# Transportation cost inequalities for SDEs with irregular drifts


**Yongqiang Suo[a)], Chenggui Yuan[b)], Shao-Qin Zhang[c)]**

[a)]School of Mathematics and Statistics, Central South University, Changsha 410083, China

Email: Suoyongqiang11@163.com

[b)]Department of Mathematics, Swansea University, Bay Campus SA1 8EN, UK.

Email: C.Yuan@swansea.ac.uk

[c)]School of Statistics and Mathematics

Central University of Finance and Economics, Beijing 100081, China

Email: zhangsq@cufe.edu.cn


November 23, 2021


**Abstract**

In this paper, we establish the quadratic transportation cost inequality for SDEs with Dini continuous drift and the $W_1$-transportation cost inequality for SDEs with singular coefficients by using the stability of the Wasserstein distance and relative entropy of measures under the homeomorphism induced by Zvonkin's transformation.



AMS Subject Classification: 60E15, 60H10, 60J60

Keywords: Stochastic differential equations, Zvonkin transformation, Transportation cost inequality.


## 1 Introduction

Let $(E,\rho)$ be a metric space equipped with a $\sigma$-field $\mathscr{B}$ such that $\rho(\cdot,\cdot)$ is $\mathscr{B}\times\mathscr{B}$ measurable and let $\mathscr{P}(E)$ be the class of all probability measures on $E$. Given $p\geq 1$, the $L^p$-Wasserstein distance between $\mu,\nu\in\mathscr{P}(E)$ is defined by

$$\mathbb{W}_p^\rho(\mu,\nu)=\inf_{\pi\in\mathscr{C}(\mu,\nu)}\Big(\int_{E\times E}\rho^p(x,y)\pi(\mathrm{d}x,\mathrm{d}y)\Big)^{\frac{1}{p}},$$

where $\mathscr{C}(\mu,\nu)$ is the space of all couplings of $\mu$ and $\nu$. The relative entropy of $\nu$ with respect to $\mu$ is given by

$$H(\nu|\mu) = \begin{cases} \int_E \log \frac{\mathrm{d}\nu}{\mathrm{d}\mu}\mathrm{d}\nu, & \text{if } \nu \ll \mu, \\ +\infty, & \text{otherwise.} \end{cases}$$

We say that the probability measure $\mu$ satisfies the $W_p$-transportation cost inequality (TCI for short) on $(E,\rho)$ if there exists a constant $C>0$ such that for any probability measure $\nu$,

$$W_p^\rho(\mu,\nu) \le \sqrt{2CH(\nu|\mu)}.$$

To be short, we write $\mu \in T_p(C)$ for this relation.

Since Talagrand's work [17], the $T_1(C)$ and the $T_2(C)$ have been intensively investigated and applied to many distributions, such as [3, 6, 23, 25] for diffusion processes, [11, 12, 15, 22] for stochastic differential equations (SDEs) with Lévy noise or fractional Brownian motion, [2, 18] for stochastic functional differential equations (SFDEs). As for $T_1(C)$, it is related to the phenomenon of Gaussian concentration, see [6, 10]. Moreover, we highlight that [6] gave an equivalent characterization of $T_1(C)$ by "Gaussian tail" on a metric space and some applications to random dynamic systems and diffusions. Using Malliavin calculus, [11, 22] proved $T_1(C)$ for the invariant probability measure and for the process-level law on the path space w.r.t. the $L^1$-metric and uniform metric of the solution to SDEs with jumps under dissipative conditions. By using the mirror coupling for the jump part and the coupling by reflection for the Brownian part, [13] extended some TCIs to non-globally dissipative SDEs with jumps. As for $T_2(C)$, it implies the dimension-free concentration of measure, see [10, 16]. However, $T_2(C)$ is stronger than $T_1(C)$ since $W_1^\rho \le W_2^\rho$, and it has been brought into relation with some functional inequalities such as Poincaré inequality and log-Sobolev inequality, see [4, 6, 7, 8, 10, 14, 19, 25] and references therein. For instance, $T_2(C)$ can be derived from the log-Sobolev inequality, and $T_2(C)$ implies the Poincaré inequality, see e.g. [4, 14]. Moreover, the $T_2(C)$ can also be established when the log-Sobolev inequality is unknown, see for instance [2, 6] and references therein.

It is worth noting that most of the above references of TCIs for laws of solutions to SDEs and SFDEs are required to meet Lipschitz condition for the drifts, some references relaxed this condition to the case with one-sided Lipschitz condition. Recently, [1] used the Girsanov transformation and the Krylov estimate to obtain $T_2(C)$ for law of solution to SDEs with measurable drift. However, the drift term does not have growth at infinity. Motivated by [20, 26, 27], we aim to establish $T_1(C)$ and $T_2(C)$ for SDEs with a Lipschitzian drift perturbed by an irregular term. To this aim, we use the Zvonkin transformation, which induces a homeomorphism on the state space. Following the idea introduced by [6], we give stability results on the Wasserstein distance and relative entropy of measures defined on a polish space under the homeomorphism, which leads to the stability of $T_p(C)$. Based on [21] and the "Gaussian tail" characterization of $T_1(C)$, $T_2(C)$ and $T_1(C)$ on the free path space are established for SDEs with irregular drifts.

The remainder of the paper is organized as follows: in Section 2, we present a general result on $T_p(C)$ for measure $\mu$ on Polish space $(E,\rho)$; in Section 3, the main results including the $T_2(C)$ for SDEs with Dini continuous drift and $T_1(C)$ for SDEs with singular coefficients are introduced; by the general results in Section 2, the $T_2(C)$ for SDE (3.1) and the $T_1(C)$ for SDE (3.5) are proved in Section 4 and Section 5, respectively.

## 2 A general result

sec:general

Let $(E,\rho)$ be a Polish space, $\mathscr{B}$ be the Borel $\sigma$-field and $\Phi$ be a homeomorphism on $E$. We can see that $\Phi$ induces a homeomorphism on $E\times E$, which is still denoted by $\Phi$:

$$\Phi(x,y)=(\Phi(x),\Phi(y))\,,\ (x,y)\in E\times E.$$

It is clear that the inverse of $\Phi$ on $E\times E$ is given by

$$\Phi^{-1}(x,y)=\left(\Phi^{-1}(x),\Phi^{-1}(y)\right)\,,\ x,y\in E.$$

We can now formulate the following result. The proof is direct, and we give the details for readers' convenience.

gthm **Lemma 2.1.** *For any $p\geq 1,\mu,\nu\in\mathscr{P}(E)$, the following assertions hold.*

*(1) $\mathbb{W}_p^\rho(\mu,\nu)=\mathbb{W}_p^{\rho\circ\Phi^{-1}}(\mu\circ\Phi^{-1},\nu\circ\Phi^{-1})$. Moreover, if there are positive constants $c_1$ and $c_2$ such that*

$$c_1\rho(x,y)\leq\rho(\Phi(x),\Phi(y))\leq c_2\rho(x,y),\ x,y\in E, \tag{2.1}$$ inq-1

*then*

$$c_1\mathbb{W}_p^\rho(\mu,\nu)\leq\mathbb{W}_p^\rho(\mu\circ\Phi^{-1},\nu\circ\Phi^{-1})\leq c_2\mathbb{W}_p^\rho(\mu,\nu). \tag{2.2}$$ inq-2

*(2) $H(\nu|\mu)=H(\nu\circ\Phi^{-1}|\mu\circ\Phi^{-1})$.*

*Proof.* (1). Let $\pi\in\mathscr{C}(\mu,\nu)$ and $A\in\mathscr{B}$. Then one has

$$\begin{aligned}\pi\circ\Phi^{-1}(A\times E)&=\pi\left(\Phi^{-1}(A\times E)\right)=\pi\left(\Phi^{-1}(A)\times\Phi^{-1}(E)\right)\\&=\pi\left(\Phi^{-1}(A)\times E\right)=\mu(\Phi^{-1}(A))\\&=\mu\circ\Phi^{-1}(A).\end{aligned}$$

Similarly, it is easy to see that

$$\pi\circ\Phi^{-1}(E\times A)=\nu\circ\Phi^{-1}(A).$$

Thus $\pi\circ\Phi^{-1}\in\mathscr{C}(\mu\circ\Phi^{-1},\nu\circ\Phi^{-1})$.

On the other hand, for any $\tilde{\pi} \in \mathscr{C}(\mu \circ \Phi^{-1}, \nu \circ \Phi^{-1})$, we similarly have $\tilde{\pi} \circ \Phi \in \mathscr{C}(\mu, \nu)$. Moreover, $(\pi \circ \Phi^{-1}) \circ \Phi = \pi$. Define

$$\left(\Phi^{-1}\right)^{\#} : \pi \to \pi \circ \Phi^{-1}, \ \pi \in \mathscr{C}(\mu, \nu),$$

then $(\Phi^{-1})^{\#}$ is a bijection from $\mathscr{C}(\mu, \nu)$ to $\mathscr{C}(\mu \circ \Phi^{-1}, \nu \circ \Phi^{-1})$ with inverse $\Phi^{\#}$.

For any $\pi \in \mathscr{C}(\mu, \nu)$, the bijection $(\Phi^{-1})^{\#}$ implies that

$$\begin{aligned} \mathbb{W}_p^{\rho}(\mu, \nu)^p &\le \int_{E \times E} \rho^p(x, y) \pi(\mathrm{d}x, \mathrm{d}y) \\ &= \int_{E \times E} \rho^p \circ \Phi^{-1}(x, y) \pi \circ \Phi^{-1}(\mathrm{d}x, \mathrm{d}y) \\ &= \int_{E \times E} \rho^p \circ \Phi^{-1}(x, y) \left((\Phi^{-1})^{\#} \pi\right)(\mathrm{d}x, \mathrm{d}y), \end{aligned}$$

which implies that

$$\begin{aligned} \mathbb{W}_p^{\rho}(\mu, \nu)^p &\le \inf_{\pi \in \mathscr{C}(\mu, \nu)} \int_{E \times E} \rho^p \circ \Phi^{-1}(x, y) \left((\Phi^{-1})^{\#} \pi\right)(\mathrm{d}x, \mathrm{d}y) \\ &= \inf_{\tilde{\pi} \in \mathscr{C}(\mu \circ \Phi^{-1}, \nu \circ \Phi^{-1})} \int_{E \times E} \rho^p \circ \Phi^{-1}(x, y) \tilde{\pi}(\mathrm{d}x, \mathrm{d}y) \\ &= \mathbb{W}_p^{\rho \circ \Phi^{-1}}(\mu \circ \Phi^{-1}, \nu \circ \Phi^{-1})^p. \end{aligned} \tag{2.3}$$

rho

Since $\Phi^{\#}$ is the inverse of $(\Phi^{-1})^{\#}$, we have

$$\mathbb{W}_p^{\rho \circ \Phi^{-1}}(\mu \circ \Phi^{-1}, \nu \circ \Phi^{-1})^p \le \mathbb{W}_p^{\rho}(\mu, \nu)^p.$$

This, together with (2.3), yields the first assertion of (1).

Since (2.1), we have that

$$c_1 \rho \circ \Phi^{-1}(x, y) \le \rho(x, y) \le c_2 \rho \circ \Phi^{-1}(x, y).$$

Then one obtains from the definition of $L^p$-Wasserstein distance that

$$\begin{aligned} c_1 \mathbb{W}_p^{\rho \circ \Phi^{-1}}(\mu \circ \Phi^{-1}, \nu \circ \Phi^{-1}) &\le \mathbb{W}_p^{\rho}(\mu \circ \Phi^{-1}, \nu \circ \Phi^{-1}) \\ &\le c_2 \mathbb{W}_p^{\rho \circ \Phi^{-1}}(\mu \circ \Phi^{-1}, \nu \circ \Phi^{-1}). \end{aligned}$$

Combining this with the first assertion, we obtain (2.2).

(2). We first assume $\nu \ll \mu$. For any $A \in \mathscr{B}$, if $\mu \circ \Phi^{-1}(A) = 0$, i.e. $\mu(\Phi^{-1}(A)) = 0$, then one has

$$\nu \circ \Phi^{-1}(A) = \nu(\Phi^{-1}(A)) = 0,$$

which implies $\nu \circ \Phi^{-1} \ll \mu \circ \Phi^{-1}$. Similarly, if $\nu \circ \Phi^{-1} \ll \mu \circ \Phi^{-1}$, then $\nu \ll \mu$. Hence, $H(\nu, \mu) < \infty$ if and only if $H(\nu \circ \Phi^{-1} | \mu \circ \Phi^{-1}) < \infty$.

By the definition of push-forward measure, one obtains that for any $\psi \in \mathscr{B}(E)$

$$\begin{aligned}\int_E \psi \frac{\mathrm{d}\nu\circ\Phi^{-1}}{\mathrm{d}\mu\circ\Phi^{-1}}\mathrm{d}\mu\circ\Phi^{-1} &= \int_E \psi\mathrm{d}\nu\circ\Phi^{-1} = \int_E \psi\circ\Phi\mathrm{d}\nu \\ &= \int_E \psi\circ\Phi\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\mathrm{d}\mu = \int_E \psi\cdot\left(\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\circ\Phi^{-1}\right)\mathrm{d}\mu\circ\Phi^{-1},\end{aligned}$$

which yields $\frac{\mathrm{d}\nu\circ\Phi^{-1}}{\mathrm{d}\mu\circ\Phi^{-1}} = \frac{\mathrm{d}\nu}{\mathrm{d}\mu}\circ\Phi^{-1}$, $\mu\circ\Phi^{-1}$-a.s. We then can see that

$$\begin{aligned}H(\nu|\mu) &= \int_E \left(\log\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right)\mathrm{d}\nu = \int_E \frac{\mathrm{d}\nu}{\mathrm{d}\mu}\left(\log\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right)\mathrm{d}\mu \\ &= \int_E \left(\left[\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\left(\log\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right)\right]\circ\Phi^{-1}\right)\mathrm{d}\mu\circ\Phi^{-1} \\ &= \int_E \left[\frac{\mathrm{d}\nu\circ\Phi^{-1}}{\mathrm{d}\mu\circ\Phi^{-1}}\left(\log\frac{\mathrm{d}\nu\circ\Phi^{-1}}{\mathrm{d}\mu\circ\Phi^{-1}}\right)\right]\mathrm{d}\mu\circ\Phi^{-1} \\ &= H(\nu\circ\Phi^{-1}|\mu\circ\Phi^{-1}).\end{aligned}$$

□

From the above lemma, we obtain the stability of $T_p(C)$ under a homeomorphism map. This result can also be derived from [6, Lemma 2.1], here we use the direct relation (2.2) and (2) of Lemma 2.1.

cor-1 **Corollary 2.2.** *Assume* (2.1). *For any* $p \geq 1$ *and* $\mu \in \mathscr{P}(E)$, *if* $\mu \in T_p(C)$, *then* $\mu\circ\Phi^{-1} \in T_p(Cc_2^2)$; *conversely, if* $\mu\circ\Phi^{-1} \in T_p(C)$, *then* $\mu \in T_p(Cc_1^{-2})$.

*Proof.* If $\mu \in T_p(C)$, we then derive from (2) of Lemma 2.1 that

$$W_p^\rho(\mu,\nu) \leq \sqrt{2CH(\nu\circ\Phi^{-1}|\mu\circ\Phi^{-1})},$$

this, together with (2.2), yields that

$$W_p^\rho(\mu\circ\Phi^{-1},\nu\circ\Phi^{-1}) \leq \sqrt{2Cc_2^2H(\nu\circ\Phi^{-1}|\mu\circ\Phi^{-1})},$$

that is, $\mu\circ\Phi^{-1} \in T_p(Cc_2^2)$ on $(E,\rho)$. Similarly, the converse statement can also be proved. □

Throughout this work, the following notation will be used. $(\mathbb{R}^d, \langle\cdot,\cdot\rangle, |\cdot|)$ denotes the $d$-dimensional Euclidean space, $\mathbb{R}^d\otimes\mathbb{R}^d$ is the family of all $d\times d$ matrices. For a vector or matrix $v$, $v^*$ denotes its transpose. Let $\|\cdot\|$ denote the usual operator norm. Fix $T>0$ and set $\|f\|_{T,\infty} := \sup_{t\in[0,T],x\in\mathbb{R}^d}\|f(t,x)\|$ for an operator or vector valued map $f$ on $[0,T]\times\mathbb{R}^d$, $C(\mathbb{R}^d;\mathbb{R}^d)$ means the set of all continuous functions $f:\mathbb{R}^d\to\mathbb{R}^d$. Let $C^2(\mathbb{R}^d;\mathbb{R}^d\otimes\mathbb{R}^d)$ be the family of all continuously twice differentiable functions $f:\mathbb{R}^d\to\mathbb{R}^d\otimes\mathbb{R}^d$. $\nabla^i, i\in\mathbb{N}$ means the $i$-th order gradient operator. Let $W_t$ be a $d$-dimensional Brownian motion defined on a complete filtration probability space $(\Omega,(\mathscr{F}_t)_{t\geq 0},\mathscr{F},\mathbb{P})$. We will use $\mathbf{0}$ to denote vectors with components 0.

# 3 TCI for SDEs with singular coefficients

sec:main

Let $T > 0$ be any fixed real number. For the process-level law of the solution to SDEs, we first present $T_2(C)$ for equations with Dini continuous drift, then $T_1(C)$ for equations with singular coefficients on the path space $C([0,T];\mathbb{R}^d)$ under the uniform metric. Throughout this paper, we denote by $\rho_T$ the uniform metric on $C([0,T];\mathbb{R}^d)$:

$$\rho_T(\xi,\eta) = \sup_{t\in[0,T]} |\xi_t - \eta_t|, \ \ \xi,\eta \in C([0,T];\mathbb{R}^d).$$

## 3.1 $T_2(C)$ for SDEs with Dini continuous drift

Consider the following SDE

$$\mathrm{d}X_t = \{B_t(X_t) + b_t(X_t)\}\mathrm{d}t + \sigma_t(X_t)\mathrm{d}W_t, \tag{3.1}$$ eq1

where $B, b : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ are measurable, and $\sigma : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d\otimes\mathbb{R}^d$ is measurable. Let

$$\mathscr{D} = \Big\{\phi : [0,\infty) \to [0,\infty) \text{ is increasing, } \phi^2 \text{ is concave, } \int_0^1 \frac{\phi(s)}{s}\mathrm{d}s < \infty\Big\}.$$

With regard to (3.1), we impose the following conditions on its coefficients.

(A1) $\|b\|_{T,\infty} < +\infty$ and there exists $\phi \in \mathscr{D}$ such that

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

(A2) $B_t(\cdot)$ satisfies Lipschitz condition and $\sup_{t\in[0,T]}|B_t(\mathbf{0})| < \infty$; for any $x\in\mathbb{R}^d$, $\sigma_t(x)$ is invertible and $\sigma_t\in C^2(\mathbb{R}^d;\mathbb{R}^d\otimes\mathbb{R}^d)$; there exists some positive increasing function $K\in C([0,\infty);(0,\infty))$ such that

$$\|\nabla B\|_{T,\infty} + \|\sigma\|_{T,\infty} + \|\nabla\sigma\|_{T,\infty} + \|\nabla^2\sigma\|_{T,\infty} + \|(\sigma\sigma^*)^{-1}\|_{T,\infty} \le K(T).$$

**Remark 3.1.** *According to [20, Theorem 1.1], for any $T > 0$, the equation* (3.1) *has a unique strong solution $(X_t)_{t\in[0,T]}$ under the assumptions (A1)-(A2). Indeed, we can choose any Hilbert space $H_0$, a cylindrical Brownian motion $(\tilde{W}_t)_{t\ge0}$ on $H_0$ independent of $(W_t)_{t\ge0}$ and a positive definite self-adjoint operator $A_0$ on $H_0$ satisfying [20,* **(a1)***]. Then $H := \mathbb{R}^d\oplus H_0$ is a Hilbert space with the natural inner product induced from that of $\mathbb{R}^d$ and $H_0$. Set*

$$A = \begin{pmatrix} I_d & \mathbf{0} \\ \mathbf{0} & A_0 \end{pmatrix}, \qquad Q_t(x_1,x_2) = \begin{pmatrix} \sigma_t(x_1) & \mathbf{0} \\ \mathbf{0} & I_{H_0} \end{pmatrix}, \ x_1\in\mathbb{R}^d, x_2\in H_0,$$

*where $I_d$ and $I_{H_0}$ are the identity on $\mathbb{R}^d$ and $H_0$ respectively. It follows from [20, Theorem 1.1] that the following equation on $H$ has a unique strong solution*

$$\mathrm{d}\begin{pmatrix} X_t \\ \tilde{X}_t \end{pmatrix} = -A\begin{pmatrix} X_t \\ \tilde{X}_t \end{pmatrix}\mathrm{d}t + \begin{pmatrix} (b_t+B_t)(X_t) + X_t \\ \mathbf{0} \end{pmatrix}\mathrm{d}t + Q_t(X_t,\tilde{X}_t)\mathrm{d}\begin{pmatrix} W_t \\ \tilde{W}_t \end{pmatrix},$$

*which yields that* (3.1) *has a unique strong solution.*

**Remark 3.2.** *The condition* $\int_0^1 \frac{\phi(s)}{s}\mathrm{d}s < \infty$ *is well known as Dini condition. If* $\phi$ *is Hölder continuous with exponent* $\alpha$*, then* $\phi$ *is Dini continuous. In fact, if* $\phi(0) = 0$ *and* $|\phi(s) - \phi(t)| \le L|s - t|^\alpha$*, then* $\int_0^1 \frac{\phi(s)}{s}\mathrm{d}s \le \frac{L}{\alpha}$ *holds. However, there are numerous Dini continuous functions which are not Hölder continuous for any* $\alpha > 0$*. For instance,*

$$\phi(s) = \begin{cases} (\log \frac{2}{(2s)\wedge 1})^{-2}, & s > 0, \\ 0, & s = 0. \end{cases}$$

*It is easy to check that* $\lim_{s\to 0^+} \frac{\phi(s)}{s^\alpha} = +\infty$ *for any* $\alpha > 0$*, so* $\phi$ *is not Hölder continuous, but* $\phi$ *is Dini continuous. Indeed,* $\phi$ *is continuous and increasing on* $[0, +\infty)$ *with* $\int_0^1 \frac{\phi(s)}{s}\mathrm{d}s < \infty$*, which implies that* $\phi$ *is Dini continuous. Moreover,* $\phi^2$ *is concave.*

We now state the first result.

Dthm **Theorem 3.1.** *Suppose the assumptions (A1)-(A2) hold.*
*(1) Let* $\mathbb{P}^x$ *be the law of the solution* $(X_t)_{t\in[0,T]}$ *to (3.1) with initial value* $X_0 = x \in \mathbb{R}^d$*. The quadratic transportation cost inequality on the path space, i.e.*

$$\mathbb{W}_2^{\rho_T}(\mathbb{Q}, \mathbb{P}^x)^2 \le CH(\mathbb{Q}|\mathbb{P}^x), \quad \mathbb{Q} \in \mathscr{P}(C([0,T];\mathbb{R}^d))$$

*holds for some constant* $C > 0$*.*
*(2) Let* $\mu \in \mathscr{P}(\mathbb{R}^d)$ *and* $\mathbb{P}^\mu$ *be the law of* $(X_t)_{t\in[0,T]}$ *with initial distribution* $\mu$*. Then*

$$\mathbb{W}_2^{\rho_T}(\mathbb{Q}, \mathbb{P}^\mu)^2 \le C_1 H(\mathbb{Q}|\mathbb{P}^\mu), \quad \mathbb{Q} \in \mathscr{P}(C([0,T];\mathbb{R}^d)) \tag{3.2}$$ eqk1

*holds for some constant* $C_1 > 0$ *if and only if*

$$\mathbb{W}_2(\nu, \mu)^2 \le C_2 H(\nu|\mu), \quad \nu \in \mathscr{P}(\mathbb{R}^d) \tag{3.3}$$ eqk2

*holds for some constant* $C_2 > 0$*.*

TCI is closely related to the concentration of measure phenomenon, and we first review the definition of measure $\mu$ satisfying the concentration property as follows:

def-1 **Definition 3.1.** *The probability measure* $\mu \in \mathscr{P}(E)$ *has concentration on* $(E, \rho)$ *with the concentration function* $\alpha(r)$*, which is defined as*

$$\alpha(r) = \sup\Big\{1 - \mu(A_r) : A \subset E, \mu(A) \ge \frac{1}{2}\Big\}, r > 0, \tag{3.4}$$ wqc

*where* $A_r$ *denotes by the* $r$*-neighbourhood of* $A$*, namely,* $A_r = \{x : \rho(x, A) \le r\}$*. The normal concentration of* $\mu$ *means that the associated concentration function* $\alpha(r) \le C\,\mathrm{e}^{-cr^2}$ *for all* $r > 0$ *with some positive constants* $C, c$*.*

**Remark 3.3.** *Based on [10, Theorem 2.4], the conclusion of this theorem implies that* $\mathbb{P}^x$ *satisfies the concentration property with*

$$\alpha(r) = \mathrm{e}^{-\frac{1}{C}(r - r_0)^2}, \ r \ge 0,$$

*where* $r_0 = \sqrt{C\log(2)}$ *and the constant* $C$ *is same as in the above theorem.*

## 3.2 $T_1(C)$ for SDEs with singular dissipative coefficients

In this subsection, we consider the following SDE

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t, \tag{3.5}$$ b-S

where $b : \mathbb{R}^d \to \mathbb{R}^d, \sigma : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ are Borel measurable functions. Assume that the coefficients $b$ and $\sigma$ satisfy the following conditions:

(B1) $b = b_1 + b_2$ such that $b_1 \in L^p(\mathbb{R}^d)$ for some $p > d$, and one of the following conditions holds for $b_2$

(1) for some $\kappa_1, \kappa_2, \kappa_3 > 0$, $r > -1$

$$\langle x, b_2(x)\rangle \le -\kappa_1|x|^{2+r} + \kappa_2, \quad \text{and} \quad |b_2(x)| \le \kappa_3(1 + |x|^{1+r}); \tag{3.6}$$ diss

(2) for some $\kappa_4 \ge 0$

$$|b_2(x)| \le \kappa_4(1 + |x|). \tag{3.7}$$ lin

(B2) $\sigma$ is uniformly continuous, and $\|\nabla\sigma\| \in L^p(\mathbb{R}^d)$ with the same $p$ in (B1). There is a constant $c_0 \ge 1$ such that

$$c_0^{-1}|\xi|^2 \le |\sigma^*(x)\xi|^2 \le c_0|\xi|^2, \quad \xi, x \in \mathbb{R}^d.$$

$b_1$ is called the singular part and $b_2$ is locally bounded. According to [27, Theorem 2.1, Theorem 3.1] (or [28, Theorem 1.1]) and the proof of [26, Theorem 2.9], (3.5) admits a unique strong solution under (B1) and (B2). We now state the $T_1(C)$ for law of the solution to SDE (3.5).

TCI-1 **Theorem 3.2.** *Assume assumptions (B1) and (B2) hold. Let $\mu \in \mathscr{P}(\mathbb{R}^d)$ with $\mu(e^{\delta_0|\cdot|^{2+r^+}}) < \infty$ for some constant $\delta_0 > 0$. Then the distribution $\mathbb{P}^\mu$ of the solution $(X_t)_{t\in[0,T]}$ to SDE (3.5) with initial distribution $\mu$ satisfies the $T_1(C)$:*

$$\mathbb{W}_1^{\rho_T}(\mathbb{Q}, \mathbb{P}^\mu) \le CH(\mathbb{Q}|\mathbb{P}^\mu), \quad \mathbb{Q} \in \mathscr{P}(C([0,T];\mathbb{R}^d)),$$

*for some constant $C > 0$.*

**Remark 3.4.** *The assumption (B1) is different from the one on the drift in [1], since our drift includes two parts such that $b = b_1 + b_2$ with $b_1 \in L^p(\mathbb{R}^d)$ for some $p > d$ while $p \ge 2(d+1)$ is needed in [1], and the condition (3.6) on $b_2$ is weaker than the condition (4.2) in [6] where the one sided Lipschitz condition is imposed on the drift.*

# 4 Proof of Theorem 3.1

sec:proof1

## 4.1 Regularization representation of the solution to (3.1)

By Lemma 2.1, we establish the $T_2(C)$ for $\mathbb{P}^x$ by constructing a differeomorphism on $C([0,T];\mathbb{R}^d)$. To this end, we will find a transform $\Phi : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ in the spirit of [20].

We first decompose $B_t$ into a smooth term and a bounded Lipschitz term.

drift-d **Lemma 4.1.** *There exist $\bar{B}_t \in C^2(\mathbb{R}^d)$ and $\hat{B}_t$ which is Lipschitz such that $B_t = \bar{B}_t + \hat{B}_t$ and*

$$\|\nabla \bar{B}\|_{T,\infty} + \|\nabla^2 \bar{B}\|_{T,\infty} < \infty,$$
$$\|\nabla \bar{B}\|_{T,\infty} \vee \|\nabla \hat{B}\|_{T,\infty} \vee \|\hat{B}\|_{T,\infty} \leq \|\nabla B\|_{T,\infty}. \tag{4.1}$$ nn-b-B

*Proof.* Let $\chi$ be a smooth function supported in $\{x \in \mathbb{R}^d \mid |x| \leq 1\}$ and $\int_{\mathbb{R}^d} \chi(x)\mathrm{d}x = 1$. Set

$$\bar{B}_t(x) = B_t * \chi(x), \qquad \hat{B}_t(x) = B_t(x) - \bar{B}_t(x).$$

Then the assertions of this lemma hold. □

From this lemma, we could explicitly decompose the drifts as follows:

$$B_t(x) + b_t(x) = \bar{B}_t(x) + \big(\hat{B}_t(x) + b_t(x)\big), \tag{4.2}$$ ad-deco1

where $\bar{B}_t$ and $\hat{B}_t$ are defined in Lemma 4.1. Let $\hat{b}_t(x) = b_t(x) + \hat{B}_t(x)$, we then have that

$$\begin{aligned} |\hat{b}_t(x) - \hat{b}_t(y)| &\leq \phi(|x-y|) + 2\|\nabla B\|_{T,\infty}(|x-y| \wedge 1) \\ &\leq \phi(|x-y|) + 2\|\nabla B\|_{T,\infty}(|x-y|^{\frac{1}{2}} \wedge 1) \\ &=: \hat{\phi}(|x-y|), \ x, y \in \mathbb{R}^d. \end{aligned} \tag{4.3}$$ ad-deco2

Moreover, we have that $\hat{\phi} \in \mathscr{D}$. For notation simplicity, we use $B_t$ and $b_t$ instead of $\bar{B}_t$ and $\hat{b}_t$, respectively. Hence, we use the following assumption instead of (A2).

(A2') $B_t(\cdot) \in C^2(\mathbb{R}^d;\mathbb{R}^d)$ with $\sup_{t\in[0,T]} |B_t(\mathbf{0})| < \infty$; $\sigma_t(x)$ is invertible and $\sigma_t \in C^2(\mathbb{R}^d;\mathbb{R}^d \otimes \mathbb{R}^d)$ with $\sup_{t\in[0,T]} \|\sigma_t(\mathbf{0})\| < \infty$; there exists some positive increasing function $K \in C([0,\infty);(0,\infty))$ such that

$$\begin{aligned} &\|\nabla B\|_{T,\infty} + \|\nabla^2 B\|_{T,\infty} + \|\sigma\|_{T,\infty} + \|\nabla\sigma\|_{T,\infty} \\ &\qquad + \|\nabla^2\sigma\|_{T,\infty} + \|(\sigma\sigma^*)^{-1}\|_{T,\infty} \leq K(T). \end{aligned}$$

**Remark 4.1.** *Before moving on, we give some comments on the case of semilinear SPDEs. According to [2, Theorem 4.1] and [9, Lemma 5.3, especially (5.33)], $T_2(C)$ holds for semilinear SPDEs under the setting of [9, Theorem 2.2]. However, the Lipschitz drift term in [9, Theorem 2.2] is imposed to be bounded since the Zvonkin type transformation used in [9] or [20] does not map the Lipschitz drift to a Lipschitzian one. In the following discussion, we use a modified transformation, see the definition of $\Phi$ and (4.5) below. Based on Lemma 4.1, (4.2) and (4.3), we can assume the drift term is twice continuously differentiable with bounded first and second order derivatives. However, this argument may fail in the infinite dimension, see e.g. [5, Subsection 2.2]. One can impose that the nonlinear regular drift term satisfies a condition as (A2') directly for SPDEs, but this assumption is strong. Hence, TCIs for semilinar SPDEs with Dini drift are prepared in another paper.*

Consider the backward PDE

$$\partial_t u_t = -L_t u_t - b_t + \lambda u_t,\ u_T = 0,\ t \in [0,T], \tag{4.4}$$ PDE1

where $\lambda > 0$ is a parameter and

$$L_t := \frac{1}{2}\mathrm{tr}(\sigma_t\sigma_t^*\nabla^2) + \nabla_{B_t} + \nabla_{b_t}.$$

Set $\Phi_t(x) = x + u_t(x)$. Then $\partial_t\Phi_t = B_t(x) - L_t\Phi_t(x) + \lambda u_t(x)$. By Itô's formula, we formally have that (see Lemma 4.4 for a proof)

$$\begin{aligned} \mathrm{d}\Phi_t(X_t) &= \{(\partial_t\Phi_t)(X_t) + L_t\Phi_t(X_t)\}\mathrm{d}t + \nabla\Phi_t(X_t)\sigma_t(X_t)\mathrm{d}W_t \\ &= (\lambda u_t(X_t) + B_t(X_t))\mathrm{d}t + \nabla\Phi_t(X_t)\sigma_t(X_t)\mathrm{d}W_t. \end{aligned} \tag{4.5}$$ eq2

The irregular term $b_t$ is canceled. $u_t$ is regular with $\|\nabla u\|_{T,\infty} < 1$ for large enough $\lambda$, see Lemma 4.3 below. Then $\Phi_t$ is a differeomorphism on $\mathbb{R}^d$.

We investigate (4.4) in a weaker form. Let $\{P^0_{s,t}\}_{0\le s\le t}$ be the semigroup associated to the SDE below

$$\mathrm{d}Z^x_{s,t} = B_t(Z^x_{s,t})\mathrm{d}t + \sigma_t(Z^x_{s,t})\mathrm{d}W_t,\quad t \ge s, Z^x_{s,s} = x. \tag{4.6}$$ eq3

It is well known that the equation (4.6) has a unique solution under assumption (A2'). Then we have

$$P^0_{s,t}f(x) = \mathbb{E}f(Z^x_{s,t}),\quad t \ge s \ge 0, x \in \mathbb{R}^d, f \in \mathscr{B}_b(\mathbb{R}^d).$$

The generator of $P^0_{s,t}$ is

$$\tilde{L}_t = \frac{1}{2}\mathrm{tr}(\sigma_t\sigma_t^*\nabla^2) + \nabla_{B_t}.$$

By using $P^0_{s,t}$, (4.4) can be rewritten into the following integral equation

$$u_s = \int_s^T \mathrm{e}^{-\lambda(t-s)}\, P^0_{s,t}\{\nabla_{b_t}u_t + b_t\}\mathrm{d}t, s \in [0,T]. \tag{4.7}$$ eq4

In the following lemma, we give the gradient estimates for semigroup $P^0_{s,t}$ defined by (4.6), which will be used to study the regularity properties of solution $u$ to (4.7). The proof of the following lemma follows from [20, Lemma 2.1] completely, and we omit it.

lem1 **Lemma 4.2.** *Fix $T > 0$. Assume (A2'). Then the following assertions hold.*

*(1) For any $f \in \mathscr{B}_b(\mathbb{R}^d)$, $P^0_{s,t}f \in C^2_b(\mathbb{R}^d)$ for any $0 \le s < t$. There exists a positive constant $c$ such that for any $0 \le s < t \le T$,*

$$|\nabla P^0_{s,t}f|^2(x) \le \frac{c}{t-s}P^0_{s,t}f^2(x), \tag{4.8}$$ eq10

$$|\nabla^2 P^0_{s,t}f|^2(x) \le \frac{c}{(t-s)^2}P^0_{s,t}f^2(x),\ x \in \mathbb{R}^d,\ f \in \mathscr{B}_b(\mathbb{R}^d). \tag{4.9}$$ ine-na2

*(2) There exist positive constants* $c_1$ *and* $c_2$ *such that for any increasing* $\phi : [0,\infty) \to [0,\infty)$ *with concave* $\phi^2$*,*

$$\|\nabla^2 P^0_{s,t} f\|_\infty := \sup_{x\in\mathbb{R}^d} \|\nabla^2 P^0_{s,t} f(x)\| \le \frac{c_1 \phi(c_2\sqrt{(t-s)})}{t-s}, \tag{4.10}$$

eq2.2

*holds for any* $f \in \mathscr{B}_b(\mathbb{R}^d)$ *satisfying*

$$|f(x) - f(y)| \le \phi(|x-y|), \quad 0 \le s < t \le T, x, y \in \mathbb{R}^d.$$

The following Lemma focuses on the existence and uniqueness of solutions to (4.7) and gradient estimates of the solution, which is essentially due to [9, Lemma 3.1] or [20, Lemma 2.3]. We include a complete proof for readers' convenience.

lu **Lemma 4.3.** *Assume* $\|b\|_{T,\infty} < \infty$ *and (A2'). Let* $T > 0$ *be fixed, then there exists a constant* $\lambda(T) > 0$ *such that the following assertions hold:*

*(1) For any* $\lambda > \lambda(T)$*, (4.7) has a unique solution* $u \in C([0,T]; C_b^1(\mathbb{R}^d;\mathbb{R}^d))$ *satisfying*

$$\lim_{\lambda\to+\infty} \{\|u\|_{T,\infty} + \|\nabla u\|_{T,\infty}\} = 0. \tag{4.11}$$

u1

*(2) Moreover, if (A1) holds, then we have*

$$\lim_{\lambda\to\infty} \|\nabla^2 u\|_{T,\infty} = 0. \tag{4.12}$$

u2

*Proof.* (1) Let $\mathscr{H} = C([0,T]; C_b^1(\mathbb{R}^d;\mathbb{R}^d))$, which is a Banach space under the norm $\|u\|_{\mathscr{H}} := \|u\|_{T,\infty} + \|\nabla u\|_{T,\infty}, u \in \mathscr{H}$.

For any $u \in \mathscr{H}$, define the mapping

$$(\Gamma u)_s(x) = \int_s^T \mathrm{e}^{-\lambda(t-s)}\, P^0_{s,t}\{\nabla_{b_t(\cdot)} u_t(\cdot) + b_t(\cdot)\}(x)\mathrm{d}t.$$

Firstly, we claim that $\Gamma\mathscr{H} \subset \mathscr{H}$. In fact, for any $u \in \mathscr{H}$, by (4.8), one has

$$\begin{aligned}\|\Gamma u\|_{T,\infty} &= \sup_{s\in[0,T],x\in\mathbb{R}^d} \Big| \int_s^T \mathrm{e}^{-\lambda(t-s)}\, P^0_{s,t}\{\nabla_{b_t(\cdot)} u_t(\cdot) + b_t(\cdot)\}(x)\mathrm{d}t \Big| \\ &\le \sup_{s\in[0,T]} \Big| \int_s^T \mathrm{e}^{-\lambda(t-s)}\, \|b\|_{T,\infty}(\|\nabla u\|_{T,\infty} + 1)\mathrm{d}t \Big| \\ &\le \frac{\|b\|_{T,\infty}(\|\nabla u\|_{T,\infty}+1)}{\lambda} < \infty, \end{aligned} \tag{4.13}$$

G1

and

$$\|\nabla \Gamma u\|_{T,\infty} = \sup_{s\in[0,T],x\in\mathbb{R}^d} \Big| \int_s^T \mathrm{e}^{-\lambda(t-s)}\, \nabla P^0_{s,t}\{\nabla_{b_t(\cdot)} u_t(\cdot) + b_t(\cdot)\}(x)\mathrm{d}t \Big|$$

$$
\begin{aligned}
&\le c \sup_{s\in[0,T]} \Big| \int_s^T \frac{\mathrm{e}^{-\lambda(t-s)}}{\sqrt{t-s}} \|b\|_{T,\infty}(\|\nabla u\|_{T,\infty}+1)\mathrm{d}t \\
&\le \frac{c\|b\|_{T,\infty}(\|\nabla u\|_{T,\infty}+1)}{\sqrt{\lambda}} < \infty. \qquad (4.14)
\end{aligned}
$$

G2

Therefore, the claim $\Gamma\mathscr{H} \subset \mathscr{H}$ holds.

Next, we will show that for large enough $\lambda > 0$, $\Gamma$ is contractive on $\mathscr{H}$. Indeed, by the similar arguments as above, it is easy to check that for any $u, \hat{u} \in \mathscr{H}$, we have

$$
\begin{aligned}
\|\Gamma u - \Gamma \hat{u}\|_{\mathscr{H}} &\le \frac{\|b\|_{T,\infty}}{\sqrt{\lambda}}(1+c)\|\nabla u - \nabla \hat{u}\|_{T,\infty} \\
&\le \frac{\|b\|_{T,\infty}}{\sqrt{\lambda}}(1+c)\|u-\hat{u}\|_{\mathscr{H}} \\
&=: C(\lambda)\|u-\hat{u}\|_{\mathscr{H}}.
\end{aligned}
$$

Choosing constant $\lambda(T)$ such that $C(\lambda) < 1$ for $\lambda > \lambda(T)$, we can see that $\Gamma$ is contractive on $\mathscr{H}$ with $\lambda > \lambda(T)$. Thus, the fixed point theorem yields that (4.7) has a unique solution $u \in \mathscr{H}$.

Finally, the estimates (4.13) and (4.14) imply that (4.11) holds.

(2) (4.9) implies that for any $f \in \mathscr{B}_b(\mathbb{R}^d)$

$$
|\nabla P^0_{s,t} f(x) - \nabla P^0_{s,t} f(y)| \le \frac{c|x-y|}{t-s}\|f\|_\infty, \quad x,y \in \mathbb{R}^d,\ 0 \le s < t \le T.
$$

This, together with (4.8), yields that

$$
|\nabla P^0_{s,t} f(x) - \nabla P^0_{s,t} f(y)| \le c\left(\frac{|x-y|}{t-s} \wedge \frac{1}{\sqrt{t-s}}\right)\|f\|_\infty, \qquad (4.15)
$$

eq2.1

where $c$ is some constant.

Combining this with (4.7), one obtains that there exists a $\tilde{\phi} \in \mathscr{D}$ such that

$$
\begin{aligned}
&|\nabla_{b_t(x)} u_t(x) + b_t(x) - \nabla_{b_t(y)} u_t(y) - b_t(y)| \\
&\le (1+\|\nabla u\|_{T,\infty})\phi(|x-y|) + \|b\|_{T,\infty}\|\nabla u_t(x) - \nabla u_t(y)\| \\
&\le (1+\|\nabla u\|_{T,\infty})\phi(|x-y|) + 2\|b\|_{T,\infty}\sqrt{|x-y|}\mathbb{1}_{\{|x-y|\ge 1\}} \\
&\quad + c\|b\|_{T,\infty}\|\nabla_b u + b\|_{T,\infty}\int_s^T \mathrm{e}^{-\lambda(t-s)}\left(\frac{|x-y|}{(t-s)} \wedge \frac{1}{\sqrt{t-s}}\right)\mathbb{1}_{\{|x-y|\le 1\}}\mathrm{d}t \\
&\le (1+\|\nabla u\|_{T,\infty})\phi(|x-y|) + 2\|b\|_{T,\infty}\sqrt{|x-y|}\mathbb{1}_{\{|x-y|\ge 1\}} \\
&\quad + c\|b\|_{T,\infty}\|\nabla_b u + b\|_{T,\infty}|x-y|\log(\mathrm{e}+\frac{1}{|x-y|})\mathbb{1}_{\{|x-y|\le 1\}} \\
&\le c\sqrt{\phi^2(|x-y|) + |x-y|} \\
&=: \tilde{\phi}(|x-y|), \qquad (4.16)
\end{aligned}
$$

eqD

the last inequality was due to the fact that for $x \in [0,1]$, $\sqrt{x}\log(\mathrm{e}+\frac{1}{x})$ is an increasing function.

Using $\|\nabla u\|_{T,\infty} + \|b\|_{T,\infty} < \infty$, (4.7), (4.10) and (4.16), we derive

$$\begin{aligned}\|\nabla^2 u\|_{T,\infty} &= \int_0^T \mathrm{e}^{-r\lambda} \sup_{x\in\mathbb{R}^d} \|\nabla^2 P^0_{0,r}\{\nabla_{b_r} u_r + b_r\}(x)\|\mathrm{d}r \\ &\le \int_0^T \mathrm{e}^{-r\lambda} \frac{c_1\tilde{\phi}(c_2\sqrt{r})}{r}\mathrm{d}r =: \delta_{\tilde{\phi}}(\lambda). \end{aligned} \tag{4.17}$$

eqh

Noting that $\tilde{\phi} \in \mathscr{D}$, we have $\int_0^T \frac{\tilde{\phi}(c_2\sqrt{r})}{r}\mathrm{d}r < \infty$, which implies that $\delta_{\tilde{\phi}}(\lambda) \to 0$ as $\lambda \to \infty$. □

We provide the regularization representation (4.5) of solution to (3.1). We borrow from [20, Proposition 2.5.] the method to prove this lemma.

lem2 **Lemma 4.4.** *Fix $T > 0$. Assume (A1) and (A2'). Then there exists a constant $\lambda(T) > 0$ such that, for any $\lambda \ge \lambda(T)$, the solution $X_t$ to* (3.1) *satisfies*

$$\begin{aligned} X_t = X_0 + u_0(X_0) - u_t(X_t) &+ \int_0^t \{\sigma_s + (\nabla u_s)\sigma_s\}(X_s)\mathrm{d}W_s \\ &+ \int_0^t \{\lambda u_s + B_s\}(X_s)\mathrm{d}s, \quad t \in [0,T],\ \mathbb{P}\text{-a.s.}, \end{aligned} \tag{4.18}$$

e2

*where $u$ solves* (4.7).

*Proof.* Let $G_r = \nabla_{b_r} u_r + b_r$, $r \ge 0$. For fixed $\delta > 0$, let

$$F^{(\delta)}_{s,r}(x) = P^0_{s,r+\delta} G_r(x), \quad 0 \le s < r \le T, x \in \mathbb{R}^d.$$

According to (A1) and (4.11), we know $G_r$ is bounded and measurable. Then, we obtain from (4.9) that

$$\sup_{0\le s<r\le T} \{\|F^{(\delta)}_{s,r}\|_\infty + \|\nabla F^{(\delta)}_{s,r}\|_\infty + \|\nabla^2 F^{(\delta)}_{s,r}\|_\infty\} < \infty. \tag{4.19}$$

ineqF

By (4.6) and Itô's formula, we derive that for any $0 \le s \le r \le T$

$$\mathrm{d}F^{(\delta)}_{s,r}(Z^x_{r,t}) = \tilde{L}_t F^{(\delta)}_{s,r}(Z^x_{r,t})\mathrm{d}t + \langle \nabla F^{(\delta)}_{s,r}(Z^x_{r,t}), \sigma_t(Z^x_{r,t})\mathrm{d}W_t\rangle, t \ge r,$$

which yields that

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}s} F^{(\delta)}_{s,r}(x) :&= -\lim_{v\downarrow 0} \frac{F^{(\delta)}_{s-v,r}(x) - F^{(\delta)}_{s,r}(x)}{v} = -\lim_{v\downarrow 0} \frac{P^0_{s-v,s}P^0_{s,r+\delta}G_r(x) - F^{(\delta)}_{s,r}(x)}{v} \\ &= -\lim_{v\downarrow 0} \frac{\mathbb{E}P^0_{s,r+\delta}G_r(Z^x_{s-v,s}) - F^{(\delta)}_{s,r}(x)}{v} \\ &= -\lim_{v\downarrow 0} \frac{\mathbb{E}F^{(\delta)}_{s,r}(Z^x_{s-v,s}) - F^{(\delta)}_{s,r}(x)}{v} \end{aligned}$$

$$
\begin{aligned}
&= -\lim_{v\downarrow 0}\frac{1}{v}\mathbb{E}\int_{s-v}^{s}(\tilde{L}_t F_{s,r}^{(\delta)})(Z_{s-v,t}^x)\mathrm{d}t \\
&= -\tilde{L}_s F_{s,r}^{(\delta)}(x), r>0, \text{a.e. } s\in[0,r].
\end{aligned}
\tag{4.20}
$$

d-F

Let

$$
u_s^{(\delta)} = \int_s^T \mathrm{e}^{-\lambda(t-s)}\, P_{s,t+\delta}^0 G_t \mathrm{d}t = \int_s^T \mathrm{e}^{-\lambda(t-s)}\, F_{s,t}^{(\delta)}\mathrm{d}t, s\in[0,T]. \tag{4.21}
$$

u-de

Then we obtain from (4.19), (4.20) and (4.21) that

$$
\partial_s u_s^{(\delta)} = (\lambda - \tilde{L}_s)u_s^{(\delta)} - P_{s,s+\delta}^0(\nabla_{b_s}u_s + b_s).
$$

By Itô's formula, we arrive at

$$
\begin{aligned}
\mathrm{d}u_s^{(\delta)}(X_s) &= \{\tilde{L}_s u_s^{(\delta)} + \nabla_{b_s}u_s^{(\delta)} + \partial_s u_s^{(\delta)}\}(X_s)\mathrm{d}s + \langle \nabla u_s^{(\delta)}(X_s), \sigma_s(X_s)\mathrm{d}W_s\rangle \\
&= \left\{\lambda u_s^{(\delta)} + \nabla_{b_s}u_s^{(\delta)} - P_{s,s+\delta}^0\{\nabla_{b_s}u_s + b_s\}\right\}(X_s)\mathrm{d}s \\
&\quad + \langle \nabla u_s^{(\delta)}(X_s), \sigma_s(X_s)\mathrm{d}W_s\rangle.
\end{aligned}
\tag{4.22}
$$

Ito-u-de

It follows from (4.7) and (4.21) that

$$
u_s^{(\delta)} - u_s = \int_s^T \mathrm{e}^{-\lambda(t-s)}(P_{s,t}^0\{P_{t,t+\delta}^0 G_t - G_t\})\mathrm{d}t,\ \ s\in[0,T]. \tag{4.23}
$$

u-de-u

By (4.16), $G_t(\cdot)$ is continuous. Then

$$
\lim_{\delta\to 0^+} P_{s,t+\delta}^0 G_t = P_{s,t}^0 G_t,
$$

which, together with the boundedness of $\|\nabla u\|$ and $b$, implies by the dominated convergence theorem that

$$
\begin{aligned}
\lim_{\delta\to 0^+}\left|u_s^{(\delta)} - u_s\right| &\le \int_s^T \mathrm{e}^{-\lambda(t-s)}\lim_{\delta\to 0^+}\left|P_{s,t}^0\{P_{t,t+\delta}^0 G_t - G_t\})\right|\mathrm{d}t \\
&= 0,\ \ s\in[0,T].
\end{aligned}
\tag{4.24}
$$

u-de-u-1

By using the boundedness of $\|\nabla u\|$ and $b$ again, we can derive from (4.8) and (4.21) that $\sup_{\delta\in(0,1)}\|\nabla u^{(\delta)}\|_{T,\infty}<\infty$. Moreover, combining (4.23) with (4.8), we obtain from the dominated convergence theorem that

$$
\begin{aligned}
\lim_{\delta\downarrow 0}\|\nabla u_s^{(\delta)} - \nabla u_s\| &= \lim_{\delta\downarrow 0}\left\|\int_s^T \mathrm{e}^{-\lambda(t-s)}\,\nabla P_{s,t}^0\{P_{t,t+\delta}^0 G_t - G_t\}\mathrm{d}t\right\| \\
&\le \lim_{\delta\downarrow 0}\left\|\int_s^T \frac{c\,\mathrm{e}^{-\lambda(t-s)}}{\sqrt{(t-s)}}\sqrt{P_{s,t}^0|P_{t,t+\delta}^0 G_t - G_t|^2\mathrm{d}t}\right\| = 0.
\end{aligned}
\tag{4.25}
$$

de-u

Combining this with (4.22), (4.24) and (3.1), we obtain (4.18). □

**Proof of Theorem 3.1**

*Proof.* (1) By Lemma 4.3, we can take $\lambda(T) > 0$ large enough such that for any $\lambda \geq \lambda(T)$, the unique solution $u$ to (4.7) satisfies

$$\|\nabla u\|_{T,\infty} < \frac{1}{2}. \tag{4.26}$$

gru

This implies that $\Phi_t(x) := x + u_t(x)$ is a differeomorphism and satisfies that for $(t,x) \in [0,T] \times \mathbb{R}^d$,

$$\frac{1}{2} \leq \|\nabla \Phi_t(x)\| \leq \frac{3}{2}, \quad \frac{2}{3} \leq \|\nabla \Phi_t^{-1}(x)\| \leq 2. \tag{4.27}$$

grp

Since $u \in C([0,T]; C_b^1(\mathbb{R}^d;\mathbb{R}^d))$, we define $\Phi : C([0,T];\mathbb{R}^d) \to C([0,T];\mathbb{R}^d)$ as

$$\Phi(\xi)(t) = \Phi_t(\xi_t), \quad \xi \in C([0,T];\mathbb{R}^d), t \in [0,T]. \tag{4.28}$$

Ph

Moreover, it follows from (4.27) that

$$\begin{aligned}
&|\Phi_{t+\triangle t}^{-1}(\xi_{t+\triangle t}) - \Phi_t^{-1}(\xi_t)| \\
&\leq |\Phi_{t+\triangle t}^{-1}(\xi_{t+\triangle t}) - \Phi_{t+\triangle t}^{-1}(\xi_t)| + |\Phi_{t+\triangle t}^{-1}(\xi_t) - \Phi_t^{-1}(\xi_t)| \\
&\leq \|\nabla \Phi_{t+\triangle t}^{-1}(\cdot)\|_\infty |\xi_{t+\triangle t} - \xi_t| + |\Phi_{t+\triangle t}^{-1}(\xi_t) - \Phi_{t+\triangle t}^{-1}(\Phi_{t+\triangle t}(\Phi_t^{-1}(\xi_t)))| \\
&\leq 2\Big\{ |\xi_{t+\triangle t} - \xi_t| + |\xi_t - \Phi_{t+\triangle t}(\Phi_t^{-1}(\xi_t))| \Big\},
\end{aligned}$$

which yields that $\Phi_\cdot^{-1}(\xi_\cdot)$ is also continuous. Hence $\Phi$ is a homeomorphisms on $C([0,T];\mathbb{R}^d)$ with

$$\Phi^{-1}(\xi)(t) = \Phi_t^{-1}(\xi_t), \quad \xi \in C([0,T];\mathbb{R}^d), \quad t \in [0,T]. \tag{4.29}$$

iPh

Then $\Phi$ induces a homeomorphism on $C([0,T];\mathbb{R}^d) \times C([0,T];\mathbb{R}^d)$ defined as in Section 2 (setting $E = C([0,T];\mathbb{R}^d)$) which is still denoted by $\Phi$, and its inverse is still denoted by $\Phi^{-1}$. Furthermore, it follows from (4.27) and (4.28) that for any $\xi, \eta \in C([0,T];\mathbb{R}^d)$

$$\frac{1}{2}\rho_T(\xi,\eta) \leq \rho_T \circ \Phi(\xi,\eta) \leq \frac{3}{2}\rho_T(\xi,\eta). \tag{4.30}$$

dis

This means that condition (2.1) hold for $\Phi$ by setting $c_1 = \frac{1}{2}, c_2 = \frac{3}{2}$.

By setting $Y_t = \Phi_t(X_t)$, it follows from Lemma 4.4 that

$$\begin{aligned}
Y_t = Y_0 &+ \int_0^t (\lambda u_s + B_s) \circ \Phi_s^{-1}(Y_s) \mathrm{d}s \\
&+ \int_0^t (\nabla \Phi_s \sigma_s) \circ \Phi_s^{-1}(Y_s) \mathrm{d}W_s, \ t \in [0,T].
\end{aligned} \tag{4.31}$$

equ-YY

Moreover, it follows from (4.1), (4.17) and (A2') that

$$\|\nabla(\lambda u + B)\|_{T,\infty} + \|\nabla(\nabla \Phi \sigma)\|_{T,\infty} < \infty. \tag{4.32}$$

lip-y

Then there exists a constant $C > 0$ (see e.g. [18, Theorem 1] or [2]) such that

$$\mathbb{W}_2^{\rho_T}(\mathbb{Q}\circ\Phi^{-1}, \mathbb{P}^x\circ\Phi^{-1})^2 \le CH(\mathbb{Q}\circ\Phi^{-1}|\mathbb{P}^x\circ\Phi^{-1}).$$

Combining this with (4.30), we derive from Corollary 2.2 that

$$\mathbb{W}_2^{\rho_T}(\mathbb{Q}, \mathbb{P}^x)^2 \le 2CH(\mathbb{Q}|\mathbb{P}^x).$$

(2) Based on [21, Theorem 2.1], it suffices to verify the following assertions respectively:

$$\mathbb{W}_2^{\rho_T}(\mathbb{Q}, \mathbb{P}^x)^2 \le C_1 H(\mathbb{Q}|\mathbb{P}^x), \quad \mathbb{Q}\in\mathscr{P}(C([0,T];\mathbb{R}^d)), \tag{4.33}$$

q1

$$\mathbb{W}_2^{\rho_T}(\mathbb{P}^x, \mathbb{P}^y)^2 \le C_2|x-y|^2, \quad x, y\in\mathbb{R}^d, \tag{4.34}$$

q2

for some constants $C_1$ and $C_2$.

Since (4.33) has been proved in (1), we only need to prove (4.34). Noting that the law of $(X_t^x, X_t^y)_{t\in[0,T]}$ is a coupling of $\mathbb{P}^x$ and $\mathbb{P}^y$, we obtain that

$$\mathbb{W}_2^{\rho_T}(\mathbb{P}^x, \mathbb{P}^y)^2 \le \mathbb{E}[\rho_T(X^x, X^y)^2] = \mathbb{E}\Big(\sup_{t\in[0,T]}|X_t^x - X_t^y|^2\Big).$$

Denote by $Y_t^{\Phi_0(x)}$ the solution of (4.31) with $Y_0 = \Phi_0(x)$. By (4.32), one can derive from the Burkholder-Davis-Gundy inequality that

$$\mathbb{E}\Big(\sup_{t\in[0,T]}|Y_t^{\Phi_0(x)} - Y_t^{\Phi_0(y)}|^2\Big) \le C|\Phi_0(x)-\Phi_0(y)|^2.$$

Combining this with (4.30), we have that

$$\mathbb{E}\Big(\sup_{t\in[0,T]}|X_t^x - X_t^y|^2\Big) \le 4\mathbb{E}\Big(\sup_{t\in[0,T]}|Y_t^{\Phi_0(x)} - Y_t^{\Phi_0(y)}|^2\Big) \le 9C|x-y|^2.$$

Thus, (4.34) holds with $C_2 = 9C$, and the proof is completed. □

## 5 Proof of Theorem 3.2

sec:proof2

For the reader's convenience, we sketch the construction of homeomorphism $\Phi$. To this end, we consider the following elliptic equation

$$\frac{1}{2}\mathrm{tr}(\sigma\sigma^*\nabla^2 u) + \nabla_{b_1}u = \lambda u - b_1. \tag{5.1}$$

e-b

Before moving on, we introduce some spaces and notations. For $(p,\alpha)\in[1,\infty]\times(0,2]-\{\infty\}\times\{1\}$, let $H_p^\alpha = (I-\Delta)^{-\frac{\alpha}{2}}(L^p(\mathbb{R}^d))$ be the usual Bessel potential space with the norm

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

where $\|\cdot\|_p$ is the usual $L^p$-norm in $\mathbb{R}^d$, $\Delta$ is the Laplace operator on $\mathbb{R}^d$, and $(I-\Delta)^{\frac{\alpha}{2}}f$ and $\Delta^{\frac{\alpha}{2}}f$ are defined through the Fourier transformation

$$(I-\Delta)^{\frac{\alpha}{2}}f := \mathcal{F}^{-1}((1+|\cdot|^2)^{\frac{\alpha}{2}}\mathcal{F}f), \quad \Delta^{\frac{\alpha}{2}}f := \mathcal{F}^{-1}(|\cdot|^\alpha \mathcal{F}f).$$

For $(p,\alpha)=(\infty,1)$, we define $H^1_\infty$ as the space of Lipschitz functions with finite norm

$$\|f\|_{1,\infty} := \|f\|_\infty + \|\nabla f\|_\infty < \infty.$$

Notice that for $n=1,2$ and $p\in(1,\infty)$, an equivalent norm in $H^n_p$ is given by

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

The following Lemma shows the solvability of equation (5.1), which is a consequence of [26, Theorem 7.6]. We remark here that the Hölder continuity assumption on $\sigma$ in [26, $(H^\sigma_\beta)$] can be replaced by the uniformly continuity in this paper due to [24, Theorem 3.2].

s-u **Lemma 5.1.** *Suppose that (B2) holds and $b_1\in L^p(\mathbb{R}^d)$ for some $p>d$. Then for some $\lambda_1\geq 1$ and for all $\lambda\geq\lambda_1$, there exists a unique solution $u\in H^2_p$ to equation* (5.1)*, and for any $p'\in[p,\infty]$ and $v\in(0,2)$ with $\frac{d}{p}<2-v+\frac{d}{p'}$, we have*

$$\lambda^{\frac{1}{2}(2-v+\frac{d}{p'}-\frac{d}{p})}\|u\|_{v,p'} + \|\nabla^2 u\|_p \leq c\|b_1\|_p. \tag{5.2}$$ u-L

Taking $p'=+\infty, v=1$ in (5.2), one can see that there exist $c,\lambda_1\geq 1$ such that for all $\lambda\geq\lambda_1$,

$$\|u\|_\infty + \|\nabla u\|_\infty \leq c\lambda^{\frac{1}{2}(\frac{d}{p}-1)}. \tag{5.3}$$ u-u-L

Define $\Phi(x)=x+u(x)$. By (5.3) with $\lambda$ large enough, the map $x\to\Phi(x)$ forms a $C^1$-diffeomorphism and

$$\frac{1}{2}\leq\|\nabla\Phi\|_\infty, \|\nabla\Phi^{-1}\|_\infty\leq 2. \tag{5.4}$$ phi-up

The Lemma below presents the regular representation of solution to (3.5) by Zvonkin's transformation. The following two lemmas are due to [26, Lemma 7.7 and Lemma 7.8].

x-phi **Lemma 5.2.** *$X_t$ solves SDE* (3.5) *if and only if $Y_t := \Phi(X_t)$ solves*

$$\mathrm{d}Y_t = \tilde{b}(Y_t)\mathrm{d}t + \tilde{\sigma}(Y_t)\mathrm{d}W_t, \tag{5.5}$$ phi-r

*with initial value $y := \Phi(x)$ and*

$$\tilde{b}(y) := (\lambda u + \nabla\Phi\cdot b_2)\circ\Phi^{-1}(y), \quad \tilde{\sigma}(y) := (\nabla\Phi\cdot\sigma)\circ\Phi^{-1}(y).$$

The following Lemma shows that the conditions for $b_2$ in (B1) are preserved under Zvonkin's transformation.

H-B **Lemma 5.3.** *Under (B1), the following assertion holds for large enough* $\lambda$.

($\tilde{B}1$) *$\tilde{b}$ satisfies one of the following conditions*
*(1) there exist* $r > -1$, $\tilde{\kappa}_1 > 0$, $\tilde{\kappa}_2 \geq 0$ *and* $\tilde{\kappa}_3 \geq 0$ *such that*

$$\langle \tilde{b}(y), y \rangle \leq -\tilde{\kappa}_1 |y|^{2+r} + \tilde{\kappa}_2, \qquad |\tilde{b}(y)| \leq \tilde{\kappa}_3 (1 + |y|^{r+1}),\ y \in \mathbb{R}^d; \tag{5.6}$$ b-dis

*(2) there exists* $\kappa_4 \geq 0$ *such that*

$$|\tilde{b}(y)| \leq \tilde{\kappa}_4 (1 + |y|),\ y \in \mathbb{R}^d. \tag{5.7}$$ b-lin

We establish $T_1(C)$ by "Gaussian tail" following [6, Theorem 2.3], and we recall the following lemma there.

G **Lemma 5.4.** *The probability measure $\mu$ on $(E, \rho)$ satisfies the $L^1$-transportation cost inequality with some constant $C$ if and only if*

$$\int\int \mathrm{e}^{\delta \rho^2(x,y)}\, \mathrm{d}\mu(x)\mathrm{d}\mu(y) < +\infty, \quad \delta \in (0, \frac{1}{4C}), \tag{5.8}$$ G-tail

*holds.*

The following two lemmas contribute to establishing (5.8) for solutions $Y_t$ to the equation (5.5). By the definition of $\tilde{\sigma}$ and (5.4), it is clear that

$$\|\tilde{\sigma}\|_{HS,\infty} := \sup_{y \in \mathbb{R}^d} \|\tilde{\sigma}(y)\|_{HS} < \infty. \tag{5.9}$$ si-hs

exp-int **Lemma 5.5.** *For the solution to* (5.5) *with random initial value* $Y_0$, *we have the following two assertions.*

*(1) Assume that $\tilde{b}$ satisfies the condition (1) in ($\tilde{B}1$). Then there is a positive constant $C$ independent of $Y_0$ such that*

$$\mathbb{E}\left[\exp\left\{\theta \int_0^T |Y_t|^{2r+2} \mathrm{d}t\right\} \Big| Y_0\right] \leq C\, \mathrm{e}^{K_\theta |Y_0|^{2+r}}$$

*holds for* $0 \leq \theta < \frac{1}{2}\tilde{\kappa}_1^2 \|\tilde{\sigma}\|_{HS,\infty}^{-2}$ *and* $K_\theta = 2^{1-r/2}\theta \tilde{\kappa}_1^{-1}(r+2)^{-1}$ *if* $r > 0$; $0 \leq \theta \leq 2^{(r-1)}\tilde{\kappa}_1^2 \|\tilde{\sigma}\|_{HS,\infty}^{-2}$ *and* $K_\theta = \sqrt{2\theta}(2+r)^{-1}\|\tilde{\sigma}\|_{HS,\infty}^{-1}$ *if* $r \in (-1, 0]$.

*(2) Assume that $\tilde{b}$ satisfies the condition (2) in ($\tilde{B}1$). Then for* $0 \leq \theta \leq \frac{e^{-(2+3\tilde{\kappa}_4 T)}}{2\|\tilde{\sigma}\|_{HS,\infty}^2 T^2}$, *there exists a constant $C > 0$ independent of $Y_0$ such that*

$$\mathbb{E}\left[\exp\left\{\theta \int_0^T |Y_t|^2 \mathrm{d}t\right\} \Big| Y_0\right] \leq C\, \mathrm{e}^{T\theta e^{2+3\tilde{\kappa}_4 T} |Y_0|^2}.$$

*Proof.* We denote by $\mathbb{E}^{Y_0}[\cdot]$ the condition expectation $\mathbb{E}[\cdot|Y_0]$.

(1) It follows from Itô's formula that

$$\mathrm{d}|Y_t|^2 \le \left(-2\tilde\kappa_1|Y_t|^{r+2}+2\tilde\kappa_2\right)\mathrm{d}t + 2\langle Y_t, \tilde\sigma(Y_t)\mathrm{d}W_t\rangle + \|\tilde\sigma(Y_t)\|_{HS}^2\mathrm{d}t. \tag{5.10}$$

Ito-Y1

Below, we prove the assertion holds for $r\in(-1,0]$ and $r\in(0,\infty)$, respectively.

a). For $r\in(-1,0]$, we have by (5.10) that

$$\begin{aligned}
\mathrm{d}\left(1+|Y_t|^2\right)^{\frac{r+2}{2}} \le\ & \frac{r+2}{2}\left(1+|Y_t|^2\right)^{\frac{r}{2}}\left(-2\tilde\kappa_1|Y_t|^{2+r}+2\tilde\kappa_2+\|\tilde\sigma\|_{HS}^2\right)\mathrm{d}t\\
&+(r+2)\left(1+|Y_t|^2\right)^{\frac{r}{2}}\langle Y_t,\tilde\sigma(Y_t)\mathrm{d}W_t\rangle\\
&+\frac{r(r+2)}{2}\left(1+|Y_t|^2\right)^{\frac{r}{2}-1}|\tilde\sigma^*(Y_t)Y_t|^2\mathrm{d}t\\
\le\ & -2^{\frac{r}{2}}(r+2)\tilde\kappa_1|Y_t|^{2+2r}\mathrm{d}t+(r+2)\left(2^{\frac{r}{2}}\tilde\kappa_1+\tilde\kappa_2+\frac{\|\tilde\sigma\|_{HS,\infty}^2}{2}\right)\mathrm{d}t\\
&+(r+2)\left(1+|Y_t|^2\right)^{\frac{r}{2}}\langle Y_t,\tilde\sigma(Y_t)\mathrm{d}W_t\rangle,
\end{aligned}\tag{5.11}$$

Ito-Yr2

where in the last inequality we use

$$(1+y^2)^{\frac{r}{2}}y^{2+r}\ge 2^{\frac{r}{2}}y^{2+2r}-2^{\frac{r}{2}},\ y\ge0,\ r\in(-1,0].$$

Let

$$\tau_n=\inf\{t>0\mid \int_0^t|Y_s|^{2+2r}\mathrm{d}s\ge n\}.$$

Then it follows from (5.11) that

$$\begin{aligned}
&\exp\left\{-\frac{\theta(1+|Y_0|^2)^{\frac{r+2}{2}}}{2^{r/2}\tilde\kappa_1(r+2)}\right\}\mathbb{E}^{Y_0}\left[\exp\left\{\theta\int_0^{T\wedge\tau_n}|Y_t|^{2+2r}\mathrm{d}t\right\}\right]\\
&\le C_\theta\mathbb{E}^{Y_0}\left[\exp\left\{\frac{\theta}{\tilde\kappa_1 2^{\frac{r}{2}}}\int_0^{T\wedge\tau_n}(1+|Y_t|^2)^{\frac{r}{2}}\langle Y_t,\tilde\sigma(Y_t)\mathrm{d}W_t\rangle\right\}\right]\\
&\le C_\theta\left(\mathbb{E}^{Y_0}\left[\exp\left\{\frac{2^{1-\frac{r}{2}}\theta}{\tilde\kappa_1}M_{T\wedge\tau_n}-\frac{2^{1-r}\theta^2}{\tilde\kappa_1^2}\langle M\rangle_{T\wedge\tau_n}\right\}\right]\right)^{\frac12}\left(\mathbb{E}^{Y_0}\left[\frac{2^{1-r}\theta^2}{\tilde\kappa_1^2}\langle M\rangle_{T\wedge\tau_n}\right]\right)^{\frac12}\\
&\le C_\theta\left(\mathbb{E}^{Y_0}\left[\exp\left\{\frac{2^{1-r}\theta^2\|\tilde\sigma\|_{HS,\infty}^2}{\tilde\kappa_1^2}\int_0^{T\wedge\tau_n}|Y_t|^{2+2r}\mathrm{d}t\right\}\right]\right)^{\frac12},
\end{aligned}$$

where

$$\begin{aligned}
C_\theta&=\exp\left\{\frac{\theta T}{\tilde\kappa_1 2^{\frac{r}{2}}}\big(2^{\frac{r}{2}}\tilde\kappa_1+\tilde\kappa_2+\frac{\|\tilde\sigma\|_{HS,\infty}^2}{2}\big)\right\},\\
M_{T\wedge\tau_n}&=\int_0^{T\wedge\tau_n}(1+|Y_t|^2)^{\frac{r}{2}}\langle Y_t,\tilde\sigma(Y_t)\mathrm{d}W_t\rangle.
\end{aligned}$$

Choosing $\theta=\tilde\kappa_1^2 2^{r-1}\|\tilde\sigma\|_{HS,\infty}^{-2}$ and letting $n\to+\infty$, we have

$$\mathbb{E}^{Y_0}\left[\exp\left\{\frac{\tilde\kappa_1^2}{2^{1-r}\|\tilde\sigma\|_{HS,\infty}^2}\int_0^T|Y_t|^{2+2r}\mathrm{d}t\right\}\right]\le C_\theta^2\exp\left\{\frac{2^{r/2}\tilde\kappa_1(1+|Y_0|^2)^{\frac{r+2}{2}}}{(r+2)\|\tilde\sigma\|_{HS,\infty}^2}\right\}$$

$$< \infty.$$

Hence, for any $0 < \theta \le \tilde{\kappa}_1^2 2^{r-1} \|\tilde{\sigma}\|_{HS,\infty}^{-2}$,

$$\mathbb{E}^{Y_0}\left[\exp\left\{\theta\int_0^T |Y_t|^{2+2r}\mathrm{d}t\right\}\right] \le C_\theta^2 \exp\left\{\frac{2^{1-r/2}\theta(1+|Y_0|^2)^{\frac{r+2}{2}}}{\tilde{\kappa}_1(r+2)}\right\}.$$

b). For $r > 0$, it follows from Itô's formula and the Hölder inequality that

$$\begin{aligned}
\mathrm{d}|Y_t|^{r+2} &\le -(r+2)\tilde{\kappa}_1|Y_t|^{2r+2}\mathrm{d}t + \frac{r+2}{2}\left(2\tilde{\kappa}_2 + \|\tilde{\sigma}\|_{HS,\infty}^2 + r\|\tilde{\sigma}\|_{HS,\infty}^2\right)|Y_t|^r\mathrm{d}t \\
&\quad + (r+2)|Y_t|^r\langle Y_t, \tilde{\sigma}(Y_t)\mathrm{d}W_t\rangle \\
&\le -(r+2)\tilde{\kappa}_1|Y_t|^{2r+2}\mathrm{d}t + \epsilon_1(r+2)\tilde{\kappa}_1|Y_t|^{2r+2}\mathrm{d}t \\
&\quad + \frac{(r+2)^2\left(2\tilde{\kappa}_2 + (r+1)\|\tilde{\sigma}\|_{HS,\infty}^2\right)\left(2r\tilde{\kappa}_2 + r(r+1)\|\tilde{\sigma}\|_{HS,\infty}^2\right)^{\frac{r}{r+2}}}{4(r+1)(4\epsilon_1\tilde{\kappa}_1(r+1))^{\frac{r}{r+2}}}\mathrm{d}t \\
&\quad + (r+2)|Y_t|^r\langle Y_t, \tilde{\sigma}(Y_t)\mathrm{d}W_t\rangle,
\end{aligned}$$

where $\epsilon_1 \in (0,1)$. For $0 < \theta < \frac{\tilde{\kappa}_1^2}{2\|\tilde{\sigma}\|_{HS,\infty}^2}$, we choose $\epsilon_1 = 1 - \sqrt{2\theta}\|\tilde{\sigma}\|_{HS,\infty}/\tilde{\kappa}_1$. Then

$$\begin{aligned}
&\exp\left\{-\frac{\theta|Y_0|^{2+r}}{\tilde{\kappa}_1(2+r)(1-\epsilon_1)}\right\}\mathbb{E}^{Y_0}\left[\exp\left\{\theta\int_0^T |Y_t|^{2r+2}\mathrm{d}t\right\}\right] \\
&\le C_{\theta 1}\mathbb{E}^{Y_0}\left[\exp\left\{\frac{\theta}{(1-\epsilon_1)\tilde{\kappa}_1}\int_0^T |Y_t|^r\langle Y_t, \tilde{\sigma}(Y_t)\mathrm{d}W_t\rangle\right\}\right] \\
&\le C_{\theta 1}\left(\mathbb{E}^{Y_0}\left[\exp\left\{\frac{2\theta^2\|\tilde{\sigma}\|_{HS,\infty}^2}{(1-\epsilon_1)^2\tilde{\kappa}_1^2}\int_0^T |Y_t|^{2r+2}\mathrm{d}t\right\}\right]\right)^{\frac{1}{2}} \\
&= C_{\theta 1}\left(\mathbb{E}^{Y_0}\left[\exp\left\{\theta\int_0^T |Y_t|^{2r+2}\mathrm{d}t\right\}\right]\right)^{\frac{1}{2}},
\end{aligned}$$

where

$$C_{\theta 1} = \exp\left\{\frac{\theta T(r+2)\left(2\tilde{\kappa}_2 + (r+1)\|\tilde{\sigma}\|_{HS,\infty}^2\right)\left(2r\tilde{\kappa}_2 + r(r+1)\|\tilde{\sigma}\|_{HS,\infty}^2\right)^{\frac{r}{r+2}}}{4\tilde{\kappa}_1(1-\epsilon_1)(r+1)(4\epsilon_1\tilde{\kappa}_1(r+1))^{\frac{r}{r+2}}}\right\}.$$

This yields that for any $\theta < \frac{1}{2}\tilde{\kappa}_1^2\|\tilde{\sigma}\|_{HS,\infty}^{-2}$

$$\begin{aligned}
\mathbb{E}^{Y_0}\left[\exp\left\{\theta\int_0^T |Y_t|^{2r+2}\mathrm{d}t\right\}\right] &\le C_{\theta 1}^2 \exp\left\{\frac{2\theta|Y_0|^{2+r}}{\tilde{\kappa}_1(2+r)(1-\epsilon_1)}\right\} \\
&= C_{\theta 1}^2 \exp\left\{\frac{\sqrt{2\theta}|Y_0|^{2+r}}{(2+r)\|\tilde{\sigma}\|_{HS,\infty}}\right\}.
\end{aligned}$$

(2) It follows from the Itô formula and the Hölder inequality that for $\alpha > 3\tilde{\kappa}_4$

$$\mathrm{d}\left(e^{-\alpha t}|Y_t|^2\right) \le e^{-\alpha t}(\tilde{\kappa}_4 + \|\tilde{\sigma}\|_{HS,\infty}^2)\mathrm{d}t - (\alpha - 3\tilde{\kappa}_4)e^{-\alpha t}|Y_t|^2\mathrm{d}t$$

$$+ 2e^{-\alpha t}\langle Y_t, \tilde{\sigma}(Y_t)\mathrm{d}W_t\rangle.$$

We then have

$$\begin{aligned}
&\exp\left\{-\frac{\theta}{\alpha-3\tilde{\kappa}_4}|Y_0|^2\right\}\mathbb{E}^{Y_0}\left[\exp\left\{\theta\int_0^T e^{-\alpha t}|Y_t|^2\mathrm{d}t\right\}\right]\\
&\leq C_{\theta 2}\mathbb{E}^{Y_0}\left[\exp\left\{\frac{2\theta}{\alpha-3\tilde{\kappa}_4}\int_0^T e^{-\alpha t}\langle Y_t, \tilde{\sigma}(Y_t)\mathrm{d}W_t\rangle\right\}\right]\\
&\leq C_{\theta 2}\left(\mathbb{E}^{Y_0}\left[\exp\left\{\frac{8\theta^2\|\tilde{\sigma}\|^2_{HS,\infty}}{(\alpha-3\tilde{\kappa}_4)^2}\int_0^T e^{-\alpha t}|Y_t|^2\mathrm{d}t\right\}\right]\right)^{\frac{1}{2}}, \qquad (5.12)
\end{aligned}$$

ad-eY0

where

$$C_{\theta 2} = \exp\left\{\frac{\theta(\tilde{\kappa}_4 + \|\tilde{\sigma}\|^2_{HS,\infty})(1-e^{-\alpha T})}{(\alpha-3\tilde{\kappa}_4)\alpha}\right\}.$$

Choosing $\theta = \frac{(\alpha-3\tilde{\kappa}_4)^2}{8\|\tilde{\sigma}\|^2_{HS,\infty}}$, we have that

$$\mathbb{E}^{Y_0}\left[\exp\left\{\frac{(\alpha-3\tilde{\kappa}_4)^2}{8\|\tilde{\sigma}\|^2_{HS,\infty}}\int_0^T e^{-\alpha t}|Y_t|^2\mathrm{d}t\right\}\right] \leq C^2_{\theta 2}\exp\left\{\frac{(\alpha-3\tilde{\kappa}_4)|Y_0|^2}{4\|\tilde{\sigma}\|^2_{HS,\infty}}\right\}.$$

By choosing the optimal $\alpha = \frac{2}{T} + 3\tilde{\kappa}_4$, we have

$$\mathbb{E}^{Y_0}\left[\exp\left\{\frac{e^{-(2+3\tilde{\kappa}_4 T)}}{2\|\tilde{\sigma}\|^2_{HS,\infty}T^2}\int_0^T |Y_t|^2\mathrm{d}t\right\}\right] \leq C^2_{\theta 2}\exp\left\{\frac{|Y_0|^2}{2\|\tilde{\sigma}\|^2_{HS,\infty}T}\right\},$$

which, together with (5.12), implies the second claim.

□

Let $W_t^{(1)}, W_t^{(2)}$ be two independent Brownian motions defined on the filtered probability $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$, and $Y_t^{(i)}$, $i=1,2$ are solutions of (5.5) driven by $W_t^{(i)}$ with independent, identically distributed initial value $Y_0^{(1)}$ and $Y_0^{(2)}$. Let

$$Z_t = Y_t^{(1)} - Y_t^{(2)}.$$

Then

$$\mathbb{E}\exp\left\{\delta\sup_{t\in[0,T]}|Z_t|^2\right\} = \iint_{C([0,T];\mathbb{R}^d)\times C([0,T];\mathbb{R}^d)} \mathrm{e}^{\delta\rho_T(\xi,\eta)}\,\tilde{\mathbb{P}}_Y(\mathrm{d}\xi)\tilde{\mathbb{P}}_Y(\mathrm{d}\eta),$$

where $\tilde{\mathbb{P}}_Y$ is the law of $Y^{(1)}$ on $C([0,T];\mathbb{R}^d)$.

exp-int2 **Lemma 5.6.** *Suppose the assumptions in Lemma 5.5 hold and there exists $\delta_0 > 0$ such that $\mathbb{E}\,\mathrm{e}^{\delta_0|Y_0^{(1)}|^{2+r^+}} < +\infty$. Then there is $\delta_1 > 0$ such that for $0 \leq \delta < \delta_1$*

$$\mathbb{E}\exp\left\{\delta\sup_{t\in[0,T]}|Z_t|^2\right\} < \infty.$$

*If* $Y_0^{(1)} = Y_0^{(2)}$ *is deterministic, then we have that*

$$\delta_1 = \begin{cases} \left(8\|\tilde\sigma\|_{HS,\infty}^2 T(1+\tilde\kappa_3^2\tilde\kappa_1^{-2}2^{r^-})\right)^{-1}, & \text{if } \tilde b \text{ satisfies (5.6)}, \\ \left(8\|\tilde\sigma\|_{HS,\infty}^2 T\left(1+T^2\tilde\kappa_4^2\,\mathrm{e}^{2+3\tilde\kappa_4 T}\right)\right)^{-1}, & \text{if } \tilde b \text{ satisfies (5.7)}. \end{cases}$$

*Proof.* Since

$$\mathbb{E}\exp\left\{\delta\sup_{t\in[0,T]}|Z_t|^2\right\} = \mathbb{E}\left\{\mathbb{E}\left[\exp\left\{\delta\sup_{t\in[0,T]}|Z_t|^2\right\}\middle|Y_0^{(1)},Y_0^{(2)}\right]\right\}, \tag{5.13}$$

ad-coE

we can first assume that the initial values of $Y^{(i)}, i=1,2$ are deterministic, i.e. $Y_0^{(1)} = y_0^{(1)} \in \mathbb{R}^d$, $Y_0^{(2)} = y_0^{(2)} \in \mathbb{R}^d$.

It follows from Itô's formula that

$$\begin{aligned} \mathrm{d}\sqrt{1+|Z_t|^2} = {} & \frac{\langle \tilde b(Y_t^{(1)}) - \tilde b(Y_t^{(2)}), Y_t^{(1)} - Y_t^{(2)}\rangle}{\sqrt{1+|Z_t|^2}}\mathrm{d}t + \frac{\|\tilde\sigma(Y_t^{(1)})\|_{HS}^2 + \|\tilde\sigma(Y_t^{(2)})\|_{HS}^2}{\sqrt{1+|Z_t|^2}}\mathrm{d}t \\ & + \frac{\langle Z_t, \tilde\sigma(Y_t^{(1)})\mathrm{d}W_t^{(1)} - \tilde\sigma(Y_t^{(2)})\mathrm{d}W_t^{(2)}\rangle}{\sqrt{1+|Z_t|^2}} \\ & - \frac{|\tilde\sigma^*(Y_t^{(1)})Z_t|^2 + |\tilde\sigma^*(Y_t^{(2)})Z_t|^2}{(1+|Z_t|^2)^{\frac{3}{2}}}\mathrm{d}t. \end{aligned} \tag{5.14}$$

Ito-Z

We first deal with the case that $\tilde b$ satisfies (5.6). In this case,

$$\frac{\langle \tilde b(Y_t^{(1)}) - \tilde b(Y_t^{(2)}), Y_t^{(1)} - Y_t^{(2)}\rangle}{\sqrt{1+|Z_t|^2}} \le \tilde\kappa_3\left(2 + |Y_t^{(1)}|^{r+1} + |Y_t^{(2)}|^{r+1}\right).$$

Putting this into (5.14), we have that

$$\mathrm{d}\sqrt{1+|Z_t|^2} \le 2(\tilde\kappa_3 + \|\tilde\sigma\|_{HS,\infty}^2)\mathrm{d}t + \tilde\kappa_3\left(|Y_t^{(1)}|^{r+1} + |Y_t^{(2)}|^{r+1}\right)\mathrm{d}t + \mathrm{d}M_t$$

with

$$M_t = \int_0^t \frac{\langle Z_s, \tilde\sigma(Y_s^{(1)})\mathrm{d}W_s^{(1)} - \tilde\sigma(Y_s^{(2)})\mathrm{d}W_s^{(2)}\rangle}{\sqrt{1+|Z_s|^2}}.$$

This, together with the Hölder inequality yields that

$$\begin{aligned} & \mathbb{E}\exp\left\{\beta\left(\sup_{t\in[0,T]}\sqrt{1+|Z_t|^2} - 2(\tilde\kappa_3 + \|\tilde\sigma\|_{HS,\infty}^2)T - \sqrt{1+|Z_0|^2}\right)\right\} \\ & \le \left(\mathbb{E}\sup_{t\in[0,T]}\mathrm{e}^{2\beta M_t}\right)^{\frac{1}{2}}\left(\mathbb{E}\,\mathrm{e}^{2\beta\tilde\kappa_3\int_0^T\left(|Y_t^{(1)}|^{r+1} + |Y_t^{(2)}|^{r+1}\right)\mathrm{d}t}\right)^{\frac{1}{2}} \\ & \le 2\left(\mathbb{E}\,\mathrm{e}^{2\beta M_T}\right)^{\frac{1}{2}}\left(\prod_{i=1}^2\mathbb{E}\,\mathrm{e}^{2\beta\tilde\kappa_3\int_0^T|Y_t^{(i)}|^{r+1}\mathrm{d}t}\right)^{\frac{1}{2}}, \end{aligned} \tag{5.15}$$

est-e

where we used the Doob's maximal inequality and the independence of $Y_t^{(1)}$ and $Y_t^{(2)}$. Since

$$\langle M\rangle_T=\int_0^T\frac{|\tilde\sigma^*(Y_t^{(1)})Z_t|^2+|\tilde\sigma^*(Y_t^{(2)})Z_t|^2}{1+|Z_t|^2}\mathrm{d}t\le 2\|\tilde\sigma\|_{HS,\infty}^2T,$$

we have

$$\mathbb{E}\,\mathrm{e}^{2\beta M_T}=\mathbb{E}\left(\mathrm{e}^{2\beta M_T-2\beta^2\langle M\rangle_T}\,\mathrm{e}^{2\beta^2\langle M\rangle_T}\right)\le \mathrm{e}^{4\beta^2\|\tilde\sigma\|_{HS,\infty}^2T}. \tag{5.16}$$

exp-m-1

Since for any $\beta>0$ and $0<\tilde\beta<\frac{\tilde\kappa_1^2}{2(r^-+1)\|\tilde\sigma\|_{HS,\infty}^2}$, we derive from Lemma 5.5 that

$$\mathbb{E}\,\mathrm{e}^{2\beta\tilde\kappa_3\int_0^T|Y_t^{(i)}|^{r+1}\mathrm{d}t}\le \mathrm{e}^{\frac{\beta^2\tilde\kappa_3^2T}{\tilde\beta}}\,\mathbb{E}\,\mathrm{e}^{\tilde\beta\int_0^T|Y_t^{(i)}|^{2r+2}\mathrm{d}t}\le C\,\mathrm{e}^{\frac{\beta^2\tilde\kappa_3^2T}{\tilde\beta}+K_{\tilde\beta}|y_0^{(i)}|^{2+r}}.$$

Combining this with (5.16), we derive that

$$\begin{aligned}&\mathbb{E}\exp\left\{\beta\left(\sup_{t\in[0,T]}\sqrt{1+|Z_t|^2}-2(\tilde\kappa_3+\|\tilde\sigma\|_{HS,\infty}^2)T-\sqrt{1+|Z_0|^2}\right)\right\}\\&\quad\le 2C\exp\left\{2\beta^2\|\tilde\sigma\|_{HS,\infty}^2T+\frac{\beta^2\tilde\kappa_3^2T}{\tilde\beta}+\frac{K_{\tilde\beta}}{2}\sum_{i=1}^2|y_0^{(i)}|^{2+r}\right\}.\end{aligned} \tag{5.17}$$

est-e1

Then, by Chebychev's inequality and an optimization of $\beta$, it yields that

$$\begin{aligned}&\mathbb{P}\left\{\sup_{t\in[0,T]}\sqrt{1+|Z_t|^2}\ge\sqrt{1+|Z_0|^2}+2(\tilde\kappa_3+\|\tilde\sigma\|_{HS,\infty}^2)T+x\right\}\\&\quad\le 2C\exp\left\{\frac{K_{\tilde\beta}}{2}\sum_{i=1}^2|y_0^{(i)}|^{2+r}-\frac{x^2}{8\|\tilde\sigma\|_{HS,\infty}^2T+4\tilde\kappa_3^2\tilde\beta^{-1}T}\right\}.\end{aligned} \tag{5.18}$$

ad-PsupZ0

Denote by $\hat C_{\tilde\beta}=\left(8\|\tilde\sigma\|_{HS,\infty}^2T+4\tilde\kappa_3^2\tilde\beta^{-1}T\right)^{-1}$, $C_{\tilde\kappa_3,\tilde\sigma,T}=2(\tilde\kappa_3+\|\tilde\sigma\|_{HS,\infty}^2)T$ and $\xi=\sup_{t\in[0,T]}(1+|Z_t|^2)^{\frac12}-(1+|Z_0|^2)^{\frac12}$. Then for $0<\delta<\hat C_{\tilde\beta}$

$$\begin{aligned}\mathbb{E}\,\mathrm{e}^{\delta\xi^2}&=\mathbb{E}\left(\int_0^{\xi^2}\delta\,\mathrm{e}^{\delta x}\,\mathrm{d}x+1\right)=\int_0^{+\infty}\delta\,\mathrm{e}^{\delta x}\,\mathbb{P}\left(\xi^2\ge x\right)\mathrm{d}x+1\\&=\int_0^{+\infty}\delta\,\mathrm{e}^{\delta x}\,\mathbb{P}\left(\sup_{t\in[0,T]}\sqrt{1+|Z_t|^2}\ge\sqrt{1+|Z_0|^2}+\sqrt{x}\right)\mathrm{d}x+1\\&\le 1+2C\delta\exp\left\{\frac{K_{\tilde\beta}}{2}\sum_{i=1}^2|y_0^{(i)}|^{2+r}\right\}\int_0^{+\infty}\mathrm{e}^{\delta x-\hat C_{\tilde\beta}(\sqrt{x}-C_{\tilde\kappa_3,\tilde\sigma,T})^2}\,\mathrm{d}x\\&\le 1+\frac{4C\delta\,\mathrm{e}^{\frac{\hat C_{\tilde\beta}+\delta}{\hat C_{\tilde\beta}-\delta}\hat C_{\tilde\beta}C_{\tilde\kappa_3,\tilde\sigma,T}^2}}{\hat C_{\tilde\beta}-\delta}\exp\left\{\frac{K_{\tilde\beta}}{2}\sum_{i=1}^2|y_0^{(i)}|^{2+r}\right\},\end{aligned} \tag{5.19}$$

ad-eZ1-0

where in the last second inequality, we have used (5.18) and $C > 0$ is the constant in (5.18) which is independent of $y_0^{(1)}, y_0^{(2)}$ and $\delta$. Since

$$\xi^2 \geq (1-\epsilon) \sup_{t\in[0,T]} (1+|Z_t|^2) - \frac{1-\epsilon}{\epsilon}(1+|Z_0|^2),\ \epsilon \in (0,1],$$

we can derive form (5.19) that for $\epsilon \in (0,1)$ and $0 \leq \delta < \hat{C}_{\tilde{\beta}}$

$$\begin{aligned}
&\mathbb{E} \exp\left\{ \delta(1-\epsilon) \sup_{t\in[0,T]} |Z_t|^2 \right\} \\
&\quad \leq \exp\left\{ \frac{\delta(1-\epsilon)}{\epsilon}(1+|Z_0|^2) \right\} \mathbb{E}\, \mathrm{e}^{\delta \xi^2} \\
&\quad \leq \bar{C}_\delta \exp\left\{ \frac{\delta(1-\epsilon)}{\epsilon}(1+|Z_0|^2) + \frac{K_{\tilde{\beta}}}{2} \sum_{i=1}^{2} |y_0^{(i)}|^{2+r} \right\}
\end{aligned} \tag{5.20}$$

ad-eZ1

where

$$\bar{C}_\delta = 1 + \frac{4C\delta}{\hat{C}_{\tilde{\beta}} - \delta} \exp\left\{ \frac{\hat{C}_{\tilde{\beta}} + \delta}{\hat{C}_{\tilde{\beta}} - \delta} \hat{C}_{\tilde{\beta}} C^2_{\tilde{\kappa}_3, \tilde{\sigma}, T} \right\}.$$

Choosing $\tilde{\beta} = \frac{\tilde{\kappa}_1^2}{2^{(r^-+1)} \|\tilde{\sigma}\|^2_{HS,\infty}}$, then $\hat{C}_{\tilde{\beta}} = \delta_1$. Then the assertion for the deterministic initial value case holds. Since $K_{\tilde{\beta}}$ is decreasing to zero as $\tilde{\beta}$ decreases to zero, there is $\delta_1 > 0$ and $\tilde{\beta}_1 > 0$ so that for any $0 < \tilde{\beta} < \tilde{\beta}_1$

$$\exp\left\{ \delta_1(1+|Z_0|^2) + \frac{K_{\tilde{\beta}}}{2} \sum_{i=1}^{2} |y_0^{(i)}|^{2+r} \right\} \leq C_{\delta_1, \tilde{\beta}}\, \mathrm{e}^{\delta_0 \left( |y_0^{(1)}|^{2+r^+} + |y_0^{(2)}|^{2+r^+} \right)}.$$

Combining this with (5.13) and (5.20), the assertion for random initial value case holds.

If $\tilde{b}$ satisfies (5.7), then

$$\frac{\langle \tilde{b}(Y_t^{(1)}) - \tilde{b}(Y_t^{(2)}), Y_t^{(1)} - Y_t^{(2)} \rangle}{\sqrt{1+|Z_t|^2}} \leq \tilde{\kappa}_4 \left( 2 + |Y_t^{(1)}| + |Y_t^{(2)}| \right).$$

Putting this into (5.14), we have that

$$\mathrm{d}\sqrt{1+|Z_t|^2} \leq 2(\tilde{\kappa}_4 + \|\tilde{\sigma}\|^2_{HS,\infty})\mathrm{d}t + \tilde{\kappa}_4 \left( |Y_t^{(1)}| + |Y_t^{(2)}| \right) \mathrm{d}t + \mathrm{d}M_t.$$

By the Hölder inequality, we derive that

$$\begin{aligned}
&\mathbb{E} \exp\left\{ \beta \left( \sup_{t\in[0,T]} \sqrt{1+|Z_t|^2} - 2(\tilde{\kappa}_4 + \|\tilde{\sigma}\|^2_{HS,\infty})T - \sqrt{1+|Z_0|^2} \right) \right\} \\
&\quad \leq \mathbb{E} \left( \sup_{t\in[0,T]} \mathrm{e}^{\beta M_t}\, \mathrm{e}^{\beta \tilde{\kappa}_4 \int_0^T \left( |Y_t^{(1)}| + |Y_t^{(2)}| \right) \mathrm{d}t} \right)
\end{aligned}$$

$$
\begin{aligned}
&\le \left(\mathbb{E}\sup_{t\in[0,T]} \mathrm{e}^{2\beta M_t}\right)^{\frac{1}{2}} \left(\mathbb{E}\,\mathrm{e}^{2\beta\tilde{\kappa}_4\int_0^T\left(|Y_t^{(1)}|+|Y_t^{(2)}|\right)\mathrm{d}t}\right)^{\frac{1}{2}}\\
&\le 2\left(\mathbb{E}\,\mathrm{e}^{2\beta M_T}\right)^{\frac{1}{2}}\left(\prod_{i=1}^{2}\mathbb{E}\,\mathrm{e}^{2\beta\tilde{\kappa}_4\int_0^T|Y_t^{(i)}|\mathrm{d}t}\right)^{\frac{1}{2}}.
\end{aligned}
$$

Combining this with (5.16) and (2) of Lemma 5.5, we can derive that for any $\tilde{\beta}\in(0,1]$,

$$
\begin{aligned}
&\mathbb{E}\exp\left\{\beta\left(\sup_{t\in[0,T]}\sqrt{1+|Z_t|^2}-2(\tilde{\kappa}_4+\|\tilde{\sigma}\|_{HS,\infty}^2)T-\sqrt{1+|Z_0|^2}\right)\right\}\\
&\le 2\exp\left\{2\beta^2\|\tilde{\sigma}\|_{HS,\infty}^2T\left(1+T^2\tilde{\kappa}_4^2\,\mathrm{e}^{2+3\tilde{\kappa}_4T}\,\tilde{\beta}^{-1}\right)\right\}\\
&\quad\times\left(\prod_{i=1}^{2}\mathbb{E}\exp\left\{\frac{\mathrm{e}^{-(2+3\tilde{\kappa}_4T)}\,\tilde{\beta}}{2\|\tilde{\sigma}\|_{HS,\infty}^2T^2}\int_0^T|Y_t^{(i)}|^2\mathrm{d}t\right\}\right)^{\frac{1}{2}}\\
&\le C\exp\left\{2\beta^2\|\tilde{\sigma}\|_{HS,\infty}^2T\left(1+\frac{T^2\tilde{\kappa}_4^2\,\mathrm{e}^{2+3\tilde{\kappa}_4T}}{\tilde{\beta}}\right)+\frac{\tilde{\beta}\left(|y_0^{(1)}|^2+|y_0^{(2)}|^2\right)}{4\|\tilde{\sigma}\|_{HS,\infty}^2T}\right\},
\end{aligned}
$$

where $C>0$ is a constant independent of $y_0^{(1)}$ and $y_0^{(2)}$. Then by Chebychev's inequality and an optimization of $\beta$, it yields that

$$
\begin{aligned}
&\mathbb{P}\left\{\sup_{t\in[0,T]}\sqrt{1+|Z_t|^2}\ge\sqrt{1+|Z_0|^2}+2(\tilde{\kappa}_4+\|\tilde{\sigma}\|_{HS,\infty}^2)T+x\right\}\\
&\quad\le C\exp\left\{\frac{\tilde{\beta}\left(|y_0^{(1)}|^2+|y_0^{(2)}|^2\right)}{4\|\tilde{\sigma}\|_{HS,\infty}^2T}-\frac{x^2}{8\|\tilde{\sigma}\|_{HS,\infty}^2T\left(1+T^2\tilde{\kappa}_4^2\,\mathrm{e}^{2+3\tilde{\kappa}_4T}\,\tilde{\beta}^{-1}\right)}\right\}.
\end{aligned}
$$

Let $\hat{C}_{\tilde{\beta}}=\left(8\|\tilde{\sigma}\|_{HS,\infty}^2T\left(1+T^2\tilde{\kappa}_4^2\,\mathrm{e}^{2+3\tilde{\kappa}_4T}\,\tilde{\beta}^{-1}\right)\right)^{-1}$ and $C_{\tilde{\kappa}_4,\tilde{\sigma},T}=2(\tilde{\kappa}_4+\|\tilde{\sigma}\|_{HS,\infty}^2)T$. Arguing as (5.19) and (5.20), we have that for $0<\delta<\hat{C}_{\tilde{\beta}}$

$$
\begin{aligned}
&\mathbb{E}\exp\left\{\delta(1-\epsilon)\sup_{t\in[0,T]}|Z_t|^2\right\}\le\exp\left\{\frac{\delta(1-\epsilon)}{\epsilon}(1+|Z_0|^2)\right\}\mathbb{E}\,\mathrm{e}^{\delta\xi^2}\\
&\le\bar{C}_\delta'\exp\left\{\frac{\delta(1-\epsilon)}{\epsilon}(1+|Z_0|^2)+\frac{\tilde{\beta}\left(|y_0^{(1)}|^2+|y_0^{(2)}|^2\right)}{4\|\tilde{\sigma}\|_{HS,\infty}^2T}\right\},\ \epsilon\in(0,1],
\end{aligned}
\tag{5.21}
$$

ad-exZ2

where

$$
\bar{C}_\delta'=1+\frac{4C\delta}{\hat{C}_{\tilde{\beta}}-\delta}\exp\left\{\frac{\hat{C}_{\tilde{\beta}}+\delta}{\hat{C}_{\tilde{\beta}}-\delta}\hat{C}_{\tilde{\beta}}C_{\tilde{\kappa}_4,\tilde{\sigma},T}^2\right\}
$$

with a constant $C>0$ independent of $y_0^{(1)}$, $y_0^{(2)}$ and $\delta$. Choosing $\tilde{\beta}=1$, one can see that $\hat{C}_{\tilde{\beta}}=\delta_1$. Hence, the assertion for the deterministic initial value case holds. For random initial value, there are $\delta_1 > 0$ and $\tilde{\beta}_1 > 0$ such that for any $\tilde{\beta} \in (0, \tilde{\beta}_1)$

$$\exp\left\{\delta_1(1+|Z_0|^2) + \frac{\tilde{\beta}\left(|y_0^{(1)}|^2 + |y_0^{(2)}|^2\right)}{4\|\tilde{\sigma}\|_{HS,\infty}^2 T}\right\} \le C_{\delta_1,\tilde{\beta}}\, \mathrm{e}^{\delta_0\left(|y_0^{(1)}|^2+|y_0^{(2)}|^2\right)}.$$

This, together with (5.13) and (5.21), yields the desired assertion.

□

**Proof of Theorem 3.2**

*Proof.* Taking the similar arguments as in the proof of Theorem 3.1, the assertions of this theorem follows from Lemma 5.6, Lemma 5.4 and (5.4). It follows from (5.4) that $\Phi$ induces a homeomorphism on $C([0,T];\mathbb{R}^d)$ by using the same argument in Theorem 3.1. Moreover,

$$\begin{aligned}
&\frac{1}{2}\rho_T(\xi,\eta) \le \rho_T\circ\Phi(\xi,\eta) \le 2\rho_T(\xi,\eta), \quad \xi,\eta \in C([0,T];\mathbb{R}^d) \qquad (5.22)\\
&\frac{1}{2}\rho_T\circ\Phi^{-1}(\xi,\eta) \le \rho_T(\xi,\eta) \le 2\rho_T\circ\Phi^{-1}(\xi,\eta), \quad \xi,\eta \in C([0,T];\mathbb{R}^d).
\end{aligned}$$

distance

Since $Y_t = \Phi_t(X_t)$, the law of $Y$ is $\mathbb{P}^\mu \circ \Phi^{-1}$ and for any $0 < \delta_0' < \delta_0$

$$\mathbb{E}\,\mathrm{e}^{\delta_0'|Y_0|^{2+r^+}} = \mathbb{E}\,\mathrm{e}^{\delta_0'(|X_0|+\|u\|_{T,\infty})^{2+r^+}} \le \mathbb{E}\,\mathrm{e}^{\delta_0|X_0|^{2+r^+}+C_{\delta_0',\delta_0}\|u\|_{T,\infty}^{2+r^+}} < +\infty.$$

Then by Lemma 5.6 and Lemma 5.4, there is a constant $C > 0$ such that for any measure $Q$ on $C([0,T];\mathbb{R}^d)$,

$$\mathbb{W}_1^{\rho_T}(\mathbb{Q}\circ\Phi^{-1}, \mathbb{P}^\mu\circ\Phi^{-1}) \le \sqrt{CH(\mathbb{Q}\circ\Phi^{-1}|\mathbb{P}^\mu\circ\Phi^{-1})}.$$

Combining this with (5.22) and Corollary 2.2, it yields that

$$\mathbb{W}_1^{\rho_T}(\mathbb{Q}, \mathbb{P}^\mu) \le \sqrt{4CH(\mathbb{Q}|\mathbb{P}^\mu)}.$$

□

**Acknowledgements**

The third author was supported by the disciplinary funding of Central University of Finance and Economics, and the National Natural Science Foundation of China (Grant No. 11901604, 11771326).

# References

BMT [1] K. Bahlali, S. Mouchtabih, L. Tangpi, Quadratic transportation inequalities for SDEs with measurable drift, *Proceedings of the American Mathematical Society.* **149** (2021), 3583-3596.

BWY [2] J. Bao, F.-Y. Wang, C. Yuan, Transportation cost inequalities for neutral functional stochastic equations, *J. Anal. Appl.* **32** (2013), 457-475.

BT [3] D. Bartl, L.Tangpi, Functional inequalities for forward and backward diffusion, *Electronic Journal of Probability.* **25** (2020),1-22.

BGL [4] S. Bobkov, I. Gentil, M.Ledoux, Hypercontractivity of Hamilton-Jacobi equations, *J. Math.Pure Appl.* **80** (2001), 669-696.

DaZ [5] G. Da Prato, J. Zabczyk, Second Order Partial Differential Equations in Hilbert Spaces, Cambridge University Press, 2002

DGW [6] H. Djellout, A. Guilin, L. Wu, Transportation cost-information inequalities for random dynamical systems and diffusions, *Ann. Probab.* **32** (2004), 2702-2732.

FS [7] S. Fang, J. Shao, Transportation cost inequalities on path and loop groups, *J. Funct. Anal.* **218** (2005), 293-317.

GRS [8] N. Gozlan, C. Roberto, P.-M. Samson, A new characterization of Talagrand's transport- entropy inequalities and applications, *Ann. Probab.* **39** (2011), 857-880.

HZ [9] X. Huang, S.-Q. Zhang, Mild Solutions and Harnack Inequality for Functional Stochastic Partial Differential Equations with Dini Drift, *J. Theor. Probab.* **32** (2019), 303-329.

L [10] M. Ledoux, The concentration of measure phenomenon, *American Mathematical Society*, 2001.

M [11] Y. Ma, Transportation inequalities for stochastic differential equations with jumps, *Stochastic Process. Appl.* **120** (2010), 2-21.

MW20 [12] Y. Ma, R. Wang, Transportation cost inequalities for stochastic reaction-diffusion equations with Lévy noises and non-Lipschitz reaction terms, *Acta Math. Sin. (Engl. Ser.)* **36** (2020), 121-136.

Maj19 [13] M. B. Majka, Transportation inequalities for non-globally dissipative SDEs with jumps via Malliavin calculus and coupling, *Ann. Inst. Henri Poincaré Probab. Stat.* **55** (2019), 2019–2057.

OV [14] F. Otto, C. Villani, Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality, *J. Funct. Anal.* **173** (2000), 361-400.

S [15] B. Saussereau, Transportation inequalities for stochastic differential equations driven by a fractional Brownian motion, *Bernoulli* **18** (2012), 1-23.

SZ [16] S. Shang, T. Zhang, Quadratic transportation cost inequalities under uniform distance for stochastic reaction diffusion equations driven by multiplicative space-time white noise. arXiv: 1904.13162, 2019.

T [17] M. Talagrand, Transportation cost for Gaussian and other product measures, *Geom. Funct. Anal.* **6**(1996), 587–600.

U [18] A.S.Üstünel, Transport cost inequalities for diffusions under uniform distance, *in Stochastic Analysis and Related Topics.* 203-214. Springer, 2012.

WF [19] F.-Y. Wang, Probability distance inequalities on Riemannian manifolds and path spaces, *J. Funct. Anal.* **206** (2004), 167-190.

WFY [20] F.-Y. Wang, Gradient estimates and applications for SDEs in Hilbert space with multiplicative noise and Dini continuous drift, *J.Differential Equations* **260** (2016), 2792-2829.

WZ [21] F.-Y. Wang, Z. Zhang, Talagrand inequality on free path space and application to stochastic reaction diffusion equations, *Acta Mathematicae Applicatae Sinica, English Series.* **36** (2020), 253-261.

W [22] L. Wu, Transportation inequalities for stochastic differential equations of pure jumps, *Ann. Inst. Henri Poincaré Probab. Stat.* **46** (2010), 465-479.

WZ1 [23] L. Wu, Z. Zhang, Talagrand's $T_2$-transportation inequality w.r.t. a uniform metric for diffusions, *Acta Math. Appl. Sin. Engl. Ser.* **20** (2004), 357-364.

XXZZ [24] P. Xia, L. Xie, X. Zhang and G. Zhao, $L^q(L^p)$-theory of stochastic differential equations, *Stoch. Proc. Appl.* **130** (2020) 5188–5211.

WZ2 [25] L. Wu, Z. Zhang, Talagrand's $T_2$-transportation inequality and log-Sobolev inequality for dissipative SPDEs and applications to reaction-diffusion equations, *Chinese Ann. Math. Ser. B* **27** (2006), 243-262.

XZ [26] L. Xie, X, Zhang, Ergodicity of stochastic differential equations with jumps and singular coefficients, *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques. Institut Henri Poincaré.* **56** (2020), 175-229.

YZ [27] C. Yuan, S.-Q. Zhang, A Zvonkin's transformation for stochastic differential equations with singular drift and related applications, *Journal of Differential Equations* **297(5)** (2021), 277–319.

Z [28] X, Zhang, Stochastic Homeomorphism Flows of SDEs with Singular Drifts and Sobolev Diffusion Coefficients, *Elect. J. Probab.*, **16** (2011), 1096-1116.