# Exponential Convergence for Functional SDEs with Hölder Continuous Drift

**Xing Huang**

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

xinghuang@tju.edu.cn

October 20, 2018

**Abstract**

Applying Zvonkin's transform, the exponential convergence in Wasserstein distance for a class of functional SDEs with Hölder continuous drift is obtained. This combining with log-Harnack inequality implies the same convergence in the sense of entropy, which also yields the convergence in total variation norm by Pinsker's inequality.

AMS subject Classification: 60H10, 60B10.
Keywords: Functional SDEs, Zvonkin's transform, Hölder continuous, Exponential convergence, Wasserstein distance.

## 1 Introduction

Consider the SDE on $\mathbb{R}^d$

$$\mathrm{d}X(t) = b(X(t))\mathrm{d}t + \mathrm{d}W(t), \tag{1.1}$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$, $W$ is a $d$-dimensional Brownian motion on some complete filtration probability space. If the dissipative condition

$$\langle b(x) - b(y), x - y\rangle \leq -\kappa_0 |x - y|^2, \quad x, y \in \mathbb{R}^d$$

holds for some $\kappa_0 > 0$, then SDE (1.1) has a unique solution and the associated semigroup has exponential convergence in Wasserstein distance. In [1, 2], the exponential convergence in the sense of Wasserstein distance and total variation norm has been obtained for a class of functional SDEs/SPDEs with regular coefficients and additive noise, where exponential convergence in total variation norm is proved due to the gradient-$L^2$ estimate

$$|\nabla P_t f|^2 \leq C P_t |f|^2, \quad t > r_0, f \in \mathscr{B}_b(\mathscr{C}),$$

see [1, 2] for more details. Recently, using Zvonkin's transform [16], the strong well-posedness of SDEs is proved for SDEs with singular drifts, see [4, 9, 12, 13, 14, 15]. For the functional SDEs with singular drift, [5] proved the existence and uniqueness. In infinite dimension, [6, 8] obtain the existence and uniqueness of the mild solution for a class of semi-linear functional SPDEs with Dini continuous drift and establish the Harnack inequality.

Recall that for two probability measures $\mu, \nu$ on some measurable space $(E, \mathscr{F})$, the entropy and total variation norm are defined as follows:

$$\mathrm{Ent}(\nu|\mu) := \begin{cases} \int (\log \frac{\mathrm{d}\nu}{\mathrm{d}\mu}) \mathrm{d}\nu, & \text{if } \nu \text{ is absolutely continuous with respect to } \mu, \\ \infty, & \text{otherwise;} \end{cases}$$

and

$$\|\mu - \nu\|_{var} := \sup_{A \in \mathscr{F}} |\mu(A) - \nu(A)|.$$

By Pinsker's inequality (see [3, 10]),

$$\|\mu - \nu\|_{var}^2 \leq \frac{1}{2} \mathrm{Ent}(\nu|\mu), \quad \mu, \nu \in \mathscr{P}(E), \tag{1.2}$$

here $\mathscr{P}(E)$ denotes all probability measures on $(E, \mathscr{F})$. Indeed, these two estimates correspond to the log-Harnack inequality for the associated semigroups, see Lemma 2.1 below for details.

When $E$ is a Polish space, in particular, $E = \mathscr{C}$ in our frame, which will be defined in the sequel, let

$$\mathscr{P}_2 := \big\{\mu \in \mathscr{P}(\mathscr{C}) : \mu(\|\cdot\|_\infty^2) < \infty\big\}.$$

It is well known that $\mathscr{P}_2$ is a Polish space under the Wasserstein distance

$$\mathbb{W}_2(\mu, \nu) := \inf_{\pi \in \mathbf{C}(\mu,\nu)} \left( \int_{\mathscr{C} \times \mathscr{C}} \|\xi - \eta\|_\infty^2 \pi(\mathrm{d}\xi, \mathrm{d}\eta) \right)^{\frac{1}{2}}, \quad \mu, \nu \in \mathscr{P}_2,$$

where $\mathbf{C}(\mu, \nu)$ is the set of all couplings of $\mu$ and $\nu$. Moreover, the topology induced by $\mathbb{W}_2$ on $\mathscr{P}_2$ coincides with the weak topology.

The purpose of this paper is to establish the exponential convergence in the sense of Wasserstein distance, the entropy and total variation norm respectively for functional SDEs with Hölder continuous drift, which is much weaker than the Lipschitz condition.

Throughout the paper, we fix $r_0 > 0$ and consider the path space $\mathscr{C} := C([-r_0, 0]; \mathbb{R}^d)$ equipped with the uniform norm $\|\xi\|_\infty := \sup_{\theta \in [-r_0.0]} |\xi(\theta)|$. For any $f \in C([-r_0, \infty); \mathbb{R}^d)$,

$t \geq 0$, let $f_t(s) = f(t+s), s \in [-r_0, 0]$. Then $f_t \in \mathscr{C}$. $\{f_t\}_{t\geq 0}$ is called the segment process of $f$. Consider the following functional SDE on $\mathbb{R}^d$:

$$\mathrm{d}X(t) = AX(t)\mathrm{d}t + \{b(X(t)) + B(X_t)\}\mathrm{d}t + \sigma \mathrm{d}W(t), \tag{1.3}$$

where $W(t)$ is a $d$-dimensional Brownian motion on a complete filtration probability space $(\Omega, \{\mathscr{F}_t\}_{t\geq 0}, \mathscr{F}, \mathbb{P})$, $\sigma \in \mathbb{R}^d \otimes \mathbb{R}^d$, and

$$b : \mathbb{R}^d \to \mathbb{R}^d, \quad B : \mathscr{C} \to \mathbb{R}^d$$

are measurable.

The remainder of the paper is organized as follows. In Section 2 we summarize the main results of the paper; In section 3, we give precise estimate for Zvonkin's transform and the main results are proved in Section 4.

# 2 Main Results

Throughout this paper, we make the following assumptions:

**(H1)** $\sigma$ is invertible and $b$ is bounded, i.e.

$$\|b\|_\infty < \infty. \tag{2.1}$$

Moreover, there exist constants $\kappa > 0$ and $\alpha \in (0, 1)$ such that

$$|b(x) - b(y)| \leq \kappa |x - y|^\alpha, \quad x, y \in \mathbb{R}^d. \tag{2.2}$$

**(H2)** $A$ is a negative definite self-adjoint operator and there exists $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_d$ such that $Ae_i = -\lambda_i e_i, i = 1, 2, \cdots, d$. Furthermore,

$$\|B\|_\infty < \infty, \quad |B(\xi) - B(\eta)| \leq \lambda_B \|\xi - \eta\|_\infty, \qquad \xi, \eta \in \mathscr{C} \tag{2.3}$$

for some constant $\lambda_B > 0$.

Since Hölder continuity is stronger than Dini continuity, according to [8, Theorem 2.1] for $\mathbb{H} = \mathbb{R}^d$, under **(H1)** and (2.3), the SDE (1.3) has a unique non-explosive solution denoted by $X_t^\xi$ with $X_0 = \xi$ and

$$\mathbb{E} \sup_{t\in[0,T]} \|X_t^\xi\|_\infty < \infty, \quad T > 0. \tag{2.4}$$

Let $P_t(\xi, \mathrm{d}\eta)$ be the distribution of $X_t^\xi$, and

$$P_t f(\xi) := \int_{\mathscr{C}} f(\eta) P_t(\xi, \mathrm{d}\eta), \quad f \in \mathscr{B}_b(\mathscr{C}).$$

Moreover, for any $\nu \in \mathscr{P}_2$, let $\nu P_t = \int_{\mathscr{C}} P_t(\xi, \cdot)\nu(\mathrm{d}\xi)$. Then $\nu P_t$ is the distribution of the solution $X_t$ to (1.3) from initial distribution $\nu$.

The lemma below gives the estimate of $\mathrm{Ent}(P_t(\xi, \cdot)|P_t(\eta, \cdot))$ and $\|P_t(\xi, \cdot) - P_t(\eta, \cdot)\|_{var}$ respectively.

**Lemma 2.1.** *Assume* **(H1)** *and* (2.3)*. Then the log-Harnack inequality holds, i.e.*

$$P_t \log f(\eta) \le \log P_t f(\xi) + \frac{C(t)}{(t-r_0)\wedge 1}\|\xi-\eta\|_\infty^2, \quad t>r_0, \xi,\eta\in\mathscr{C}, f\in\mathscr{B}_b^+(\mathbb{R}^d) \tag{2.5}$$

*for some function* $C:(r_0,\infty)\to(0,\infty)$*. Thus, for any* $t>r_0$*,* $P_t(\xi,\cdot)$ *is equivalent to* $P_t(\eta,\cdot)$*. Moreover,*

$$\mathrm{Ent}(P_t(\xi,\cdot)|P_t(\eta,\cdot)) = P_t\left\{\log\frac{\mathrm{d}P_t(\xi,\cdot)}{\mathrm{d}P_t(\eta,\cdot)}\right\}(\xi) \le \frac{C(t)}{(t-r_0)\wedge 1}\|\xi-\eta\|_\infty^2, \tag{2.6}$$

*and*

$$\|P_t(\xi,\cdot)-P_t(\eta,\cdot)\|_{var}^2 \le \frac{C(t)}{2(t-r_0)\wedge 2}\|\xi-\eta\|_\infty^2. \tag{2.7}$$

*Proof.* The log-Harnack inequality (2.5) is a known result in [8, Theorem 2.2] with $\mathbb{H}=\mathbb{R}^d$, see also [7, Theorem 2.4] for log-Harnack inequality of the path-distribution dependent SDEs with Dini continuous drift. Combining the definition of $\mathrm{Ent}(P_t(\xi,\cdot)|P_t(\eta,\cdot))$ and [11, Theorem 1.4.2], we obtain (2.6). Finally, (2.7) follows from (1.2) and (2.6). □

**Remark 2.2.** *For any* $\nu,\tilde\nu\in\mathscr{P}_2$ *and* $\pi\in\mathbf{C}(\nu,\tilde\nu)$*, taking expectation on both sides of* (2.5) *with respect to* $\pi$*, we have for any* $t>r_0$*,*

$$\int_{\mathscr{C}^2} P_t\log f(\eta)\pi(\mathrm{d}\xi,\mathrm{d}\eta) \le \int_{\mathscr{C}^2}\log P_t f(\xi)\pi(\mathrm{d}\xi,\mathrm{d}\eta) + \frac{C(t)}{(t-r_0)\wedge 1}\int_{\mathscr{C}^2}\|\xi-\eta\|_\infty^2\pi(\mathrm{d}\xi,\mathrm{d}\eta).$$

*Jensen's inequality and the definition of* $\mathbb{W}_2$ *imply that*

$$(\tilde\nu P_t)(\log f) \le \log(\nu P_t)(f) + \frac{C(t)}{(t-r_0)\wedge 1}\mathbb{W}_2(\nu,\tilde\nu)^2, \quad t>r_0. \tag{2.8}$$

*Then we have*

$$\mathrm{Ent}(\nu P_t|\tilde\nu P_t) \le \frac{C(t)}{(t-r_0)\wedge 1}\mathbb{W}_2(\nu,\tilde\nu)^2, \quad t>r_0, \tag{2.9}$$

*and*

$$\|\nu P_t-\tilde\nu P_t\|_{var}^2 \le \frac{C(t)}{2(t-r_0)\wedge 2}\mathbb{W}_2(\nu,\tilde\nu)^2, \quad t>r_0. \tag{2.10}$$

The main result in this paper is the following theorem.

**Theorem 2.3.** *Assume* **(H1)-(H2)***, and let*

$$\lambda_0 = \left(3\sqrt{\pi}\|\sigma^{-1}\|\|b\|_\infty + \sqrt{9\pi\|\sigma^{-1}\|^2\|b\|_\infty^2 + 6\|b\|_\infty}\right)^2, \tag{2.11}$$

*then the following assertions hold.*

*(1) The following estimate holds:*

$$\mathbb{E}\|X_t^\xi - X_t^\eta\|_\infty^2 \le \inf_{\varepsilon\in(0,1),\lambda\ge\lambda_0} \mathrm{e}^{\frac{25}{8}\lambda_1 r_0} \exp\left\{\mathrm{e}^{\frac{25}{8}\lambda_1 r_0}\left(\Lambda(\lambda,\varepsilon) - \frac{25}{8}\lambda_1 \mathrm{e}^{-\frac{25}{8}\lambda_1 r_0}\right) t\right\} \frac{\|\xi-\eta\|_\infty^2}{1-\varepsilon}$$

*with*

$$\Lambda(\lambda,\varepsilon) := \frac{\frac{3}{4}\lambda + \frac{25}{16}\left\{\frac{2}{5}\lambda_d + \frac{72}{25}\lambda_B + \frac{12}{25}\Upsilon_{b,\sigma,\lambda,\alpha}\|B\|_\infty\right\} + \left(\frac{d}{16} + \frac{25}{16\varepsilon}\right)\|\sigma\|^2\Upsilon_{b,\sigma,\lambda,\alpha}^2}{1-\varepsilon},$$

*and*

(2.12)
$$\Upsilon_{b,\sigma,\lambda,\alpha} = \Gamma(\frac{\alpha}{2})\lambda^{-\frac{\alpha}{2}}\|\sigma\|^\alpha\|\sigma^{-1}\|^2\left(12\kappa + 4\|b\|_\infty + 48(\|\sigma^{-1}\|^2 + \|\sigma^{-1}\|)(3+\lambda^{-1})\|b\|_\infty^2\right).$$

*(2) If there exists $\tilde\varepsilon \in (0,1), \tilde\lambda \ge \lambda_0$ such that*

$$\Lambda(\tilde\lambda,\tilde\varepsilon) < \frac{25}{8}\lambda_1 \mathrm{e}^{-\frac{25}{8}\lambda_1 r_0},$$

*then* (1.3) *has a unique invariant probability measure $\mu$ and for any $t_0 > r_0$,*

$$\max\{\mathbb{W}_2(P_t(\xi,\cdot),\mu), \mathrm{Ent}(P_t(\xi,\cdot)|\mu), \|P_t(\xi,\cdot)-\mu\|_{var}\} \le \kappa_1(\xi)\mathrm{e}^{-\kappa_2 t}, \quad \xi\in\mathscr{C}, t>t_0$$

*for some constants $\kappa_1(\xi), \kappa_2 > 0$, here $\kappa_1(\xi)$ means it depends on $\xi$.*

**Remark 2.4.** *Since $\sup_{\xi\in\mathscr{C}}\kappa_1(\xi) = \infty$ from the proof of Theorem 2.3 below, Theorem 2.3 can not imply the strong exponential ergodicity, i.e. there exist constants $\kappa_1, \kappa_2 > 0$ such that*

$$\sup_{\xi\in\mathscr{C}} \mathbb{W}_2(P_t(\xi,\cdot),\mu) \le \kappa_1 \mathrm{e}^{-\kappa_2 t}.$$

*On the other hand, there are a lot of examples where $\mathbb{W}_2(\mu_n,\mu_0)$ goes to 0 as $n$ goes to infinity, but $\mu_n$ is singular with respect to $\mu_0$ such that $\mathrm{Ent}(\mu_n|\mu_0) = \infty$ and $\|\mu_n - \mu_0\|_{var} =$ 1. Thus, the assertion in (2) is not trivial. Finally, [1, Theorem 1.1 (2),(3))] obtained the exponential convergence in relative entropy and $L^2(\mu)$ by the hyper-contractivity of $P_t$ for large enough $t$. However, it is difficult to establish hyper-contractivity of $P_t$ since the coupling property*

$$\|X_t^\xi - X_t^\eta\|_\infty^2 \le C\mathrm{e}^{-\lambda t}\|\xi-\eta\|_\infty^2$$

*does not hold due to the singularity of $b$.*

# 3 Precise Estimate for Zvonkin's Transform

Since $b$ is singular, we need to construct a regular transform to remove $b$. To this end, for any $\lambda > 0$, consider the following equation

$$u = \int_0^\infty \mathrm{e}^{-\lambda t} P_t^0 \{b + \nabla_b u\} \mathrm{d}t, \tag{3.1}$$

where the semigroup $(P_t^0)_{t>0}$ is generated by $(Z_t^x)_{t\geq 0}$ which solves the SDE

$$\mathrm{d}Z_t^x = AZ_t^x + \sigma \mathrm{d}W(t), \quad Z_0^x = x. \tag{3.2}$$

The following lemma gives a precise estimate for the solution to (3.1), and it is very important in the proof of the exponential convergence.

**Lemma 3.1.** Under **(H1)**, for any $\lambda \geq \lambda_0$ with $\lambda_0$ defined in (2.11),

**(i)** The equation (3.1) has a unique strong solution $\mathbf{u}^\lambda \in C_b^1(\mathbb{R}^d; \mathbb{R}^d)$;

**(ii)** $\|\nabla \mathbf{u}^\lambda\|_\infty \leq \frac{1}{5}$;

**(iii)** $\|\nabla^2 \mathbf{u}^\lambda\|_\infty \leq \frac{1}{5} \Upsilon_{b,\sigma,\lambda,\alpha}$ with $\Upsilon_{b,\sigma,\lambda,\alpha}$ defined in (2.12).

*Proof.* Firstly, it is easy to see that

$$\nabla_\eta Z_t^x = \mathrm{e}^{At}\eta, \quad \nabla_{\eta'} \nabla_\eta Z_t^x = 0, \quad t \geq 0.$$

This combining the Bismut formula [12, (2.8)] implies that

$$\begin{aligned} \nabla_\eta P_t^0 f(x) &= \mathbb{E}\Big( \frac{f(Z_t^x)}{t} \int_0^t \langle \sigma^{-1} \nabla_\eta Z_r^x, \mathrm{d}W(r) \rangle \Big) \\ &= \mathbb{E}\Big( \frac{f(Z_t^x)}{t} \int_0^t \langle \sigma^{-1} \mathrm{e}^{Ar}\eta, \mathrm{d}W(r) \rangle \Big), \quad f \in \mathscr{B}_b(\mathbb{R}^d), t > 0. \end{aligned} \tag{3.3}$$

By the Cauchy-Schwartz inequality and Itô's isometry, we obtain that

$$|\nabla_\eta P_t^0 f|^2(x) \leq \frac{\|\sigma^{-1}\|^2 |\eta|^2 P_t^0 f^2(x)}{t}, \quad f \in \mathscr{B}_b(\mathbb{R}^d), t > 0. \tag{3.4}$$

**(i)** Let $\mathscr{H} = C_b^1(\mathbb{R}^d; \mathbb{R}^d))$, which is a Banach space under the norm

$$\|u\|_{\mathscr{H}} := \|u\|_\infty + \|\nabla u\|_\infty.$$

For any $\lambda > 0$, $u \in \mathscr{H}$, define

$$(\Gamma^\lambda u)(x) = \int_0^\infty \mathrm{e}^{-\lambda t} P_t^0 (\nabla_b u + b)(x) \mathrm{d}t.$$

Then we claim $\Gamma^\lambda \mathscr{H} \subset \mathscr{H}$ for any $\lambda > 0$. In fact, for any $u \in \mathscr{H}$, it holds that

$$
\begin{aligned}
\|\Gamma^\lambda u\|_\infty &= \sup_{x\in\mathbb{R}^d} \left| \int_0^\infty \mathrm{e}^{-\lambda t} P_t^0(\nabla_b u + b)(x)\mathrm{d}t \right| \\
&\le (\|b\|_\infty \|\nabla u\|_\infty + \|b\|_\infty) \int_0^\infty \mathrm{e}^{-\lambda t}\mathrm{d}t \\
&\le \frac{\|b\|_\infty \|\nabla u\|_\infty + \|b\|_\infty}{\lambda} < \infty.
\end{aligned}
$$

By (3.4), we have

$$
\begin{aligned}
\|\nabla\Gamma^\lambda u\|_\infty &= \sup_{x\in\mathbb{R}^d, |\eta|\le 1} \left| \int_0^\infty \mathrm{e}^{-\lambda t} \nabla_\eta P_t^0(\nabla_b u + b)(x)\mathrm{d}t \right| \\
&\le \|\sigma^{-1}\| \int_0^\infty \frac{\mathrm{e}^{-\lambda t}}{\sqrt{t}} (\|b\|_\infty \|\nabla u\|_\infty + \|b\|_\infty)\mathrm{d}t \\
&\le \|\sigma^{-1}\| (\|b\|_\infty \|\nabla u\|_\infty + \|b\|_\infty) \int_0^\infty \frac{\mathrm{e}^{-\lambda t}}{\sqrt{t}}\mathrm{d}t \\
&\le \frac{\sqrt{\pi}\|\sigma^{-1}\|(\|b\|_\infty \|\nabla u\|_\infty + \|b\|_\infty)}{\sqrt{\lambda}} < \infty.
\end{aligned}
$$

So, $\Gamma^\lambda \mathscr{H} \subset \mathscr{H}$ for any $\lambda > 0$. Next, by the fixed-point theorem, it suffices to show that for large enough $\lambda > 0$, $\Gamma^\lambda$ is contractive on $\mathscr{H}$. To do this, for any $u$, $\tilde{u} \in \mathscr{H}$, similarly to the estimates of $\|\Gamma^\lambda u\|_\infty$ and $\|\nabla\Gamma^\lambda u\|_\infty$ above, we obtain that

$$
\begin{aligned}
&\|\Gamma^\lambda u - \Gamma^\lambda \tilde{u}\|_\infty \le \frac{\|b\|_\infty}{\lambda} \|\nabla u - \nabla\tilde{u}\|_\infty, \\
&\|\nabla(\Gamma^\lambda u - \Gamma^\lambda \tilde{u})\|_\infty \le \frac{\sqrt{\pi}\|\sigma^{-1}\|\|b\|_\infty}{\sqrt{\lambda}} \|\nabla u - \nabla\tilde{u}\|_\infty.
\end{aligned} \tag{3.5}
$$

Taking $\lambda > 0$ satisfying

$$
\frac{\sqrt{\pi}\|\sigma^{-1}\|\|b\|_\infty}{\sqrt{\lambda}} + \frac{\|b\|_\infty}{\lambda} \le \frac{1}{6}, \tag{3.6}
$$

then $\Gamma^\lambda$ is contractive on $\mathscr{H}$, which implies that (3.1) has a unique solution $\mathbf{u}^\lambda \in C_b^1(\mathbb{R}^d; \mathbb{R}^d))$ by the fixed-point theorem. Thus, from (3.6), **(i)** holds for

$$
\lambda \ge \lambda_0 = \left(3\sqrt{\pi}\|\sigma^{-1}\|\|b\|_\infty + \sqrt{9\pi\|\sigma^{-1}\|^2\|b\|_\infty^2 + 6\|b\|_\infty}\right)^2.
$$

**(ii)** For any $\lambda \ge \lambda_0$, one infers from (3.1) and (3.4) that

$$
\begin{aligned}
\|\nabla\mathbf{u}^\lambda\|_\infty &\le \int_0^\infty \mathrm{e}^{-\lambda t} \|\nabla P_t^0\{b + \nabla_b \mathbf{u}^\lambda\}\|_\infty \mathrm{d}t \\
&\le \|\sigma^{-1}\|(1 + \|\nabla\mathbf{u}^\lambda\|_\infty)\|b\|_\infty \int_0^\infty \frac{\mathrm{e}^{-\lambda t}}{\sqrt{t}}\mathrm{d}t \\
&\le \lambda^{-\frac{1}{2}} \sqrt{\pi}\|\sigma^{-1}\|\|b\|_\infty(1 + \|\nabla\mathbf{u}^\lambda\|_\infty).
\end{aligned}
$$

This and (3.6) yield **(ii)**.

In the sequel, we intend to verify **(iii)**. From (3.3) and the semigroup property, we have

$$\begin{aligned}\nabla_\eta P_t^0 f(x) &= \nabla_\eta P_{t/2}^0 (P_{t/2}^0 f)(x)\\ &= \mathbb{E}\bigg(\frac{(P_{t/2}^0 f)(Z_{t/2}^x)}{t/2}\int_0^{\frac{t}{2}}\langle\sigma^{-1}\mathrm{e}^{Ar}\eta,\mathrm{d}W(r)\rangle\bigg),\quad t>0, f\in\mathscr{B}_b(\mathbb{R}^d).\end{aligned}$$

This further gives that

$$\frac{1}{2}(\nabla_{\eta'}\nabla_\eta P_t^0 f)(x) = \mathbb{E}\bigg(\frac{(\nabla_{\mathrm{e}^{At/2}\eta'}P_{t/2}^0 f)(Z_{t/2}^x)}{t}\int_0^{t/2}\langle\sigma^{-1}\mathrm{e}^{Ar}\eta,\mathrm{d}W(r)\rangle\bigg),\quad t>0, f\in\mathscr{B}_b(\mathbb{R}^d),$$

where we have used $\nabla_{\eta'}Z_{t/2}^x = \mathrm{e}^{At/2}\eta'$. Thus, applying Cauchy-Schwartz's inequality and Itô's isometry and taking (3.4) into consideration, we derive that

$$\begin{aligned}|\nabla_{\eta'}\nabla_\eta P_t^0 f|^2(x) &\le \frac{4}{t}\|\sigma^{-1}\|^2|\eta|^2\frac{\|\sigma^{-1}\|^2|\eta'|^2 P_t^0 f^2(x)}{t}\\ &= \frac{4\|\sigma^{-1}\|^4|\eta|^2|\eta'|^2 P_t^0 f^2(x)}{t^2},\quad t>0, f\in\mathscr{B}_b(\mathbb{R}^d).\end{aligned}\tag{3.7}$$

Set $\tilde h(\cdot) := h(\cdot)-h(x)$ for fixed $x\in\mathbb{R}^d$ and $h\in\mathscr{B}_b(\mathbb{R}^d)$ which verifies

$$|h(x)-h(y)|\le\tilde\kappa|x-y|^{\tilde\alpha},\qquad x,y\in\mathbb{R}^d\tag{3.8}$$

for some $\tilde\kappa>0$ and $\tilde\alpha\in(0,1)$. Then (3.7) implies that

$$\begin{aligned}|\nabla_{\eta'}\nabla_\eta P_t^0 h|^2(x) = |\nabla_{\eta'}\nabla_\eta P_t^0 \tilde h|^2(x) &\le \frac{4\|\sigma^{-1}\|^4|\eta|^2|\eta'|^2}{t^2}\mathbb{E}|h(Z_t^x)-h(x)|^2\\ &\le \frac{4\|\sigma^{-1}\|^4|\eta|^2|\eta'|^2}{t^2}\tilde\kappa^2\|\sigma\|^{2\tilde\alpha}t^{\tilde\alpha},\quad t>0.\end{aligned}\tag{3.9}$$

where in the second display we have used that

$$Z_t^x - x = \int_0^t \sigma\mathrm{d}W(r),$$

and utilized Jensen's inequality as well as Itô's isometry. Thus, if we can prove that

$$|(b+\nabla_b\mathbf{u}^\lambda)(x)-(b+\nabla_b\mathbf{u}^\lambda)(y)|\le\tilde\kappa|x-y|^{\tilde\alpha},\qquad x,y\in\mathbb{R}^d\tag{3.10}$$

for some $\tilde\kappa>0$ and $\tilde\alpha\in(0,1)$, we get from (3.1) and (3.9) that

$$\begin{aligned}\|\nabla^2\mathbf{u}^\lambda\|_\infty &\le 2\|\sigma^{-1}\|^2\tilde\kappa\|\sigma\|^{\tilde\alpha}\int_0^\infty\frac{\mathrm{e}^{-\lambda t}}{t}t^{\frac{\tilde\alpha}{2}}\mathrm{d}t\\ &= 2\|\sigma^{-1}\|^2\tilde\kappa\|\sigma\|^{\tilde\alpha}\lambda^{-\frac{\tilde\alpha}{2}}\int_0^\infty\frac{\mathrm{e}^{-t}}{t^{1-\frac{\tilde\alpha}{2}}}\mathrm{d}t\\ &= 2\|\sigma^{-1}\|^2\tilde\kappa\|\sigma\|^{\tilde\alpha}\Gamma(\frac{\tilde\alpha}{2})\lambda^{-\frac{\tilde\alpha}{2}}.\end{aligned}\tag{3.11}$$

In the remaining, we intend to prove (3.10). Combining (3.4) and (3.7), we arrive at

$$\begin{aligned}&\|\nabla P_t^0 f(x)-\nabla P_t^0 f(y)\|\\ &\le\left(\frac{2\|\sigma^{-1}\|^2|x-y|}{t}\wedge\frac{2\|\sigma^{-1}\|}{\sqrt{t}}\right)\|f\|_\infty,\quad f\in\mathscr{B}_b(\mathbb{R}^d),t>0,x,y\in\mathbb{R}^d.\end{aligned}\tag{3.12}$$

Thus, for any $f\in\mathscr{B}_b(\mathbb{R}^d)$, $\lambda>0,x,y\in\mathbb{R}^d$, it holds that

$$\begin{aligned}&\left\|\int_0^\infty e^{-\lambda t}\left(\nabla P_t^0 f(x)-\nabla P_t^0 f(y)\right)dt\right\|\\ &\le(2\|\sigma^{-1}\|^2+2\|\sigma^{-1}\|)\|f\|_\infty\int_0^\infty e^{-\lambda t}\left(\frac{|x-y|}{t}\wedge\frac{1}{\sqrt{t}}\right)dt\\ &\le(2\|\sigma^{-1}\|^2+2\|\sigma^{-1}\|)\|f\|_\infty\Bigg(\int_0^{|x-y|^2\wedge e^{-1}}\frac{1}{\sqrt{t}}dt\\ &\qquad+|x-y|\int_{|x-y|^2\wedge e^{-1}}^1\frac{1}{t}dt+|x-y|\int_1^\infty e^{-\lambda t}dt\Bigg)\\ &\le(2\|\sigma^{-1}\|^2+2\|\sigma^{-1}\|)\|f\|_\infty\left((2+\lambda^{-1})|x-y|+|x-y|\log(|x-y|^{-2}\vee e)\right)\\ &\le(2\|\sigma^{-1}\|^2+2\|\sigma^{-1}\|)\|f\|_\infty(3+\lambda^{-1})|x-y|\log(|x-y|^{-2}+e).\end{aligned}\tag{3.13}$$

For any $\lambda\ge\lambda_0$, note from (2.2), **(ii)**, (3.1), (3.4), (3.7) and (3.13) that

$$\begin{aligned}&|(b+\nabla_b\mathbf{u}^\lambda)(x)-(b+\nabla_b\mathbf{u}^\lambda)(y)|\\ &\le(1+\|\nabla\mathbf{u}^\lambda\|_\infty)\kappa|x-y|^\alpha+\|b\|_\infty\|\nabla\mathbf{u}^\lambda(x)-\nabla\mathbf{u}^\lambda(y)\|\mathbf{1}_{\{|x-y|\ge1\}}\\ &\quad+\|b\|_\infty\|\nabla\mathbf{u}^\lambda(x)-\nabla\mathbf{u}^\lambda(y)\|\mathbf{1}_{\{|x-y|\le1\}}\\ &\le\frac{6}{5}\kappa|x-y|^\alpha+\frac{2}{5}\|b\|_\infty|x-y|^\alpha\mathbf{1}_{\{|x-y|\ge1\}}\\ &\quad+\frac{6}{5}(2\|\sigma^{-1}\|^2+2\|\sigma^{-1}\|)(3+\lambda^{-1})\|b\|_\infty^2|x-y|^\alpha\\ &\qquad\times|x-y|^{1-\alpha}\log\left(e+\frac{1}{|x-y|^2}\right)\mathbf{1}_{\{|x-y|\le1\}}\\ &\le\left(\frac{6}{5}\kappa+\frac{2}{5}\|b\|_\infty+\frac{24}{5}(\|\sigma^{-1}\|^2+\|\sigma^{-1}\|)(3+\lambda^{-1})\|b\|_\infty^2\right)|x-y|^\alpha,\end{aligned}$$

where in the third inequality we have used the fact that the function $[0,1]\ni x\mapsto x^{1-\alpha}\log(e+\frac{1}{x^2})$ is non-decreasing. Thus, (3.10) holds for $\tilde\kappa=\frac{6}{5}\kappa+\frac{2}{5}\|b\|_\infty+\frac{24}{5}(\|\sigma^{-1}\|^2+\|\sigma^{-1}\|)(3+\lambda^{-1})\|b\|_\infty^2$ and $\tilde\alpha=\alpha$. From (3.11), we finish the proof. □

**Remark 3.2.** *In the multiplicative noise case, $\nabla_\eta Z_t^x$ is a random variable, precise estimate for the solution of (3.1), especially $\|\nabla^2\mathbf{u}^\lambda\|_\infty$, is so sophisticated that we only consider the additive noise case in this paper.*

# 4 Proof of Theorem 2.3

**Lemma 4.1.** *Assume* **(H1)**-**(H2)** *and*

$$\lambda_B < \lambda_1 \mathrm{e}^{-2\lambda_1 r_0},$$

*then*

$$\sup_{t\geq 0} \mathbb{E}\|X_t^\xi\|_\infty^2 < \infty \tag{4.1}$$

*Proof.* For simplicity, we denote $X^\xi(t)$ by $X(t)$. Itô's formula implies that

$$\begin{aligned}
\mathrm{d}|X(t)|^2 =& 2\langle X(t), AX(t))\rangle \,\mathrm{d}t + 2\langle X(t), b(X(t))\rangle \,\mathrm{d}t + 2\langle X(t), \sigma\,\mathrm{d}W(t)\rangle \\
&+ 2\langle X(t), B(X_t)\rangle \,\mathrm{d}t + \|\sigma\|_{HS}^2 \,\mathrm{d}t.
\end{aligned}$$

Let $\xi_0(s) = 0, s \in [-r_0, 0]$. It follows from (2.3) that

$$\langle X(t), B(X_t)\rangle \leq |X(t)||B(X_t)| \leq \lambda_B \|X_t\|_\infty^2 + |B(\xi_0)||X(t)|$$

Since $\lambda_B < \lambda_1 \mathrm{e}^{-2\lambda_1 r_0}$, we can take small enough $\varepsilon \in (0,1)$ and $\epsilon \in (0, 2\lambda_1)$ such that

$$\frac{2\lambda_B}{1-\varepsilon} < (2\lambda_1 - \epsilon)\mathrm{e}^{-(2\lambda_1-\epsilon)r_0}.$$

This together with Young's inequality and **(H2)** yields that

$$\begin{aligned}
\mathrm{d}|X(t)|^2 \leq& -2\lambda_1|X(t)|^2\mathrm{d}t + \{2(\|b\|_\infty + |B(\xi_0)|)|X(t)| + \|\sigma\|_{HS}^2\}\mathrm{d}t \\
&+ 2\lambda_B\|X_t\|_\infty^2\mathrm{d}t + 2\langle X(t), \sigma\,\mathrm{d}W(t)\rangle \\
\leq& (-2\lambda_1 + \epsilon)|X(t)|^2\mathrm{d}t + \left\{\frac{1}{\epsilon}(\|b\|_\infty + |B(\xi_0)|)^2 + \|\sigma\|_{HS}^2\right\}\mathrm{d}t \\
&+ 2\lambda_B\|X_t\|_\infty^2\mathrm{d}t + 2\langle X(t), \sigma\,\mathrm{d}W(t)\rangle.
\end{aligned} \tag{4.2}$$

Thus it is not difficult to see that

$$\begin{aligned}
\mathrm{d}\mathrm{e}^{(2\lambda_1-\epsilon)t}|X(t)|^2\mathrm{d}t \leq& 2\lambda_B\mathrm{e}^{(2\lambda_1-\epsilon)t}\|X_t\|_\infty^2\mathrm{d}t \\
&+ \mathrm{e}^{(2\lambda_1-\epsilon)t}\left\{\frac{1}{\epsilon}(\|b\|_\infty + |B(\xi_0)|)^2 + \|\sigma\|_{HS}^2\right\}\mathrm{d}t \\
&+ 2\mathrm{e}^{(2\lambda_1-\epsilon)t}\langle X(t), \sigma\,\mathrm{d}W(t)\rangle.
\end{aligned} \tag{4.3}$$

Let $\eta_r = \sup_{s\in[-r_0,r]} \mathrm{e}^{(2\lambda_1-\epsilon)s^+}|X(s)|^2$, then

$$\begin{aligned}
\mathbb{E}\eta_r \leq& \|\xi\|_\infty^2 + 2\lambda_B\mathrm{e}^{(2\lambda_1-\epsilon)r_0}\mathbb{E}\int_0^r \eta_t\mathrm{d}t \\
&+ \int_0^r \mathrm{e}^{(2\lambda_1-\epsilon)t}\left\{\frac{1}{\epsilon}(\|b\|_\infty + |B(\xi_0)|)^2 + \|\sigma\|_{HS}^2\right\}\mathrm{d}t \\
&+ \mathbb{E}\sup_{s\in[0,r]}\int_0^s 2\mathrm{e}^{(2\lambda_1-\epsilon)t}\langle X(t), \sigma\,\mathrm{d}W(t)\rangle
\end{aligned} \tag{4.4}$$

On the other hand, BDG inequality and Young's inequality imply that

$$\begin{aligned}
&\mathbb{E}\sup_{s\in[0,r]}\int_0^s 2\mathrm{e}^{(2\lambda_1-\epsilon)t}\left\langle X(t),\sigma\,\mathrm{d}W(t)\right\rangle\\
&\le \mathbb{E}\left\{\int_0^r 16\mathrm{e}^{(4\lambda_1-2\epsilon)t}\|\sigma\|^2\left|X(t)\right|^2\mathrm{d}t\right\}^{\frac{1}{2}}\\
&\le \varepsilon\mathbb{E}\eta_r+\int_0^r \mathrm{e}^{(2\lambda_1-\epsilon)t}\times\frac{4\|\sigma\|^2}{\varepsilon}\mathrm{d}t.
\end{aligned}\tag{4.5}$$

Let $\beta:=\frac{\frac{1}{\epsilon}(\|b\|_\infty+|B(\xi_0)|)^2+\|\sigma\|_{HS}^2+\frac{4\|\sigma\|^2}{\varepsilon}}{1-\varepsilon}$. Combining (4.4) and (4.5), we have

$$\mathbb{E}\eta_r\le\frac{\|\xi\|_\infty^2}{1-\varepsilon}+\frac{2\lambda_B\mathrm{e}^{(2\lambda_1-\epsilon)r_0}}{1-\varepsilon}\mathbb{E}\int_0^r\eta_t\mathrm{d}t+\beta\int_0^r\mathrm{e}^{(2\lambda_1-\epsilon)t}\mathrm{d}t.\tag{4.6}$$

By (2.4), Gronwall's inequality implies that

$$\begin{aligned}
\mathbb{E}\eta_t\le&\exp\left\{\frac{2\lambda_B\mathrm{e}^{(2\lambda_1-\epsilon)r_0}}{1-\varepsilon}t\right\}\frac{\|\xi\|_\infty^2}{1-\varepsilon}\\
&+\beta\int_0^t\exp\left\{\frac{2\lambda_B\mathrm{e}^{(2\lambda_1-\epsilon)r_0}}{1-\varepsilon}(t-s)+(2\lambda_1-\epsilon)s\right\}\mathrm{d}s.
\end{aligned}\tag{4.7}$$

Noting that $\mathbb{E}\eta_t\ge\mathrm{e}^{(t-r_0)(2\lambda_1-\epsilon)}\mathbb{E}\|X_t\|_\infty^2$, we obtain that

$$\begin{aligned}
\mathbb{E}\|X_t\|_\infty^2\le&\,\mathrm{e}^{r_0(2\lambda_1-\epsilon)}\exp\left\{\left(\frac{2\lambda_B\mathrm{e}^{(2\lambda_1-\epsilon)r_0}}{1-\varepsilon}-(2\lambda_1-\epsilon)\right)t\right\}\frac{\|\xi\|_\infty^2}{1-\varepsilon}\\
&+\mathrm{e}^{r_0(2\lambda_1-\epsilon)}\beta\int_0^t\exp\left\{\left(\frac{2\lambda_B\mathrm{e}^{(2\lambda_1-\epsilon)r_0}}{1-\varepsilon}-(2\lambda_1-\epsilon)\right)(t-s)\right\}\mathrm{d}s.
\end{aligned}\tag{4.8}$$

Since $\frac{2\lambda_B}{1-\varepsilon}<(2\lambda_1-\epsilon)\mathrm{e}^{-(2\lambda_1-\epsilon)r_0}$, we conclude that (4.1) holds. □

*Proof of Theorem 2.3.* (1) Let $X$ and $\bar{X}$ be solutions to (1.3) with $X_0=\xi$, $\bar{X}_0=\eta$, then

$$\begin{aligned}
\mathrm{d}X(t)&=\{AX(t)+b(X(t))+B(X_t)\}\mathrm{d}t+\sigma\mathrm{d}W(t),\quad X_0=\xi,\\
\mathrm{d}\bar{X}(t)&=\{AX(t)+b(\bar{X}(t))+B(\bar{X}_t)\}\mathrm{d}t+\sigma\mathrm{d}W(t),\quad \bar{X}_0=\eta.
\end{aligned}\tag{4.9}$$

For any $\lambda\ge\lambda_0$, let $\theta^\lambda(x)=x+\mathbf{u}^\lambda(x)$. Combining (3.1) and Lemma 3.1, we have

$$\frac{1}{2}\mathrm{Tr}(\sigma\sigma^*\nabla^2\mathbf{u}^\lambda)+\nabla_b\mathbf{u}^\lambda+b+\nabla_{A\cdot}\mathbf{u}^\lambda=\lambda\mathbf{u}^\lambda.\tag{4.10}$$

By (4.9), (4.10) and Itô's formula, we have

$$\mathrm{d}\theta^\lambda(X(t))=\{AX(t)+\lambda\mathbf{u}^\lambda(X(t))+\nabla\theta^\lambda(X(t))B(X_t)\}\mathrm{d}t+\nabla\theta^\lambda(X(t))\sigma\,\mathrm{d}W(t),$$

$$\mathrm{d}\theta^\lambda(\bar X(t)) = \{A\bar X(t) + \lambda \mathbf{u}^\lambda(\bar X(t)) + \nabla\theta^\lambda(\bar X(t))B(\bar X_t)\}\mathrm{d}t + \nabla\theta^\lambda(\bar X(t))\sigma\,\mathrm{d}W(t).$$

So, letting $\xi(t) = \theta^\lambda(X(t)) - \theta^\lambda(\bar X(t))$, we arrive at

$$\begin{aligned}
\mathrm{d}|\xi(t)|^2 =& 2\langle AX(t) - A\bar X(t), \theta^\lambda(X(t)) - \theta^\lambda(\bar X(t))\rangle \\
& 2\lambda \left\langle \xi(t), \mathbf{u}^\lambda(X(t)) - \mathbf{u}^\lambda(\bar X(t))\right\rangle \mathrm{d}t \\
& + 2\left\langle \xi(t), [\nabla\theta^\lambda(X(t))B(X_t) - \nabla\theta^\lambda(\bar X(t))B(\bar X_t)]\right\rangle \mathrm{d}t \\
& + 2\left\langle \xi(t), [\nabla\theta^\lambda(X(t)) - \nabla\theta^\lambda(\bar X(t))]\sigma \mathrm{d}W(t)\right\rangle \\
& + \left\|[\nabla\theta^\lambda(X(t)) - \nabla\theta^\lambda(\bar X(t))]\sigma\right\|_{HS}^2 \mathrm{d}t
\end{aligned} \tag{4.11}$$

Firstly, it is easy to see that

$$\begin{aligned}
& 2\langle AX(t) - A\bar X(t), \theta^\lambda(X(t)) - \theta^\lambda(\bar X(t))\rangle \\
=& 2\langle AX(t) - A\bar X(t), X(t) - \bar X(t) + \mathbf{u}^\lambda(X(t)) - \mathbf{u}^\lambda(\bar X(t))\rangle \\
\le& -2\lambda_1|X(t) - \bar X(t)|^2 + \frac{2\lambda_d}{5}|X(t) - \bar X(t)|^2
\end{aligned}$$

By Lemma 3.1 and **(H2)**, it holds that

$$\begin{aligned}
& 2\left\langle \xi(t), [\nabla\theta^\lambda(X(t))B(X_t) - \nabla\theta^\lambda(\bar X(t))B(\bar X_t)]\right\rangle \\
& \le \frac{12}{5}|X(t) - \bar X(t)|\|\nabla\theta^\lambda(X(t)) - \nabla\theta^\lambda(\bar X(t))\||B(X_t)| \\
& + \frac{12}{5}|X(t) - \bar X(t)|\|\nabla\theta^\lambda(\bar X(t))\||B(X_t) - B(\bar X_t)| \\
& \le \frac{12}{25}\Upsilon_{b,\sigma,\lambda,\alpha}\|B\|_\infty|X(t) - \bar X(t)|^2 + \frac{72}{25}\lambda_B\|X_t - \bar X_t\|_\infty^2,
\end{aligned}$$

and

$$2\lambda|\xi(t)| \cdot |\mathbf{u}^\lambda(X(t)) - \mathbf{u}^\lambda(\bar X(t))| \le \frac{12}{25}\lambda\ |X(t) - \bar X(t)|^2. \tag{4.12}$$

Moreover,

$$\left\|[\nabla\theta^\lambda(X(t)) - \nabla\theta^\lambda(\bar X(t))]\sigma\right\|_{HS}^2 \le \frac{1}{25}\Upsilon_{b,\sigma,\lambda,\alpha}^2\|\sigma\|_{HS}^2|X(t) - \bar X(t)|^2. \tag{4.13}$$

Let

$$\Lambda_1(\lambda) := \frac{3}{4}\lambda + \frac{25}{16}\left\{\frac{2}{5}\lambda_d + \frac{72}{25}\lambda_B + \frac{12}{25}\Upsilon_{b,\sigma,\lambda,\alpha}\|B\|_\infty + \frac{1}{25}\Upsilon_{b,\sigma,\lambda,\alpha}^2\|\sigma\|_{HS}^2\right\}.$$

Since $\|\nabla\theta^\lambda(x)\| \ge \frac{4}{5}$ for any $x \in \mathbb{R}^d$, we have

$$\begin{aligned}
\mathrm{d}|X(t) - \bar X(t)|^2 \le& -\frac{25}{8}\lambda_1|X(t) - \bar X(t)|^2\mathrm{d}t \\
& + \Lambda_1(\lambda)\|X_t - \bar X_t\|_\infty^2\mathrm{d}t \\
& + \frac{25}{8}\left\langle \xi(t), [\nabla\theta^\lambda(X(t)) - \nabla\theta^\lambda(\bar X(t))]\sigma\mathrm{d}W(t)\right\rangle.
\end{aligned} \tag{4.14}$$

Similarly to (4.3), it holds that

$$\mathrm{d}e^{\frac{25}{8}\lambda_1 t}|X(t)-\bar{X}(t)|^2\mathrm{d}t \leq \Lambda_1(\lambda)e^{\frac{25}{8}\lambda_1 t}\|X_t-\bar{X}_t\|_\infty^2 + \frac{25}{8}e^{\frac{25}{8}\lambda_1 t}\left\langle \xi(t), [\nabla\theta^\lambda(X(t))-\nabla\theta^\lambda(\bar{X}(t))]\sigma\mathrm{d}W(t)\right\rangle.$$

Set $\gamma_r = \sup_{s\in[-r_0,r]} e^{\frac{25}{8}\lambda_1 s^+}|X(s)-\bar{X}(s)|^2$ and we get

$$\begin{aligned}\mathbb{E}\gamma_r &\leq \|\xi-\eta\|_\infty^2 + \Lambda_1(\lambda)e^{\frac{25}{8}\lambda_1 r_0}\mathbb{E}\int_0^r \gamma_t\mathrm{d}t \\ &\quad + \mathbb{E}\sup_{s\in[0,r]}\int_0^s \frac{25}{8}e^{\frac{25}{8}\lambda_1 t}\left\langle \xi(t), [\nabla\theta^\lambda(X(t))-\nabla\theta^\lambda(\bar{X}(t))]\sigma\mathrm{d}W(t)\right\rangle.\end{aligned} \tag{4.15}$$

Again, BDG inequality and Young's inequality yield that for any $\varepsilon\in(0,1)$,

$$\begin{aligned}&\mathbb{E}\sup_{s\in[0,r]}\int_0^s \frac{25}{8}e^{\frac{25}{8}\lambda_1 t}\left\langle \xi(t), [\nabla\theta^\lambda(X(t))-\nabla\theta^\lambda(\bar{X}(t))]\sigma\mathrm{d}W(t)\right\rangle \\ &\leq 2\mathbb{E}\left\{\int_0^r \frac{625}{64}e^{\frac{25}{4}\lambda_1 t}\left|\sigma^*[\nabla\theta^\lambda(X(t))-\nabla\theta^\lambda(\bar{X}(t))]^*\xi(t)\right|^2\mathrm{d}t\right\}^{\frac{1}{2}} \\ &\leq 2\mathbb{E}\sqrt{\frac{25}{16}\Upsilon_{b,\sigma,\lambda,\alpha}^2\|\sigma\|^2\int_0^r e^{\frac{25}{4}\lambda_1 t}|X(t)-\bar{X}(t)|^4\mathrm{d}t} \\ &\leq \varepsilon\mathbb{E}\gamma_r + \frac{25\Upsilon_{b,\sigma,\lambda,\alpha}^2}{16\varepsilon}\|\sigma\|^2\mathbb{E}\int_0^r\gamma_t\mathrm{d}t.\end{aligned} \tag{4.16}$$

Substituting (4.16) into (4.15), we obtain

$$\mathbb{E}\gamma_r \leq \frac{\|\xi-\eta\|_\infty^2}{1-\varepsilon} + \frac{\Lambda_1(\lambda)e^{\frac{25}{8}\lambda_1 r_0} + \frac{25\Upsilon_{b,\sigma,\lambda,\alpha}^2}{16\varepsilon}\|\sigma\|^2}{1-\varepsilon}\mathbb{E}\int_0^r\gamma_t\mathrm{d}t. \tag{4.17}$$

Again by (2.4), Gronwall's inequality implies that

$$\mathbb{E}\gamma_t \leq \exp\left\{\frac{\Lambda_1(\lambda)e^{\frac{25}{8}\lambda_1 r_0} + \frac{25\Upsilon_{b,\sigma,\lambda,\alpha}^2}{16\varepsilon}\|\sigma\|^2}{1-\varepsilon}t\right\}\frac{\|\xi-\eta\|_\infty^2}{1-\varepsilon}. \tag{4.18}$$

Noting that $\mathbb{E}\gamma_t \geq e^{(t-r_0)\frac{25}{8}\lambda_1}\mathbb{E}\|X_t-\bar{X}_t\|_\infty^2$, we obtain

$$\begin{aligned}&\mathbb{E}\|X_t-\bar{X}_t\|_\infty^2 \\ &\leq e^{\frac{25}{8}\lambda_1 r_0}\exp\left\{e^{\frac{25}{8}\lambda_1 r_0}\left(\frac{\Lambda_1(\lambda) + \frac{25\Upsilon_{b,\sigma,\lambda,\alpha}^2}{16\varepsilon}\|\sigma\|^2}{1-\varepsilon} - \frac{25}{8}\lambda_1 e^{-\frac{25}{8}\lambda_1 r_0}\right)t\right\}\frac{\|\xi-\eta\|_\infty^2}{1-\varepsilon}.\end{aligned} \tag{4.19}$$

Thus, letting

$$\Lambda(\lambda,\varepsilon) := \frac{\Lambda_1(\lambda) + \frac{25\Upsilon^2_{b,\sigma,\lambda,\alpha}}{16\varepsilon}\|\sigma\|^2}{1-\varepsilon},$$

combining the definition of $\Lambda_1(\lambda)$ and noting $\|\sigma\|^2_{HS} \le d\|\sigma\|^2$, we prove (1).

(2) Now, if there exists $\tilde{\varepsilon} \in (0,1), \tilde{\lambda} \ge \lambda_0$ such that

$$\Lambda(\tilde{\lambda},\tilde{\varepsilon}) < \frac{25}{8}\lambda_1 \mathrm{e}^{-\frac{25}{8}\lambda_1 r_0},$$

then there exists constants $\kappa_0, \kappa_2 > 0$ such that

$$\mathbb{W}_2(P_t(\xi,\cdot,P_t(\eta,\cdot))^2 \le \mathbb{E}\|X_t - \bar{X}_t\|^2_\infty \le \kappa_0 \mathrm{e}^{-\kappa_2 t}\|\xi-\eta\|^2_\infty. \tag{4.20}$$

This yields that for any $0 < t < s$,

$$\mathbb{E}\|X^\xi_t - X^\xi_s\|^2_\infty = \mathbb{E}\left\|X^\xi_t - X_t^{X^\xi_{s-t}}\right\|^2_\infty \le \kappa_0\mathrm{e}^{-\kappa_2 t}\mathbb{E}\|\xi - X^\xi_{s-t}\|^2_\infty, \tag{4.21}$$

Noting that $\Lambda(\tilde{\lambda},\tilde{\varepsilon}) \ge \frac{\Lambda_1(\lambda)}{1-\tilde{\varepsilon}} \ge \frac{\frac{25}{8}\lambda_B}{1-\tilde{\varepsilon}}$ and

$$\Lambda(\tilde{\lambda},\tilde{\varepsilon}) < \frac{25}{8}\lambda_1 \mathrm{e}^{-\frac{25}{8}\lambda_1 r_0},$$

we conclude that

$$\lambda_B < \lambda_1 \mathrm{e}^{-2\lambda_1 r_0}.$$

According to Lemma 4.1, it holds that

$$\sup_{r\ge 0}\mathbb{E}\|X^\xi_r\|^2_\infty < \infty.$$

Thus,

$$\mathbb{W}_2(P_t(\xi,\cdot),P_s(\xi,\cdot))^2 \le \mathbb{E}\|X^\xi_t - X^\xi_s\|^2_\infty \le \kappa_1(\xi)\mathrm{e}^{-\kappa_2 t} \tag{4.22}$$

holds for some positive constant $\kappa_1(\xi)$ depending on $\xi$. So, there exists a probability measure $\mu_\xi$ such that

$$\mathbb{W}_2(P_t(\xi,\cdot),\mu_\xi)^2 \le \kappa_1(\xi)\mathrm{e}^{-\kappa_2 t}. \tag{4.23}$$

It remains to prove that $\mu_\xi$ does not depend on $\xi$. For any $\xi,\eta \in \mathscr{C}$,

$$\begin{aligned}\mathbb{W}_2(\mu_\xi,\mu_\eta)^2 &\le \mathbb{W}_2(P_t(\xi,\cdot),\mu_\xi)^2 + \mathbb{W}_2(P_t(\xi,\cdot),P_t(\eta,\cdot))^2 + \mathbb{W}_2(P_t(\eta,\cdot),\mu_\eta)^2 \\ &\le \kappa_1(\xi)\mathrm{e}^{-\kappa_2 t} + \kappa_0\mathrm{e}^{-\kappa_2 t}\|\xi-\eta\|^2_\infty + \kappa_1(\eta)\mathrm{e}^{-\kappa_2 t}.\end{aligned} \tag{4.24}$$

Letting $t \to \infty$, we obtain $\mu_\xi = \mu_\eta$.

Next, for any $t_0 > r_0$ and $t > t_0$, by the semigroup property, (2.9) and (4.22), we have

$$\begin{aligned}\mathrm{Ent}(P_t(\xi,\cdot)|P_t(\eta,\cdot)) &= \mathrm{Ent}(P_{t-t_0}(\xi,\cdot)P_{t_0}|P_{t-t_0}(\eta,\cdot)P_{t_0})\\ &\le \frac{C(t_0)}{(t_0-r_0)\wedge 1}\mathbb{W}_2(P_{t-t_0}(\xi,\cdot),P_{t-t_0}(\eta,\cdot))^2\\ &\le \frac{C(t_0)}{(t_0-r_0)\wedge 1}\kappa_0 \mathrm{e}^{\kappa_2 t_0}\mathrm{e}^{-\kappa_2 t}\|\xi-\eta\|_\infty^2.\end{aligned}$$

Thus (1.2) implies that for any $t > t_0$,

$$\|P_t(\xi,\cdot)-P_t(\eta,\cdot)\|_{var}^2 \le \frac{C(t_0)}{2(t_0-r_0)\wedge 2}\kappa_0 \mathrm{e}^{\kappa_2 t_0}\mathrm{e}^{-\kappa_2 t}\|\xi-\eta\|_\infty^2. \tag{4.25}$$

Combining (2.9), (2.10), (4.23) and the semigroup property, since $\mu$ is the invariant probability measure, we have

$$\begin{aligned}\mathrm{Ent}(P_t(\xi,\cdot)|\mu) &= \mathrm{Ent}(P_{t-t_0}(\xi,\cdot)P_{t_0}|\mu P_{t_0})\\ &\le \frac{C(t_0)}{(t_0-r_0)\wedge 1}\mathbb{W}_2(P_{t-t_0}(\xi,\cdot),\mu)^2\\ &\le \frac{C(t_0)}{(t_0-r_0)\wedge 1}\kappa_1(\xi)\mathrm{e}^{\kappa_2 t_0}\mathrm{e}^{-\kappa_2 t}.\end{aligned}$$

and

$$\|P_t(\xi,\cdot)-\mu\|_{var}^2 \le \frac{C(t_0)}{2(t_0-r_0)\wedge 2}\kappa_1(\xi)\mathrm{e}^{\kappa_2 t_0}\mathrm{e}^{-\kappa_2 t}.$$

Thus, we complete the proof. □

**Acknowledgement.** The author would like to thank Professor Feng-Yu Wang for helpful comments.

# References

[1] J. Bao, F.-Y. Wang, C. Yuan, *Hypercontractivity for functional stochastic differential equations,* Stochastic Process. Appl. 125 (2015), 3636-3656.

[2] J. Bao, F.-Y. Wang, C. Yuan, *Hypercontractivity for Functional Stochastic Partial Differential Equations,* Electron. J. Probab. 20 (2015), 1-15.

[3] I. Csiszár, J. Körne, *Information Theory: Coding Theorems for Discrete Memoryless Systems,* Academic Press, New York, 1981.

[4] L. Gyongy, T. Martinez, *On stochastic differential equations with locally unbounded drift,* Czechoslovak Math. J. 51(2001), 763-783.

[5] Xing Huang, *Strong Solutions for Functional SDEs with Singular Drift,* Stoch. Dyn., 18, 1850015 (2018), 19 pages. https://doi.org/10.1142/S0219493718500156.

[6] X. Huang, F.-Y. Wang, *Functional SPDE with Multiplicative Noise and Dini Drift,* Ann. Fac. Sci. Toulouse Math. 6 (2017), 519-537.

[7] X. Huang, F.-Y. Wang, *Distribution Dependent SDEs with Singular Coefficients,* arXiv:1805.01682.

[8] X. Huang, S.-Q. Zhang, *Mild Solutions and Harnack Inequality for Functional S-tochastic Partial Differential Equations with Dini Drift,* J Theor Probab (2018). https://doi.org/10.1007/s10959-018-0830-4.

[9] N. V. Krylov, M. Röckner,*Strong solutions of stochastic equations with singular time dependent drift,* Probab. Theory Related Fields 131 (2005), 154-196.

[10] M. S. Pinsker, *Information and Information Stability of Random Variables and Processes,* Holden-Day, San Francisco, 1964.

[11] F.-Y. Wang, *Harnack Inequality and Applications for Stochastic Partial Differential Equations,* Springer, New York, 2013.

[12] F.-Y. Wang, *Gradient estimate and applications for SDEs in Hilbert space with multiplicative noise and Dini continuous drift,* J. Differential Equations 260 (2016), 2792-2829.

[13] X. Zhang, *Strong solutions of SDEs with singural drift and Sobolev diffusion coefficients,* Stoch. Proc. Appl. 115(2005), 1805–1818.

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

[15] X. Zhang, *Stochastic differential equations with Sobolev didiffusion and singular drifts and applications,* Ann. Appl. Probab. 26 (2016), 2697-2732

[16] A. K. Zvonkin, *A transformation of the phase space of a diffusion process that removes the drift,* Math. Sb. 93(1974), 129-149, 152.