_\infty, \quad 0\le s<t\le T, \phi\in C_b^i(\bar D),\ i=1,2.$$

Theorem 4.1. *Assume* (B) *and let* $\partial D \in C_b^{2,L}$. *Then the assertions in Theorem 1.1 hold for* (4.1) *replacing* (1.1).

Proof. Let $\gamma \in \mathscr{P}_k$ and consider the initial value X_0 with $\mathscr{L}_{X_0} = \gamma$. It suffices to prove that Lemmas 2.1-2.3 hold for $\Phi^\gamma \mu := \mathscr{L}_{X^\mu}$ with the following reflecting SDE replacing (2.1):

$$\mathrm{d}X_t^\mu = b_t(X_t^\mu, \mu_t)\mathrm{d}t + \sigma_t(X_t^\mu, \mathscr{L}_{X_t^\mu})\mathrm{d}W_t + \mathbf{n}(X_t^\mu)\mathrm{d}l_t^\mu, \quad t \in [0,T], X_0^\mu = X_0. \tag{4.3}$$

(a) Assertions in Lemma 2.1. For $\gamma^i \in \mathscr{P}_k, \mu^i \in \mathscr{C}_k^{\gamma^i,N}$ and $\nu^i \in \mathscr{C}_k^{\gamma^i}$, $i = 1,2$, instead of (2.7) we consider the reflecting SDEs

$$\mathrm{d}X_t^i = b_t^{\mu^i}(X_t^i)\mathrm{d}t + \sigma_t^{\nu^i}(X_t^i)\mathrm{d}W_t + \mathbf{n}(X_t^i)\mathrm{d}l_t^i, \quad \mathscr{L}_{X_0^i} = \gamma^i, t \in [0,T], i = 1,2.$$

By [21, Theorem 2.2(ii)], (B) implies the well-posedness of these reflecting SDEs.

Next, according to the proof of [21, Theorem 2.2(ii)], there exists a semimartingale H_t such that

$$C^{-1}|X_t^1 - X_t^2|^2 \le H_t \le C|X_t^1 - X_t^2|^2, \quad t \in [0,T]$$

holds for some constant $C > 1$, and instead of (2.10),

$$\mathrm{d}H_t^j \le c_2 \eta_t^{2j} \mathrm{d}\{A_t + l_t^1 + l_t^2\} + c_2(g_t^2 + \tilde{g}_t)\big\{\mathbb{W}_k(\mu_t^1, \mu_t^2)^{2j} + \mathbb{W}_k(\nu_t^1, \nu_t^2)^{2j}\big\}\mathrm{d}t + \mathrm{d}M_t$$

holds for some constant $c_2 > 0$ and all $t \in [0,T]$.

Then the desired assertions can be proved as in the proof of Lemma 2.1 by using Khasminskii's estimate in [21, Lemma 2.7], as well as the estimate

$$\mathbb{E}\big[\mathrm{e}^{\lambda(l_T^1 + l_T^2)}\big] \le \mathrm{e}^{c(1+\lambda^2)}, \quad \lambda > 0 \tag{4.4}$$

for some constant $c > 0$ presented in [21, Lemma 2.5], where condition $(A_0^{a,b})$ follows from (A_0) included in (B), according to [21, Lemma 2.6].

(b) Proof of Lemma 2.2. In the present case (2.14) becomes

$$|X_t|^{2j} - |X_0|^{2j} \le c_1 \int_0^t \big\{1 + |X_s|^{2j} + |X_s|^{2j-1} f_0(s, X_s) + \|\mu_s\|_k^{2j}\big\}\mathrm{d}s + c_1 \int_0^t |X_s|^{2j-1}\mathrm{d}l_s + M_t,$$

such that (2.16) reduces to

$$\mathbb{E}\big(\eta_t\big|\mathscr{F}_0\big) \le c_5 + c_5|X_0|^{2j} + c_5 \int_0^t \big\{\mathbb{E}(\eta_s|\mathscr{F}_0) + \|\mu_s\|_k^{2j}\big\}\mathrm{d}s + c_5 \int_0^t \mathbb{E}(\eta_s|\mathscr{F}_0)\mathrm{d}l_s, \quad t \in [0,T].$$

Combining this with (4.4) for $l_T^1 = l_T$ and using Gronwall's inequality, we derive (2.17). Then the remainder of the proof is as same as in the proof of Lemma 2.2.

(c) Proof of Lemma 2.3. According to [21, Lemma 2.7], under (B) the estimate (2.20) holds for the solution to the reflecting SDE:

$$\mathrm{d}\bar{X}_t = b_t^{(1)}(\bar{X}_t)\mathrm{d}t + \sigma_t(\bar{X}_t, \Phi_t^\gamma \mu)\mathrm{d}W_t + \mathbf{n}(\bar{X}_t)\mathrm{d}l_t \quad \bar{X}_0 = X_0, t \in [0,T].$$

Then the desired assertion follows as in the original proof.

□

Acknowledgement. The authors would like to thank the referee and editors for helpful comments and corrections.

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,* Electron. J. Probab. 23(2018), 1-35.

[3] I. V. Bogachev, N. V. Krylov, M. Röckner, S. V. Shaposhnikov, *Fokker-Planck-Kolmogorov equations,* American Math. Soc. 2015.

[4] 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.

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

[6] 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.

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

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

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

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

[11] X. Huang, F.-Y. Wang, *Derivative estimates on distributions of McKean-Vlasov SDEs,* Electron. J. Probab. 26(2021), 1-12.

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

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

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

[15] P. Ren, *Singular McKean-Vlasov SDEs: well-posedness, regularities and Wang's Harnack inequality,* arXiv:2110.08846.

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

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

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

[19] F.-Y. Wang, *Exponential ergodicity for singular reflecting McKean-Vlasov SDEs,* arXiv:2108.03908.

[20] F.-Y. Wang, *Regularity estimates and intrinsic-Lions derivative formula for singular McKean-Vlasov SDEs,* arXiv:2109.02030.

[21] F.-Y. Wang, *Distribution dependent reflecting stochastic differential equations,* arXiv:2106.12737.

[22] P. Xia, L. Xie, X. Zhang, G. Zhao, *$L^q(L^p)$-theory of stochastic differential equations,* Stochastic Process. Appl. 130(2020), 5188-5211.

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

[24] S.-Q. Zhang, C. Yuan, *A Zvonkin's transformation for stochastic differential equations with singular drift and applications,* J. Differential Equations 297(2021), 277-319.

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

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

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