Path-Distribution Dependent SDEs: Well-Posedness and Asymptotic Log-Harnack Inequality*

Feng-Yu Wang[a)], **Chenggui Yuan**[b)], **Xiao-Yu Zhao**[a)]

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

wangfy@tju.edu.cn, zhxy_0628@tju.edu.cn

b)Department of Mathematics, Swansea University, Bay campus, SA1 8EN, UK

c.yuan@swansea.ac.uk

July 12, 2025

Abstract

We consider stochastic different equations on $\mathbb{R}^d$ with coefficients depending on the path and distribution for the whole history. Under a local integrability condition on the time-spatial singular drift, the well-posedness and Lipschitz continuity in initial values are proved, which is new even in the distribution independent case. Moreover, under a monotone condition, the asymptotic log-Harnack inequality is established, which extends the corresponding result of [5] derived in the distribution independent case.

Keywords: Path-distribution dependent SDEs, well-posedness, asymptotic log-Harnack inequality, gradient estimate.

1 Introduction

The dimension-free Harnack inequality with power was first introduced in [11] to study the log-Sobolev inequality on Riemanian manifolds, and has been intensively extended and applied to derive regularity estimates many for SDEs (stochastic differential equations), SPDEs (stochastic partial differential equations), path dependent SDEs, and distribution-dependent SDEs, see [13, 10, 9] and references therein. As a limit version when the power goes to infinite, the log-Harnack inequality has been introduced in [12] to characterize the curvature lower bounded and entropy-cost estimates, and has been extended to metric measure spaces [1, 8]. When the noise of a stochastic system is weak such that the log-Harnack inequality is not available, the asymptotic log-Harnack inequality has been studied in [5, 16] for path-dependent SDEs and SPDEs, which in particular implies an asymptotic gradient estimate.

In this paper, we study the well-posedness and asymptotic log-Harnack inequality for path-distribution dependent SDEs with infinite memory.

*Supported in part by the National Key R&D Program of China (No. 2022YFA1006000, 2020YFA0712900).

Let $(\mathbb{R}^d, |\cdot|)$ be the d-dimensional Euclidean space for some $d \in \mathbb{N}$. Denote by $\mathbb{R}^d \otimes \mathbb{R}^d$ the family of all $d \times d$-matrices with real entries, which is equipped with the operator norm $\|\cdot\|$. Let A^* denote the transpose of $A \in \mathbb{R}^d \otimes \mathbb{R}^d$, and let $\|\cdot\|_\infty$ be the uniform norm for functions taking values in $\mathbb{R}, \mathbb{R}^d$ or $\mathbb{R}^d \otimes \mathbb{R}^d$.

To describe the path dependence with exponential decay memory, let $\mathscr{C} := C((-\infty, 0]; \mathbb{R}^d)$ and for $\tau > 0$, set

$$\mathscr{C}_\tau = \left\{ \xi \in \mathscr{C} : \|\xi\|_\tau := \sup_{s \in (-\infty, 0]} (\mathrm{e}^{\tau s} |\xi(s)|) < \infty \right\}. \tag{1.1}$$

It is well known that $(C((-\infty, 0]; \mathbb{R}^d), \|\cdot\|_\infty)$ is complete but not separable, so is $(\mathscr{C}_\tau, \|\cdot\|_\tau)$ due to the isometric

$$\mathscr{C}_\tau \ni \xi := (\xi_s)_{s \in (-\infty, 0]} \mapsto \mathrm{e}^{\tau \cdot} \xi := (\mathrm{e}^{\tau s} \xi_s)_{s \in (-\infty, 0]} \in C((-\infty, 0]; \mathbb{R}^d).$$

Let $\mathscr{P}$ and $\mathscr{P}(\mathscr{C}_\tau)$ be the set of all probability measures on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ and $(\mathscr{C}_\tau, \mathscr{B}(\mathscr{C}_\tau))$, respectively, equipped with the weak topology. Let $\mathscr{B}_b(\mathscr{C}_\tau)$ be the class of bounded measurable functions on $\mathscr{C}_\tau$, and $\mathscr{B}_b^+(\mathscr{C}_\tau)$ the set of strictly positive functions in $\mathscr{B}_b(\mathscr{C}_\tau)$.

Let $(W(t))_{t \geq 0}$ be a d-dimensional Brownian motion defined on a complete filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t \geq 0}, \mathbb{P})$. For an $\mathscr{F}_0$-measurable random variable $X_0 := ((-\infty, 0] \ni r \mapsto X(r))$ on $\mathscr{C}_\tau$, we consider the following path-distribution dependent SDE with infinite memory:

$$\mathrm{d}X(t) = b(t, X_t, \mathscr{L}_{X_t})\mathrm{d}t + \sigma(t, X_t)\mathrm{d}W(t), \quad t \geq 0, \tag{1.2}$$

where for each fixed $t \geq 0$, $X_t(\cdot) \in \mathscr{C}_\tau$ is defined by

$$X_t(r) := X(t + r), \quad r \in (-\infty, 0],$$

which is called the segment process of $X(t)$, $\mathscr{L}_{X_t} \in \mathscr{P}(\mathscr{C}_\tau)$ is the distribution of X_t, and

$$b : \mathbb{R}^+ \times \mathscr{C}_\tau \times \mathscr{P}(\mathscr{C}_\tau) \to \mathbb{R}^d, \quad \sigma : \mathbb{R}^+ \times \mathscr{C}_\tau \to \mathbb{R}^d \otimes \mathbb{R}^d$$

are measurable. When different probability spaces are concerned, we use $\mathscr{L}_{X_t|\mathbb{P}}$ in place of $\mathscr{L}_{X_t}$ to emphasize the underline probability.

For any constant $k \geq 0$, let

$$\mathscr{P}_k(\mathscr{C}_\tau) := \left\{ \mu \in \mathscr{P}(\mathscr{C}_\tau) : \|\mu\|_k := \mu(\|\cdot\|_\tau^k)^{\frac{1}{k}} < \infty \right\},$$

where for $k = 0$ we set $\mu(\|\cdot\|_\tau^k)^{\frac{1}{k}} = 1$ such that $\mathscr{P}_0(\mathscr{C}_\tau) = \mathscr{P}(\mathscr{C}_\tau)$. When $k > 0$, $\mathscr{P}_k(\mathscr{C}_\tau)$ is a complete metric space under the L^k-Wasserstein distance,

$$\mathbb{W}_k(\mu, \nu) := \sup_{N \geq 1} \inf_{\pi \in C(\mu, \nu)} \left(\int_{\mathscr{C}_\tau \times \mathscr{C}_\tau} \|\xi - \eta\|_{N,\tau}^k \, \pi(\mathrm{d}\xi, \mathrm{d}\eta) \right)^{\frac{1}{1 \vee k}}, \quad \mu, \nu \in \mathscr{P}_k(\mathscr{C}_\tau),$$

where $C(\mu, \nu)$ is the set of all coupling of μ and ν, and

$$\|\xi\|_{N,\tau} := \sup_{s \in [-N, 0]} (\mathrm{e}^{\tau s} |\xi(s)|), \quad N \in \mathbb{N}.$$

To see this, let $\mathscr{P}_k(\mathscr{C}_{N,\tau})$ be the space of all probability measures on $\mathscr{C}_{N,\tau} := C([-N,0];\mathbb{R}^d)$ with finite k-moment of the uniform norm, let $\mu_N \in \mathscr{P}_k(\mathscr{C}_{N,\tau})$ be the marginal distribution on $\mathscr{C}_{N,\tau}$ of μ, and let

$$\mathbb{W}_k(\mu_N,\nu_N) := \inf_{\pi\in C(\mu_N,\nu_N)} \left(\int_{\mathscr{C}_{N,\tau}\times\mathscr{C}_{N,\tau}} \|\xi-\eta\|_{N,\tau}^k \,\pi(\mathrm{d}\xi,\mathrm{d}\eta)\right)^{\frac{1}{1\vee k}}.$$

Since $\mathscr{C}_{N,\tau}$ is Polish under the norm $\|\cdot\|_{N,\tau}$, $(\mathscr{P}_k(\mathscr{C}_{N,\tau}),\mathbb{W}_k)$ is a Polish space as well, and

$$\mathbb{W}_k(\mu_N,\nu_N) = \tilde{\mathbb{W}}_k(\mu_N,\nu_N) := \inf_{\pi\in C(\mu,\nu)} \left(\int_{\mathscr{C}_\tau\times\mathscr{C}_\tau} \|\xi-\eta\|_{N,\tau}^k \,\pi(\mathrm{d}\xi,\mathrm{d}\eta)\right)^{\frac{1}{1\vee k}}. \tag{1.3}$$

Since the marginal distribution on $\mathscr{C}_{N,\tau}\times\mathscr{C}_{N,\tau}$ of a coupling for μ and ν is a coupling of μ_N and ν_N, we have

$$\mathbb{W}_k(\mu_N,\nu_N) \le \tilde{\mathbb{W}}_k(\mu_N,\nu_N).$$

On the other hand, let $\pi_N \in \mathscr{C}(\mu_N,\nu_N)$ such that

$$\mathbb{W}_k(\mu_N,\nu_N) = \left(\int_{\mathscr{C}_{N,\tau}\times\mathscr{C}_{N,\tau}} \|\xi_N-\eta_N\|_{N,\tau}^k \,\pi_N(\mathrm{d}\xi_N,\mathrm{d}\eta_N)\right)^{\frac{1}{1\vee k}}.$$

Noting that $\mathscr{C}_\tau$ is separable under norm $\|\cdot\|_{\tau+1}$, the completeness $\overline{\mathscr{C}_\tau}$ of $\mathscr{C}_\tau$ under this norm becomes a Polish space. Since

$$\|\cdot\|_\tau = \lim_{N\to\infty} \|\cdot\|_{N,\tau}, \quad \tau\ge 0,$$

and $\|\cdot\|_{N,\tau}$ is continuous with respect to $\|\cdot\|_{\tau'}$ for any $\tau,\tau'\ge 0$, we conclude that $\|\cdot\|_\tau$ and $\|\cdot\|_{tau+1}$ induce the same Borel σ-field on $\overline{\mathscr{C}_\tau}$. So, by extending $\mu,\nu\in\mathscr{P}_k(\mathscr{C}_\tau)$ as probability measures on the Polish space $\overline{\mathscr{C}_\tau}$ such that $\overline{\mathscr{C}_\tau}\setminus\mathscr{C}_\tau$ is a null set, μ and ν have regular conditional distributions $\mu(\cdot|\xi_N)$ and $\nu(\cdot|\xi_N)$ on $C((-\infty,-N);\mathbb{R}^d)$ given $\xi_N\in\mathscr{C}_{N,\tau}$. So, for any $\xi\in\mathscr{C}_\tau$ letting

$$\xi_N := \xi|_{[-N,0]}, \quad \xi_N^c := \xi|_{(-\infty,-N)},$$

the measure

$$\pi(\mathrm{d}\xi,\mathrm{d}\eta) := \pi_N(\mathrm{d}\xi_N,\mathrm{d}\eta_N)\mu(\mathrm{d}\xi_N^c|\xi_N)\nu(\mathrm{d}\eta_N^c|\eta_N)$$

is a coupling of μ and ν, and

$$\begin{aligned}\tilde{\mathbb{W}}_k(\mu_N,\nu_N) &\le \left(\int_{\mathscr{C}_\tau\times\mathscr{C}_\tau} \|\xi-\eta\|_{N,\tau}^k \,\pi(\mathrm{d}\xi,\mathrm{d}\eta)\right)^{\frac{1}{1\vee k}} \\ &= \left(\int_{\mathscr{C}_\tau\times\mathscr{C}_\tau} \|\xi-\eta\|_{N,\tau}^k \,\pi_N(\mathrm{d}\xi,\mathrm{d}\eta)\right)^{\frac{1}{1\vee k}} = \mathbb{W}_k(\mu_N,\nu_N).\end{aligned}$$

Thus, (1.3) holds, so that $\mathbb{W}_k$ is a complete metric on $\mathscr{P}_k(\mathscr{C}_{N,\tau})$, which trivially implies the triangle inequality of $\mathbb{W}_k$ on $\mathscr{P}_k(\mathscr{C}_\tau)$. If $\{\mu^{(n)}\}_{n\ge1}$ is a $\mathbb{W}_k$-Cauchy sequence in $\mathscr{P}_k(\mathscr{C}_\tau)$, then so is $\{\mu_N^{(n)}\}_{n\ge1}$ for every $N\in\mathbb{N}$. Hence, $\mu_N^{(n)}$ has a unique limit μ_N in $\mathscr{P}_k(\mathscr{C}_{N,\tau})$ under $\mathbb{W}_k$, and the family $\{\mu_N\}_{N\geq 1}$ is consistent, so that by Kolmogorov's extension theorem, there exists a unique $\mu \in \mathscr{P}_k(\mathscr{C}_\tau)$ with $\{\mu_N\}_{N\geq 1}$ as marginal distributions, so that

$$\begin{aligned}&\lim_{n\to\infty} \mathbb{W}_k(\mu^{(n)},\mu) = \lim_{n\to\infty}\sup_{N\geq 1} \mathbb{W}_k(\mu_N^{(n)},\mu_N) = \lim_{n\to\infty}\sup_{N\geq 1}\lim_{m\to\infty} \mathbb{W}_k(\mu_N^{(n)},\mu_N^{(m)})\\ &\leq \lim_{n\to\infty}\lim_{m\to\infty}\sup_{N\geq 1} \mathbb{W}_k(\mu_N^{(n)},\mu_N^{(m)}) = \lim_{n\to\infty}\lim_{m\to\infty} \mathbb{W}_k(\mu^{(n)},\mu^{(m)}) = 0.\end{aligned}$$

Hence, $\mathscr{P}_k(\mathscr{C}_\tau)$ is complete under $\mathbb{W}_k$.

Moreover, for any $k \geq 0$, $\mathscr{P}_k(\mathscr{C}_\tau)$ is a complete (but not separable) metric space under the weighted variation distance

$$\|\mu-\nu\|_{k,var} := \sup_{f\in\mathscr{B}_b(\mathscr{C}_\tau),|f|\leq 1+\|\cdot\|_\tau^k} |\mu(f)-\nu(f)| = |\mu-\nu|(1+\|\cdot\|_\tau^k),$$

where $|\mu-\nu|$ is the total variation of $\mu-\nu$. According to [10, Remark 3.2.1], for any $k > 0$, there exists a constant $c > 0$ such that

$$\|\mu-\nu\|_{var} + \mathbb{W}_k(\mu,\nu)^{1\vee k} \leq c\|\mu-\nu\|_{k,var}, \quad \mu,\nu\in\mathscr{P}_k(\mathscr{C}_\tau). \tag{1.4}$$

Denote $\mathbb{W}_{k,var} = \mathbb{W}_k + \|\cdot\|_{k,var}$ for simplicity.

We will solve (1.2) for distribution $\mathscr{L}_{X_t}$ belonging to a subclass $\tilde{\mathscr{P}}_k(\mathscr{C}_\tau) \subset \mathscr{P}_k(\mathscr{C}_\tau)$ such that $\mathscr{L}_{X_t}$ is weakly continuous in $t \geq 0$.

Definition 1.1. (1) An adapted continuous process $(X_t)_{t\geq 0}$ on $\mathscr{C}_\tau$ is called a segment solution of (1.2) with initial value X_0, if X_0 is an $\mathscr{F}_0$-measurable random variable on $\mathscr{C}_\tau$, and $(X(t) := X_t(0))_{t\geq 0}$ satisfies

$$\int_0^t \mathbb{E}\big[|b(r,X_r,\mathscr{L}_{X_r})| + \|\sigma(r,X_r)\|^2\big|\mathscr{F}_0\big]\,\mathrm{d}r < \infty, \quad t\geq 0,$$

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

$$X(t) = X(0) + \int_0^t b(r,X_r,\mathscr{L}_{X_r})\,\mathrm{d}r + \int_0^t \sigma(r,X_r)\,\mathrm{d}W(r), \quad t\geq 0.$$

In this case $(X(t))_{t\geq 0}$ is called the (strong) solution. The SDE (1.2) is called strongly well-posed for distributions in $\tilde{\mathscr{P}}_k(\mathscr{C}_\tau)$, if for any $\mathscr{F}_0$-measurable X_0 with $\mathscr{L}_{X_0} \in \tilde{\mathscr{P}}_k(\mathscr{C}_\tau)$, it has a unique segment solution with $\mathscr{L}_{X_t} \in \tilde{\mathscr{P}}_k(\mathscr{C}_\tau)$ for $t \geq 0$.

(2) A couple $(X_t, W(t))_{t\geq 0}$ is called a weak segment solution of (1.2) with initial distribution $\mu \in \tilde{\mathscr{P}}_k(\mathscr{C}_\tau)$, if there exists a probability space under which $W(t)$ is d-dimensional Brownian motion and $\mathscr{L}_{X_0} = \mu$, such that $(X(t) := X_t(0))_{t\geq 0}$ solves the SDE (1.2). We call (1.2) weakly unique, if for any two weak segment solutions $(X_t^i, W^i(t))_{t\geq 0}$ under probabilities $\mathbb{P}^i$ with common initial distribution, we have $\mathscr{L}_{X_t^1|\mathbb{P}^1} = \mathscr{L}_{X_t^2|\mathbb{P}^2}$ for all $t \geq 0$. We call (1.2) weakly well-posed for distributions in $\tilde{\mathscr{P}}_k(\mathscr{C}_\tau)$, if for any initial distribution it has a unique weak segment solution.

(3) The SDE (1.2) is called well-posed for distributions in $\tilde{\mathscr{P}}_k(\mathscr{C}_\tau)$, if it is both strongly and weakly well-posed for distributions in $\tilde{\mathscr{P}}_k(\mathscr{C}_\tau)$. In this case, for any $\xi \in \mathscr{C}_\tau$ such that $\mathscr{L}_\xi \in \tilde{\mathscr{P}}_k(\mathscr{C}_\tau)$, let

$$P_t^*\gamma = \mathscr{L}_{X_t^\xi}, \quad \gamma = \mathscr{L}_\xi$$

and denote

$$P_t f(\gamma) := \mathbb{E}[f(X_t^\xi)] = \int_{\mathscr{C}_\tau} f \, \mathrm{d}\{P_t^*\gamma\}, \quad \gamma = \mathscr{L}_\xi, \, t \geq 0, \, f \in \mathscr{B}_b(\mathscr{C}_\tau),$$

where $(X_t^\xi)_{t\geq 0}$ is the unique solution to (1.2) with $X_0^\xi = \xi$.

To characterize the singularity of coefficients in time–space variables, we recall some functional spaces introduced in [14]. For any $p \geq 1$, $L^p(\mathbb{R}^d)$ is the class of measurable functions f on $\mathbb{R}^d$ such that

$$\|f\|_{L^p(\mathbb{R}^d)} := \left(\int_{\mathbb{R}^d} |f(x)|^p \, \mathrm{d}x \right)^{\frac{1}{p}} < \infty.$$

For any $p, q \geq 1$ and $0 \leq s < t$, let $\tilde{L}_q^p(s,t)$ denote the class of measurable functions f on $[s,t] \times \mathbb{R}^d$ such that

$$\|f\|_{\tilde{L}_q^p(s,t)} := \sup_{z\in\mathbb{R}^d} \left(\int_s^t \|1_{B(z,1)} f_r\|_{L^p(\mathbb{R}^d)}^q \, \mathrm{d}r \right)^{\frac{1}{q}} < \infty, \tag{1.5}$$

where $B(z,1) := \{x \in \mathbb{R}^d : |x - z| \leq 1\}$.

When $s = 0$, we simply denote

$$\tilde{L}_q^p(t) = \tilde{L}_q^p(0,t), \quad t > 0.$$

We will take (p,q) from the class

$$\mathscr{K} := \Big\{ (p,q) : p, q \in (2, \infty), \ \frac{d}{p} + \frac{2}{q} < 1 \Big\}. \tag{1.6}$$

2 The Singular Case: Well-Posedness and Lipschitz Continuity in Initial Value

In this section, we let $k \geq 0$ and consider (1.2) with singular drifts and $\sigma(t,\xi) = \sigma(t, \xi(0))$.

2.1 Path Dependent SDEs with Infinite Memory

In this part, we consider the following path dependent SDE with infinite memory on $\mathbb{R}^d$:

$$\mathrm{d}X(t) = b(t, X_t)\mathrm{d}t + \sigma(t, X(t))\mathrm{d}W(t), \quad t > 0, \quad X_0 = \xi \in \mathscr{C}_\tau. \tag{2.1}$$

To ensure the existence and uniqueness of solutions to (2.1), we decompose b as

$$b(t,\xi) = b^{(0)}(t, \xi(0)) + b^{(1)}(t,\xi), \quad t \geq 0, \ \xi \in \mathscr{C}_\tau$$

and make the following assumptions on $b^{(0)}, b^{(1)}$ and σ. For any $\xi \in \mathscr{C}_\tau$, let $\xi^0 \in \mathscr{C}_\tau$ be defined as

$$\xi^0(r) = \xi(0), \quad r \le 0.$$

(A_1) $a := \sigma\sigma^*$ is invertible with $\|a\|_\infty + \|(a)^{-1}\|_\infty < K$ for some constant $K > 0$ and

$$\lim_{\varepsilon \downarrow 0} \sup_{|x-y|\le\varepsilon, t\in[0,T]} \|a_t(x) - a_t(y)\| = 0, \quad T \in (0,\infty),\ x, y \in \mathbb{R}^d.$$

(A_2) There exist constants $\{(p_i, q_i)\}_{0\le i\le l} \in \mathscr{K}$ with $l \ge 1$, $p_i > 2$, and functions $0 \le f_i \in \cap_{n\in\mathbb{N}} \tilde{L}_{q_i}^{p_i}(n)$, $0 \le i \le l$, such that

$$|b^{(0)}| \le f_0, \quad \|\nabla\sigma\| \le \sum_{i=0}^{l} f_i.$$

(A_3) For every $n > 0$, there exists a constant $K_n > 0$ such that

$$|b^{(1)}(t,\xi) - b^{(1)}(t,\eta)| \le K_n \|\xi - \eta\|_\tau, \quad t \ge 0, \|\xi\|_\tau, \|\eta\|_\tau \le n. \tag{2.2}$$

Moreover, there exists a constant $K > 0$ such that

$$|b^{(1)}(t,\xi) - b^{(1)}(t,\xi^0)| \le K(1 + \|\xi\|_\tau), \quad t \ge 0, \xi \in \mathscr{C}_\tau. \tag{2.3}$$

Theorem 2.1. *Assume (A_1)–(A_3). Then* (2.1) *is well-posed for any initial value in $\mathscr{C}_\tau$, and for any constants $k, T > 0$, there exists a constant $c > 0$, such that*

$$\mathbb{E}\left[\sup_{t\in[0,T]} \left(1 + \|X_t\|_\tau^k\right) \middle| X_0\right] \le c\left(1 + \|X_0\|_\tau^k\right). \tag{2.4}$$

Let X_t^ξ be the segment solution with $X_0^\xi = \xi$. To ensure the Lipschitz continuity of X_t^ξ in ξ, we strengthen (A_3) to

(A_3') $\sup_{t\ge0} |b^{(1)}(t,\mathbf{0})| < \infty$, and there exist constants $K > 0$ and $\alpha \in [0,1]$ such that

$$|b^{(1)}(t,\xi) - b^{(1)}(t,\eta)| \le K\|\xi - \eta\|_\tau,$$
$$|b^{(1)}(t,\xi) - b^{(1)}(t,\xi^0)| \le K\left(1 + \|\xi\|_\tau^\alpha\right), \quad t \ge 0,\ \xi, \eta \in \mathscr{C}_\tau.$$

Theorem 2.2. *Assume (A_1), (A_2) and (A_3'). Then for any constants $\varepsilon \in (0,1)$ and $k, T \ge 1$, there exists a constant $c > 0$ such that*

$$\mathbb{E}\left[\sup_{t\in[0,T]} \|X_t^\xi - X_t^\eta\|_\tau^k\right] \le c e^{\varepsilon\left(\|\xi\|_\tau^{2\alpha} + \|\eta\|_\tau^{2\alpha}\right)} \|\xi - \eta\|_\tau^k, \quad \xi, \eta \in \mathscr{C}_\tau. \tag{2.5}$$

2.2 Path-Distribution Dependent SDEs with Infinite Memory

To ensure the well-posedness of solutions to (1.2) with singular drift for distributions in $\mathscr{P}_k(\mathscr{C}_\tau)$, we decompose b as

$$b(t,\xi,\mu) = b^{(0)}(t,\xi(0)) + b^{(1)}(t,\xi,\mu), \quad t \geq 0,\ \xi \in \mathscr{C}_\tau,\ \mu \in \mathscr{P}_k(\mathscr{C}_\tau),$$

and impose the following assumptions on $b^{(0)}, b^{(1)}$ and σ.

(H_1) $b^{(0)}$ and $a := \sigma\sigma^*$ satisfy (A_1) and (A_2).

(H_2) $\sup_{t\geq 0} |b^{(1)}(t,\mathbf{0},\delta_{\mathbf{0}})| < \infty$, and there exists a function $H \in L^1_{loc}([0,\infty);(0,\infty))$ and constants $K > 0$, $\alpha \in [0,1]$ such that

$$|b^{(1)}(t,\xi,\mu) - b^{(1)}(t,\eta,\nu)|^2 \leq K\|\xi-\eta\|_\tau^2 + H(t)\mathbb{W}_{k,var}(\mu,\nu)^2, \tag{2.6}$$

$$\begin{aligned}&|b^{(1)}(t,\xi,\mu) - b^{(1)}(t,\xi^0,\mu)| \leq K\left\{1 + \|\xi\|_\tau^\alpha + \|\mu\|_k\right\},\\ &\qquad t \geq 0,\ \xi,\eta \in \mathscr{C}_\tau,\ \mu,\nu \in \mathscr{P}(\mathscr{C}_\tau).\end{aligned} \tag{2.7}$$

Theorem 2.3. *Assume* (H_1)-(H_2). *Then the SDE* (1.2) *is well-posed for distributions in* $\mathscr{P}^\alpha_{k,\mathrm{e}}(\mathscr{C}_\tau)$, *where*

$$\mathscr{P}^\alpha_{k,\mathrm{e}}(\mathscr{C}_\tau) := \left\{\mu \in \mathscr{P}_k(\mathscr{C}_\tau) : \mu\left(\mathrm{e}^{\varepsilon\|\cdot\|_\tau^{2\alpha}}\right) < \infty \text{ for some } \varepsilon \in (0,1)\right\}. \tag{2.8}$$

Moreover, for any $n \geq 1$ *and* $T \in (0,\infty)$, *there exists a constant* $c > 0$ *such that any solution to* (1.2) *satisfies*

$$\mathbb{E}\left[\sup_{t\in[0,T]} \|X_t\|_\tau^n \middle| X_0\right] \leq c\left(1 + \|X_0\|_\tau^n\right). \tag{2.9}$$

3 The Monotone Case: Well-Posedness and Asymptotic log-Harnack inequality

In this section, we consider (1.2) with monotone coefficients and establish the well-posedness and asymptotic log-Harnack inequality.

3.1 Path Dependent SDEs with Infinite Memory

Note that when the SDE

$$\mathrm{d}X(t) = b(t,X_t)\mathrm{d}t + \sigma(t,X_t)\mathrm{d}W(t), \quad t > 0, \quad X_0 = \xi \in \mathscr{C}_\tau. \tag{3.1}$$

is time-homogenous and monotone, the asymptotic log-Harnack inequality has been derived in [5, Theorem 3.1] for the Markov semigroup

$$P_t f(\xi) := \mathbb{E}\big[f(X_t^\xi)\big], \qquad t \geq 0,\ \xi \in \mathscr{C}_\tau,\ f \in \mathscr{B}_b(\mathscr{C}_\tau).$$

We call μ an invariant probability measure of P_t, if it is a probability measure on $\mathscr{C}_\tau$ such that

$$\int_{\mathscr{C}_\tau} P_t f \, \mathrm{d}\mu = \int_{\mathscr{C}_\tau} f \, \mathrm{d}\mu, \quad f \in \mathscr{B}_b(\mathscr{C}_\tau).$$

An important application of this inequality is the gradient estimate. For any function f on $\mathscr{C}_\tau$, let

$$\|\nabla f(\xi)\|_\infty := \sup_{\eta \in \mathscr{C}_\tau \setminus \{\xi\}} \frac{|f(\xi) - f(\eta)|}{\|\xi - \eta\|_\tau}$$

be the Lipschitz constant of a function f at point $\xi \in \mathscr{C}_\tau$, which coincides with the norm of the gradient $\nabla f(\xi)$ if f is Gâteaux differentiable at ξ. We denote $f \in C_{b,L}(\mathscr{C}_\tau)$ if

$$\|f\|_\infty + \|\nabla f\|_\infty < \infty.$$

Since the proof of [5, Theorem 3.1 and Theorem A.1] apply also to the time-inhomogenous case which is crucial in the study of distribution dependent SDEs, we reformulate this result and its consequence for (3.1) with time dependent coefficients without proof.

(*B*) $b(t, \cdot) \in C(\mathscr{C}_\tau; \mathbb{R}^d)$ for each $t \geq 0$, σ is invertible, and there exists a constant $K > 0$ such that

$$\begin{aligned} &\|\sigma(t,\xi) - \sigma(t,\eta)\|^2 + \langle b(t,\xi) - b(t,\eta), \xi(0) - \eta(0)\rangle^+ \leq K\|\xi - \eta\|_\tau^2, \\ &\|\sigma(t,\xi)\| + \|\sigma(t,\xi)^{-1}\| + |b(t,\mathbf{0})| \leq K, \quad t \geq 0, \ \xi, \eta \in \mathscr{C}_\tau, \end{aligned}$$

where $\mathbf{0} \in \mathscr{C}_\tau$ with $\mathbf{0}(r) = 0, r \leq 0$.

Theorem 3.1. *Under assumption* (*B*), *for any* $\xi \in \mathscr{C}_\tau$, *the SDE* (3.1) *has a unique solution with* $X_0 = \xi$ *such that* X_t^ξ *is a Markov process on* $\mathscr{C}_\tau$. *Moreover, for any* $\tau_0 \in (0, \tau)$, *there exists a constant* $c > 0$ *such that for any* $t \geq 0$, $\xi, \eta \in \mathscr{C}_\tau$ *and* $f \in \mathscr{B}_b^+(\mathscr{C}_\tau)$ *with* $\|\nabla \log f\|_\infty < \infty$,

$$\begin{aligned} &\mathbb{E}\big[\|X_t^\xi\|_\tau^2\big] \leq c\mathrm{e}^{ct}\big(1 + \|\xi\|_\tau^2\big), \\ &P_t \log f(\eta) \leq \log P_t f(\xi) + c\|\xi - \eta\|_\tau^2 + c\mathrm{e}^{-\tau_0 t}\|\nabla \log f\|_\infty \|\xi - \eta\|_\tau. \end{aligned} \tag{3.2}$$

The following result is a direct consequence of Theorem 3.1 and [5, Theorem 2.1] with

$$E = \mathscr{C}_\tau, \quad \rho(\xi, \eta) = \|\xi - \eta\|_\tau, \quad \Gamma_t(\xi) = c\mathrm{e}^{-\tau_0 t}, \quad \Lambda(\xi) = c. \tag{3.3}$$

Corollary 3.2. *In the situation of Theorem* 3.1, *the following assertions hold.*

(1) *For any* $t \geq 0, \xi \in \mathscr{C}_\tau$ *and* $f \in C_{b,L}(\mathscr{C}_\tau)$,

$$|\nabla P_t f(\xi)| \leq \sqrt{2c[P_t f^2 - (P_t f)^2]}(\xi) + c\mathrm{e}^{-\tau_0 t}\|\nabla f\|_\infty.$$

(2) *When the coefficients do not depend on* t, P_t *has at most one invariant probability measure, and if* μ *is its invariant probability measure, then*

$$\limsup_{t\to\infty} P_t f(\xi) \leq \log\left(\frac{\mu(\mathrm{e}^f)}{\int_{\mathscr{C}_\tau} \mathrm{e}^{-c\|\xi-\eta\|_\tau^2}\mu(\mathrm{d}\eta)}\right), \quad \xi \in \mathscr{C}_\tau, \ f \in \mathscr{B}_b^+(\mathscr{C}_\tau).$$

Consequently, for any closed set $A \subset \mathscr{C}_\tau$ *with* $\mu(A) = 0$,

$$\lim_{t\to\infty} P_t \mathbf{1}_A(\xi) = 0. \ \ \xi \in \mathscr{C}_\tau.$$

(3) *Let* $\xi \in \mathscr{C}_\tau$ *and* $A \subset \mathscr{C}_\tau$ *be a measurable set such that*

$$\delta(\xi, A) := \liminf_{t\to\infty} P_t(\xi, A) > 0.$$

Then for any $\varepsilon > 0$ *and* $A_\varepsilon := \{\eta \in \mathscr{C}_\tau : \|\eta - \xi\| < \varepsilon, \xi \in A\}$,

$$\liminf_{t\to\infty} P_t(\eta, A_\varepsilon) > 0, \quad \eta \in \mathscr{C}_\tau.$$

Moreover precisely, for any $\varepsilon_0 \in (0, \delta(\xi, A))$*, there exists a constant* $t_0 > 0$ *such that for any* $\eta \in \mathscr{C}_\tau$ *and* $\varepsilon > 0$,

$$\inf\left\{P_t(\eta, A_\varepsilon) : \ t > t_0 \vee \Big(\frac{1}{\tau_0}\log\frac{\Lambda(\eta)\|\xi-\eta\|_\tau}{\varepsilon\varepsilon_0}\Big)\right\} > 0, \quad \eta \in \mathscr{C}_\tau.$$

3.2 Path-Distribution Dependent SDEs with Infinite Memory

we assume that the following monotone assumption holds, as in [5] in the distribution-free case.

(H') $\sup_{t\ge0}|b(t,\mathbf{0},\delta_{\mathbf{0}})| < \infty$, and $b(t,\cdot,\mu) \in C(\mathscr{C}_\tau;\mathbb{R}^d)$ holds for any $t \ge 0$, $\mu \in \mathscr{P}(\mathscr{C}_\tau)$. Moreover, there exist constants $K_1, K_2 > 0$ such that and

$$\begin{aligned}&\langle b(t,\xi,\mu) - b(t,\eta,\nu), \xi(0) - \eta(0)\rangle^+ + \|\sigma(t,\xi) - \sigma(t,\eta)\|^2\\ &\le K_1\|\xi-\eta\|_\tau^2 + K_2\mathbb{W}_2(\mu,\nu)^2, \quad t \ge 0,\ \xi,\eta \in \mathscr{C}_\tau,\ \mu,\nu \in \mathscr{P}_2(\mathscr{C}_\tau).\end{aligned}$$

Theorem 3.3. *Assume* (H')*. Then the following assertions hold.*

(1) *The SDE* (1.2) *is well-posed for distributions in* $\mathscr{P}_2(\mathscr{C}_\tau)$*, and there exists a constant* $c > 0$ *such that*
$$\mathbb{W}_2(P_t^*\mu_0, P_t^*\nu_0) \le c\mathrm{e}^{ct}\mathbb{W}_2(\mu_0,\nu_0), \quad t \ge 0,\ \mu_0,\nu_0 \in \mathscr{P}_2(\mathscr{C}_\tau). \tag{3.4}$$

(2) *If* σ *is invertible with* $\|\sigma\|_\infty + \|\sigma^{-1}\|_\infty < \infty$*, then for any* $\tau_0 \in (0,\tau)$*, there exists a constant* $c \ge 1$ *such that*

$$P_t \log f(\nu) \le \log P_t f(\mu) + cK_2\mathrm{e}^{ct}\mathbb{W}_2(\mu,\nu)^2 + c\mathrm{e}^{-\tau_0 t}\|\nabla \log f\|_\infty \mathbb{W}_2(\mu,\nu) \tag{3.5}$$

holds for $t \ge 0$, $\mu,\nu \in \mathscr{P}_2(\mathscr{C}_\tau)$ *and* $f \in \mathscr{B}_b^+(\mathscr{C}_\tau)$ *with* $\|\nabla \log f\|_\infty < \infty$*. Consequently, for any* $t > 0$ *and* $f \in C_{b,L}(\mathscr{C}_\tau)$,

$$|P_t f(\mu) - P_t f(\nu)| \le \mathbb{W}_2(\mu,\nu)\big[c\mathrm{e}^{-\tau_0 t}\|\nabla f\|_\infty + 2\sqrt{cK_2\mathrm{e}^{ct}}\|f\|_\infty\big]. \tag{3.6}$$

Remark 3.4. *Note that* (3.5) *implies* (3.2) *for* $\mu = \delta_\xi$, $\nu = \delta_\eta$ *and* $K_2 = 0$.

4 Proof of Theorem 2.1 Theorem 2.2

To prove these two theorems , we present some lemmas below.

Lemma 4.1. *Assume* (A_1), (A_2), (A_3') *and* $\|b^{(1)}\|_\infty < \infty$. *Then for any initial value* $\xi \in \mathscr{C}_\tau$, (2.1) *has a unique non-explosive strong solution satisfying* (2.4), *and for any* $T \in (0, \infty)$ *there exists a constant* $c > 0$ *such that*

$$\mathbb{E}\left[\sup_{t\in[0,T]} \left|X^\xi(t) - X^\eta(t)\right|\right] \le c\|\xi - \eta\|_\tau, \quad \xi, \eta \in \mathscr{C}_\tau. \tag{4.1}$$

Proof. The desired estimate follows from the proof of [18, Lemma 4.1] with $\|\cdot\|_\tau$ in place of uniform norm $\|\cdot\|_{\mathscr{C}}$ on the path space with finite time interval. □

Next, consider the local Hardy-Littlewood maximal function for a nonnegative function f on $\mathbb{R}^d$:

$$\mathscr{M}f(x) := \sup_{r\in(0,1)} \frac{1}{|B(0,r)|} \int_{B(0,r)} f(x+y)\,\mathrm{d}y, \quad x \in \mathbb{R}^d, \tag{4.2}$$

where $|B(0,r)|$ is the volume of $B(0,r) := \{y : |y| < r\}$. The following result is taken from [15, Lemma 2.1].

Lemma 4.2. (1) *There exists a constant* $c > 0$ *such that for any* $f \in C_b(\mathbb{R}^d)$ *with* $|\nabla f| \in L^1_{loc}(\mathbb{R}^d)$,

$$|f(x) - f(y)| \le c|x-y|(\mathscr{M}|\nabla f|(x) + \mathscr{M}|\nabla f|(y) + \|f\|_\infty), \quad x, y \in \mathbb{R}^d. \tag{4.3}$$

(2) *For any* $T, p, q \in (1, \infty)$, *there exists a constant* $c > 0$ *such that*

$$\|\mathscr{M}f\|_{\tilde{L}^p_q(T)} \le c\|f\|_{\tilde{L}^p_q(T)}, \quad f \in \tilde{L}^p_q(T). \tag{4.4}$$

The lemma below examines the exponential integrability of functionals for segment process.

Lemma 4.3. *Assume* (A_1), (A_2) *and* (A_3'). *For any* $T \in (1, \infty)$, *there exist constants* $\beta_T, k_T > 0$ *such that for any solution* $(X_t)_{t\in[0,T]}$ *to* (2.1),

$$\mathbb{E}\left[\exp\left(\beta \int_s^t \|X_u\|_\tau^{2\alpha}\,\mathrm{d}u\right)\middle|\mathscr{F}_s\right] \le \mathrm{e}^{k_T\beta\left(1+\|X_s\|_\tau^{2\alpha}\right)} \tag{4.5}$$

holds for any $0 \le \beta \le \beta_T$ *and* $0 \le s \le t \le T$.

Proof. Obviously, (4.5) holds for $\alpha = 0$. By shifting the starting time from 0 to s, we may and do assume that $s = 0$. So, by Jensen's inequality, it suffices to find constants $\beta_T, k_T > 0$ such that

$$\mathbb{E}\left[\exp\left(\beta_T \int_0^T \|X_u\|_\tau^{2\alpha}\,\mathrm{d}u\right)\middle|\mathscr{F}_0\right] \le \mathrm{e}^{k_T\beta_T\left(1+\|X_0\|_\tau^{2\alpha}\right)}. \tag{4.6}$$

By [14, Theorem 3.2] and [17, Theorem 2.1], there exists a constant $\lambda_0 > 0$ such that for any $\lambda \geq \lambda_0$, the PDE for $u(t, \cdot) : \mathbb{R}^d \to \mathbb{R}^d$

$$(\partial_t + L^0(t))u(t, \cdot) = \lambda u(t, \cdot) - b^{(0)}(t, \cdot), \quad t \in [0, T], \quad u_T = 0 \tag{4.7}$$

has a unique solution in $\tilde{H}^{2,p_0}_{q_0}(T)$, where $L^0(t) := \nabla_{b^{(0)}(t,\cdot)} + \frac{1}{2}\mathrm{tr}(\{\sigma\sigma^*\}(t, \cdot)\nabla^2)$, and there exist constants $c, \theta > 0$ such that

$$\lambda^\theta(\|u\|_\infty + \|\nabla u\|_\infty) + \|\partial_t u\|_{\tilde{L}^{p_0}_{q_0}(T)} + \|\nabla^2 u\|_{\tilde{L}^{p_0}_{q_0}(T)} \leq c, \quad \lambda \geq \lambda_0. \tag{4.8}$$

We may take $\lambda \geq \lambda_0$ such that

$$\|u\|_\infty + \|\nabla u\|_\infty \leq \frac{1}{2}, \tag{4.9}$$

so that $\Theta(t, x) := x + u(t, x)$ satisfies

$$\frac{1}{2}|x - y|^2 \leq |\Theta(t, x) - \Theta(t, y)|^2 \leq 2|x - y|^2, \quad t \in [0, T], \ x, y \in \mathbb{R}^d. \tag{4.10}$$

For any $t \in [0, T]$, let $Y(t) = \Theta(t, X(t))$. By Itô's formula in [10, Theorem 1.2.3(3)] and (4.7), we obtain

$$\mathrm{d}Y(t) = \big\{\lambda u(t, X(t)) + \nabla\Theta(t, X(t))b^{(1)}(t, X_t)\big\}\mathrm{d}t + \big(\{\nabla\Theta\}\sigma\big)(t, X(t))\mathrm{d}W(t).$$

Combining this with (A_1), (A_2), the second inequality in (A_3'), (4.9) and Itô's formula, we find a constant $c_0 > 0$ such that

$$\mathrm{d}\{|Y(t)|^2 + 1\}^\alpha \leq c_0\{|Y(t)|^2 + 1\}^{\alpha - \frac{1}{2}}\{1 + \|Y_t\|^\alpha_\tau\}\mathrm{d}t + \mathrm{d}M(t), \quad t \in [0, T], \tag{4.11}$$

where

$$\mathrm{d}M(t) = 2\alpha\{|Y(t)|^2 + 1\}^{(\alpha-1)}\langle Y(t), \big(\{\nabla\Theta\}\sigma\big)(t, X(t))\mathrm{d}W(t)\rangle.$$

Note that

$$\begin{aligned} \mathrm{e}^{p\tau t}\|X_t\|^p_\tau &:= \sup_{s\in(-\infty,0]} \Big(\mathrm{e}^{p\tau(t+s)}|X(t+s)|^p\Big) \\ &\leq \sup_{s\in(-\infty,0]} (\mathrm{e}^{p\tau s}|X(s)|^p) + \sup_{s\in[0,t]} (\mathrm{e}^{p\tau s}|X(s)|^p) \\ &= \|X_0\|^p_\tau + \sup_{s\in[0,t]} (\mathrm{e}^{p\tau s}|X(s)|^p), \quad t \geq 0. \end{aligned} \tag{4.12}$$

Combining this with (4.11), Young's inequality, Itô's formula, and $(s + 1)^r \leq 1$ for $s \geq 0$ and $r \leq 0$, we find a constant $c_1 > 0$ such that

$$\begin{aligned} \mathrm{d}\big\{\mathrm{e}^{2\alpha\tau t}\{|Y(t)|^2 + 1\}^\alpha\big\} &\leq \mathrm{e}^{2\alpha\tau t}\mathrm{d}M(t) + c_1\left(1 + \|Y_0\|^{2\alpha}_\tau\right)\mathrm{d}t \\ &\quad + \sup_{s\in[0,t]} \mathrm{e}^{2\alpha\tau s}\left(|Y(s)|^2 + 1\right)^\alpha \mathrm{d}t, \quad t \in [0, T]. \end{aligned} \tag{4.13}$$

So, letting

$$
\begin{aligned}
&l(t) = \mathrm{e}^{2\alpha\tau t}\{|Y(t)|^2+1\}^\alpha + c_1\left(1+\|Y_0\|_\tau^{2\alpha}\right),\\
&\bar{l}_t = \sup_{s\in[0,t]} l(s) \quad t\in[0,T],
\end{aligned}
$$

we find a constant $c>0$ and a martingale $\tilde{M}(t)$ such that

$$
\mathrm{d}l(t) \le c\,\bar{l}_t\mathrm{d}t + \mathrm{d}\tilde{M}(t). \tag{4.14}
$$

Then by [6, Lemma 3.1], we find a constant $c_2>0$ such that

$$
\begin{aligned}
&\mathbb{E}\left[\exp\left(\frac{\varepsilon}{T\mathrm{e}^{1+cT}}\int_0^t \bar{l}_s\,\mathrm{d}s\right)\bigg|\mathscr{F}_0\right] \le \mathrm{e}^{\varepsilon l(s)+1}\left(\mathbb{E}\left[\exp\left(2\varepsilon^2\langle\tilde{M}\rangle(t)\right)|\mathscr{F}_0\right]\right)^{\frac{1}{2}}\\
&\le \mathrm{e}^{\varepsilon l(0)+1}\left(\mathbb{E}\left[\exp\left(c_2\varepsilon^2\int_0^t l(s)\,\mathrm{d}s\right)\bigg|\mathscr{F}_0\right]\right)^{\frac{1}{2}},\quad \varepsilon>0,\ t\in[0,T].
\end{aligned}
$$

Taking $\varepsilon = \frac{1}{c_2T\mathrm{e}^{1+cT}}$, we derive

$$
\mathbb{E}\left[\exp\left(\frac{\varepsilon}{T\mathrm{e}^{1+cT}}\int_0^t \bar{l}_s\,\mathrm{d}s\right)\bigg|\mathscr{F}_0\right] \le \mathrm{e}^{2(\varepsilon l(0)+1)},\quad t\in[0,T].
$$

By combining this with (4.9) and (4.12), we derive (4.6) for $\beta_T = \frac{\varepsilon}{T\mathrm{e}^{1+cT}}$ and some constant $k_T>0$. □

In the next result, we present Krylov-Khasminskii estimates for SDEs with memory.

Lemma 4.4. *Assume* (A_1), (A_2) *and* (A_3'). *For any constants* $T,p,q>1$ *such that* $(2p,2q)\in\mathscr{K}$ *and* $\varepsilon>0$, *there exist constants* $c=c(T,p,q,\varepsilon)>0$ *and* $\theta=\theta(T,p,q)>0$ *such that the solution to* (2.1) *satisfies*

$$
\mathbb{E}\left[\int_s^t |f_u(X(u))|\,\mathrm{d}u\bigg|\mathscr{F}_s\right] \le (c+\varepsilon\|X_s\|_\tau^\alpha)\,\|f\|^2_{\tilde{L}^p_q(s,t)}, \tag{4.15}
$$

$$
\mathbb{E}\left[\exp\left(\int_s^t |f_u(X(u))|\,\mathrm{d}u\right)\bigg|\mathscr{F}_s\right] \le \exp\left[\varepsilon\|X_s\|_\tau^{2\alpha} + c\left(1+\|f\|^\theta_{\tilde{L}^p_q(s,t)}\right)\right] \tag{4.16}
$$

for any $0\le s\le t\le T$ *and* $f\in\tilde{L}^p_q(T)$.

Proof. Without loss of generality, we may simply prove these estimates for $s=0$.

For any $x\in\mathbb{R}^d$, let $\phi^x\in\mathscr{C}_\tau$ be defined as

$$
\phi^x(r) := x,\quad r\le 0.
$$

For any $t\in[0,T]$, $x\in\mathbb{R}^d$ and $\xi\in\mathscr{C}_\tau$, let

$$
\bar{b}(t,x) := b^{(0)}(t,x) + b^{(1)}(t,\phi^x),\quad \hat{b}(t,\xi) := b^{(1)}(t,\xi) - b^{(1)}(t,\xi^0).
$$

By (4.9), (4.12), (4.13) and the Burkholder-Davis-Gundy (BDG) inequality, we find a constant $k_1 > 0$ such that

$$\mathbb{E}\left[\sup_{t\in[0,T]} \|X_t\|_\tau^{2\alpha}\Big|\mathscr{F}_0\right] \le k_1\left(1+\|X_0\|_\tau^{2\alpha}\right). \tag{4.17}$$

For any $n \ge 1$, let

$$\tau_n := \inf\{t\in[0,T] : \|X_t\|_\tau \ge n\}, \quad \inf\emptyset := T. \tag{4.18}$$

By (4.17), (A_1), (A_2), the second inequality in (A_3'), and applying Krylov's estimate in [17, Theorem 3.1] for $\bar{b}(t,x)$ in place of $b(t,x)$, see also [10, Theorem 1.2.3(2)], we find constants k_2 depending on ε and $k_3 > 0$ independent of ε such that

$$\begin{aligned}
&\mathbb{E}\left[\int_0^{t\wedge\tau_n} |f_s|\,(X(s))\,\mathrm{d}s\Big|\mathscr{F}_0\right]\\
\le&\left\{k_2+\varepsilon\left[\mathbb{E}\left(\int_0^{t\wedge\tau_n}|\hat{b}(s,X_s)|^2\,\mathrm{d}s\Big|\mathscr{F}_0\right)\right]^{\frac{1}{2}}\right\}\|f\|_{\tilde{L}_q^p(0,t)}\\
\le&k_3\left(k_2+\varepsilon\|X_0\|_\tau^\alpha\right)\|f\|_{\tilde{L}_q^p(0,t)}, \quad t\in[0,T].
\end{aligned}\tag{4.19}$$

Letting $n\to\infty$, we derive (4.15) since $\varepsilon > 0$ is arbitrary.

To prove (4.16), we shall use Girsanov's transform. Let

$$\gamma_t = \{\sigma^*(\sigma\sigma^*)^{-1}\}(t,X(t))\hat{b}(t,X_t), \quad t\in[0,T]. \tag{4.20}$$

Then for any $n\ge 1$ and τ_n in (4.18),

$$R_n(t) := \exp\left[-\int_0^{t\wedge\tau_n}\langle\gamma_r,\mathrm{d}W(r)\rangle - \frac{1}{2}\int_0^{t\wedge\tau_n}|\gamma_r|^2\,\mathrm{d}r\right], \quad t\in[0,T]$$

is a martingale. By Girsanov's theorem,

$$W_n(t) := W(t) + \int_0^{t\wedge\tau_n}\gamma_s\,\mathrm{d}s, \quad t\in[0,T]$$

is a Brownian motion under the probability measure $\mathrm{d}\mathbb{Q}_n := R_n(T)\mathrm{d}\mathbb{P}$, and X solves the equation

$$X(t) = X(0) + \int_0^t \bar{b}(s,X(s))\,\mathrm{d}s + \int_0^t \sigma(s,X(s))\,\mathrm{d}W_n(s), \quad t\in[0,\tau_n]. \tag{4.21}$$

Since the Krylov estimate in [10, Theorem 1.2.4] holds for (4.21), letting $f_r^n(x) = f_r(x)\mathbf{1}_{\{|x|<n\}}$, by the argument in the proof of [18, Lemma 4.4], we find a constant $k > 0$ independent of s such that for any $t\in[0,T]$,

$$\mathbb{E}_{\mathbb{Q}_n}\left[\exp\left(\lambda\int_0^t\left|f_r^n\left(X(r\wedge\tau_n)\right)\right|\,\mathrm{d}r\right)\Big|\mathscr{F}_0\right] \le \exp\left[k + k\|f\|_{\tilde{L}_q^p(0,t)}^k\right]. \tag{4.22}$$

On the other hand, for constants K in (A_1) (A_2), (A_3'), and β_T in Lemma 4.3, let

$$p_0 := \frac{3+\sqrt{1+8\beta_T K^{-3}}}{4}.$$

We have

$$k(p) := (p-1)(2p-1)K^3 \in (0,\beta_T], \quad p \in (1,p_0].$$

Thus, taking $\tilde{p} \in [1,p_0]$ with $\tilde{p}-1$ small enough such that $k_T k(\tilde{p}) \le 2\varepsilon$, by (A_1) (A_2), (A_3') and (4.5), we find a constant $\beta = \beta(p,q,\varepsilon) > 0$ such that

$$\begin{aligned}
&\mathbb{E}\big[|R_n(T)|^{-(\tilde{p}-1)}\big|\mathscr{F}_0\big]\\
=&\mathbb{E}\left[\exp\left((\tilde{p}-1)\int_0^{T\wedge\tau_n}\langle\gamma_s, \mathrm{d}W(s)\rangle + \frac{\tilde{p}-1}{2}\int_0^{T\wedge\tau_n}|\gamma_s|^2\,\mathrm{d}s\right)\bigg|\mathscr{F}_0\right]\\
\le&\left(\mathbb{E}\left[\exp\left((\tilde{p}-1)(2\tilde{p}-1)\int_0^{T\wedge\tau_n}|\gamma_s|^2\,\mathrm{d}s\right)\bigg|\mathscr{F}_0\right]\right)^{\frac{1}{2}}\\
&\quad\times\left(\mathbb{E}\left[\exp\left(2(\tilde{p}-1)\int_0^{T\wedge\tau_n}\langle\gamma_s, \mathrm{d}W(s)\rangle - 2(\tilde{p}-1)^2\int_0^{T\wedge\tau_n}|\gamma_s|^2\,\mathrm{d}s\right)\bigg|\mathscr{F}_0\right]\right)^{\frac{1}{2}}\\
\le&\exp\left[\beta+\varepsilon\|X_0\|_\tau^{2\alpha}\right]\times 1 = \exp\left[\beta+\varepsilon\|X_0\|_\tau^{2\alpha}\right].
\end{aligned} \tag{4.23}$$

Combining this with (4.22) and applying Hölder's inequality, we find constants $c = c(p,q,\varepsilon) > 0$ and $\theta = \theta(p,q) > 0$ such that

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(\int_0^t |f_s^n(X(s\wedge\tau_n))|\,\mathrm{d}s\right)\bigg|\mathscr{F}_0\right]\\
=&\mathbb{E}_{\mathbb{Q}_n}\left[R_n(T)^{-1}\exp\left(\int_0^t |f_s^n(X(s\wedge\tau_n))|\,\mathrm{d}s\right)\bigg|\mathscr{F}_0\right]\\
\le&\left(\mathbb{E}\big[R_n(T)^{-(\tilde{p}-1)}\big|\mathscr{F}_0\big]\right)^{\frac{1}{\tilde{p}}}\left(\mathbb{E}_{\mathbb{Q}_n}\left[\exp\left(\frac{\tilde{p}}{\tilde{p}-1}\int_0^t |f_r^n(X(s\wedge\tau_n))|\,\mathrm{d}s\right)\bigg|\mathscr{F}_0\right]\right)^{\frac{\tilde{p}-1}{\tilde{p}}}\\
\le&\exp\left[\varepsilon\|X_0\|_\tau^{2\alpha} + c\left(1+\|f\|^\theta_{\tilde{L}_p^q(0,t)}\right)\right].
\end{aligned}$$

This with $n\to\infty$ implies (4.16). □

By Lemma 4.3 and (4.16), we have the following corollary which is a key in the proofs of Theorem 2.1 and 2.2.

Corollary 4.5. *Assume (A_1), (A_2) and (A_3').*

(1) *For any $T\in(1,\infty)$, $(p,q)\in\mathscr{K}$ and $\varepsilon\in(0,1)$, there exist constants $k = k(T,p,q,\varepsilon) > 0$ and $\theta = \theta(T,p,q) > 0$ such that any solution X_s to (2.1) satisfies*

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(\int_s^t |f_u(X(u))|\|X_u\|_\tau^\alpha\,\mathrm{d}u\right)\bigg|\mathscr{F}_s\right] \le \exp\left[\varepsilon\|X_s\|_\tau^{2\alpha} + k\left(1+\|f\|^\theta_{\tilde{L}_q^p(s,t)}\right)\right],\\
&\qquad 0\le s\le t\le T,\ f\in\tilde{L}_q^p(T).
\end{aligned} \tag{4.24}$$

(2) *For any* $T \in (1,\infty)$, *there exist constants* $\beta_T, k_T > 0$ *such that any solution to* (2.1) *satisfies*

$$\sup_{t\in[0,T]} \mathbb{E}\Big[\mathrm{e}^{\beta\|X_t\|_\tau^{2\alpha}}\Big|\mathscr{F}_0\Big] \le \mathrm{e}^{k_T\beta(1+\|X_0\|_\tau^{2\alpha})}, \quad \beta\in[0,\beta_T]. \tag{4.25}$$

Proof. (a) Let $\beta_T, k_T > 0$ be in Lemma 4.3, and let $\varepsilon, c > 0$ be in Lemma 4.4. We choose $\beta \in (0,\beta_T]$ such that $k_T\beta \le 2\varepsilon$, so that by Lemma 4.3, (4.16) and Young's and Hölder's inequalities, we find constants $\theta = \theta(T,p,q), k = k(T,p,q,\varepsilon) > 0$ such that

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(\int_s^t |f_u(X(u))|\cdot\|X_u\|_\tau^\alpha\,\mathrm{d}u\right)\middle|\mathscr{F}_s\right]\\
\le&\mathbb{E}\left[\exp\left(\int_s^t \left(\frac{1}{2\beta}|f_u(X(u))|^2+\frac{\beta}{2}\cdot\|X_u\|_\tau^{2\alpha}\right)\mathrm{d}u\right)\middle|\mathscr{F}_s\right]\\
\le&\left(\mathbb{E}\left[\exp\left(\frac{1}{\beta}\int_s^t|f_u(X(u))|^2\,\mathrm{d}u\right)\middle|\mathscr{F}_s\right]\right)^{\frac12}\left(\mathbb{E}\left[\exp\left(\beta\int_s^t\|X_u\|_\tau^{2\alpha}\,\mathrm{d}u\right)\middle|\mathscr{F}_s\right]\right)^{\frac12}\\
\le&\exp\left[\varepsilon\|X_s\|_\tau^{2\alpha}+k\left(1+\|f^2\|_{\tilde L_{q/2}^{p/2}(s,t)}^{\theta/2}\right)\right]=\exp\left[\varepsilon\|X_s\|_\tau^{2\alpha}+k\left(1+\|f\|_{\tilde L_q^p(s,t)}^{\theta}\right)\right].
\end{aligned}$$

(b) In the case where $b^{(1)} = 0$, (4.25) can be obtained by repeating the proof of [4, Lemma 2.4] and [3, Corollary 2.5], so that by Girsanov's theorem, for the segment solution to (4.21) and $\mathbb{Q}_n$ in the proof of Lemma 4.4, there exists constants $c_0, c_1 > 0$ such that

$$\mathbb{E}_{\mathbb{Q}_n}\left[\mathrm{e}^{c_0\|X_{t\wedge\tau_n}\|_\tau^{2\alpha}}\middle|\mathscr{F}_0\right] \le \mathrm{e}^{c_1+c_1\|X_0\|_\tau^{2\alpha}}, \quad t\in[0,T],\ n\ge1. \tag{4.26}$$

Letting $\tilde p > 1$ be in (4.23) and taking

$$\beta_T := \frac{\tilde p c_0}{\tilde p - 1},$$

by combining (4.23) with (4.26) and using Hölder's inequality, we find a constant $c > 0$ such that

$$\begin{aligned}
&\mathbb{E}\big[\mathrm{e}^{\beta_T\|X_{t\wedge\tau_n}\|_\tau^{2\alpha}}\big|\mathscr{F}_0\big] \le \mathbb{E}_{\mathbb{Q}_n}\left[R_n(T)^{-1}\mathrm{e}^{\beta_T\|X_{t\wedge\tau_n}\|_\tau^{2\alpha}}\big|\mathscr{F}_0\right]\\
\le&\left(\mathbb{E}\big[R_n(T)^{-(\tilde p-1)}\big|\mathscr{F}_0\big]\right)^{\frac{1}{\tilde p}}\left(\mathbb{E}_{\mathbb{Q}_n}\left[\exp\left(\frac{\tilde p\beta_T}{\tilde p-1}\|X_{t\wedge\tau_n}\|_\tau^{2\alpha}\right)\middle|\mathscr{F}_0\right]\right)^{\frac{\tilde p-1}{\tilde p}}\\
\le&\exp\left(c+c\|X_0\|_\tau^{2\alpha}\right),\quad t\in[0,T],\ n\ge1.
\end{aligned}\tag{4.27}$$

By letting $n\to\infty$ we derive

$$\mathbb{E}\big[\mathrm{e}^{\beta_T\|X_t\|_\tau^{2\alpha}}\big|\mathscr{F}_0\big] \le \exp\left(c+c\|X_0\|_\tau^{2\alpha}\right),\quad t\in[0,T],$$

which implies (4.25) for $k_T := \frac{c}{\beta_T}$ by Jensen's inequality. □

Proof of Theorem 2.1. Let X_0 be an $\mathscr{F}_0$-measurable random variable in $\mathscr{C}_\tau$. Take $\psi \in C_b^\infty([0,\infty))$ such that $0\le\psi\le1$, $\psi(u)=1$ for $u\in[0,1]$, and $\psi(u)=0$ for $u\in[2,\infty)$. Define

$$b_n^{(1)}(t,\xi) = b^{(1)}(t,\xi)\psi\left(\frac{\|\xi\|_\tau}{n}\right),\quad t\ge0,\ \xi\in\mathscr{C}_\tau. \tag{4.28}$$

By (2.2), for any $n \ge 1$, $b_n^{(1)}$ is bounded and satisfies (A_3'). So, Lemma 4.1 ensures that the SDE

$$\mathrm{d}X^n(t) = b_n^{(1)}(t, X_t^n)\mathrm{d}t + b^{(0)}(t, X^n(t))\mathrm{d}t + \sigma(t, X^n(t))\mathrm{d}W(t), \quad X_0^n = X_0,\ t \in [0,T]$$

has a unique strong solution. Let τ_n be in (4.18). It suffices to prove that for any $k \ge 2$ there exists a constant $c(T,k) > 0$ such that

$$\mathbb{E}\bigg[\sup_{t\in[0,T]} \|X_t^n\|_\tau^k \bigg| \mathscr{F}_0\bigg] \le c(T,k)\big(1 + \|X_0\|_\tau^k\big), \quad n \ge 1. \tag{4.29}$$

This implies $\mathbb{P}(\tau_n = T) \to 1$ as $n \to \infty$ and

$$X(t) := 1_{\{t=0\}}X(0) + \sum_{n=1}^{\infty} X_n(t) 1_{(\tau_{n-1},\tau_n]}(t), \quad t \in [0,T]$$

is the unique solution to (2.1) and (2.4) holds, where $\tau_0 := 0$. To prove (4.29), we use Zvonkin's transform.

For $\lambda \ge \lambda_0$ and u solving the PDE (4.7) such that (4.8) and (4.9) hold, let

$$\Theta(t,x) := x + u(t,x), \quad Y^n(t) := \Theta(t, X^n(t)), \quad t \in [0,T].$$

By Itô's formula, we obtain

$$\begin{aligned} \mathrm{d}Y^n(t) =& \{\nabla\Theta(t, X^n(t)) b_n^{(1)}(t, X_t^n) + \lambda u(t, X^n(t))\}\mathrm{d}t \\ &+ \{(\nabla\Theta\sigma)(t, X^n(t))\}\mathrm{d}W(t). \end{aligned} \tag{4.30}$$

Let

$$Z(t) = \mathrm{e}^{k\tau t}|Y^n(t)|^k + 1.$$

By (A_1)–(A_3), (4.9) and Itô's formula, we find a constant $c_1 > 0$ such that

$$\mathrm{d}Z(t) \le c_1 \Big(1 + \|Y_0^n\|_\tau^k\Big)\,\mathrm{d}t + c_1 \sup_{s\in[0,t]} Z(s)\mathrm{d}t + \mathrm{d}M(t), \quad t \le \tau_n, \tag{4.31}$$

where

$$\mathrm{d}M(t) = k\mathrm{e}^{k\tau t}|Y^n(t)|^{j-2}\langle Y^n(t), \big(\{\nabla\Theta\}\sigma\big)(t, X^n(t))\mathrm{d}W(t)\rangle.$$

By the BDG inequality and the Young inequality, we find constants $c_2, c_3 > 0$ such that

$$\begin{aligned} &\mathbb{E}\bigg[\sup_{s\in[0,t\wedge\tau_n]} M(s)\bigg|\mathscr{F}_0\bigg] \le c_2\mathbb{E}\bigg[\int_0^t Z(s\wedge\tau_n)^2\,\mathrm{d}s\bigg|\mathscr{F}_0\bigg]^{\frac{1}{2}} \\ \le& \frac{1}{2}\mathbb{E}\bigg[\sup_{s\in[0,t\wedge\tau_n]} Z(s)\bigg|\mathscr{F}_0\bigg] + c_3\int_0^t \mathbb{E}\big[Z(s\wedge\tau_n)\big|\mathscr{F}_0\big]\,\mathrm{d}s, \quad t \in [0,T]. \end{aligned}$$

Combining this with (4.12), (4.31) and

$$|X^n(t) - Y^n(t)| \le \|u\|_\infty < \infty,$$

we may apply Gronwall's inequality to derive the desired (4.29) for some constant $c = c(T,k) > 0$. □

Proof of Theorem 2.2. Let $\Theta(t,x) := x + u(t,x)$ for u solving the PDE (4.7) such that (4.8) and (4.9) holds, and denote

$$\begin{aligned}
&Y^\xi(t) := \Theta(t, X^\xi(t)), \quad Y^\eta(t) := \Theta(t, X^\eta(t)),\\
&\Phi(t) := \mathrm{e}^{\tau t}(X^\xi(t) - X^\eta(t)), \quad \Xi(t) := \mathrm{e}^{2k\tau t}|Y^\xi(t) - Y^\eta(t)|^{2k},\\
&g_1(t) := \mathscr{M}\|\nabla^2 u\|(t, X^\xi(t)) + \mathscr{M}\|\nabla^2 u\|(t, X^\eta(t)),\\
&g_2(t) := g_1(t) + \mathscr{M}\|\nabla\sigma\|(t, X^\xi(t)) + \mathscr{M}\|\nabla\sigma\|(t, X^\eta(t)),\\
&A(t) := \int_0^t \Big(1 + g_2(s)^2 + g_1(s)\|X_s^\xi\|_\tau^\alpha\Big)\,\mathrm{d}s, \quad t \in [0,T].
\end{aligned}$$

By Lemma 4.2, Lemma 4.4, Corollary 4.5 and Holder's inequality, for any $\beta > 0$ and $\varepsilon \in (0,1)$ we find a constants $c(\beta,\varepsilon) > 0$ such that

$$\mathbb{E}\left[\mathrm{e}^{\beta A(T)}\right] \le \mathrm{e}^{c(\beta,\varepsilon)+\varepsilon(\|\xi\|_\tau^{2\alpha}+\|\eta\|_\tau^{2\alpha})}. \tag{4.32}$$

By (A_1), (A_2), (A_3'), (4.9) and Lemma 4.2, we find a constant $k_0 > 0$ such that

$$\begin{aligned}
&\Big|[\nabla\Theta(t, X^\xi(t))]b^{(1)}(t, X_t^\xi) - [\nabla\Theta(t, X^\eta(t))]b^{(1)}(t, X_t^\eta)\Big|\\
&\le k_0\|X_t^\xi - X_t^\eta\|_\tau + k_0\left(1 + g_1(t)\right)\|X_t^\xi\|_\tau^\alpha |X^\xi(t) - X^\eta(t)|,
\end{aligned}$$

$$\Big\|\left([\nabla\Theta]\sigma\right)(t, X^\xi(t)) - \left([\nabla\Theta]\sigma\right)(t, X^\eta(t))\Big\|_{HS}^2 \le k_0\left(1 + g_2(t)^2\right)|X^\xi(t) - X^\eta(t)|^2.$$

Combining this with (4.10), (4.12), (4.30) and Itô's formula, for any $k \ge 1$, we find a martingale $M(t)$ and a constant $c_1(T,k) > 0$ such that for any $t \in [0,T]$,

$$\mathrm{d}\Xi(t) \le c_1(T,k)\left\{\Xi(t)\mathrm{d}A(t) + \sup_{s\in[0,t]}\Xi(s)\mathrm{d}t + \|Y_0^\xi - Y_0^\eta\|_\tau^{2k}\mathrm{d}t\right\} + \mathrm{d}M(t).$$

Using the Itô formula, the stochastic Grönwall inequality in [3, Lemma A.5] and (4.10), there exists a constant $c_2(T,k) > 0$ such that

$$\left(\mathbb{E}\left[\sup_{s\in[0,t]}\left(\mathrm{e}^{-c_1(T,k)A(s)}|\Phi(s)|^{2k}\right)^{\frac{2}{3}}\right]\right)^{\frac{3}{2}} \le c_2(T,k)\|\xi - \eta\|_\tau^{2k}, \quad t \in [0,T].$$

By combining this with (4.32) and Hölder's inequality, we find a constant $c_3(\varepsilon,T,k) > 0$, such that

$$\begin{aligned}
&\mathbb{E}\left[\sup_{s\in[0,t]}|\Phi(s)|^k\right] \le \mathbb{E}\left[\mathrm{e}^{\frac{1}{2}c_1(T,k)A(t)}\left(\sup_{s\in[0,t]}\mathrm{e}^{-\frac{1}{2}c_1(T,k)A(s)}|\Phi(s)|^k\right)\right]\\
&\le \left(\mathbb{E}\left[\mathrm{e}^{2c_1(T,k)A(t)}\right]\right)^{\frac{1}{4}}\left(\mathbb{E}\left[\sup_{s\in[0,t]}\left(\mathrm{e}^{-c_1(T,k)A(s)}|\Phi(s)|^{2k}\right)^{\frac{2}{3}}\right]\right)^{\frac{3}{4}}\\
&\le c_3(\varepsilon,T,k)\mathrm{e}^{\varepsilon(\|\xi\|_\tau^{2\alpha}+\|\eta\|_\tau^{2\alpha})}\|\xi - \eta\|_\tau^k, \quad t \in [0,T].
\end{aligned} \tag{4.33}$$

This together with (4.12) yields the desired estimate (2.5) for some constant $c > 0$ depending on ε, T and k. □

5 Proof of Theorem 2.3

It suffices to prove the assertions up to any fixed time $T \in (0, \infty)$. To this end, we use the fixed point argument as in [7].

Let X_0 be an $\mathscr{F}_0$-measurable random variable on $\mathscr{C}_\tau$ with $\gamma := \mathscr{L}_{X_0} \in \mathscr{P}^\alpha_{k,\mathrm{e}}(\mathscr{C}_\tau)$. For any constant $N \geq 2$, let

$$\mathcal{C}_k^{\gamma,N} := \left\{\mu \in C_b^w([0,T]; \mathscr{P}^\alpha_{k,\mathrm{e}}(\mathscr{C}_\tau)) : \mu_0 = \gamma, \sup_{t\in[0,T]} \mathrm{e}^{-Nt}(1 + \mu_t(\|\cdot\|_\tau^k)) \leq N\right\},$$

where

$$C^w([0,T]; \mathscr{P}^\alpha_{k,\mathrm{e}}(\mathscr{C}_\tau)) := \{\mu : [0,T] \to \mathscr{P}^\alpha_{k,\mathrm{e}}(\mathscr{C}_\tau)) \text{ is weakly continuous}\},$$
$$C_b^w([0,T]; \mathscr{P}^\alpha_{k,\mathrm{e}}(\mathscr{C}_\tau)) := \Big\{\mu \in C^w([0,T]; \mathscr{P}^\alpha_{k,\mathrm{e}}(\mathscr{C}_\tau)) : \sup_{t\in[0,T]} \mathbb{W}_{k,var}(\mu_t, \mu_0) < \infty\Big\}.$$

Then as $N \uparrow \infty$,

$$\mathcal{C}_k^{\gamma,N} \uparrow \mathcal{C}_k^\gamma := \left\{\mu \in C_b^w([0,T]; \mathscr{P}^\alpha_{k,\mathrm{e}}(\mathscr{C}_\tau)) : \mu_0 = \gamma\right\}. \tag{5.1}$$

According to Theorem 2.1 and Corollary 4.5, under the assumptions (H_1)–(H_2), for any $\mu \in \mathcal{C}_k^\gamma$, the SDE

$$\mathrm{d}X^\mu(t) = b(t, X_t^\mu, \mu_t)\mathrm{d}t + \sigma(t, X^\mu(t))\mathrm{d}W(t), \quad t \in [0,T], X_0^\mu = X_0 \tag{5.2}$$

has a unique segment solution with

$$\Phi^\gamma_\cdot \mu := \mathscr{L}_{X^\mu_\cdot} \in \mathcal{C}_k^\gamma.$$

Then the well-posedness of (1.2) follows if the map Φ^γ has a unique fixed point in $\mathcal{C}_k^\gamma$. To this end, we need to verify that there exists a constant $N_0 \geq 2$ such that for any $N \geq N_0$, the following two assertions hold:

(a) $\Phi^\gamma : \mathcal{C}_k^{\gamma,N} \to \mathcal{C}_k^{\gamma,N}$, i.e.

$$\sup_{t\in[0,T]} \mathrm{e}^{-Nt}\left(1 + \|\Phi_t^\gamma\mu\|_k^k\right) \leq N, \quad \mu \in \mathcal{C}_k^{\gamma,N}. \tag{5.3}$$

(b) Φ^γ has a unique fixed point in $\mathcal{C}_k^{\gamma,N}$.

Once these two assertions are confirmed, Φ^γ has a unique fixed point $\mu \in \mathcal{C}_k^\gamma$, and $X_t = X_t^\mu$ is the unique segment solution of (1.2) with initial value X_0, so that (2.9) follows from (5.8).

In the following two subsections, we prove assertions (a) and (b) respectively.

5.1 Proof of (a)

It suffices to prove (5.3) for $k > 0$ and large N. By Lemma 4.4, (H_1)-(H_2) implies that for some constant $c_1 > 0$ we have

$$\mathbb{E}\left[\exp\left(\int_0^T |f_0(s, X^\mu(s))|^2 \, \mathrm{d}s\right)\right] \le c_1 \mathbb{E}[\mathrm{e}^{\varepsilon \|X_0^\mu\|_\tau^{2\alpha}}] = c_1 \gamma(\mathrm{e}^{\varepsilon \|\cdot\|_\tau^{2\alpha}}) < \infty. \tag{5.4}$$

Hence,

$$\sup_{\mu \in \mathcal{C}_k^\gamma} \mathbb{E}\bigg(\int_0^T |f_0(s, X^\mu(s))|^2 \, \mathrm{d}s\bigg)^k < \infty.$$

Combining this with (H_1)-(H_2), $\gamma := \mathscr{L}_{X_0} \in \mathscr{P}_{k,\mathrm{e}}^\alpha(\mathscr{C}_\tau)$, the Itô formula and the BDG inequality, we find constants $c_2, c_3 > 0$ such that

$$\begin{aligned}
&\mathbb{E}\left[\sup_{s \in [0,t]} \left(1 + \mathrm{e}^{k\tau s} |X^\mu(s)|^k\right)\right] \le c_2 \mathbb{E}(1 + |X(0)|^k) \\
&+ c_2 \mathbb{E}\left(\int_0^t \mathrm{e}^{\tau s} \left\{\|X_s^\mu\|_\tau + |f_0(s, X^\mu(s))| + \|\mu_s\|_k\right\} \mathrm{d}s\right)^k \\
&\le c_3 + c_3 \mathbb{E}\left(\int_0^t \left\{\mathrm{e}^{2\tau s} \|X_s^\mu\|_\tau^2 + \mathrm{e}^{2\tau s} \|\mu_s\|_k^2\right\} \mathrm{d}s\right)^{\frac{k}{2}}, \quad t \in [0, T].
\end{aligned} \tag{5.5}$$

Below we verify (5.3) by considering $k \ge 2$ and $k \in (0, 2)$ respectively.

(a_1) Let $k \ge 2$. By (4.12) and (5.5), there exists a constant $c_4 > 0$ such that for any $t \in [0, T]$,

$$\mathbb{E}\left[\sup_{s \in [0,t]} \left(1 + \mathrm{e}^{k\tau s} \|X_s^\mu\|_\tau^k\right)\right] \le c_4 + c_4 \int_0^t \left\{\mathbb{E}\left[\mathrm{e}^{k\tau s} \|X_s^\mu\|_\tau^k\right] + \mathrm{e}^{k\tau s} \|\mu_s\|_k^k\right\} \mathrm{d}s.$$

Using Grönwall's lemma, it holds

$$\mathbb{E}\left[\sup_{s \in [0,t]} \left(1 + \mathrm{e}^{k\tau s} \|X_s^\mu\|_\tau^k\right)\right] \le c_4 \mathrm{e}^{c_4 T} \left(1 + \int_0^t \mathrm{e}^{k\tau s} \|\mu_s\|_k^k \, \mathrm{d}s\right), \quad t \in [0, T].$$

Hence, by noting that $\mu \in \mathcal{C}_k^{\gamma,N}$, we can find a constant c_5 such that

$$\begin{aligned}
&\mathbb{E}\left[1 + \|X_t^\mu\|_\tau^k\right] \le 1 + \mathrm{e}^{-k\tau t} \mathbb{E}\left[\sup_{s \in [0,t]} \left(1 + \mathrm{e}^{k\tau s} \|X_s^\mu\|_\tau^k\right)\right] \\
&\le c_5 + c_5 \int_0^t \mathrm{e}^{-k\tau(t-s)} \|\mu_s\|_k^k \, \mathrm{d}s \\
&\le c_5 + c_5 N \mathrm{e}^{Nt} \int_0^t \mathrm{e}^{-N(t-s)} \mathrm{d}s \le 2c_5 \mathrm{e}^{Nt}, \quad t \in [0, T].
\end{aligned} \tag{5.6}$$

Taking $N_0 = 2c_5$, we derive

$$\sup_{t \in [0,T]} \mathrm{e}^{-Nt} \left(1 + \|\Phi_t^\gamma \mu\|_k^k\right) = \sup_{t \in [0,T]} \mathrm{e}^{-Nt} \mathbb{E}\left(1 + \|X_t^\mu\|_\tau^k\right) \le N, \quad N \ge N_0, \ \mu \in \mathcal{C}_k^{\gamma,N}.$$

(a_2) Let $k \in (0,2)$. By (4.12) and (5.5), we find constants $c_6, c_7, c_8 > 0$ such that

$$
\begin{aligned}
U_t &:= \mathbb{E}\left[\sup_{s\in[0,t]}\left(1+\mathrm{e}^{k\tau s}\|X_s^\mu\|_\tau^k\right)\right] \le c_6 + c_6\mathbb{E}\left(\int_0^t \left\{\mathrm{e}^{2\tau s}\|X_s^\mu\|_\tau^2 + \mathrm{e}^{2\tau s}\|\mu_s\|_k^2\right\}\mathrm{d}s\right)^{\frac{k}{2}}, \\
&\le c_7 + c_7\left(\int_0^t \mathrm{e}^{2\tau s}\|\mu_s\|_k^2\,\mathrm{d}s\right)^{\frac{k}{2}} + c_7\mathbb{E}\left\{\left[\sup_{s\in[0,t]}\mathrm{e}^{k\tau s}\|X_s^\mu\|_\tau^k\right]^{1-\frac{k}{2}}\left(\int_0^t \mathrm{e}^{k\tau s}\|X_s^\mu\|_\tau^k\,\mathrm{d}s\right)^{\frac{k}{2}}\right\} \\
&\le c_7 + c_7\left(\int_0^t \mathrm{e}^{2\tau s}\|\mu_s\|_k^2\,\mathrm{d}s\right)^{\frac{k}{2}} + \frac{1}{2}U_t + c_8\int_0^t U_s\,\mathrm{d}s, \quad t\in[0,T],
\end{aligned}
$$

then Grönwall's lemma implies that

$$
U_t \le 2c_7\mathrm{e}^{2c_8 T}\left\{1+\left(\int_0^t \mathrm{e}^{2\tau s}\|\mu_s\|_k^2\,\mathrm{d}s\right)^{\frac{k}{2}}\right\}, \quad t\in[0,T].
$$

Thus, there exist constants $c_9, c_{10} > 0$ such that for any $\mu \in \mathcal{C}_k^{\gamma,N}$,

$$
\begin{aligned}
&\mathbb{E}\left(1+\|X_t^\mu\|_\tau^k\right) \le 1+\mathrm{e}^{-k\tau t}U_t \le c_9 + c_9\left(\int_0^t \mathrm{e}^{-2\tau(t-s)}\|\mu_s\|_k^2\,\mathrm{d}s\right)^{\frac{k}{2}} \\
&\le c_9 + c_9 N e^{Nt}\left(\int_0^t \mathrm{e}^{-2N(t-s)}\,\mathrm{d}s\right)^{\frac{k}{2}} \le c_9 + c_{10}N^{1-\frac{k}{2}}\mathrm{e}^{Nt}, \quad t\in[0,T].
\end{aligned}
\tag{5.7}
$$

Thus, there exists a constant $N_0 > 2$ such that for any $N \ge N_0$,

$$
\begin{aligned}
&\sup_{t\in[0,T]}\mathrm{e}^{-Nt}\left(1+\|\Phi_t^\gamma\mu\|_k^k\right) = \sup_{t\in[0,T]}\mathrm{e}^{-Nt}\mathbb{E}\left(1+\|X_t^\mu\|_\tau^k\right) \\
&\le c_9 + c_{10}N^{1-\frac{k}{2}} \le N, \quad \mu\in\mathcal{C}_k^{\gamma,N}.
\end{aligned}
$$

5.2 Proof of (b)

To ensure that Φ^γ has a unique fixed point in $\mathcal{C}_k^{\gamma,N}$ for $N \ge N_0$, we shall prove that it is contractive under a complete metric.

For any $\theta > 0$, let

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

Then the metric $\tilde{\mathbb{W}}_{k,\theta,var} = \mathbb{W}_{k,\theta,var} + \mathbb{W}_{k,\theta}$ is complete on $\mathcal{C}_k^{\gamma,N}$. We intend to show that Φ^γ is contractive in $C_k^{\gamma,N}$ under the metric $\tilde{\mathbb{W}}_{k,\theta,var}$ when θ is large enough.

By Theorem 2.1, (H_1)-(H_2), (5.3) and $\gamma := \mathscr{L}_{X_0} \in \mathscr{P}_k(\mathscr{C}_\tau)$, we find a constant $C_0(N) > 0$ such that

$$
\sup_{\mu\in C_k^{\gamma,N}}\mathbb{E}\Big(\mathbb{E}\Big[\sup_{t\in[0,T]}\|X_t^\mu\|_\tau^{2k}\Big|\mathscr{F}_0\Big]\Big)^{\frac{1}{2}} \le C_0(N)\mathbb{E}\left(1+\|X_0\|_\tau^k\right) < \infty. \tag{5.8}
$$

Next, since $\mu, \nu \in \mathcal{C}_k^{\gamma,N}$, we find a constant $C_0(N) > 0$ such that

$$\sup_{t\in[0,T]} \mathbb{W}_{k,var}(\mu_t, \nu_t) \le C_0(N).$$

Thus, by (H_1) and (H_2), we find a constant $C_1(N) > 0$ such that for any $\mu, \nu \in \mathcal{C}_k^{\gamma,N}$,

$$\zeta_s := \left\{\sigma^*(\sigma\sigma^*)^{-1}\right\}(s, X^\mu(s)) \left[b^{(1)}(s, X_s^\mu, \mu_s) - b^{(1)}(s, X_s^\mu, \nu_s)\right], \quad s \in [0,T]$$

satisfies

$$|\zeta_s|^2 \le C_1(N) H(s) \left(1 \wedge \mathbb{W}_{k,var}(\mu_s, \nu_s)^2\right), \quad s \in [0,T]. \tag{5.9}$$

Recalling that $H \in L^1_{loc}([0,\infty); (0,\infty))$, we have

$$\int_0^t H(s) \left(1 \wedge \mathbb{W}_{k,var}(\mu_s, \nu_s)^2\right) \mathrm{d}s \le \int_0^t H(s)\, \mathrm{d}s < \infty,$$

then by (H_2) and Girsanov's theorem, for any $t \in [0,T]$,

$$R_t := \exp\left(\int_0^t \langle \zeta_s, \mathrm{d}W(s)\rangle - \frac{1}{2}\int_0^t |\zeta_s|^2\, \mathrm{d}s\right) \tag{5.10}$$

is a martingale and

$$\tilde{W}_u := W_u - \int_0^u \zeta_s\, \mathrm{d}s, \quad u \in [0,t]$$

is a Brownian motion under the probability measure $\mathbb{Q}_t := R(t)\mathbb{P}$. Since $\mathrm{e}^s - 1 \le s\mathrm{e}^s$ for $s \ge 0$, we find constants $C_2(N), C_3(N) > 0$ such that

$$\begin{aligned}
&\mathbb{E}\left(|R_t - 1|^2 \big| \mathscr{F}_0\right) = \mathbb{E}\left[\mathrm{e}^{2\int_0^t \langle \zeta_s, \mathrm{d}W(s)\rangle - 2\int_0^t |\zeta_s|^2\, \mathrm{d}s + \int_0^t |\zeta_s|^2\, \mathrm{d}s} \middle| \mathscr{F}_0\right] - 1 \\
&\le \mathrm{e}^{C_1(N)\int_0^t H(s)\left(1 \wedge \mathbb{W}_{k,var}(\mu_s,\nu_s)^2\right)\mathrm{d}s} - 1 \le C_2(N) \int_0^t H(s) \mathbb{W}_{k,var}(\mu_s, \nu_s)^2\, \mathrm{d}s \\
&\le C_3(N) \mathrm{e}^{2\theta t} \tilde{\mathbb{W}}_{k,\theta,var}(\mu, \nu)^2 \int_0^t H(s) \mathrm{e}^{-2\theta(t-s)}\, \mathrm{d}s, \quad \mu, \nu \in \mathcal{C}_k^{\gamma,N}.
\end{aligned} \tag{5.11}$$

Reformulating (5.2) as

$$\mathrm{d}X^\mu(r) = b(r, X_r^\mu, \nu_r)\mathrm{d}r + \sigma(r, X^\mu(r))\mathrm{d}\tilde{W}(r), \quad \mathscr{L}_{X_0^\mu} = \gamma, r \in [0,t]. \tag{5.12}$$

By the uniqueness, we obtain

$$\Phi_t^\gamma \nu = \mathscr{L}_{X_t^\nu} = \mathscr{L}_{X_t^\mu | \mathbb{Q}_t},$$

where $\mathscr{L}_{X_t^\mu|\mathbb{Q}_t}$ stands for the distribution of X_t^μ under $\mathbb{Q}_t$. Thus, by (5.8), (5.9), (5.11) and Hölder's inequality, we find constant $C_4(N) > 0$ such that

$$\|\Phi_t^\gamma \mu - \Phi_t^\gamma \nu\|_{k,var} = \sup_{|f| \le 1 + \|\cdot\|_\tau^k} \left|\mathbb{E}\left[f(X_t^\nu) - f(X_t^\mu)\right]\right|$$

$$
\begin{aligned}
&= \sup_{|f|\le 1+\|\cdot\|_\tau^k} |\mathbb{E}\left[(R_t-1)f(X_t^\mu)\right]| \le \mathbb{E}\Big[(1+\|X_t^\mu\|_\tau^k)|R_t-1|\Big]\\
&\le \mathbb{E}\left[\Big\{\mathbb{E}\Big((1+\|X_t^\mu\|_\tau^k)^2\big|\mathscr{F}_0\Big)\Big\}^{\frac{1}{2}}\left\{\mathbb{E}\left(|R_t-1|^2\big|\mathscr{F}_0\right)\right\}^{\frac{1}{2}}\right]\\
&\le C_4(N)\mathrm{e}^{\theta t}\tilde{\mathbb{W}}_{k,\theta,var}(\mu,\nu)\left(\int_0^t H(s)\mathrm{e}^{-2\theta(t-s)}\,\mathrm{d}s\right)^{\frac{1}{2}}, \quad \mu,\nu\in\mathcal{C}_k^{\gamma,N}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&\mathbb{W}_{k,\theta,var}(\Phi^\gamma\mu,\Phi^\gamma\nu) = \sup_{t\in[0,T]} \mathrm{e}^{-\theta t}\|\Phi_t^\gamma\mu-\Phi_t^\gamma\nu\|_{k,var}\\
&\le C_4(N)\sup_{t\in[0,T]}\left(\int_0^t H(s)\mathrm{e}^{-2\theta(t-s)}\,\mathrm{d}s\right)^{\frac{1}{2}}\tilde{\mathbb{W}}_{k,\theta,var}(\mu,\nu), \quad \mu,\nu\in\mathcal{C}_k^{\gamma,N}.
\end{aligned}
\tag{5.13}
$$

We will finish the proof by considering $k\le 1$ and $k>1$ respectively.

(b_1) Let $k\le 1$. By (1.4) and (5.13), we obtain

$$
\begin{aligned}
&\tilde{\mathbb{W}}_{k,\theta,var}(\Phi^\gamma\mu,\Phi^\gamma\nu) \le (1+c)\mathbb{W}_{k,\theta,var}(\Phi^\gamma\mu,\Phi^\gamma\nu)\\
&\le (1+c)C_4(N)\left(\int_0^t H(s)\mathrm{e}^{-2\theta(t-s)}\,\mathrm{d}s\right)^{\frac{1}{2}}\tilde{\mathbb{W}}_{k,\theta,var}(\mu,\nu), \quad \mu,\nu\in\mathcal{C}_k^{\gamma,N}.
\end{aligned}
$$

Noting that $H\in L^1_{loc}([0,\infty);(0,\infty))$ yields that

$$
\lim_{\theta\to\infty}\sup_{t\in[0,T]}\left(\int_0^t H(s)\mathrm{e}^{-2\theta(t-s)}\,\mathrm{d}s\right)^{\frac{1}{2}} = 0, \tag{5.14}
$$

we may choose $\theta>0$ large enough such that

$$
\tilde{\mathbb{W}}_{k,\theta,var}(\Phi^\gamma\mu,\Phi^\gamma\nu) \le \frac{1}{2}\tilde{\mathbb{W}}_{k,\theta,var}(\mu,\nu), \quad \mu,\nu\in\mathcal{C}_k^{\gamma,N}.
$$

This together with (a) implies that Φ^γ has a unique fixed point in $C_k^{\gamma,N}$.

(b_2) Let $k>1$. Let $\Theta(t,\cdot):=id+u(t,\cdot)$ for u solving (4.7) such that (4.8) and (4.9) holds. Let

$$
\begin{aligned}
&Z(t) := X^\mu(t)+\Theta(t,X^\mu(t))-X^\nu(t)-\Theta(t,X^\nu(t)),\\
&g_1(t) := \mathscr{M}\|\nabla^2 u\|(t,X^\mu(t))+\mathscr{M}\|\nabla^2 u\|(t,X^\nu(t)),\\
&g_2(t) := g_1(t)+\mathscr{M}\|\nabla\sigma\|(t,X^\mu(t))+\mathscr{M}\|\nabla\sigma\|(t,X^\nu(t)), \quad t\in[0,T].
\end{aligned}
$$

By (H_1)-(H_3), (4.9), (4.10) and Itô's formula, we find a constant $c_1(k)>0$ such that

$$
\begin{aligned}
&\mathrm{d}\big\{\mathrm{e}^{2k\tau t}|Z(t)|^{2k}\big\} \le c_1(k)\left\{\mathrm{e}^{2k\tau t}\|Z_t\|_\tau^{2k}+H(t)\mathrm{e}^{2k\tau t}\mathbb{W}_{k,var}(\mu_t,\nu_t)^{2k}\right\}\mathrm{d}t\\
&+\mathrm{e}^{2k\tau t}|Z(t)|^{2k}\mathrm{d}A(t)+\mathrm{d}M(t), \quad t\in[0,T],
\end{aligned}
\tag{5.15}
$$

where

$$\begin{aligned}&A(t) := c_1(k)\int_0^t \left(1+g_2(s)^2+g_1(s)\|X_s^\mu\|_\tau^\alpha\right)\,\mathrm{d}s,\\&\mathrm{d}\langle M\rangle(t)\le c_1(k)\left(1+g_2(t)\right)^2\mathrm{e}^{4k\tau t}|Z(t)|^{4k}\,\mathrm{d}t.\end{aligned}\tag{5.16}$$

Combining this with (4.10), (4.12), the stochastic Grönwall inequality in [3, Lemma A.5] , Lemma 4.4, Corollary 4.5 and the fact that $X_0^\mu = X_0^\nu = X_0$, we find a constant $c_2(k) > 0$ such that

$$\begin{aligned}&\mathbb{E}\left[\sup_{s\in[0,t]}\mathrm{e}^{k\tau s}\|X_s^\mu - X_s^\nu\|_\tau^k\right] = \mathbb{E}\left[\sup_{s\in[0,t]}\mathrm{e}^{k\tau s}|X^\mu(s)-X^\nu(s)|^k\right]\\&\le\left(\mathbb{E}\left[\mathrm{e}^{\frac{3}{2}A(t)}\right]\right)^{\frac{1}{3}}\left(\mathbb{E}\left[\sup_{s\in[0,t]}\mathrm{e}^{-A(s)}\mathrm{e}^{2k\tau s}|X^\mu(s)-X^\nu(s)|^{2k}\right]^{\frac{3}{4}}\right)^{\frac{2}{3}}\\&\le c_2(k)\left(\int_0^t H(s)\mathrm{e}^{2k\tau s}\mathbb{W}_{k,var}(\mu_s,\nu_s)^{2k}\,\mathrm{d}s\right)^{\frac{1}{2}}.\end{aligned}$$

Since $\Phi_t^\gamma\mu = \mathscr{L}_{X_t^\mu}$ and $\Phi_t^\gamma\nu = \mathscr{L}_{X_t^\nu}$, by the inequalities above, we find a constant $c_3(k) > 0$ such that

$$\begin{aligned}&\mathbb{W}_{k,\theta}(\Phi^\gamma\mu,\Phi^\gamma\nu) = \sup_{t\in[0,T]}\mathrm{e}^{-\theta t}\mathbb{W}_k(\Phi_t^\gamma\mu,\Phi_t^\gamma\nu)\le\sup_{t\in[0,T]}\mathrm{e}^{-\theta t}\left(\mathbb{E}\left[\|X_t^\mu - X_t^\nu\|_\tau^k\right]\right)^{\frac{1}{k}}\\&\le c_2(k)\sup_{t\in[0,T]}\left(\int_0^t H(s)\mathrm{e}^{-2k\theta t}\mathbb{W}_{k,var}(\mu_s,\nu_s)^{2k}\,\mathrm{d}s\right)^{\frac{1}{2k}}\\&\le c_3(k)\tilde{\mathbb{W}}_{k,\theta,var}(\mu,\nu)\sup_{t\in[0,T]}\left(\int_0^t H(s)\mathrm{e}^{-2k\theta(t-s)}\,\mathrm{d}s\right)^{\frac{1}{2k}},\quad \mu,\nu\in\mathcal{C}_k^{\gamma,N},\ \theta>0.\end{aligned}$$

Combining this with (5.13) and (5.14), we may choose $\theta > 0$ such that

$$\tilde{\mathbb{W}}_{k,\theta,var}(\Phi^\gamma\mu,\Phi^\gamma\nu)\le\frac{1}{2}\tilde{\mathbb{W}}_{k,\theta,var}(\mu,\nu),\quad \mu,\nu\in\mathcal{C}_k^{\gamma,N}.\tag{5.17}$$

Thus, Φ^γ has a unique fixed point in $\mathcal{C}_k^{\gamma,N}$.

6 Proof of Theorem 3.3

Proof of Theorem 3.3(1). The well-posedness can be proved by using a standard fixed point theorem. Let X_0 be $\mathscr{F}_0$-measurable with $\gamma := \mathscr{L}_{X_0} \in \mathscr{P}_2(\mathscr{C}_\tau)$. For any constant $\theta, T > 0$, the path space

$$\mathcal{C}_2^\gamma := \left\{\mu\in C^w([0,T];\mathscr{P}_2(\mathscr{C}_\tau)):\ \mu_0=\gamma\right\}$$

is complete under the metric

$$\mathbb{W}_{2,\theta}(\mu,\nu) := \sup_{t\in[0,T]}\mathrm{e}^{-\theta t}\mathbb{W}_2(\mu_t,\nu_t).$$

According to Theorem 3.1, (H') implies that for any $\mu \in \mathcal{C}_2^\gamma$, the SDE

$$\mathrm{d}X^\mu(t) = b(t, X_t^\mu, \mu_t)\mathrm{d}t + \sigma(t, X_t^\mu)\mathrm{d}W(t), \quad t \in [0,T], X_0^\mu = X_0 \tag{6.1}$$

is well-posed with

$$\Phi_\cdot^\gamma \mu := \mathscr{L}_{X_\cdot^\mu} \in \mathcal{C}_2^\gamma.$$

For the well-posedness of (1.2), it suffices to show that Φ^γ has a unique fixed point in $\mathcal{C}_2^\gamma$.

Now, for $\nu_0 \in \mathscr{P}_2(\mathscr{C}_\tau)$, for simplicity, we assume that we can choose $\mathscr{F}_0$-measurable random variables X_0^μ and X_0^ν on $\mathscr{C}_\tau$ such that

$$\mathbb{W}_2(\mu_0, \nu_0)^2 = \mathbb{E}\big[\|X_0^\mu - X_0^\nu\|_\tau^2\big]. \tag{6.2}$$

Otherwise, in the following it suffices to first replace (X_0^μ, X_0^ν) be the sequences $(X_0^{\mu,n}, X_0^{\nu,n})$ such that

$$n^{-1} + \mathbb{W}_2(\mu_0, \nu_0)^2 \geq \mathbb{E}\big[\|X_0^{n,\mu} - X_0^{n,\nu}\|_\tau^2\big], \quad n \geq 1,$$

then let $n \to \infty$.

For any $\nu \in \mathcal{C}_2^{\nu_0}$ which is defined as $\mathcal{C}_2^\gamma$ for ν_0 replacing γ, let X_t^ν be the unique solution to

$$\mathrm{d}X^\nu(t) = b(t, X_t^\nu, \nu_t)\mathrm{d}t + \sigma(t, X_t^\nu)\mathrm{d}W(t), \quad t \in [0,T] \tag{6.3}$$

with initial value X_0^ν. By (H') and Itô's formula, we obtain

$$\mathrm{d}\mathrm{e}^{2\tau t}|X^\mu(t) - X^\nu(t)|^2 \leq K\mathrm{e}^{2\tau t}\big(\|X_t^\mu - X_t^\nu\|_\tau^2 + \mathbb{W}_2(\mu_t, \nu_t)^2\big)\mathrm{d}t + \mathrm{d}M(t)$$

for some constant $K > 0$, where M_t is a martingale with

$$\mathrm{d}\langle M\rangle(t) \leq K\mathrm{e}^{4\tau t}\|X_t^\mu - X_t^\nu\|_\tau^4\mathrm{d}t.$$

Combining this with (4.12), Itô's isometry and Young's inequality, we find a constant $c_1 > 0$ such that

$$\begin{aligned}
\mathbb{E}\bigg[\sup_{s\in[0,t]} \mathrm{e}^{2\tau s}\|X_s^\mu - X_s^\nu\|_\tau^2\bigg] &\leq c_1\mathbb{E}\|X_0^\mu - X_0^\nu\|_\tau^2 + c_1\mathbb{E}\bigg[\Big(\int_0^t \mathrm{e}^{4\tau s}\|X_s^\mu - X_s^\nu\|_\tau^4\,\mathrm{d}s\Big)^{\frac{1}{2}}\bigg] \\
&\quad + c_1\int_0^t \mathrm{e}^{2\tau s}\mathbb{E}\big[\|X_s^\mu - X_s^\nu\|_\tau^2 + \mathbb{W}_2(\mu_s, \nu_s)^2\big]\,\mathrm{d}s \\
&\leq c_1\mathbb{E}\|X_0^\mu - X_0^\nu\|_\tau^2 + \Big(c_1 + \frac{c_1^2}{2}\Big)\int_0^t \mathrm{e}^{2\tau s}\mathbb{E}\big[\|X_s^\mu - X_s^\nu\|_\tau^2 + \mathbb{W}_2(\mu_s, \nu_s)^2\big]\,\mathrm{d}s \\
&\quad + \frac{1}{2}\mathbb{E}\bigg[\sup_{s\in[0,t]} \mathrm{e}^{2\tau s}\|X_s^\mu - X_s^\nu\|_\tau^2\bigg], \quad t \in [0,T].
\end{aligned}$$

By an approximation argument with stopping times, we may and do assume that

$$\mathbb{E}\bigg[\sup_{s\in[0,t]} \mathrm{e}^{2\tau s}\|X_s^\mu - X_s^\nu\|_\tau^2\bigg] < \infty,$$

so that the above estimate together with (6.2) and Grönwall's inequality yields

$$\mathbb{E}\bigg[\sup_{s\in[0,t]}\|X_s^\mu - X_s^\nu\|_\tau^2\bigg] \le c_2(T)\mathbb{W}_2(\mu_0,\nu_0)^2 + c_2(T)\int_0^t \mathbb{W}_2(\mu_s,\nu_s)^2\,\mathrm{d}s \tag{6.4}$$

for any $t\in[0,T]$ and some constant $c_2(T)>0$. In particular, when $\mu_0=\nu_0=\gamma$, we derive

$$\begin{aligned}\mathbb{W}_{2,\theta}(\Phi^\gamma\mu,\Phi^\gamma\nu)^2 &\le \sup_{t\in[0,T]} \mathrm{e}^{-\theta t}\mathbb{E}\big[\|X_t^\mu - X_t^\nu\|_\tau^2\big]\\ &\le c_2(T)\sup_{t\in[0,T]}\mathrm{e}^{-\theta t}\int_0^t \mathbb{W}_2(\mu_s,\nu_s)^2\,\mathrm{d}s \le \frac{c_2}{\theta}\mathbb{W}_{2,\theta}(\mu,\nu)^2,\quad \mu,\nu\in\mathcal{C}_2^\gamma.\end{aligned}$$

So, when $\theta>c_2$, Φ^γ is contractive in the complete metric space $(\mathcal{C}_2^\gamma,\mathbb{W}_{2,\theta})$, hence it has a unique fixed point.

Moreover, letting $\mu_t=P_t^*\mu,\nu_t=P_t^*\nu$, we have $\mathscr{L}_{X_t^\mu}=P_t^*\mu$ and $\mathscr{L}_{X_t^\nu}=P_t^*\nu$, so that (6.4) and Grönwall's inequality yield (3.4) for some constant $c>0$. □

Now, let $\tau_0\in(0,\tau)$ and $\|\sigma\|_\infty+\|\sigma^{-1}\|_\infty<\infty$, it remains to verify (3.5), which implies (3.6) through repeating the proof of [5, Theorem 2.1(1)] for

$$E=\mathscr{P}_{2,\mathrm{e}}^\alpha(\mathscr{C}_\tau),\ \rho(\mu,\nu)=\mathbb{W}_2(\mu,\nu),\ \Gamma_t=c\mathrm{e}^{-\tau_0 t}, \tag{6.5}$$

and $\Lambda_t=c\mathrm{e}^{ct}$ in place of Λ.

Proof of (3.5). Let X_t^μ be the unique solution to (1.2) with the initial distribution μ and denote $\mu_t=P_t^*\mu$, $\nu_t=P_t^*\nu$,

$$\begin{aligned}\bar\zeta_s &:= \big\{\sigma^*(\sigma\sigma^*)^{-1}\big\}(s,X_s^\mu)\left[b(s,X_s^\mu,\mu_s)-b(s,X_s^\mu,\nu_s)\right],\\ \bar R_s &:= \exp\left[-\int_0^s\langle\bar\zeta_u,\mathrm{d}W(u)\rangle - \frac12\int_0^s|\bar\zeta_u|^2\,\mathrm{d}u\right],\quad s\in[0,t].\end{aligned}$$

Then (H') and (3.4) implies that

$$|\bar\zeta_s|^2 \le c_1K_2\mathbb{W}_2(\mu_s,\nu_s)^2 \le c_2K_2\mathrm{e}^{c_2 s}\mathbb{W}_2(\mu,\nu)^2 \tag{6.6}$$

for some constants $c_1,c_2>0$. Thus by Girsanov's theorem, $\bar R_s$ is a martingale and

$$\bar W(s)=W(s)+\int_0^s\bar\zeta_r\,\mathrm{d}r,\quad s\in[0,t]$$

is a Brownian motion under the probability measure $\bar{\mathbb{P}}:=\bar R_t\mathbb{P}$. Then (1.2) can be reformulated as

$$\mathrm{d}X^\mu(s)=b(s,X_s^\mu,\nu_s)\mathrm{d}s+\sigma(s,X_s^\mu)\mathrm{d}\bar W(s),\quad \mathscr{L}_{X_0^\mu}=\mu,\ s\in[0,t]. \tag{6.7}$$

For any $\kappa>\tau$, where $\tau>0$ is given in (1.1), consider the following SDE:

$$\begin{aligned}\mathrm{d}Y(s)=&\{b(s,Y_s,\nu_s)+\kappa\sigma(s,Y_s)\sigma(s,X_s^\mu)^{-1}(X^\mu(s)-Y(s))\}\mathrm{d}s\\ &+\sigma(s,Y_s)\mathrm{d}\bar W(s),\quad s\in[0,t],\quad Y_0=X_0^\nu.\end{aligned} \tag{6.8}$$

Let

$$\tilde{\zeta}_s := \kappa\sigma(s, X_s^\mu)^{-1}(X^\mu(s) - Y(s)),$$
$$\tilde{R}_s := \exp\left[-\int_0^s \langle\tilde{\zeta}_r, \mathrm{d}\bar{W}(r)\rangle - \frac{1}{2}\int_0^s |\tilde{\zeta}_r|^2\,\mathrm{d}r\right], \quad s \in [0,t].$$

Due to [5, Lemma 3.2] with $\mathbb{P}$ replaced by $\bar{\mathbb{P}}$, under the assumption (H'),

$$\tilde{W}(s) = \bar{W}(s) + \int_0^s \tilde{\zeta}_r\,\mathrm{d}r = W(s) + \int_0^s \left(\tilde{\zeta}_r + \bar{\zeta}_r\right)\mathrm{d}r, \quad s \in [0,t]$$

is a Brownian motion under the probability measure $\mathbb{Q} = \tilde{R}_t\bar{\mathbb{P}} = \tilde{R}_t\bar{R}_t\mathbb{P}$. Hence (6.8) can be reformulated as

$$\mathrm{d}Y(s) = b(s, Y_s, \nu_s)\mathrm{d}s + \sigma(s, Y_s)\mathrm{d}\tilde{W}(s), \quad s \in [0,t], \quad Y_0 = X_0^\nu, \tag{6.9}$$

which together with the uniqueness of (6.3) derives that $\mathscr{L}_{Y_t|\mathbb{Q}} = \mathscr{L}_{X_t^\nu}$. Moreover, if we choose $\mathscr{F}_0$-measurable random variables X_0^μ and X_0^ν on $\mathscr{C}_\tau$ such that

$$\mathbb{W}_2(\mu,\nu)^2 = \mathbb{E}\big[\|X_0^\mu - X_0^\nu\|_\tau^2\big]. \tag{6.10}$$

By (6.7), (H') and the proof of [5, Lemma 3.3], for any $p > 0$ and $\tau_0 \in (0,\tau)$ we find a constant $\kappa > \tau$ to define Y_s in (6.8) such that

$$\mathbb{E}_{\mathbb{Q}}\left[\|X_t^\mu - Y_t\|_\tau^p|\mathscr{F}_0\right] \le c\mathrm{e}^{-p\tau_0 t}\|X_0^\mu - X_0^\nu\|_\tau^p, \quad t \ge 0 \tag{6.11}$$

holds for some constant $c > 0$. Therefore, applying Young's inequality in [2, Lemma 2.4],

$$\begin{aligned}
P_t\log f(\nu) &= \mathbb{E}_{\mathbb{Q}}\left[\log f(Y_t)\right] = \mathbb{E}_{\mathbb{Q}}\left[\log f(X_t^\mu)\right] + \mathbb{E}_{\mathbb{Q}}\left[\log f(Y_t) - \log f(X_t^\mu)\right]\\
&\le \mathbb{E}\left[\bar{R}_t\tilde{R}_t\log f(X_t^\mu)\right] + \|\nabla\log f\|_\infty\mathbb{E}_{\mathbb{Q}}\|X_t^\mu - Y_t\|_\tau\\
&\le \mathbb{E}\left[\bar{R}_t\tilde{R}_t\log\left(\bar{R}_t\tilde{R}_t\right)\right] + \log P_t f(\mu) + c\|\nabla\log f\|_\infty\mathrm{e}^{-\tau_0 t}\mathbb{W}_2(\mu,\nu), \quad t \ge 0
\end{aligned} \tag{6.12}$$

holds for any $f \in \mathscr{B}_b^+(\mathscr{C}_r)$ with $\|\nabla\log f\|_\infty < \infty$. Next, denote $R_t = \bar{R}_t\tilde{R}_t$, it follows from (6.6), (6.11) that for some positive constants c_3, c_4,

$$\begin{aligned}
\mathbb{E}\left[R_t\log R_t\right] &\le \frac{1}{2}\mathbb{E}_{\mathbb{Q}}\int_0^t|\bar{\zeta}_s + \tilde{\zeta}_s|^2\,\mathrm{d}s \le \mathbb{E}_{\mathbb{Q}}\int_0^t|\bar{\zeta}_s|^2\,\mathrm{d}s + \mathbb{E}_{\mathbb{Q}}\int_0^t|\tilde{\zeta}_s|^2\,\mathrm{d}s\\
&\le c_3K_2\mathrm{e}^{c_3t}\mathbb{W}_2(\mu,\nu)^2 + c_3\int_0^t\mathbb{E}_{\mathbb{Q}}\|X_s^\mu - Y_s\|_\tau^2\,\mathrm{d}s \le \left(c_3K_2\mathrm{e}^{c_3t} + c_4\right)\mathbb{W}_2(\mu,\nu)^2.
\end{aligned}$$

Substituting this back into (6.12) yields (3.5). □

Acknowledgements

The authors would like to thank the associated editor and referees for their helpful comments and suggestions.

References

[1] L. Ambrosio, N. Gigli, G. Savaré, *Bakry-Emery curvature-dimension condition and Riemannian Ricci curvature bounds,* Ann. Probab. **43**, 339–440(2015)

[2] Arnaudon, M., Thalmaier, A., Wang, F.-Y.: Gradient estimates and Harnack inequalities on noncompact Riemannian manifolds. Stoch. Process. Appl. **119**, 3653–3670(2009)

[3] Bachmann, S.: Well-posedness and stability for a class of stochastic delay differential equations with singular drift. Stoch. Dyn. **18**, 1850019(2018)

[4] Bao, J., Shao J.: Weak convergence of path-dependent SDEs with irregular coefficients. Numerical Algebra, Control and Optimization. **15**(1), 108–129(2025),

[5] Bao, J., Wang, F.-Y., Yuan, C.:Asymptotic log-Harnack inequality and applications for stochastic systems of infinite memory. Stoch. Process. Appl. **129**, 4576–4596(2019)

[6] Bao, J., Wang, F.-Y., Yuan, C.: Derivative formula and Harnack inequality for degenerate functional SDEs. Stoch. Dyn. **13**, 1250013(2013)

[7] Huang, X., Wang, F.-Y.: Singular McKean–Vlasov (reflecting) SDEs with distribution dependent noise. J. Math. Anal. Appl. **514**, 126301(2022)

[8] E. Kopfer, K.-Th. Sturm, *Functional inequalities for the heat flow on time-dependent metric measure space,* J. Lond. Math. Soc. **104**, 926–955(2021)

[9] Ren, P.: Singular McKean-Vlasov SDEs: well-posedness, regularities and Wang's Harnack inequality. Stoch. Process. Appl. **156**, 291–311(2023)

[10] Ren, P., Wang, F.-Y.: Distribution Dependent Stochastic Differential Equations. World Scientific, 2024, Singapore.

[11] Wang, F.-Y., Logarithmic Sobolev inequalities on noncompact Riemannian manifolds, *Probab. Theory Related Fields*, **109**, 417–424(1997)

[12] Wang, F.-Y., Harnack inequalities on manifolds with boundary and applications, *J. Math. Pures Appl.*, **94**, 304–321(2010)

[13] Wang, F.-Y., *Harnack inequalities and Applications for Stochastic Partial Differential Equations*, Springer, 2013, Berlin.

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

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

[16] Xu, L. : A modified log-Harnack inequality and asymptotically strong Feller property. J. Evol. Equ. **11**, 925–942 (2011).

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

[18] X-Y. Zhao, Well-Posedness for Path-Distribution Dependent Stochastic Differential Equations with Singular Drifts. J. Theor. Probab. **37**(4), 3654–3687(2024)