A Zvonkin's transformation for stochastic differential equations with singular drift and applications

Shao-Qin Zhang[a] and Chenggui Yuan[b]

[a]School of Statistics and Mathematics
Central University of Finance and Economics, Beijing 100081, China
Email: zhangsq@cufe.edu.cn

[b]Department of Mathematics, Swansea University, Bay Campus SA1 8EN, UK.
Email: C.Yuan@swansea.ac.uk

September 9, 2021

Abstract

In this paper, by establishing the localized L^p-L^q estimate and Sobolev estimates for parabolic partial differential equations with a singular first order term and a Lipschitz first order term, a new Zvonkin-type transformation is given for stochastic differential equations with singular and Lipschitz drifts. The associated Krylov's estimate is established. As applications, dimension-free Harnack inequalities are established for stochastic equations with Hölder continuous diffusion coefficient and singular drift term without regularity assumption.

AMS Subject Classification (2020): 60H10, 60J60
Keywords: Zvonkin's transformation; singular diffusion processes; Krylov's estimate; Harnack inequality

1 Introduction

In [36], a transformation that removes the drift of stochastic differential equation (in short SDE) was introduced by Zvonkin. This transformation of the phase space together with Krylov's estimate (see [10]) gives a powerful tool in studying SDEs with irregular coefficients. For instance, in [23] the author first proved the existence and uniqueness of strong solutions to SDEs with bounded measurable drift; [6] proved the uniqueness of strong solution to SDEs with locally Lipschitz and strong elliptic diffusion coefficients and integrable drifts; [32] extended results to equations with local integral drifts which has linear growth and Sobolev diffusion coefficients. [13] obtained the existence and uniqueness of strong solutions to SDEs with additive noise and time dependent drifts satisfying the L^p-L^q integration condition. Krylov and Röckner's results were extended by [33] to the case of multiplicative noise, and stochastic homeomorphism flow property of singular SDEs were studied therein. The ergodicity of SDEs with singular or even distribution coefficients is also investigated by using Zvonkin's transformation, see e.g. [31, 35]. It is clear that the bounded drifts do not satisfy the L^p-L^q condition, and the Zvonkin's transformation used in [31, 33, 35] can not be applied the bounded drift. Recently, some localized integrable space are introduced in [29] to treat the bounded coefficients and the L^p-L^q coefficients in the same framework, and the weak differentiability as well as Bismut-Elworthy-Li's derivative formula are also established therein. For more properties of singular SDEs investigated by using Zvonkin's transformation and Krylov's estimate, see [8, 14, 30, 31, 29, 34] and references therein.

In this paper, we consider the following equation

$$\mathrm{d}X_t = b_1(t, X_t)\mathrm{d}t + b_0(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}W_t, \tag{1.1}$$

where $\{W_t\}_{t\geq 0}$ is a Brownian motion w.r.t. a complete probability space with filtration $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, \mathbb{P})$, $b_1(t,\cdot) : \mathbb{R}^d \to \mathbb{R}^d$ is Lipschitz uniformly w.r.t $t \geq 0$, $\sigma : [0,\infty)\times\mathbb{R}^d \to \mathbb{R}^d\otimes\mathbb{R}^d$ is non-degenerate, $b_0 : [0,\infty)\times\mathbb{R}^d \to \mathbb{R}^d$ is singular term satisfying the local L^p-L^q condition as in [29] with $p, q \in (1,\infty)$ and $\frac{d}{p}+\frac{2}{q} < 1$. We shall give a new Zvonkin-type transformation $\Phi_t(\cdot) : \mathbb{R}^d \to \mathbb{R}^d$ by solving a PDE associated with (1.1). Precisely, let $\phi(t,x) = (\phi^1(t,x),\cdots,\phi^d(t,x))$ satisfy

$$\partial_t\phi^i + \frac{1}{2}\mathrm{tr}(\sigma\sigma^*\nabla^2\phi^i) + \langle b_1 + b_0, \nabla\phi^i\rangle = -b_0^i + \lambda\phi^i,\ \ i = 1,\cdots,d. \tag{1.2}$$

Then $\Phi_t(x) := \phi(t,x) + x$ satisfies the following equation equivalently

$$\partial_t\Phi^i + \frac{1}{2}\mathrm{tr}(\sigma\sigma^*\nabla^2\Phi^i) + \langle b_1 + b_0, \nabla\Phi^i\rangle = b_1^i + \lambda\phi^i,\ \ i = 1,\cdots,d.$$

The equation (1.2) is different from the parabolic equation considered in [29, 30, 31, 32, 33, 34] and can not be covered by their studies since the coefficient b_1 here is allowed to have linear growth. In fact, the L^p-L^q estimate established in [13, (10.3)] or [29, (3.2)] fails for $\partial_t\phi$, see Theorem 2.1 and Remark 2.2 below. We solve (1.2) in a localized weighted space, and more details on the well-posedness and a priori estimates are available in Theorem 2.1 below.

We prove that Φ_t is a homeomorphism on $\mathbb{R}^d$ by choosing λ large enough, see Theorem 2.1 and (4.2) below. Let Φ_t^{-1} is the inverse of Φ_t. By Itô's formula (see Lemma 3.3 for a proof), we have

$$\mathrm{d}\Phi_t(X_t) = b_1(t, X_t)\mathrm{d}t + \lambda\phi(t, X_t)\mathrm{d}t + \text{martingale part}$$

$$= b_1(t, \Phi_t^{-1}(\Phi_t(X_t)))\mathrm{d}t + \lambda\phi_t(\Phi_t^{-1}(\Phi_t(X_t)))\mathrm{d}t + \text{martingale part.}$$

Then $b_1(t, \Phi_t^{-1}(\cdot))$, as a drift term of a SDE for $\Phi_t(X_t)$, remains to be Lipschitz. Moreover, if b is monotone and satisfies

$$\langle b_1(t,x) - b_1(t,y), x-y\rangle \le K|x-y|^2,\ x,y \in \mathbb{R}^d, \tag{1.3}$$

then $b(t, \Phi_t^{-1}(\cdot))$ also satisfies (1.3) with another constant of the same sign as K. However, by applying Zvonkin's transformation used in [30, 31, 32, 33, 34] to (1.1), one gets a SDE with a locally Lipschitz drift term. This property allows us to establish Harnack inequalities for (1.1). The dimension-free Harnack inequality was first introduced in [24], and the log-Harnack inequality was introduced in [19, 25]. These type of inequalities are called Wang's Harnack inequality and Wang's log-Harnack inequality in references to emphasize the essential difference between these inequality and the classical ones. Harnack inequalities are established for various stochastic models, and various applications of theses inequalities are investigated. One can consult [27] and references therein for more details. Wang's Harnack inequalities for SDEs with singular drifts have been investigated in [7, 9, 14, 20]. In [14], only log-Harnack inequality is established for SDEs with the drift satisfying the L^p-L^q condition. [20] obtains Harnack inequalities with an extra constant. In [7], the author imposes extra regularities on space variable, which turns out to require that the drift term should be Hölder continuous, see Remark 4.1. However, the drift term of the SDEs discussed in [7, 14, 20] can not include a Lipschitz drift. [28] introduced a transformation for SDEs with Dini-continuous drift that retains the linear drift which automatically is Lipschitzian. Following this transformation, [9] obtained Harnack inequalities for stochastic functional partial differential equations with Dini-continuous drift. We establish Harnack inequality with power for (1.1) under localized L^p-L^q integral condition with $\frac{d}{p} + \frac{2}{q} < 1$ and the diffusion coefficient that can be Hölder continuous with order in $[\frac{1}{2}, 1]$. Moreover, if $\frac{d}{p} + \frac{2}{q} < \frac{1}{2}$ and the Hölder continuity order of the diffusion coefficient is in $(\frac{1}{2}, 1]$, then the Harnack inequality without extra constant is established. We use a coupling modified from [26, 27] so that the diffusion coefficient can be Hölder continuity with index in $(\frac{1}{2}, 1]$. This is new even in the case that the drift is regular.

This paper is structured as follows. In Section 2, we investigate well-posedness and a priori estimates of a parabolic equation which covers (1.2). Then Krylov's estimates for the solution of (1.1) will be given in Section 3. In Section 4, we study Harnack inequality for the associated transition semigroup generated by (1.1).

We finish this section by introduce some spaces and notations which will be used throughout this paper.

We denote $B_r(x) = \{y \in \mathbb{R}^d \mid |x-y| \le r\}$. Let $\{e_j\}_{j=1}^d$ be the orthonormal basis of $\mathbb{R}^d$. For any $A \in \mathbb{R}^d \otimes \mathbb{R}^d$ which denotes all $d \times d$ matrix, we set $A^i_j = \langle Ae_j, e_i\rangle$. For any $g \in C^1(\mathbb{R}^d)$, we denote by $\nabla g(x)$ the gradient of g at x with

$$(\nabla g)^j(x) := \langle \nabla g(x), e_j\rangle := \nabla_{e_j} g(x).$$

For any $g \in C^1(\mathbb{R}^d, \mathbb{R}^d)$, we denote by $\nabla g(x) \in \mathbb{R}^d \otimes \mathbb{R}^d$ the gradient of g with

$$g^j(x) := \langle g(x), e_j \rangle, \qquad (\nabla g)_i^j(x) := \langle \nabla g(x) e_i, e_j \rangle = \nabla_{e_i} g^j(x).$$

Particularly, for any $g \in C^2(\mathbb{R}^d)$, we denote by $\nabla^2 g(x) \in \mathbb{R}^d \otimes \mathbb{R}^d$ the Hessen matrix of g at x with

$$(\nabla^2 g)_j^i(x) := \langle \nabla^2 g(x) e_j, e_i \rangle := \nabla_{e_j} \nabla_{e_i} g(x).$$

We denote by $\mathcal{C}_b^2$ the continuous function on $\mathbb{R}^d$ with bounded first and second order derivates.
Denote by $\|\cdot\|$ the operator norm of matrixes. For a (real, vector or matrix value) function on $[0, T] \times \mathbb{R}^d$, we denote

$$\|f\|_{T,\infty} = \sup_{t \in [0,T], x \in \mathbb{R}^d} \|f(t, x)\|, \quad \|f(t, \cdot)\|_\infty = \sup_{x \in \mathbb{R}^d} \|f(t, x)\|.$$

We denote $L^p = L^p(\mathbb{R}^d, \mathrm{d}x)$ and by $\|\cdot\|_p$ the usual L^p-norm. Let w be a positive function. Define the weighted L^p-space $L^p_w = L^p(\mathbb{R}^d, w(x)\mathrm{d}x)$ and denote by $\|\cdot\|_{p,w}$ the L^p-norm under the measure $w(x)\mathrm{d}x$. If $w \equiv 1$, then $\|\cdot\|_p = \|\cdot\|_{p,1}$.
For $(\theta, p) \in [0, 2] \times (1, +\infty)$, let $H^{\theta,p} = (\mathbb{1} - \Delta)^{-\frac{\theta}{2}}(L^p)$ be the usual Bessel potential space with norm

$$\|f\|_{\theta,p} = \|(\mathbb{1} - \Delta)^{\frac{\theta}{2}} f\|_p.$$

Let $\chi \in C_c^\infty(\mathbb{R}^d)$ with $\mathbb{1}_{[|x| \le 1]} \le \chi \le \mathbb{1}_{[|x| \le 2]}$. We define

$$\chi_r(x) = \chi(\frac{x}{r}), \qquad \chi_r^z(x) = \chi\left(\frac{x - z}{r}\right), \; r > 0, x, z \in \mathbb{R}^d.$$

We denote by $\tilde{H}^{\theta,p}$ the localized $H^{\theta,p}$-space introduced in [29]:

$$\tilde{H}^{\theta,p} := \left\{ f \in H^{\theta,p}_{loc}(\mathbb{R}^d) \mid \|f\|_{\tilde{H}^{\theta,p}} := \sup_z \|\chi_r^z f\|_{\theta,p} < \infty \right\}.$$

Given $0 \le t \le T$, $p, q \in [1, +\infty]$. we denote $L^p_q(t, T) = L^q([t, T], L^p)$ and $L^{p,w}_q(t, T) = L^q([t, T], L^p_w)$, and denote by $\|\cdot\|_{L^p_q(t,T)}$ and $\|\cdot\|_{L^{p,w}_q(t,T)}$ the norms on these spaces respectively. We denote by $\|\cdot\|_{W^{\theta,p}_q(t,T)}$ the norm of $W^{\theta,p}_q(t, T) := L^q([t, T], H^{\theta,p})$. We also use the following Sobolev space

$$W^{p,w}_{1,q}(t, T) = \left\{ f \in L^p_q(t, T) \mid \partial_t f \in L^{p,w}_q(t, T) \right\},$$
$$\|f\|_{W^{p,w}_{1,q}(t,T)} = \|\partial_t f\|_{L^{p,w}_q(t,T)} + \|f\|_{L^p_q(t,T)},$$

where $\partial_t f$ is denoted as the derivative w.r.t. the first variable of f.
The localized space of $W^{\theta,p}_q(t, T)$ is

$$\tilde{W}^{\theta,p}_q(t, T) = \left\{ f \in L^q([t, T], H^{\theta,p}_{loc}) \mid \|f\|_{\tilde{W}^{\theta,p}_q(t,T)} := \sup_{z \in \mathbb{R}^d} \|\chi_r^z f\|_{W^{\theta,p}_q(t,T)} < \infty \right\}.$$

If $\theta = 0$, we denote $\tilde{L}_q^p(t,T) = \tilde{W}_q^{0,p}(t,T)$, and

$$\|f\|_{\tilde{L}_q^p(t,T)} = \sup_{z\in\mathbb{R}^d} \|\chi_r^z f\|_{L_q^p(t,T)}. \tag{1.4}$$

The localized space of $W_{1,q}^{p,w}(t,T)$ is

$$\tilde{W}_q^{p,w}(t,T) = \left\{ f \in \tilde{L}_q^p([t,T]) \mid \sup_z \|\chi_r^z \partial_t f\|_{L_q^{p,w}(t,T)} < \infty \right\},$$

and the associated norm

$$\|f\|_{\tilde{W}_q^{p,w}(t,T)} := \sup_z \|\chi_r^z \partial_t f\|_{L_q^{p,w}(t,T)} + \|f\|_{\tilde{L}_q^p(t,T)}.$$

Let $\tilde{\mathcal{W}}_{1,q}^{2,p,w}(t,T) = \tilde{W}_q^{p,w}(t,T) \bigcap \tilde{W}_q^{2,p}(t,T)$, and

$$\|f\|_{\tilde{W}_{1,q}^{2,p,w}(t,T)} = \|f\|_{\tilde{W}_{1,q}^{p,w}(t,T)} + \|f\|_{\tilde{W}_q^{2,p}(t,T)}.$$

We denote by $\tilde{\mathcal{W}}_{1,q}^{2,p}(t,T)$ the case that $w \equiv 1$. If $t = 0$, then $L_q^{p,w}(0,T)$, $L_q^p(0,T)$ etc. will be denoted by $L_q^{p,w}(T)$, $L_q^p(T)$ etc., respectively.

2 L^p-L^q estimates for parabolic equations

We first study the L^p-L^q estimates of the following parabolic equation on $[0,T]$

$$(\partial_t + L)u := \partial_t u + \mathrm{tr}\big(a\nabla^2 u\big) + (b_1 + b_0)\cdot \nabla u = \lambda u + f,\ u(T,\cdot) = 0, \tag{2.1}$$

where the derivatives of u are understood in the weak sense, and $a : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$, $b_1, b_0 : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ and $f : [0,T]\times\mathbb{R}^d \to \mathbb{R}$ are measurable. We assume that a, b_1 satisfy the following hypothesis.

(H1) a is uniformly continuous in x uniformly w.r.t. t, i.e. for any $T > 0$ and $\epsilon > 0$, there exists $\delta > 0$ such that for any $x, y \in \mathbb{R}^d$ with $|x - y| < \delta$

$$\sup_{t\in[0,T]} ||a(t,x) - a(t,y)|| < \epsilon.$$

For each $T > 0$, there exist positive constants κ_1, κ_2 with $\kappa_1 \le \kappa_2$ such that

$$\kappa_1|v|^2 \le \langle a(t,x)v, v\rangle \le \kappa_2|v|^2,\ (t,x) \in [0,T]\times\mathbb{R}^d, v \in \mathbb{R}^d.$$

(H2) For every $T > 0$ and $t \in [0,T]$, $b_1(t,\cdot)$ is Lipschitz continuous with Lipschitz constant $\|\nabla b_1(t,\cdot)\|_\infty$, and

$$\sup_{t\in[0,T]} (|b_1|(t,0) + \|\nabla b_1(t,\cdot)\|_\infty) < \infty.$$

The condition **(H2)** implies that b_1 has linear growth: there exists $K_0 > 0$ such that

$$\sup_{t\in[0,T]} |b_1(t,x)| \le K_0(1+|x|),\ x\in\mathbb{R}^d. \tag{2.2}$$

Remark 2.1. *Let $b_1 : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ satisfy $\|\nabla b_1\|_{T,\infty} < \infty$. Then there exist $\tilde b_1(t,\cdot) \in \mathcal{C}_b^2$ and bounded $\tilde b_2$ such that $b_1 = \tilde b_1 + \tilde b_2$. In fact, for any nonnegative $\eta \in C_c^\infty(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d}\eta = 1$, we set*

$$\tilde b_1(t,x) = (b_1(t,\cdot) * \eta)(x), \qquad \tilde b_2(t,x) = b_1(t,x) - \tilde b_1(t,x).$$

Then it is clear that $\tilde b_1(t,\cdot) \in C^2(\mathbb{R}^d)$ and

$$\|\nabla \tilde b_1\|_{T,\infty} + \|\nabla^2 \tilde b_1\|_{T,\infty} + \|\tilde b_2\|_{T,\infty} \le C\|\nabla b_1\|_{T,\infty}.$$

Consequently, $\tilde b_2 \in \tilde L^p_q(T)$ for any $p,q \in [1,+\infty]$, and $b_0 + \tilde b_2 \in \tilde L^p_q(T)$ for some $p,q \in (1,+\infty)$ if so is b_0.

Due to this remark, $b_1 + b_0 = \tilde b_1 + (b_0 + \tilde b_2)$, and we can use the following **(H2')** to replace **(H2)** if $b_0 \in \tilde L^p_q(T)$:

(H2') For every $T > 0$ and $t \in [0,T]$, $b_1(t,\cdot) \in \mathcal{C}_b^2$ and

$$\sup_{t\in[0,T]} \big(|b_1|(t,0) + \|\nabla b_1(t,\cdot)\|_\infty + \|\nabla^2 b_1(t,\cdot)\|_\infty\big) < \infty.$$

We denote by Ξ the parameter set $p, q, d, \kappa_1, \kappa_2, T, \|\nabla b_1\|_{T,\infty}$ and the continuity modulus of a

Theorem 2.1. *Let $p,q \in (1,\infty)$, $w(x) = (1+|x|^2)^{\frac{p}{2}}$. Assume that* **(H1)** *and* **(H2)** *hold and $b_0 \in \tilde L^{p_1}_{q_1}(T)$ for some $p_1 \in [p,+\infty]$ and $q_1 \in [q,+\infty]$ with $\frac{d}{p_1} + \frac{2}{q_1} < 1$. Then for any $f \in \tilde L^p_q(T)$, there exists a unique strong solution $u \in \tilde{\mathcal{W}}^{2,p,w}_{1,q}$ to (2.1). Moreover, for any $T > 0$, there exists $\lambda_0 > 0$ depending on Ξ and $\|b_0\|_{\tilde L^{p_1}_{q_1}(T)}$ such that for any $\theta \in [0,2)$, $p_2 \in [p,+\infty]$ and $q_2 \in [q,+\infty]$ with $\frac{d}{p} + \frac{2}{q} < 2 - \theta + \frac{d}{p_2} + \frac{2}{q_2}$, there is a constant $C > 0$ such that for any $\lambda \ge \lambda_0$*

$$\lambda^{\frac{1}{2}(2-\theta+\frac{d}{p_2}+\frac{1}{q_2}-\frac{d}{p}-\frac{2}{q})}\|u\|_{\tilde W^{\theta,p_2}_{q_2}(T)} + \|(\partial_t + b_1\cdot\nabla)u\|_{\tilde L^p_q(T)} + \|u\|_{\tilde W^{2,p}_q(T)} \le C\|f\|_{\tilde L^p_q(T)}. \tag{2.3}$$

Remark 2.2. *The L^p-L^q estimate, or generally the maximal L^q-regularity for parabolic equations, is crucial to establish the Zvonkin transformation and the Krylov estimate. The (2.1) with $u(T,\cdot) = 0$ is said to have maximal L^q-regularity on $[t,T]$ (see [1, 5, 18] for more general definition) if there exists a constant $C > 0$ such that for any $f \in L^p_q(t,T)$, there is a unique solution satisfying*

$$\|\partial_t u\|_{L^p_q(t,T)} + \|u\|_{W^{2,p}_q(t,T)} \le C\|f\|_{L^p_q(t,T)}.$$

[5, 29] extend a maximal L^q-regularity result of [11] to parabolic equations with the highest order coefficients assumed to be measurable in time and continuous in the space variables, and this is an important step to obtain the L^p-L^q estimate for equations with singular coefficients under the assumption **(H1)** *on the coefficient of second order term. However, the maximal L^q-regularity for evolution equations with time independent operator implies that the operator generates an analytic semigroup, see [18, Proposition 2.2]. While the generator of O-U semigroup can not generate an analytic semigroup in $L^p(\mathbb{R}^d)$, see [15]. Note that b_1 with linear growth is not necessary in $L^p(\mathbb{R}^d)$. The elliptic operator in* (2.1) *covers the generator of O-U semigroup as an example. One can not expect to derive the maximal regularity for* (2.1) *in $\mathcal{W}^{2,p}_{1,q}(T)$. In the cases of localized spaces, it is also easy to see that $b_1(t,\cdot)$ is not necessary in $\tilde{L}^p(\mathbb{R}^d)$ if b_1 has linear growth. According to our proof, see* (2.11)*, we have that $\partial_t u + b_1\nabla u = \nabla v$ for some $v \in \tilde{\mathcal{W}}^{2,p}_{1,q}(T)$, and $\partial_t u$ can not be controlled by $\|f\|_{\tilde{L}^p_q}$. That is why we have to take the sum of $\partial_t u$ and $b_1 \cdot \nabla u$ into consideration in* (2.3).

2.1 Proof of Theorem 2.1

We first give a transform to remove b_1 under the assumption **(H2')**. This transformation has been used in [4, 15, 16, 17] to investigate elliptic operators with unbounded coefficients. Consider the following ordinary differential equation (ODE)

$$\frac{\mathrm{d}\psi}{\mathrm{d}t}(t,x) = b_1(t,\psi(t,x)),\ \ \psi(T,x) = x. \tag{2.4}$$

For the solution of (2.4), we have the following lemma.

Lemma 2.2. *Assume that* **(H2')** *holds. Then, for any $t \in [0,T]$, $\psi(t,\cdot) \in \mathcal{C}^2_b$ and $\psi(t,\cdot)$ is a diffeomorphism on $\mathbb{R}^d$. Denote by $\psi^{-1}(t,\cdot)$ with the inverse of $\psi(t,\cdot)$. Then ψ^{-1} satisfies the following ODE*

$$\frac{\mathrm{d}\psi^{-1}}{\mathrm{d}t}(t,x) = -(\nabla\psi)^{-1}(t,\psi^{-1}(t,x))b_1(t,x), \qquad \psi^{-1}(T,x) = x \tag{2.5}$$

and $\psi^{-1}(t,\cdot) \in \mathcal{C}^2_b$. Moreover,

$$\begin{aligned}&\|\nabla\psi\|_{T,\infty} + \|(\nabla\psi)^{-1}\|_{T,\infty} + \|\nabla\psi^{-1}\|_{T,\infty}\\ &\qquad + \sup_{1\le l\le d}\left(\|\nabla^2\psi^l\|_{T,\infty} + \|\nabla^2[(\psi^{-1})^l]\|_{T,\infty}\right) < \infty,\end{aligned} \tag{2.6}$$

where $(\nabla\psi)^{-1}(t,x)$ is the inverse of the matrix $\nabla\psi(t,x)$ and satisfies

$$(\nabla\psi)^{-1}(t,x) = (\nabla\psi^{-1})(t,\psi(t,x)),\ \ (t,x) \in [0,T]\times\mathbb{R}^d, \tag{2.7}$$

and the upperbound of (2.6) *only depends on $T, \|\nabla b_1\|_{T,\infty}, \|\nabla^2 b_1\|_{T,\infty}, d$.*

The proof of this lemma is similar to that of [16, Section 2] for the case that b_1 is independent of t, and we omit it here.

For any $u \in C^{1,2}([0,T]\times\mathbb{R}^d)$, set $v(t,x) = u(t,\psi(t,x))$, by a direct calculus, we have that

$$\begin{aligned}
(\nabla u)(t,\psi(t,x)) &= \left\{[(\nabla\psi)^{-1}]^*\nabla v\right\}(t,x),\\
(\nabla^2 u)(t,\psi(t,x)) &= \left\{[(\nabla\psi)^{-1}]^*\nabla^2 v\left[(\nabla\psi)^{-1}\right]\right\}(t,x)\\
&\quad -\sum_{j=1}^d \left\{[(\nabla\psi)^{-1}]^*(\nabla^2\psi^j)(\nabla\psi)^{-1}\right\}\left([(\nabla\psi)^{-1}]^*\nabla v\right)^j(t,x).
\end{aligned}$$

This implies

$$\begin{aligned}
\partial_t v(t,x) &= \partial_t u(t,\psi(t,x)) + (b_1\cdot\nabla u)(t,\psi(t,x)),\\
(Lu)(t,\psi(t,x)) &= \mathrm{tr}\left(a(t,\psi)[(\nabla\psi)^{-1}]^*\nabla^2 v(\nabla\psi)^{-1}\right)(t,x)\\
&\quad -\left\{\sum_{j=1}^d \mathrm{tr}\left(a(t,\psi)[(\nabla\psi)^{-1}]^*(\nabla^2\psi^j)(\nabla\psi)^{-1}\right)\left([(\nabla\psi)^{-1}]^*\nabla v\right)^j\right\}(t,x)\\
&\quad +\left\{\left[(\nabla\psi)^{-1}(b_0+b_1)(t,\psi)\right]\cdot\nabla v\right\}(t,x).
\end{aligned}$$

Let $\bar L$ be a differential operator defined as follows

$$\bar L g = \mathrm{tr}(\bar a\nabla^2 g) + \bar b_0\cdot\nabla g,\ \ g\in C^2([0,T]\times\mathbb{R}^d),$$

where

$$\begin{aligned}
\bar a(t,x) &= \left\{(\nabla\psi)^{-1}a(t,\psi)[(\nabla\psi)^{-1}]^*\right\}(t,x),\\
\bar b_0(t,x) &= (\nabla\psi)^{-1}(t,x)b_0(t,\psi(t,x))\\
&\quad -\sum_{j=1}^d\left\{\mathrm{tr}\left(a(t,\psi)[(\nabla\psi)^{-1}]^*(\nabla^2\psi^j)(\nabla\psi)^{-1}\right)[(\nabla\psi)^{-1}]_j\right\}(t,x).
\end{aligned}$$

Then, by setting $\bar f(t,x) = f(t,\psi(t,x))$, we have that

$$\begin{aligned}
&[(\partial_t+\bar L-\lambda)v](t,x) - \bar f(t,x)\\
&\qquad = [(\partial_t + L-\lambda)u](t,\psi(t,x)) - f(t,\psi(t,x)).
\end{aligned}\tag{2.8}$$

It follows from Lemma 2.2 that if $b_0\in\tilde L_q^p(T)$, then $\bar b_0\in\tilde L_q^p(T)$. Moreover, since **(H1)**, (2.6) and (2.7), it is clear that $\bar a$ is uniformly continuous and uniformly elliptic with some $\bar\kappa_1$, $\bar\kappa_2$ such that

$$\bar\kappa_1|w|^2\le\langle\bar a(t,x)w,w\rangle\le\bar\kappa_2|w|^2,\ \ (t,x)\in[0,T]\times\mathbb{R}^d, w\in\mathbb{R}^d.$$

For any $g\in\mathscr{B}([0,T]\times\mathbb{R}^d)$, we define a mapping J as follows

$$(Jg)(t,x) = g(t,\psi^{-1}(t,x)),\ \ (t,x)\in[0,T]\times\mathbb{R}^d.$$

Hence, due to (2.8), we can investigate (2.1) by applying [29, Theorem 3.2] to

$$(\partial_t + \bar{L} - \lambda)v = \bar{f},\ v(T,\cdot) = 0, \tag{2.9}$$

and showing that $u = Jv$ satisfies (2.1). Then the assertion of Theorem 2.1 holds. Let

$$\tilde{\mathcal{W}}^{2,p}_{1,q,b_1}(T) = \left\{ g \in \tilde{\mathcal{W}}^{2,p,w}_{1,q}(T) \,\middle|\, (\partial_t + b_1\cdot\nabla)g \in \tilde{L}^p_q(T) \right\}$$

equipped the norm

$$\|\cdot\|_{\tilde{\mathcal{W}}^{2,p,w}_{1,q,b_1}(T)} := \|(\partial_t + b_1\cdot\nabla)\cdot\|_{\tilde{L}^p_q(T)} + \|\cdot\|_{\tilde{\mathcal{W}}^{2,p,w}_{1,q}(T)}.$$

The following lemma shows that J has nice properties.

Lemma 2.3. *Let $p, q \in (1,\infty)$ and $w(x) = (1+|x|^2)^{-\frac{p}{2}}$.*
(1) *Let $\theta \in [0,2]$, and let*

$$(J_t g)(x) = (Jg)(t,x),\ (t,x) \in [0,T]\times\mathbb{R}^d, g \in \mathscr{B}(\mathbb{R}^d) \subset \mathscr{B}([0,T]\times\mathbb{R}^d).$$

Then $\{J_t\}_{t\in[0,T]}$ are uniformly bounded linear operators in $H^{\theta,p}$ and $\tilde{H}^{\theta,p}$. Consequently, J is a homeomorphism on $W^{\theta,p}_q(T)$ and $\tilde{W}^{\theta,p}_q(T)$. Moreover, for any $g \in \tilde{W}^{\theta,p}_q(T)$ and a.e. $t\in[0,T]$,

$$(\nabla Jg)(t,\cdot) = \left[J[(\nabla\psi)^{-1}]^*\nabla g\right](t,\cdot). \tag{2.10}$$

(2) *The operator J is bounded from $\tilde{\mathcal{W}}^{2,p}_{1,q}(T)$ to $\tilde{\mathcal{W}}^{2,p,w}_{1,q}(T)$, and*

$$(\partial_t + b_1\cdot\nabla)Jg = J\partial_t g,\ g \in \tilde{\mathcal{W}}^{2,p}_{1,q}(T). \tag{2.11}$$

Moreover, J is a homeomorphism from $\tilde{\mathcal{W}}^{2,p}_{1,q}(T)$ to $\tilde{\mathcal{W}}^{2,p}_{1,q,b_1}(T)$, and there exist positive constants c_1, c_2 such that

$$c_1\|\partial_t g\|_{\tilde{L}^p_q(T)} \le \|(\partial_t + b_1\cdot\nabla)Jg\|_{\tilde{L}^p_q(T)} \le c_2\|\partial_t g\|_{\tilde{L}^p_q(T)},\ g \in \tilde{\mathcal{W}}^{2,p}_{1,q}(T). \tag{2.12}$$

(3) *If $\frac{d}{p}+\frac{2}{q} < 1$, then for every $g \in \tilde{\mathcal{W}}^{2,p}_{1,q,b_1}(T)$ and $\delta \in (0, 1-\frac{d}{p}-\frac{2}{q})$, there exists $C>0$ depending on $\|\nabla^2\psi^{-1}\|_{T,\infty}$ and $\|\nabla\psi^{-1}\|_{T,\infty}$ and δ such that*

$$\|\nabla g\|_{T,\infty} + \frac{|\nabla g(t,x) - \nabla g(t,y)|}{|x-y|^\delta} \le C\|g\|_{\tilde{\mathcal{W}}^{2,p}_{1,q,b_1}(T)},\ x,y\in\mathbb{R}^d. \tag{2.13}$$

Remark 2.3. *We remark here that all the assertions of (2) and (3) of this lemma hold with the localized spaces $\tilde{L}^p_q(T)$, $\tilde{\mathcal{W}}^{2,p}_{1,q}(T)$, $\tilde{\mathcal{W}}^{2,p,w}_{1,q}(T)$ and $\tilde{\mathcal{W}}^{2,p}_{1,q,b_1}(T)$ replaced by the original spaces $L^p_q(T)$, $\mathcal{W}^{2,p}_{1,q}(T)$, $\mathcal{W}^{2,p,w}_{1,q}(T)$ and $\mathcal{W}^{2,p}_{1,q,b_1}(T)$ respectively.*

Proof. (1) For any $h\in C_c^2(\mathbb{R}^d)$ and $j\in\{0,1,2\}$, it follows from (2.6) that

$$\nabla^j(J_th)(x)=h(\psi^{-1}(t,x))\mathbb{1}_{[j=0]}+\sum_{i=1}^d\left[(\nabla^j(\psi^{-1})^i)\partial_i h(\psi^{-1})\right](t,x)\mathbb{1}_{[j\geq 1]}$$
$$+\left[(\nabla\psi^{-1})^*\nabla^2 h(\psi^{-1})(\nabla\psi^{-1})\right](t,x)\mathbb{1}_{[j=2]}. \quad (2.14)$$

Then

$$\int_{\mathbb{R}^d}|\nabla^j(J_th)(x)|^p\mathrm{d}x=\int_{\mathbb{R}^d}|\nabla^j(J_th)(\psi(t,x))|^p|\det\nabla\psi(t,x)|\mathrm{d}x$$
$$\leq C_{T,\|\nabla b_1\|_\infty,p,j}\sum_{k=0}^j\int_{\mathbb{R}^d}|\nabla^k h(x)|^p\mathrm{d}x.$$

Since $C_c^2(\mathbb{R}^d)$ is dense in $H^{j,p}$, it is easy to see that J_t is a bounded operator in $H^{j,p}$ with $\sup_{t\in[0,T]}\|J_t\|_{j,p}<\infty$ for any $j\in\{0,1,2\}$. Moreover, (2.14) holds for any $h\in H^{j,p}$. By the interpolation theorem (see e.g. [3]), it is clear that $\{J_t\}_{t\in[0,T]}$ is uniformly bounded in $H^{\theta,p}$.

Since for any $g\in W_q^{\theta,p}(T)$, $(Jg)(t,x)=(J_tg(t,\cdot))(x)$. Then J is a bounded operator in $W_q^{\theta,p}(T)$. It is easy to see that J_t and J are invertible, and

$$(J^{-1}g)(t,x)=g(t,\psi(t,x)),\ g\in\mathscr{B}(\mathbb{R}^d).$$

It is easy to see that J^{-1} is a bounded linear operator. By (2.14), (2.7) and that for any $g\in W_q^{2,p}(T)$, $g(t,\cdot)\in H^{2,p}$ for a.e. $t\in[0,T]$. Then (2.10) follows from (2.14) and the approximation argument.

For every $g\in\tilde{H}^{\theta,p}$,

$$(\chi_r^zJ_tg)(x)=(\chi_r^z(\psi(t,\cdot))g)(\psi^{-1}(t,x))=J_t(\chi_r^z(\psi(t,\cdot))g)(x). \quad (2.15)$$

Since

$$|x-\psi^{-1}(t,z)|=|\psi^{-1}(t,\psi(t,x))-\psi^{-1}(t,z)|\leq\|\nabla(\psi^{-1})\|_{T,\infty}|\psi(t,x)-z|,$$

we have that $\mathrm{supp}(\chi_r^z(\psi(t,\cdot)))\subset B_{\|\nabla(\psi^{-1})\|_{T,\infty}}(\psi^{-1}(t,z))$ and

$$\chi_r^z(\psi(t,x))g(x)=\chi_r^z(\psi(t,x))\chi_{2r\|\nabla\psi^{-1}\|_{T,\infty}}^{\psi^{-1}(t,z)}(x)g(x). \quad (2.16)$$

Due to Lemma 2.2,

$$\sup_{z\in\mathbb{R}^d}\|\nabla^j(\chi_r^z(\psi(t,\cdot)))\|_{T,\infty}$$
$$\leq\frac{C}{r^j}\sum_{k=0}^j\|\nabla^k\chi\|_\infty\left(\mathbb{1}_{[j=0]}+\sum_{k=1}^j\|\nabla^k\psi\|_{T,\infty}^{j-k+1}\mathbb{1}_{[j\geq 1]}\right),\ 0\leq j\leq 2.$$

Thus, the multipliers $\{\chi_r^z(\psi(t,\cdot))\}_{t\in[0,T],z\in\mathbb{R}^d}$ induce uniformly bounded operators in $H^{\theta,p}$. Then, taking into account (2.16), we have that

$$\sup_{z\in\mathbb{R}^d,t\in[0,T]}\|\chi_r^z(\psi(t,\cdot))g\|_{\theta,p}\leq C_{\psi,r,T}\sup_{z\in\mathbb{R}^d}\|\chi_{2r\|\nabla\psi^{-1}\|_{T,\infty}}^{\psi^{-1}(t,z)}g\|_{\theta,p}$$

$$\leq C_{\psi,r,T} \sup_{z\in\mathbb{R}^d} \|\chi^z_{2r\|\nabla\psi^{-1}\|_{T,\infty}} g\|_{\theta,p}.$$

Due to that $\{J_t\}_{t\in[0,T]}$ is uniformly bounded in $H^{\theta,p}$, we have by (2.15) that

$$\begin{aligned}\sup_{t\in[0,T]} \|J_t g\|_{\tilde{H}^{\theta,p}} &= \sup_{t\in[0,T],z\in\mathbb{R}^d} \|\chi^z_r J_t g\|_{\theta,p}\\ &\leq C_{\psi,r,T} \sup_{z\in\mathbb{R}^d} \|\chi^z_{2r\|\nabla\psi^{-1}\|_{T,\infty}} g\|_{\theta,p}\\ &\leq \tilde{C}_{\psi,r,T}\|g\|_{\tilde{H}^{\theta,p}}.\end{aligned}$$

Hence, $\{J_t\}_{t\in[0,T]}$ are uniformly bounded linear operators in $\tilde{H}^{\theta,p}$, and J is a homeomorphism on $\tilde{W}^{\theta,p}_q(T)$. Moreover, for each $R>0$, $\chi_R g \in W^{2,p}_q(T)$. Then

$$(\nabla J(\chi_R g))(t,\cdot) = \left[J[(\nabla\psi)^{-1}]^*\nabla(\chi_R g)\right](t,\cdot).$$

Since $R>0$ is arbitrary, we get (2.10) a.e. on $[0,T]$.

(2) For any $R>0$, by Lemma 2.2, there exists $C_{T,R}>0$ such that

$$\chi_R(\psi(t+\epsilon,\cdot)) \leq \chi_{C_{T,R}}(\cdot),\ t\in[0,T], |\epsilon|\leq 1. \tag{2.17}$$

For every $g \in \tilde{\mathcal{W}}^{2,p}_{1,q}(T)$ and $R>0$, $\chi_R g \in \mathcal{W}^{2,p}_{1,q}(T)$. Then we have by [2, Theorem 1.17] and (2.17) that

$$\begin{aligned}&\overline{\lim_{\epsilon\to 0}} \left\| \left(\frac{g(t+\epsilon,\psi^{-1}(t+\epsilon,\cdot)) - g(t,\psi^{-1}(t+\epsilon,\cdot))}{\epsilon} - \partial_t g(t,\psi^{-1}(t+\epsilon,\cdot))\right)\chi_R\right\|_p\\ &\quad\leq C_{T,\psi}\overline{\lim_{\epsilon\to 0}} \left\| \left(\frac{g(t+\epsilon,\cdot)-g(t,\cdot)}{\epsilon} - \partial_t g(t,\cdot)\right)\chi_R(\psi(t+\epsilon,\cdot))\right\|_p\\ &\quad\leq C_{T,\psi}\overline{\lim_{\epsilon\to 0}} \left\| \left(\frac{g(t+\epsilon,\cdot)-g(t,\cdot)}{\epsilon} - \partial_t g(t,\cdot)\right)\chi_{C_{T,R}}\right\|_p\\ &\quad= 0,\ \text{a.e. } t\in[0,T].\end{aligned} \tag{2.18}$$

Since $\partial_t g(t,\cdot)\chi_{C_{T,R}} \in L^p$, there exist $g_\delta \in \mathcal{C}^2_b$ such that

$$\|(g_\delta(\cdot) - \partial_t g(t,\cdot))\chi_{C_{T,R}}\|_p \leq \delta.$$

Then

$$\begin{aligned}&\|(\partial_t g(t,\psi^{-1}(t+\epsilon,\cdot)) - \partial_t g(t,\psi^{-1}(t,\cdot)))\chi_R\|_p\\ &\quad\leq C_T\|(\partial_t g(t,\cdot) - \partial_t g(t,\cdot))\chi_R(\psi(t+\epsilon,\cdot))\|_p\\ &\qquad+ \|g_\delta(\psi^{-1}(t+\epsilon,\cdot)) - g_\delta(\psi^{-1}(t,\cdot))\|_p\\ &\qquad+ C_T\|(\partial_t g(t,\cdot) - \partial_t g(t,\cdot))\chi_R(\psi(t,\cdot))\|_p\\ &\quad\leq 2C_T\|(\partial_t g(t,\cdot) - g_\delta(t,\cdot))\chi_{C_{T,R}}\|_p\\ &\qquad+ \|(g_\delta(\psi^{-1}(t+\epsilon,\cdot)) - g_\delta(\psi^{-1}(t,\cdot)))\chi_R\|_p\\ &\quad\leq 2C_T\delta + \|(g_\delta(\psi^{-1}(t+\epsilon,\cdot)) - g_\delta(\psi^{-1}(t,\cdot)))\chi_R\|_p.\end{aligned}$$

It is clear that by Lemma 2.2

$$\begin{aligned}
&\big|(g_\delta(\psi^{-1}(t+\epsilon,x)) - g_\delta(\psi^{-1}(t,x)))\big| \\
&\quad \le \|\nabla g_\delta\|_\infty \left|\psi^{-1}(t+\epsilon,x) - \psi^{-1}(t,x)\right| \\
&\quad \le \|\nabla g_\delta\|_\infty \|(\nabla\psi)^{-1}\|_\infty \left|\int_t^{t+\epsilon} |b_1(r,x)|\mathrm{d}r\right| \\
&\quad \le \|\nabla g_\delta\|_\infty \|(\nabla\psi)^{-1}\|_\infty |\epsilon| K_0(1+|x|) \\
&\quad = C_{T,R,\delta}|\epsilon|(1+|x|).
\end{aligned}$$

Thus

$$\lim_{\epsilon\to 0} \|(\partial_t g(t,\psi^{-1}(t+\epsilon,\cdot)) - \partial_t g(t,\psi^{-1}(t,\cdot)))\chi_R\|_p = 0.$$

Combining this with (2.18) and that for a.e. $t \in [0,T]$, $g(t,\cdot) \in H^{2,p}_{loc}$, we have in L^p_{loc} that

$$\begin{aligned}
&\lim_{\epsilon\to 0} \frac{1}{\epsilon}\left(g(t+\epsilon,\psi^{-1}(t+\epsilon,\cdot)) - g(t,\psi^{-1}(t,\cdot))\right) \\
&\quad = \lim_{\epsilon\to 0} \frac{1}{\epsilon}\left(g(t+\epsilon,\psi^{-1}(t+\epsilon,\cdot)) - g(t,\psi^{-1}(t+\epsilon,\cdot))\right) \\
&\qquad + \lim_{\epsilon\to 0} \frac{1}{\epsilon}\left(g(t,\psi^{-1}(t+\epsilon,\cdot)) - g(t,\psi^{-1}(t,\cdot))\right) \\
&\quad = \partial_t g(t,\psi^{-1}(t,\cdot)) - \langle \nabla g(t,\psi^{-1}(t,\cdot)), (\nabla\psi)^{-1}(t,\psi^{-1}(t,\cdot))b_1(t,\cdot)\rangle \\
&\quad = \partial_t g(t,\psi^{-1}(t,\cdot)) - \big\langle (\nabla g)(t,\psi^{-1}(t,\cdot)), (\nabla\psi^{-1})(t,\cdot)b_1(t,\cdot)\big\rangle \\
&\quad = \partial_t g(t,\psi^{-1}(t,\cdot)) - \big\langle (\nabla\psi^{-1})^*(t,\cdot)(\nabla g)(t,\psi^{-1}(t,\cdot)), b_1(t,\cdot)\big\rangle \\
&\quad = (J\partial_t g)(t,\cdot) - \langle (\nabla J g(t,\cdot)), b_1(t,\cdot)\rangle. \qquad (2.19)
\end{aligned}$$

Thus

$$\partial_t Jg = J\partial_t g - b_1\cdot\nabla Jg,\ \ g\in \tilde{\mathcal{W}}^{2,p}_{1,q}(T).$$

Consequently, for $g \in \mathcal{W}^{2,p}_{1,q}(T)$, we have by the assertion (1) of this lemma that

$$\begin{aligned}
\|\partial_t Jg\|_{L^{p,w}_q(T)} &\le \|J\partial_t g\|_{L^{p,w}_q(T)} + \|b_1\cdot\nabla Jg\|_{L^{p,w}_q(T)} \\
&\le \|J\partial_t g\|_{L^p_q(T)} + K_0\|\nabla Jg\|_{L^p_q(T)} \\
&\le C_T\|\partial_t g\|_{L^p_q(T)} + K_0\|Jg\|_{W^{1,p}_q(T)} \\
&\le C_T\|\partial_t g\|_{L^p_q(T)} + C_T K_0\|g\|_{W^{2,p}_q(T)} \\
&\le C\|g\|_{\mathcal{W}^{2,p}_{1,q}(T)}.
\end{aligned}$$

Combining this with that J is a homeomorphism on $W^{s,p}_q(T)$, we get that J is bounded from $\mathcal{W}^{2,p}_{1,q}(T)$ to $\mathcal{W}^{2,p,w}_{1,q}(T)$. Similarly,

$$\begin{aligned}
\|\chi^z_r\partial_t Jg\|_{L^{p,w}_q(T)} &\le \|\chi^z_r J\partial_t g\|_{L^{p,w}_q(T)} + \|\chi^z_r b_1\cdot\nabla Jg\|_{L^{p,w}_q(T)} \\
&\le \|\chi^z_r J\partial_t g\|_{L^p_q(T)} + K_0\|\chi^z_r\nabla Jg\|_{L^p_q(T)}
\end{aligned}$$

$$\leq C_{r,T}\|\partial_t g\|_{\tilde{L}^p_q(T)} + K_0\|\nabla(\chi^z_r Jg)\|_{L^p_q(T)} + K_0\|\chi^z_{r+1}Jg\|_{L^p_q(T)}$$
$$\leq C_T\|\partial_t g\|_{\tilde{L}^p_q(T)} + C_T K_0\|g\|_{\tilde{W}^{2,p}_q(T)}$$
$$\leq C\|g\|_{\tilde{\mathcal{W}}^{2,p}_{1,q}(T)}.$$

Then J is bounded from $\tilde{\mathcal{W}}^{2,p}_{1,q}(T)$ to $\tilde{\mathcal{W}}^{2,p,w}_{1,q}(T)$.

By Lemma 2.2, we have similarly that

$$\partial_t J^{-1}g = \left(J^{-1}\left(\partial_t + b_1\cdot\nabla\right)g\right),\ g\in\tilde{\mathcal{W}}^{2,p,w}_{1,q}(T). \tag{2.20}$$

Moreover, for $g\in\tilde{\mathcal{W}}^{2,p}_{1,q,b_1}(T)$, $\partial_t J^{-1}g\in\tilde{\mathcal{W}}^{2,p}_{1,q}(T)$. Hence, J is a homeomorphism from $\tilde{\mathcal{W}}^{2,p}_{1,q}(T)$ to $\tilde{\mathcal{W}}^{2,p}_{1,q,b_1}(T)$. Moreover, (2.12) follows from (2.11) and (2.20).

(3) We first prove that for every $h\in\tilde{\mathcal{W}}^{2,p}_{1,q}(T)$ and $\delta\in(0,1-\frac{d}{p}-\frac{2}{q})$, there is $C>0$ depending on T,δ,d such that

$$|\nabla h|_{T,\infty} + \frac{|\nabla h(t,x)-\nabla h(t,y)|}{|x-y|^\delta} \leq C\|h\|_{\tilde{\mathcal{W}}^{2,p}_{1,q}(T)},\ x,y\in\mathbb{R}^d, t\in[0,T].$$

For any $r>0$ and $z\in\mathbb{R}^d$, $\chi^z_r h\in\mathcal{W}^{2,p}_{1,q}(T)$. Then it follows from [13, Lemma 10.2] or [12, Theorem 7.3] that

$$|\nabla(\chi^z_r h)|_{T,\infty} + \frac{|\nabla(\chi^z_r h)(t,x)-\nabla(\chi^z_r h)(t,y)|}{|x-y|^\delta} \leq C\|\chi^z_r h\|_{\mathcal{W}^{2,p}_{1,q}(T)},\ x,y\in\mathbb{R}^d.$$

Thus

$$\begin{aligned}
&|\nabla h|_{T,\infty} + \sup_{|x-y|<r}\frac{|\nabla h(t,x)-\nabla h(t,y)|}{|x-y|^\delta}\\
&\leq \sup_{z\in\mathbb{R}^d}\|\nabla(\chi^z_r h)\|_{T,\infty} + \sup_{z\in\mathbb{R}^d,|x-y|<r}\frac{|\nabla(\chi^z_r h)(t,x)-\nabla(\chi^z_r h)(t,y)|}{|x-y|^\delta}\\
&\leq C\sup_{z\in\mathbb{R}^d}\|\chi^z_r h\|_{\tilde{\mathcal{W}}^{2,p}_{1,q}(T)}\\
&= C\|h\|_{\tilde{\mathcal{W}}^{2,p}_{1,q}(T)}.
\end{aligned}$$

For $|x-y|\geq 2r$, it is clear that

$$\frac{|\nabla h(t,x)-\nabla h(t,y)|}{|x-y|^\delta} \leq \frac{\|\nabla h\|_{T,\infty}}{2^\delta r^\delta} \leq \frac{C\|h\|_{\tilde{\mathcal{W}}^{2,p}_{1,q}(T)}}{2^\delta r^\delta}.$$

Hence, the assertion follows.

Now, we can prove the assertion for g. Let $h=J^{-1}g$. Then $h\in\mathcal{W}^{2,p}_{1,q}(T)$. Thus, it is easy to see from (2) that

$$\|\nabla g\|_{T,\infty} \leq \|\nabla\psi\|_{T,\infty}\|\nabla h\|_{T,\infty} \leq C\|h\|_{\tilde{\mathcal{W}}^{2,p}_{1,q}(T)} \leq C\|g\|_{\tilde{\mathcal{W}}^{2,p}_{1,q,b_1}(T)}.$$

For ∇g, we have by (2.10) that

$$\|\nabla g(t,x)-\nabla g(t,y)\| = \left\|\left[J[(\nabla\psi)^{-1}]^*\nabla g\right](t,x) - \left[J[(\nabla\psi)^{-1}]^*\nabla g\right](t,y)\right\|$$

$$\begin{aligned}&\leq \|\nabla g\|_{T,\infty}\|[(\nabla\psi)^{-1}](t,\psi^{-1}(t,x)-[(\nabla\psi)^{-1}](t,\psi^{-1}(t,y)\|\\&\quad+\|(\nabla\psi)^{-1}\|_{T,\infty}|\psi^{-1}(t,x)-\psi^{-1}(t,y)|^\delta\\&\leq C\|g\|_{\tilde{\mathcal{W}}^{2,p}_{1,q,b_1}(T)}|x-y|^\delta.\end{aligned}$$

□

Proof of Theorem 2.1:

By [29, Theorem 3.2], the PDE (2.9) has a unique strong solution v in $\tilde{\mathcal{W}}^{2,p}_{1,q}$, and for $\lambda\geq\lambda_0$ with λ_0 depending on Ξ and $\|b_0\|_{\tilde{L}^{p_1}_{q_1}(T)}$, there is $C>0$ depending on Ξ, $\|b_0\|_{\tilde{L}^{p_1}_{q_1}(T)}, p_2, q_2, \theta$ so that

$$\lambda^{\frac{1}{2}(2-\theta+\frac{d}{p_2}+\frac{1}{q_2}-\frac{d}{p}-\frac{2}{q})}\|v\|_{\tilde{W}^{\theta,p_2}_{q_2}(T)}+\|\partial_t v\|_{\tilde{L}^p_q(T)}+\|v\|_{\tilde{W}^{2,p}_q(T)}\leq C\|\bar{f}\|_{\tilde{L}^p_q(T)}.$$

Let $u=Jv$. Then assertions of this theorem with **(H2)** replaced by **(H2')** follows from Lemma 2.3 directly. By Lemma 2.3, with b_0,b_1 replacing by $b_0+\tilde{b}_2,\tilde{b}_1$ respectively, we can see that assertions of this theorem hold. □

3 Krylov's estimate

Let X_t satisfy the following equation

$$X_t=X_0+\int_0^t(b_1+b_0)(s,X_s)\mathrm{d}s+\int_0^t\sigma(s,X_s)\mathrm{d}W_s+\int_0^t\xi(s)\mathrm{d}s, \tag{3.1}$$

where $\xi(t)$ is an $\mathscr{F}_t$-adapted process, and

$$\mathbb{P}\left(\int_0^t|\xi(s)|\mathrm{d}s<\infty\right)=1,\ t\geq 0.$$

We investigate Krylov's estimate for X_t in this section. Let $a=\frac{1}{2}\sigma\sigma^*$. By using (2.1) and (2.9), we can prove the following Krylov's estimate for X_t.

Theorem 3.1. *Let $p,q,p_1,q_1\in(1,\infty)$ with*

$$\frac{d}{p}+\frac{2}{q}<2,\qquad \frac{d}{p_1}+\frac{2}{q_1}<1.$$

Assume **(H1)**, **(H2)** *and $b_0\in\tilde{L}^{p_1}_{q_1}(T)$. Let τ be a stopping time and $0\leq t_0<t_1\leq T$. Then for any $f\in\tilde{L}^p_q(T)$ and $\delta\in(0,1)$, there exist a constant $C_\delta>0$ depending on Ξ and $\|b_0\|_{\tilde{L}^{p_1}_{q_1}(T)}$ such that*

$$\begin{aligned}&\mathbb{E}\left(\int_{t_0\wedge\tau}^{t_1\wedge\tau}f(s,X_s)\mathrm{d}s\Big|\mathscr{F}_{t_0}\right)\\&\qquad\leq\left\{C_\delta+\delta\left(\mathbb{E}\left(\int_{t_0\wedge\tau}^{t_1\wedge\tau}|\xi(s)|^2\mathrm{d}s\Big|\mathscr{F}_{t_0}\right)\right)^{\frac{1}{2}}\right\}^{k_{p,q}}\|f\|_{\tilde{L}^p_q(t_0,t_1)},\end{aligned}$$

where $k_{p,q}$ is the smallest integer that greater than $\log_2\left(\frac{2}{2-\frac{d}{p}-\frac{2}{q}}\right)$.

Remark 3.1. *Comparing with [29, 31], it is new to allow an integrable process ξ in Krylov's estimate for* (3.1)*. The Krylov's estimate in [31, Lemma 7.4] is established for SDEs with drifts that can have some growth condition, while the right hand side of the Krylov's estimate there depends on X_0. In this theorem, the constant C_δ is independent of X_0 although b_1 has linear growth.*

If $\xi = 0$, under the assumption of Theorem 3.1 and $\nabla\sigma \in \tilde{L}^{p_2}_{q_2}(T)$, it follows from [29, Theorem 1.1] that (3.1) *has a unique local strong solution. It can be proved that the solution of* (3.1) *is nonexplosive, see* (3.2)*,* (3.3) *and* (3.4) *in the proof of this theorem. We will use this Krylov's estimate to establish Harnack inequalities in Section 4.*

3.1 Proofs of Theorem 3.1

To prove Theorem 3.1, we need the following two lemmas. Denote

$$\eta_1(\tau) = \mathbb{E}\left(\int_{t_0\wedge\tau}^{t_1\wedge\tau} |\xi(s)|\mathrm{d}s\Big|\mathscr{F}_{t_0}\right),\ \eta_2(\tau) = \left(\mathbb{E}\left(\int_{t_0\wedge\tau}^{t_1\wedge\tau} |\xi(s)|^2\mathrm{d}s\Big|\mathscr{F}_{t_0}\right)\right)^{\frac{1}{2}}.$$

Lemma 3.2. *Assume* **(H1)**, **(H2')** *and $p, q \in (1,\infty)$. Let τ be a stopping time and $0 \le t_0 \le t_1 \le T$.*
(1) If b_0 is bounded, then for any $f \in \tilde{L}^p_q(T)$ with $\frac{d}{p} + \frac{2}{q} < 2$ and any $\delta \in (0,1)$, there exist positive constants C depending on Ξ and $\|b_0\|_{\tilde{L}^p_q(T)}$ such that

$$\mathbb{E}\left(\int_{t_0\wedge\tau}^{t_1\wedge\tau} f(s,X_s)\mathrm{d}s\Big|\mathscr{F}_{t_0}\right) \le (C_\delta + \delta\eta_2(\tau))^{k_{p,q}}\, \|f\|_{\tilde{L}^p_q(t_0,t_1)}.$$

If, $\frac{d}{p} + \frac{2}{q} < 1$ furthermore, $\eta_2(\tau)$ can be replaced by $\eta_1(\tau)$, and $k_{p,q} = 1$.
(2) If $\frac{d}{p} + \frac{2}{q} < 1$ and $b_0 \in \tilde{L}^p_q(T)$, then for any $f \in L^p_q(T)$ and $\delta \in (0,1)$, there exists a positive constant C_de depending on Ξ, $\|b_0\|_{\tilde{L}^p_q(T)}$ and δ such that

$$\mathbb{E}\left(\int_{t_0\wedge\tau}^{t_1\wedge\tau} f(s,X_s)\mathrm{d}s\Big|\mathscr{F}_{t_0}\right) \le (C_\delta + \delta\eta_1(\tau))\, \|f\|_{\tilde{L}^p_q(t_0,t_1)}.$$

By using Krylov's estimate in Lemma 3.2, we can establish a generalized Itô's formula for $u \in \mathcal{W}^{2,\tilde{p},w}_{1,\tilde{q},\tilde{b}_1}(T)$, where $\tilde{b}_1$ satisfies **(H2')** which can be different from b_1 and $\frac{d}{\tilde{p}} + \frac{2}{\tilde{q}} < 1$. Let $\tilde{J}$ be defined as J with b_1 replaced by $\tilde{b}_1$. Then we have the following lemma.

Lemma 3.3. *Let X_t be a solution of* (3.1) *and $u \in \tilde{\mathcal{W}}^{2,\tilde{p}}_{1,\tilde{q},\tilde{b}_1}(T)$ with $\frac{d}{\tilde{p}} + \frac{2}{\tilde{q}} < 1$. Then $\mathbb{P}$-a.s.*

$$\begin{aligned} u(t,X_t) = u(t_0,X_{t_0}) &+ \int_{t_0}^t (\partial_t + L)\, u(s,X_s)\mathrm{d}s + \int_{t_0}^t \langle\xi(s), \nabla u(s,X_s)\rangle\mathrm{d}s \\ &+ \int_{t_0}^t \langle\nabla u(s,X_s), \sigma(s,X_s)\mathrm{d}W_s\rangle,\ 0 \le t_0 \le t \le T. \end{aligned}$$

The proofs of Lemma 3.2 and Lemma 3.3 will be given in the next subsection. Now, we prove Theorem 3.1 by using these lemmas.

Proof of Theorem 3.1: In the definition of the operator L, let $a = \frac{1}{2}\sigma\sigma^*$. By Theorem 2.1, we have $u = (u^1, \cdots, u^d) \in \tilde{\mathcal{W}}^{2,p_1,w}_{1,q_1,b_1}(t_1)$, where u_i solves the PDE on $[0, t_1]$:

$$\partial_t u^i + Lu^i = -b_0^i + \lambda u^i,\ u^i(t_1, \cdot) = 0,\ i = 1, \cdots, d.$$

Let $U(t, x) = x + u(t, x)$. By Lemma 3.3, we have

$$\begin{aligned} \mathrm{d}U(t, X_t) &= (b_1(t, X_t) + b_0(t, X_t) + \xi(t))\mathrm{d}t + (\partial_t u(t, X_t) + Lu(t, X_t))\,\mathrm{d}t \\ &\quad + (I + \nabla u(t, X_t))\,\sigma(t, X_t)\mathrm{d}W_t + \nabla u(t, X_t)\xi(t)\mathrm{d}t \\ &= (b_1 + \lambda u)(t, X_t)\mathrm{d}t + (I + \nabla u(t, X_t))\,\sigma(t, X_t)\mathrm{d}W_t \\ &\quad + (I + \nabla u(t, X_t))\,\xi(t)\mathrm{d}t. \end{aligned} \tag{3.2}$$

By Theorem 2.1, for large enough λ, we have that $\|\nabla u\|_{t_1,\infty} < 1$. Then $U(s, \cdot) : \mathbb{R}^d \to \mathbb{R}^d$ is a diffeomorphism on $\mathbb{R}^d$ for any $0 \le s \le t_1$, and

$$\|\nabla U^{-1}\|_{t_1,\infty} + \|\nabla U\|_{t_1,\infty} < \infty. \tag{3.3}$$

Thus, letting $Y_t = U(t, X_t)$, we have

$$\begin{aligned} \mathrm{d}Y_t &= \big(b_1(t, U^{-1}(t, Y_t)) + \lambda u(t, U^{-1}(t, Y_t)) + (I + \nabla u(t, X_t))\,\xi(t)\big)\,\mathrm{d}t \\ &\quad + \big(I + \nabla u(t, U^{-1}(t, Y_t))\big)\,\sigma(t, U^{-1}(t, Y_t))\mathrm{d}W_t. \end{aligned} \tag{3.4}$$

Since $(b_1 + \lambda u)(t, U^{-1}(t, x))$ satisfies **(H2)**, it follows from Remark 2.1 that we can apply (1) of Lemma 3.2 to Y_t. Taking into account that

$$\begin{aligned} \int_{\mathbb{R}^d} \big|f(s, U^{-1}(s, x))\big|^p\,\mathrm{d}x &= \int_{\mathbb{R}^d} |f(s, y)|^p\,|\det(\nabla U^{-1}(s, y))|\mathrm{d}x \\ &\le C \int_{\mathbb{R}^d} |f(s, y)|^p\,\mathrm{d}y,\ f \in L^p_q(t_0, t_1) \end{aligned}$$
$$\text{and } |\,(I + \nabla u(t, X_t))\,\xi(t)| \le 2|\xi(t)|,\ t \in (t_0, t_1),$$

we have that for any $\delta \in (0, 1)$, there is $C'_\delta > 0$ depending on Ξ, $\|b_0\|_{\tilde{L}^{p_1}_{q_1}(T)}$ and δ so that

$$\begin{aligned} \mathbb{E}\left(\int_{t_0\wedge\tau}^{t_1\wedge\tau} f(s, X_s)\mathrm{d}s\Big|\mathscr{F}_{t_0}\right) &= \mathbb{E}\left(\int_{t_0\wedge\tau}^{t_1\wedge\tau} f(s, U^{-1}(s, Y_s))\mathrm{d}s\Big|\mathscr{F}_{t_0}\right) \\ &\le \{C_\delta + \delta\eta_2(\tau)\}^{k_{p,q}}\,\|f\|_{\tilde{L}^p_q(t_0,t_1)}. \end{aligned}$$

Therefore, we complete the proof.

3.2 Proofs of Lemma 3.2 and Lemma 3.3

Proof of Lemma 3.2:

(1) Let $r = d + 3$ and $f \in \tilde{L}^p_q(T) \cap \tilde{L}^r_r(T)$ with any $p, q \in (1, +\infty)$. By Theorem 2.1, the following equation has a unique solution

$$(\partial_t + L - \lambda)u = f,\ u(t_1, \cdot) = 0.$$

Let ψ be the solution of (2.4) with T replaced by t_1 and J being the mapping induced by ψ^{-1} as in Lemma 2.3. Then $v := J^{-1}u$ satisfies (2.9) with T replaced by t_1. Let ρ be a non-negative smooth function on $\mathbb{R}^{d+1}$ with compact support in the unit ball centre at zero and $\int_{\mathbb{R}^{d+1}} \rho(t, x)\mathrm{d}t\mathrm{d}x = 1$. Set $\rho_n(t, x) = n^{d+1}\rho(nt, nx)$. Extending v by zero for $t \geq t_1$ and by $v(0, \cdot)$ for $t < 0$. We define

$$\begin{aligned} v_n(t, x) &= \int_{\mathbb{R}^{d+1}} v(t - s, x - y)\rho_n(s, y)\mathrm{d}s\mathrm{d}y, \\ \bar{f}_n &= (\partial_t + \bar{L} - \lambda)v_n. \end{aligned}$$

Recalling $\bar{f}(t, x) = f(t, \psi(t, x))$, noting that $\bar{b}_0$ is bounded and $v_n \in \tilde{\mathcal{W}}^{2,p}_{1,q}$, one has as the proof of [33, Theorem 2.1] that

$$\lim_{n\to+\infty} \|(\bar{f}_n - \bar{f})\chi_R\|_{L^r_r(t_1)} = 0,\ R > 0.$$

Let $f_n = J\bar{f}_n$. Owing to Lemma 2.3, there exists $C > 0$ and $\tilde{R} > 0$ such that

$$\overline{\lim_{n\to+\infty}} \|(f_n - f)\chi_R\|_{L^r_r(t_1)} \leq C \overline{\lim_{n\to+\infty}} \|(\bar{f}_n - \bar{f})\chi_{\tilde{R}}\|_{L^r_r(t_1)} = 0.$$

This together with [6, Lemma 3.1] yields that

$$\begin{aligned} &\overline{\lim_{n\to+\infty}} \mathbb{E}\left(\int_0^{\tau\wedge t_1\wedge\tau_R} |f_n(s, X_s) - f(s, X_s)|\,\mathrm{d}s\right) \\ &\leq \overline{\lim_{n\to+\infty}} \mathbb{E}\left(\int_0^{t_1\wedge\tau_R} |f_n(s, X_s) - f(s, X_s)|\,\chi_R(X_s)\mathrm{d}s\right) \\ &\leq C \overline{\lim_{n\to+\infty}} \|(f_n - f)\chi_R\|_{L^r_r(t_1)} \\ &= 0, \end{aligned} \tag{3.5}$$

where $\tau_R = \inf\{t > 0 \mid |X_t| \geq R\}$. By (2.5) and **(H2')**, $\psi^{-1}(t, x)$ is absolutely continuous in t and $\frac{\mathrm{d}\psi^{-1}}{\mathrm{d}t}(t, x)$ is continuous in x uniformly w.r.t $t \in [0, t_1]$. Then $u_n := Jv_n$ is absolutely continuous in t and $(\partial_t u_n)(t, x)$ is continuous in x uniformly w.r.t $t \in [0, t_1]$. Moreover, due to Lemma 2.2 and that

$$\|\nabla v_n\|_{t_1,\infty} + \|\nabla^2 v_n\|_{t_1,\infty} < \infty,$$

we have that $u_n \in C^{0,2}_b([0, t_1] \times \mathbb{R}^d)$. Then, we can apply the Itô formula to $u_n(t, X_t)$ and obtain that

$$u_n(t, X_t) = u_n(0, X_0) + \int_0^t (\partial_s + L)\,u_n(s, X_s)\mathrm{d}s + \int_0^t \xi(s) \cdot \nabla u_n(s, X_s)\mathrm{d}s$$

$$
\begin{aligned}
&+\int_0^t \langle \nabla u_n(s,X_s), \sigma(s,X_s)\mathrm{d}W_s\rangle \\
&= u_n(0,X_0) + \int_0^t (f_n+\lambda u_n)(s,X_s)\mathrm{d}s + \int_0^t \xi(s)\cdot\nabla u_n(s,X_s)\mathrm{d}s \\
&+\int_0^t \langle \nabla u_n(s,X_s), \sigma(s,X_s)\mathrm{d}W_s\rangle. \qquad (3.6)
\end{aligned}
$$

Since $\|\nabla u_n\|_{t_1,\infty}<\infty$, Doob's optional theorem yields

$$
\mathbb{E}\left(\int_{t_0\wedge\tau\wedge\tau_R}^{t_1\wedge\tau\wedge\tau_R} \langle \nabla u_n(s,X_s), \sigma(s,X_s)\mathrm{d}W_s\rangle \Big|\mathscr{F}_{t_0}\right)=0.
$$

Then

$$
\begin{aligned}
&\mathbb{E}\left(u_n(t_1\wedge\tau, X_{t_1\wedge\tau\wedge\tau_R}) - u_n(t_0\wedge\tau, X_{t_0\wedge\tau\wedge\tau_R})\Big|\mathscr{F}_{t_0}\right) \\
&\geq \mathbb{E}\left(\int_{t_0\wedge\tau\wedge\tau_R}^{t_1\wedge\tau\wedge\tau_R} f_n(s,X_s)\mathrm{d}s\Big|\mathscr{F}_{t_0}\right) - \lambda \sup_{s\in[t_0,t_1]}\|u_n(s,\cdot)\|_\infty \\
&- \sup_{s\in[t_0,t_1]}\|\nabla u_n(s)\|_\infty \mathbb{E}\left(\int_{t_0\wedge\tau\wedge\tau_R}^{t_1\wedge\tau\wedge\tau_R} |\xi(s)|\mathrm{d}s\Big|\mathscr{F}_{t_0}\right). \qquad (3.7)
\end{aligned}
$$

By Theorem 2.1 and Lemma 2.3, for large enough λ it holds that

$$
\begin{aligned}
\sup_{n\geq 1, t\in[t_0,t_1]}\|u_n(t,\cdot)\|_\infty &= \sup_{n\geq 1, t\in[t_0,t_1]}\|v_n(t,\cdot)\|_\infty \leq \sup_{t\in[t_0,t_1]}\|v(t,\cdot)\|_\infty \\
&\leq C\sup_{t\in[t_0,t_1]}\|u(t,\cdot)\|_\infty \leq C\lambda^{-\frac{1}{2}(2-\frac{d}{p}-\frac{2}{q})}\|f\|_{\tilde{L}^p_q(t_0,t_1)}, \qquad (3.8)
\end{aligned}
$$

where the constant C depends on Ξ and $\|b_0\|_{\tilde{L}^p_q(T)}$.

We first assume that $\frac{d}{p}+\frac{2}{q}<1$. Then, by Theorem 2.1 and Lemma 2.3, we have that

$$
\begin{aligned}
\sup_{n\geq 1, t\in[t_0,t_1]}\|\nabla u_n\|_\infty &= \sup_{n\geq 1, t\in[t_0,t_1]}\|(\nabla\psi^{-1})^*\nabla v_n(t,\cdot)\|_\infty \\
&\leq C_1 \sup_{n\geq 1, t\in[t_0,t_1]}\|\nabla v_n(t,\cdot)\|_\infty \\
&\leq C_1 \sup_{t\in[t_0,t_1]}\|\nabla v(t,\cdot)\|_\infty \\
&\leq C_2\lambda^{-\frac{1}{2}(1-\frac{d}{p}-\frac{2}{q})}\|f\|_{\tilde{L}^p_q(t_0,t_1)}
\end{aligned}
$$

for some C_1, C_2 independent of λ. Putting this and (3.8) into (3.7), we have by that

$$
\mathbb{E}\left(\int_{t_0\wedge\tau\wedge\tau_R}^{t_1\wedge\tau\wedge\tau_R} f_n(s,X_s)\mathrm{d}s\Big|\mathscr{F}_{t_0}\right) \leq \left(\frac{C(2+\lambda)}{\lambda^{\frac{1}{2}(2-\frac{d}{p}-\frac{2}{q})}} + \frac{C_2\eta_1(\tau\wedge\tau_R)}{\lambda^{\frac{1}{2}(1-\frac{d}{p}-\frac{2}{q})}}\right)\|f\|_{\tilde{L}^p_q(t_0,t_1)}. \qquad (3.9)
$$

For any $\delta > 0$ we can choose $\lambda > 0$ such that $C_2\lambda^{-\frac{1}{2}(1-\frac{d}{p}-\frac{2}{q})} \leq \delta$. Then we derive from (3.5) that

$$\mathbb{E}\left(\int_{t_0\wedge\tau}^{t_1\wedge\tau} f(s,X_s)\mathrm{d}s\Big|\mathscr{F}_{t_0}\right) \leq (C_\delta + \delta\eta_1(\tau))\, \|f\|_{\tilde{L}^p_q(t_0,t_1)}, \tag{3.10}$$

where C_δ depends on Ξ, $\|b_0\|_{\tilde{L}^p_q(T)}$ and δ.

Let $\gamma_k = 2(1-2^{-k})$, $k \in \mathbb{N}$. We prove the following assertion by using induction:

for all $k \geq 1$, $\delta \in (0,1)$ and $p, q \in (1,+\infty)$ with $\frac{d}{p}+\frac{2}{q} < \gamma_k$, there is $C_\delta > 0$ such that for every $f \in \tilde{L}^p_q(t_0,t_1)$,

$$\mathbb{E}\left(\int_{t_0\wedge\tau}^{t_1\wedge\tau} f(s,X_s)\mathrm{d}s\Big|\mathscr{F}_{t_0}\right) \leq (C_\delta + \delta\eta_2(\tau))^k\, \|f\|_{\tilde{L}^p_q(t_0,t_1)}. \tag{3.11}$$

Noticing that $\lim_{k\to+\infty}\gamma_k = 2$, the assertion of (1) follows.

It is clear that (3.11) holds for $k=1$ since (3.10). Let $f_n, \bar{f}, u_n, u, v_n, v, \psi$ be defined as above. Suppose $\frac{d}{p}+\frac{2}{q} < \gamma_{k+1}$. Let $p' = \frac{\gamma_{k+1}}{\gamma_k}p$, $q' = \frac{\gamma_{k+1}}{\gamma_k}q$. Then $\frac{d}{p'}+\frac{2}{q'} < \gamma_k$. Thus, we have by (3.11) that

$$\begin{aligned}\mathbb{E}\left(\int_{t_0\wedge\tau}^{t_1\wedge\tau} \xi(s)\cdot\nabla u_n(s,X_s)\mathrm{d}s\Big|\mathscr{F}_{t_0}\right) &\leq \eta_2(\tau)\left(\mathbb{E}\left(\int_{t_0\wedge\tau}^{t_1\wedge\tau} |\nabla u_n(s,X_s)|^2\mathrm{d}s\Big|\mathscr{F}_{t_0}\right)\right)^{\frac{1}{2}}\\ &\leq \eta_2(\tau)(C_\delta+\delta\eta_2(\tau))^k \||\nabla u_n|^2\|^{\frac{1}{2}}_{\tilde{L}^{p'}_{q'}(t_0,t_1)}\\ &\leq C\eta_2(\tau)(C_\delta+\delta\eta_2(\tau))^k\|\nabla u_n\|_{\tilde{L}^{2p'}_{2q'}(t_0,t_1)}\end{aligned} \tag{3.12}$$

Since

$$\begin{aligned}1+\frac{d}{2p'}+\frac{2}{2q'}-\frac{d}{p}-\frac{2}{q} &= 1+\frac{\gamma_k}{2\gamma_{k+1}}\left(\frac{d}{p}+\frac{2}{q}\right)-\left(\frac{d}{p}+\frac{2}{q}\right)\\ &= 1-\frac{2\gamma_{k+1}-\gamma_k}{2\gamma_{k+1}}\left(\frac{d}{p}+\frac{2}{q}\right)\\ &> 1-\gamma_{k+1}+\frac{1}{2}\gamma_k\\ &= 0,\end{aligned}$$

it follows from Theorem 2.1, Lemma 2.3, $v_n = v * \rho_n$ and [29, (2.6)] that

$$\begin{aligned}\overline{\lim_{n\to+\infty}}\|\nabla u_n\|_{\tilde{L}^{2p'}_{2q'}(t_0,t_1)} &\leq C_1\overline{\lim_{n\to+\infty}}\|\nabla v_n\|_{\tilde{L}^{2p'}_{2q'}(t_0,t_1)} \leq C_2\overline{\lim_{n\to+\infty}}\|v_n\|_{\tilde{\mathcal{W}}^{1,2p'}_{2q'}(t_0,t_1)}\\ &\leq C_2\|v\|_{\tilde{W}^{1,2p'}_{2q'}(t_0,t_1)} \leq C_3\|u\|_{\tilde{W}^{1,2p'}_{2q'}(t_0,t_1)}\\ &\leq C_4\lambda^{-\frac{1}{2}(1+\frac{d}{2p'}+\frac{2}{2q'}-\frac{d}{p}-\frac{2}{q})}\|f\|_{\tilde{L}^p_q(t_0,t_1)}.\end{aligned}$$

Putting this and (3.12) into (3.6), and arguing as from (3.7) to (3.10), we can see for large enough λ it holds that

$$\begin{aligned}\mathbb{E}\left(\int_{t_0\wedge\tau}^{t_1\wedge\tau} f(s,X_s)\mathrm{d}s\Big|\mathscr{F}_{t_0}\right) &\leq (2+\lambda)\sup_{n\geq 1, t\in[t_0,t_1]}\|u_n(t,\cdot)\|_\infty\\ &\quad+\frac{CC_4(C_\delta+\delta\eta_2(\tau))^k\eta_2(\tau)}{\lambda^{\frac{1}{2}(1+\frac{d}{2p'}+\frac{2}{2q'}-\frac{d}{p}-\frac{2}{q})}}\|f\|_{\tilde{L}^p_q(t_0,t_1)}\\ &\leq \left(\tilde{C}_\delta+\delta\eta_2(\tau)\right)^{k+1}\|f\|_{\tilde{L}^p_q(t_0,t_1)},\end{aligned}$$

where $\tilde{C}_\delta$ is another constant depending on δ. By the definition of γ_k, it is clear that $k>\log_2(2/(2-\frac{d}{p}-\frac{2}{q}))$, then we can set $k_{p,q}$ as required.

(2) We investigate the case that $b_0\in\tilde{L}^p_q(T)$ with $\frac{d}{p}+\frac{2}{q}<1$. Set $m>0$, $M>0$, and let

$$\tau_M=\inf\{t\geq 0 \mid \int_0^t|b_0(s,X_s)|\mathrm{d}s\geq M\}.$$

It is clear that

$$\begin{aligned}X_t=X_0&+\int_0^t\left(b_1+b_0\mathbb{1}_{[|b_0|\leq m]}\right)(s,X_s)\mathrm{d}s+\int_0^t\left((b_0\mathbb{1}_{[|b_0|>m]})(s,X_s)+\xi(s)\right)\mathrm{d}s\\ &+\int_0^t\sigma(s,X_s)\mathrm{d}W_s,\ t>0.\end{aligned}$$

Let $L_t^{[m]}=L_t-(b_0\mathbb{1}_{[|b_0|\geq m]})\cdot\nabla$. Consider

$$(\partial_t+L_t^{[m]}-\lambda)u=f,\ u(t_1,\cdot)=0.$$

Note that C_δ in (3.10) depends on $\|b_0\|_{\tilde{L}^p_q}$ instead of $\|b_0\|_{T,\infty}$ and

$$\|b_0\mathbb{1}_{[|b_0|\leq m]}\|_{\tilde{L}^p_q(T)}\leq\|b_0\|_{\tilde{L}^p_q(T)},$$

we have by (3.10) that

$$\begin{aligned}&\mathbb{E}\left(\int_{t_0\wedge\tau\wedge\tau_M}^{t_1\wedge\tau\wedge\tau_M} f(s,X_s)\mathrm{d}s\Big|\mathscr{F}_{t_0}\right)\\ &\leq (C_\delta+\delta\eta_1(\tau\wedge\tau_M))\|f\|_{\tilde{L}^p_q(t_0,t_1)}\\ &\quad+\delta\mathbb{E}\left(\int_{t_0\wedge\tau\wedge\tau_M}^{t_1\wedge\tau\wedge\tau_M}(|b_0|\mathbb{1}_{[|b_0|>m]})(s,X_s)\mathrm{d}s\Big|\mathscr{F}_{t_0}\right)\|f\|_{\tilde{L}^p_q(t_0,t_1)},\end{aligned}\tag{3.13}$$

where C_δ is a constant independent of m. For fixed $M>0$, we have that

$$\lim_{m\to\infty}\mathbb{E}\int_{t_0\wedge\tau\wedge\tau_M}^{t_1\wedge\tau\wedge\tau_M}|b_0(s,X_s)|\mathbb{1}_{[|b_0(s,X_s)|\geq m]}\mathrm{d}s=0.$$

Thus, taking $m \to +\infty$ first and then $M \to +\infty$ in (3.13), we complete the proof of the second assertion of this lemma.

Proof of Lemma 3.3:

Let ρ be a non-negative smooth function on $\mathbb{R}^d$ with compact support in $B_1(0)$ and $\int_{\mathbb{R}^d} \rho(x)\mathrm{d}x = 1$, and let $\rho_n(x) = n^d \rho(nx)$. For any $n \geq 1$, define

$$u_n(t,x) = \int_{\mathbb{R}^d} u(t,y)\rho_n(x-y)\mathrm{d}y, \ x \in \mathbb{R}^d.$$

We now divide the proof into two steps.

Step (i): we are going to show that the conclusion holds for $u_n(t,x)$.

For any $m \in \mathbb{N}$, Let $t_k = \frac{kt}{m}$, $k = 0, 1, \cdots, m$. Then

$$\begin{aligned} u_n(t,X_t) - u_n(0,X_0) &= \sum_{k=0}^{m-1} \big(u_n(t_{k+1}, X_{t_{k+1}}) - u_n(t_k, X_{t_{k+1}})\big) \\ &\quad + \sum_{k=0}^{m-1} \big(u_n(t_k, X_{t_{k+1}}) - u_n(t_k, X_{t_k})\big) \\ &=: I_{1,m} + I_{2,m}. \end{aligned} \tag{3.14}$$

We first study $I_{1,m}$. Let

$$\partial_t u_n(t,x) = \int_{\mathbb{R}^d} \partial_t u(t,y)\rho_n(x-y)\mathrm{d}y.$$

Then for any $M > 0$, we have by Jensen's inequality that

$$\begin{aligned} \sup_{|x|\leq M} |\partial_t u_n(t,x)| &\leq \sup_{|x|\leq M} \left(\int_{\mathbb{R}^d} |\partial_t u(t,y)|^{\tilde{p}} \rho_n(x-y)\mathrm{d}y \right)^{\frac{1}{\tilde{p}}} \\ &\leq \left(\sup_{|x|\leq M, y\in\mathbb{R}^d} (1+|y|^2)^{\frac{1}{2}} \rho_n^{\frac{1}{\tilde{p}}}(x-y) \right) \|\partial_t u(t,\cdot)\chi_{n+M}\|_{\tilde{p},w} \\ &\equiv C_{M,n,\tilde{p}} \|\partial_t u(t,\cdot)\chi_{n+M}\|_{\tilde{p},w}. \end{aligned} \tag{3.15}$$

This implies that $\sup_{|x|\leq M} |\partial_t u_n(\cdot,x)| \in L^{\tilde{q}}([0,T])$ since that $u \in \tilde{\mathcal{W}}^{2,\tilde{p}}_{1,\tilde{q},\tilde{b}_1}(T) \subset \tilde{\mathcal{W}}^{2,\tilde{p},w}_{1,\tilde{q}}(T)$. Moreover,

$$\begin{aligned} &\sup_{|x|\leq M} \left| \frac{u_n(t+\epsilon,x) - u_n(t,x)}{\epsilon} - \partial_t u_n(t,x) \right| \\ &= \sup_{|x|\leq M} \left| \int_{\mathbb{R}^d} \left(\frac{u(t+\epsilon,y) - u(t,y)}{\epsilon} - \partial_t u(t,y) \right) \rho_n(x-y)\mathrm{d}y \right| \\ &\leq \sup_{|x|\leq M} \left(\int_{\mathbb{R}^d} \left| \frac{u(t+\epsilon,y) - u(t,y)}{\epsilon} - \partial_t u(t,y) \right|^{\tilde{p}} \rho_n(x-y)\mathrm{d}y \right)^{\frac{1}{\tilde{p}}} \end{aligned}$$

$$\leq \tilde{C}_{M,n,\tilde{p}} \left\| \left(\frac{u(t+\epsilon,\cdot)-u(t,\cdot)}{\epsilon} - \partial_t u(t,\cdot) \right) \chi_{n+M} \right\|_{\tilde{p},w}.$$

Because $u \in \tilde{\mathcal{W}}^{2,\tilde{p}}_{1,\tilde{q},\tilde{b}_1}(T) \subset \tilde{\mathcal{W}}^{2,\tilde{p},w}_{1,\tilde{q}}(T)$, $\chi_{n+M}u$ is absolutely continuous in $L^{\tilde{p}}_w$. Then for a.e. $t \in [0,T]$, we have

$$\overline{\lim_{\epsilon\to 0}} \sup_{|x|\leq M} \left| \frac{u_n(t+\epsilon,x)-u_n(t,x)}{\epsilon} - \partial_t u_n(t,x) \right| = 0.$$

Hence,

$$\begin{aligned} I_{1,m} &= \sum_{k=0}^{m-1} \int_0^t (\partial_s u_n)(s, X_{t_{k+1}}) \mathbb{1}_{[t_k \leq s < t_{k+1}]}(s) \mathrm{d}s \\ &= \int_0^t \sum_{k=0}^{m-1} (\partial_s u_n)(s, X_{t_{k+1}}) \mathbb{1}_{[t_k \leq s < t_{k+1}]}(s) \mathrm{d}s. \end{aligned}$$

By the Hölder inequality, for any $|x_1| \leq M$ and $|x_2| \leq M$, we have

$$\begin{aligned} &|\partial_t u_n(t,x_1) - \partial_t u_n(t,x_2)| \\ &\quad\leq \int_{\mathbb{R}^d} |\partial_t u(t,y)| \int_0^1 |\nabla\rho_n(x_2+\theta(x_1-x_2)-y)| \mathrm{d}\theta \mathrm{d}y |x_1-x_2| \\ &\quad\leq |x_1-x_2| \|\partial_t u(t,\cdot)\chi_{n+M}\|_{\tilde{p},w} \\ &\qquad\times \left(\int_{\mathbb{R}^d} \int_0^1 |\nabla\rho_n(x_2+\theta(x_1-x_2)-y)|^{\frac{\tilde{p}}{\tilde{p}-1}} (1+|y|^2)^{\frac{\tilde{p}}{2(\tilde{p}-1)}} \mathrm{d}\theta \mathrm{d}y \right)^{\frac{\tilde{p}-1}{\tilde{p}}} \\ &\quad\leq C_{M,\tilde{p},n} |x_1-x_2| \|\partial_t u(t,\cdot)\chi_{n+M}\|_{\tilde{p},w}, \ \text{a.e. } t \in [0,T]. \end{aligned}$$

Combining this with the pathwise continuity X_t, we have that

$$\lim_{m\to\infty} \sum_{k=0}^{m-1} (\partial_s u_n)(s, X_{t_{k+1}}) \mathbb{1}_{[t_k \leq s < t_{k+1}]}(s) = \partial_s u_n(s, X_s), \ \text{a.e. } s \in [0,T].$$

Due to that $\mathbb{P}$-a.s. $\tilde{M} := \sup_{s\in[0,T]} |X_s| < \infty$ and (3.15), we have that

$$\begin{aligned} &\left| \sum_{k=0}^{m-1} (\partial_s u_n)(s, X_{t_{k+1}}) \mathbb{1}_{[t_k \leq s < t_{k+1}]}(s) \right| \\ &\quad\leq \sum_{k=0}^{m-1} |\sup_{0\leq k\leq m-1} (\partial_s u_n)(s, X_{t_{k+1}})| \mathbb{1}_{[t_k \leq s < t_{k+1}]}(s) \\ &\quad\leq C_{\tilde{M},n,\tilde{p}} \sum_{k=0}^{m-1} \|\chi_{n+M}\partial_s u(s,\cdot)\|_{L^{\tilde{p}}_w} \mathbb{1}_{[t_k \leq s < t_{k+1}]}(s) \\ &\quad= C_{\tilde{M},n,\tilde{p}} \|\chi_{n+M}\partial_s u(s,\cdot)\|_{L^p_w}. \end{aligned}$$

Hence, it follows from the dominated convergence theorem that $\mathbb{P}$-a.s.

$$\lim_{m\to\infty} I_{1,m} = \int_0^t \partial_s u_n(s, X_s)\mathrm{d}s,\ t \in [0, T].$$

Next, we investigate $I_{2,m}$. Since $u \in \tilde{\mathcal{W}}^{2,\tilde{p}}_{1,\tilde{q},\tilde{b}_1}(T)$, we have $\tilde{J}^{-1}u \in \mathcal{W}^{2,\tilde{p}}_{1,\tilde{q}}(T)$. Then it follows from [13, Lemma 10.2] and (2.5) that $u \in C(\mathbb{R} \times \mathbb{R}^d)$. For any $t \in [0, T]$, by the definition of u_n, [13, Lemma 10.2] and Lemma 2.3, we have $u_n(t, \cdot) \in C^2(\mathbb{R}^d)$ and for any $j \in \{0, 1, 2\}$

$$\begin{aligned}\|\nabla^j u_n\|_{T,\infty} &\le \|u\|_{T,\infty} \int_{\mathbb{R}^d} |\nabla^j \rho_n(x-y)|\mathrm{d}y = \|\tilde{J}^{-1}u\|_{T,\infty} \int_{\mathbb{R}^d} |\nabla^j \rho_n(y)|\mathrm{d}y \\ &\le C_{n,j} C_{T,\tilde{p},\tilde{q},d} \left(\|\partial_t \tilde{J}^{-1}u\|_{\tilde{L}^{\tilde{p}}_{\tilde{q}}(T)} + T\|\tilde{J}^{-1}u\|_{\tilde{W}^{2,\tilde{p}}_{\tilde{q}}(T)} \right). \end{aligned} \tag{3.16}$$

For $j \in \{0, 1, 2\}$, any $t, s \in [0, T]$ and $M > 0$,

$$\begin{aligned}&\sup_{|x|\le M} \left\|\nabla^j u_n(t,x) - \nabla^j u_n(s,x)\right\| \\ &\le \sup_{|x|\le M} \int_{\mathbb{R}^d} |u(t,y) - u(s,y)||\nabla^j \rho_n(x-y)|\mathrm{d}y \\ &\le |t-s| \sup_{|x|\le M} \int_0^1 \int_{\mathbb{R}^d} |\partial_t u(s+\theta(t-s), y)||\nabla^j \rho_n(x-y)|\mathrm{d}y\mathrm{d}\theta \\ &\le C_{M,n,\tilde{q},j}|t-s| \left(\int_0^1 \|\partial_t u(s+\theta(t-s),\cdot)\chi_{n+M}\|^{\tilde{q}}_{\tilde{L}^{\tilde{p}}_w} \mathrm{d}\theta \right)^{\frac{1}{\tilde{q}}} \\ &\le C_{M,n,\tilde{q},j} \|u\|_{\tilde{W}^{\tilde{p},w}_{1,\tilde{q}}(T)} |t-s|^{\frac{\tilde{q}-1}{\tilde{q}}}. \end{aligned} \tag{3.17}$$

By Itô's formula, we have

$$\begin{aligned} I_{2,m} &= \sum_{k=0}^{m-1} \left(\int_{t_k}^{t_{k+1}} \langle (b_1+b_0)(s,X_s) + \xi(s), \nabla u_n(t_k, X_s)\rangle \mathrm{d}s \right. \\ &\quad + \frac{1}{2} \int_{t_k}^{t_{k+1}} \mathrm{tr}\left(\sigma\sigma^*(s,X_s)\nabla^2 u_n(t_k,X_s)\right)\mathrm{d}s \\ &\quad \left. + \int_{t_k}^{t_{k+1}} \langle \nabla u_n(t_k, X_s), \sigma(s,X_s)\mathrm{d}W_s\rangle \right) \\ &= \int_0^t \langle (b_1+b_0)(s,X_s) + \xi(s), \sum_{k=0}^{m-1} \nabla u_n(t_k,X_s) \mathbb{1}_{[t_k \le s < t_{k+1}]}(s)\rangle \mathrm{d}s \\ &\quad + \frac{1}{2} \int_0^t \sum_{k=0}^{m-1} \mathrm{tr}\left(\sigma\sigma^*(s,X_s)\nabla^2 u_n(t_k,X_s)\right) \mathbb{1}_{[t_k \le s < t_{k+1}]}(s)\mathrm{d}s \\ &\quad + \int_0^t \langle \sum_{k=0}^{m-1} \nabla u_n(t_k,X_s)\mathbb{1}_{[t_k \le s < t_{k+1}]}(s), \sigma(s,X_s)\mathrm{d}W_s\rangle. \end{aligned}$$

It follows from (3.17) that

$$\begin{aligned}&\left|\sum_{k=0}^{m-1}\left\{(\nabla u_n(t_k,X_s)-\nabla u_n(s,X_s))+\left(\nabla^2 u_n(t_k,X_s)-\nabla^2 u_n(s,X_s)\right)\right\}\mathbb{1}_{[t_k\le s<t_{k+1}]}(s)\right|\\&\quad\le C_{M,n,q}\left(\frac{t}{m}\right)^{\frac{q-1}{q}},\end{aligned}$$

where $M=\sup_{s\in[0,t]}|X_s|$. Then

$$\begin{aligned}&\lim_{m\to\infty}\sum_{k=0}^{m-1}\left(\nabla u_n(t_k,X_s)+\nabla^2 u_n(t_k,X_s)\right)\mathbb{1}_{[t_k\le s<t_{k+1}]}(s)\\&\quad=\nabla u_n(s,X_s)+\nabla^2 u_n(s,X_s).\end{aligned}$$

Since $\int_0^T(|b_1(s,X_s)|+|b_0(s,X_s)|+|\xi(s)|)\,\mathrm{d}s<\infty$ and $\|\sigma\|_\infty<\infty$, it follows from (3.16) and the dominated convergence theorem that $\mathbb{P}$-a.s.

$$\begin{aligned}&\lim_{m\to\infty}\Bigg(\int_0^t\langle(b_1+b_0)(s,X_s)+\xi(s),\sum_{k=0}^{m-1}\nabla u_n(t_k,X_s)\mathbb{1}_{[t_k\le s<t_{k+1}]}(s)\rangle\mathrm{d}s\\&\qquad+\frac{1}{2}\int_0^t\sum_{k=0}^{m-1}\mathrm{tr}\left(\sigma\sigma^*(s,X_s)\nabla^2 u_n(t_k,X_s)\right)\mathbb{1}_{[t_k\le s<t_{k+1}]}(s)\mathrm{d}s\Bigg)\\&=\Bigg(\int_0^t\langle(b_1+b_0)(s,X_s)+\xi(s),\nabla u_n(s,X_s)\rangle\mathrm{d}s\\&\qquad+\frac{1}{2}\int_0^t\mathrm{tr}\left(\sigma\sigma^*(s,X_s)\nabla^2 u_n(s,X_s)\right)\mathrm{d}s\Bigg),\ t\in[0,T].\end{aligned}$$

Let $\tilde\tau_M=\inf\{s>0\mid|X_s|\ge M\}$. Then (3.17) yields that

$$\begin{aligned}&\overline{\lim_{m\to\infty}}\,\mathbb{E}\int_0^{t\wedge\tilde\tau_M}\left|\sigma^*(s,X_s)\sum_{k=0}^{m-1}(\nabla u_n(t_k,X_s)-\nabla u_n(s,X_s))\mathbb{1}_{[t_k\le s<t_{k+1}]}(s)\right|^2\mathrm{d}s\\&\quad\le\lim_{m\to\infty}C_{M,n}\|\sigma\|^2\left(\frac{t}{m}\right)^{\frac{2(q-1)}{q}}t\\&\quad=0.\end{aligned}$$

Hence, $\mathbb{P}$-a.s.

$$\begin{aligned}\lim_{m\to\infty}I_{2,m}&=\int_0^t\langle(b_1+b_0)(s,X_s)+\xi(s),\nabla u_n(s,X_s)\rangle\mathrm{d}s\\&\quad+\frac{1}{2}\int_0^t\mathrm{tr}\left(\sigma\sigma^*\nabla^2 u_n\right)(s,X_s)\mathrm{d}s+\int_0^t\langle\nabla u_n(s,X_s),\sigma(s,X_s)\mathrm{d}W_s\rangle.\end{aligned}$$

By letting $m\to\infty$ in (3.14), we obtain the Itô formula for $u_n(t,x)$:

$$u_n(t,X_t)-u_n(0,X_0)$$

$$\begin{aligned}&= \int_0^t \partial_t u_n(s,X_s)\mathrm{d}s + \int_0^t \langle (b_1+b_0)(s,X_s)+\xi(s), \nabla u_n(s,X_s)\rangle \mathrm{d}s\\
&\quad + \frac{1}{2}\int_0^t \mathrm{tr}\left(\sigma\sigma^*\nabla^2 u_n\right)(s,X_s)\mathrm{d}s + \int_0^t \langle \nabla u_n(s,X_s), \sigma(s,X_s)\mathrm{d}W_s\rangle\\
&=: I_1(n)+I_2(n)+I_3(n)+I_4(n). \qquad (3.18)\end{aligned}$$

Step (ii): we shall complete the proof of this lemma by using the approximation argument.

It is clear that for any $j\in\{0,1\}$, we have $\|\nabla^j u_n\|_{T,\infty} \le \|\nabla^j u\|_{T,\infty}$. It follows from (2.10), $u \in \mathcal{W}^{2,\tilde p}_{1,\tilde q,\tilde b_1}(T)$, [13, Lemma 10.2] and Lemma 2.3 that, there is $C>0$ depend on $T, \|\nabla \tilde b_1\|_{T,\infty}, \|u\|_{\tilde{\mathcal{W}}^{2,\tilde p}_{1,\tilde q,\tilde b_1}(T)}$

$$\sup_{t\in[0,T]} |u(t,x)-u(t,y)| \le C|x-y|.$$

By (2.13), for any $\delta' \in (0, 1-\frac{d}{\tilde p}-\frac{2}{\tilde q})$, there is $C>0$ depend on $\|u\|_{\tilde{\mathcal{W}}^{2,\tilde p}_{1,\tilde q,\tilde b_1}(T)}$, $\|\nabla\tilde b_1\|_{T,\infty}$, T, δ such that

$$\sup_{t\in[0,T]} |\nabla u(t,x)-\nabla u(t,y)| \le C\left(|x-y|^{\delta'}+|x-y|\right).$$

Then for any $j\in\{0,1\}$

$$\begin{aligned}&\|\nabla^j u_n - \nabla^j u\|_{T,\infty}\\
&\quad= \sup_{x\in\mathbb{R}^d, t\in[0,T]} \int_{\mathbb{R}^d} \left|\nabla^j u(t,x-y)-\nabla^j u(t,x)\right| \rho_n(y)\mathrm{d}y\\
&\quad= \sup_{x\in\mathbb{R}^d, t\in[0,T]} \int_{\mathbb{R}^d} \left|\nabla^j u(t,x-\frac{y}{n})-\nabla^j u(t,x)\right| \rho(y)\mathrm{d}y\\
&\quad\le \frac{C_{T,\|\nabla\tilde b_1\|_{T,\infty},\|u\|_{\tilde{\mathcal{W}}^{2,\tilde p}_{1,\tilde q,\tilde b_1}(T)}} \int_{\mathbb{R}^d}\left(|u|\vee|u|^{\delta'}\right)\rho(u)\mathrm{d}u}{n^{\delta'}},\end{aligned}$$

which implies $\lim_{m\to\infty}\|\nabla^j u_n - \nabla^j u\|_{T,\infty}=0$. Applying the dominated convergence theorem to $I_2(n)$ and $I_4(n)$, we get $\mathbb{P}$-a.s. (by a subsequence if necessary)

$$\begin{aligned}\lim_{n\to\infty}(I_2(n)+I_4(n)) = \Bigg(&\int_0^t \langle (b_1+b_0)(s,X_s)+\xi(s), \nabla u(s,X_s)\rangle\\
&+\int_0^t \langle \nabla u(s,X_s), \sigma(s,X_s)\mathrm{d}W_s\rangle,\ t\in[0,T].\end{aligned}$$

Let $R>0$ and $\tau_R = \inf\{t\in[0,T] \mid \int_0^t |\xi(s)|\mathrm{d}s + |X_t| > R\}$. Then it follows from (2) of Lemma 3.2 that

$$\mathbb{E}\int_0^{T\wedge\tau_R} \|\sigma\sigma^*(\nabla^2 u_n - \nabla^2 u)(s,X_s)\|_{HS}\mathrm{d}s$$

$$\leq \|\sigma\|_\infty^2 \mathbb{E}\int_0^{T\wedge\tau_R} \|(\nabla^2 u_n - \nabla^2 u)\chi_R(s,X_s)\|_{HS} \mathrm{d}s$$
$$\leq C\left(1+\mathbb{E}\int_0^{T\wedge\tau_R} |\xi(s)|\mathrm{d}s\right) \|\left(\nabla^2 u_n - \nabla^2 u\right)\chi_R\|_{L^p_q(T)}.$$

It follows from [29, (2.6)] that

$$\lim_{n\to\infty}\int_0^T \|(\nabla^2 u_n(t,\cdot) - \nabla^2 u(t,\cdot))\chi_R\|_{L^{\tilde p}}^{\tilde q}\mathrm{d}t = 0,$$

which implies

$$\lim_{n\to\infty}\mathbb{E}\int_0^{T\wedge\tau_R} \|\sigma\sigma^*(\nabla^2 u_n - \nabla^2 u)(s,X_s)\|_{HS}\mathrm{d}s = 0.$$

Thus there exists a subsequence u_{n_k} such that on $\{\tau_R > T\}$

$$\lim_{k\to\infty}\int_0^t \mathrm{tr}\left(\sigma\sigma^*\nabla^2 u_{n_k}\right)(s,X_s)\mathrm{d}s = \int_0^t \mathrm{tr}\left(\sigma\sigma^*\nabla^2 u\right)(s,X_s)\mathrm{d}s,\ t\in[0,T]. \tag{3.19}$$

Since $\lim_{R\to\infty}\mathbb{P}(\tau_R \leq T) = 0$, by Cantor's diagonal argument, there exists a subsequence, which we also denote by u_{n_k}, such that (3.19) holds $\mathbb{P}$-a.s.

For $I_1(n)$, it follows from the property of convolution that

$$\begin{aligned}\left|\frac{\partial_s u_n(s,x) - \partial_s u(s,x)}{\sqrt{1+|x|^2}}\right| &\leq \int_{\mathbb{R}^d}\frac{|\partial_s u(s,x-y)|}{\sqrt{1+|x-y|^2}}\frac{\left|\sqrt{1+|x-y|^2}-\sqrt{1+|x|^2}\right|}{\sqrt{1+|x|^2}}\rho_n(y)\mathrm{d}y\\
&\quad+\int_{\mathbb{R}^d}\left|\frac{\partial_s u(s,x-y)}{\sqrt{1+|x-y|^2}} - \frac{\partial_s u(s,x)}{\sqrt{1+|x|^2}}\right|\rho_n(y)\mathrm{d}y\\
&\leq \int_{\mathbb{R}^d}\frac{|\partial_s u(s,x-y)|}{\sqrt{1+|x-y|^2}}\frac{|y|}{\sqrt{1+|x|^2}}\rho_n(y)\mathrm{d}y\\
&\quad+\int_{\mathbb{R}^d}\left|\frac{\partial_s u(s,x-y)}{\sqrt{1+|x-y|^2}} - \frac{\partial_t u(s,x)}{\sqrt{1+|x|^2}}\right|\rho_n(y)\mathrm{d}y\\
&\leq \frac{1}{n}\int_{\mathbb{R}^d}\frac{|\partial_s u(s,x-y)|}{\sqrt{1+|x-y|^2}}\rho_n(y)\mathrm{d}y\\
&\quad+\int_{\mathbb{R}^d}\left|\frac{\partial_s u(s,x-y)}{\sqrt{1+|x-y|^2}} - \frac{\partial_t u(s,x)}{\sqrt{1+|x|^2}}\right|\rho_n(y)\mathrm{d}y\\
&\equiv J_{1,n}(s) + J_{2,n}(s).\end{aligned}$$

By convolution inequality,

$$\left(\int_0^T\left(\int_{\mathbb{R}^d}\left(\chi_R\int_{\mathbb{R}^d}\frac{|\partial_s u(s,x-y)|}{\sqrt{1+|x-y|^2}}\rho_n(y)\mathrm{d}y\right)^{\tilde p}\mathrm{d}x\right)^{\frac{\tilde q}{\tilde p}}\mathrm{d}s\right)^{\frac{1}{\tilde q}}$$

$$
\begin{aligned}
&\le \int_{\mathbb{R}^d} \left(\int_0^T \left(\int_{\mathbb{R}^d} \left(\frac{|\partial_s u(s,x-y)|}{\sqrt{1+|x-y|^2}} \right)^{\tilde{p}} \chi_R(x)\mathrm{d}x \right)^{\frac{\tilde{q}}{\tilde{p}}} \mathrm{d}s \right)^{\frac{1}{\tilde{q}}} \rho_n(y)\mathrm{d}y \\
&\le \int_{\mathbb{R}^d} \left(\int_0^T \left(\int_{\mathbb{R}^d} \left(\frac{|\partial_s u(s,x)|}{\sqrt{1+|x|^2}} \right)^{\tilde{p}} \chi_R^{-y}(x)\mathrm{d}x \right)^{\frac{\tilde{q}}{\tilde{p}}} \mathrm{d}s \right)^{\frac{1}{\tilde{q}}} \rho_n(y)\mathrm{d}y \\
&\le \int_{\mathbb{R}^d} \sup_{y\in\mathbb{R}^d} \left(\int_0^T \left(\int_{\mathbb{R}^d} \left(\frac{|\partial_s u(s,x)|}{\sqrt{1+|x|^2}} \right)^{\tilde{p}} \chi_R^{-y}(x)\mathrm{d}x \right)^{\frac{\tilde{q}}{\tilde{p}}} \mathrm{d}s \right)^{\frac{1}{\tilde{q}}} \rho_n(y)\mathrm{d}y \\
&\le C_R \|\partial_s u\|_{\tilde{L}^{\tilde{p},w}_{\tilde{q}}(T)}. \qquad (3.20)
\end{aligned}
$$

Thus

$$
\|J_{1,n}\chi_R\|_{L^{\tilde{p}}_{\tilde{q}}(T)} \le \frac{1}{n}\|\partial_s u\|_{\tilde{L}^{\tilde{p},w}_{\tilde{q}}(T)}.
$$

Similarly to (3.20), we have that

$$
\|J_{2,n}\chi_R\|_{L^{\tilde{p}}_{\tilde{q}}(T)} \le \int_{\mathbb{R}^d} \left\| \left(\frac{\partial_s u(s,\cdot-\frac{y}{n})}{\sqrt{1+|\cdot-\frac{y}{n}|^2}} - \frac{\partial_s u(s,\cdot)}{\sqrt{1+|\cdot|^2}} \right)\chi_R \right\|_{L^{\tilde{p}}_{\tilde{q}}(T)} \rho(y)\mathrm{d}y.
$$

Due to (3.20) again and that ρ has compact support, the dominated convergence theorem yields that

$$
\begin{aligned}
&\overline{\lim_{n\to+\infty}} \|J_{2,n}\chi_R\|_{L^{\tilde{p}}_{\tilde{q}}(T)} \\
&\quad \le \int_{\mathbb{R}^d} \overline{\lim_{n\to+\infty}} \left\| \left(\frac{\partial_s u(s,\cdot-\frac{y}{n})}{\sqrt{1+|\cdot-\frac{y}{n}|^2}} - \frac{\partial_s u(s,\cdot)}{\sqrt{1+|\cdot|^2}} \right)\chi_R \right\|_{L^{\tilde{p}}_{\tilde{q}}(T)} \rho(y)\mathrm{d}y \\
&\quad = 0.
\end{aligned}
$$

Thus

$$
\lim_{n\to\infty} \left\| \frac{\partial_t u_n - \partial_t u}{\sqrt{1+|\cdot|^2}}\chi_R \right\|_{L^{\tilde{p}}_{\tilde{q}}(T)} = 0. \qquad (3.21)
$$

Combining this with (2) of Lemma 3.2, we have that

$$
\begin{aligned}
&\overline{\lim_{n\to+\infty}} \mathbb{E}\int_0^{T\wedge\tau_R} |\partial_s u_n(s,X_s) - \partial_s u(s,X_s)|\,\mathrm{d}s \\
&\quad \le \sqrt{1+R^2}\, \overline{\lim_{n\to+\infty}} \mathbb{E}\int_0^{T\wedge\tau_R} \frac{|\partial_s u_n(s,X_s) - \partial_s u(s,X_s)|}{\sqrt{1+|X_s|^2}}\mathrm{d}s \\
&\quad \le C_R \overline{\lim_{n\to+\infty}} \left(1 + \mathbb{E}\int_0^{T\wedge\tau_R} |\xi(s)|\mathrm{d}s \right) \|(\partial_s u_n - \partial_s u)\chi_R\|_{L^{\tilde{p},w}_{\tilde{q}}(T)}
\end{aligned}
$$

$$= 0.$$

Hence, by the Cantor's diagonal argument, there exists a subsequence, denoted also by u_n, such that $\mathbb{P}$-a.s.

$$\lim_{n\to\infty}\int_0^t |\partial_s u_n(s,X_s) - \partial_s u(s,X_s)|\,\mathrm{d}s = 0,\ t\in[0,T].$$

Therefore, combining all these together, we complete the proof by taking $n\to\infty$ in (3.18).

4 Applications

Consider (1.1) with $a = \frac{1}{2}\sigma\sigma^*$ satisfying **(H1)** , and

(H3) for every $T > 0$, $\sup_{t\in[0,T]}|b_1(t,0)| < \infty$ and $\|\nabla b_1\|_{T,\infty} < \infty$, and $b_0 \in \tilde{L}^{p_1}_{q_1}(T), \nabla\sigma \in \tilde{L}^{p_2}_{q_2}(T)$ for some $p_i, q_i \in (1,\infty)$ with $\frac{d}{p_i} + \frac{2}{q_i} < 1$, $i = 1, 2$.

By Remark 2.1, **(H1)**, **(H3)** and Theorem 3.1, we can follow the proofs of [29, Theorem 1.1] to prove that (1.1) has a unique strong solution. Let P_t be the associated semigroup generated by X_t. In this section, we shall investigate Harnack inequalities for (1.1).

We first establish log-Harnack inequality following the methodology of [14].

Theorem 4.1. *Assume* **(H1)** *and* **(H3)**. *Then there exists* $K_1 > 0$ *such that*

$$P_T\log f(x) \le \log P_T f(y) + \frac{K_1|x-y|^2}{\kappa_1 T},\ \ x,y\in\mathbb{R}^d, f\in\mathscr{B}_b^+(\mathbb{R}^d). \tag{4.1}$$

Fix some $T > 0$. By Remark 2.1, **(H1)** and **(H3)**, it follows from Theorem 2.1 that (1.2) with $\phi(T,x)\equiv 0$ has a unique solution. Let $\phi = (\phi^1,\cdots,\phi^d)$ and $\Phi_s(x) = x+\phi(s,x)$, $s\in[0,t]$. Since $\frac{d}{p_1}+\frac{2}{q_1} < 1$, it follows from (2.3) that we can choose large enough λ such that $\|\nabla\phi\|_{T,\infty} < \frac{1}{2}$. Then for $s\in[0,T],\ x,y\in\mathbb{R}^d$

$$\frac{1}{2}|x-y| \le |\Phi_s(x)-\Phi_s(y)| \le \frac{3}{2}|x-y|, \tag{4.2}$$

$$\frac{2}{3}|x-y| \le |\Phi_s^{-1}(x)-\Phi_s^{-1}(y)| \le 2|x-y|. \tag{4.3}$$

By Lemma 3.3, we have

$$\begin{aligned}\mathrm{d}\Phi_s(X_s) &= (\partial_s + L)\,\Phi_s(X_s)\mathrm{d}s + (I+\nabla\phi(s,X_s))\,\sigma(s,X_s)\mathrm{d}W_s\\ &= (b_1(s,X_s)+\lambda\phi(s,X_s))\,\mathrm{d}s + (I+\nabla\phi(s,X_s))\,\sigma(s,X_s)\mathrm{d}W_s.\end{aligned}$$

Let $Y_s = \Phi_s(X_s)$ and $X_s = \Phi_s^{-1}(Y_s)$. Define

$$Z(s,y) = (b_1+\lambda u)(s,\Phi_s^{-1}(y)),\quad \Sigma(s,y) = \left(I+\nabla\phi(s,\Phi_s^{-1}(y))\right)\sigma\left(s,\Phi_s^{-1}(y)\right).$$

Then we transform (1.1) to

$$\mathrm{d}Y_s = Z(s,Y_s)\mathrm{d}s + \Sigma(s,Y_s)\mathrm{d}W_s, \qquad Y_0 = \Phi_0(X_0). \tag{4.4}$$

By Theorem 2.1, it is clear that

$$\|\nabla Z(s,y)\| \le C_T \left(\|\sigma\|_\infty \|\nabla^2\phi(s,\Psi_s^{-1}(y))\| + \|\nabla\sigma(s,\Psi_s^{-1}(y))\|\right). \tag{4.5}$$
$$\frac{1}{4}\kappa_1|x|^2 \le |\Sigma^*(s,y)x|^2 \le \frac{9}{4}\kappa_2|x|^2,\ s\in[0,T],\ x,y\in\mathbb{R}^d,$$

and there exists $\bar{K}_1 > 0$ such that

$$|Z(s,y_1) - Z(s,y_2)| \le \bar{K}_1|y_1-y_2|,\ y_1,y_2\in\mathbb{R}^d,\ s\in[0,T].$$

Let $\mathcal{T}_s f(x) = \mathbb{E}f(Y_s^x)$ with $Y_0^x \equiv x$. Hence, we have the following log-Harnack inequality for (4.4). Since $\|\nabla\sigma\| \in \tilde{L}_{q_2}^{p_2}(T)$, $\|\nabla^2\phi\| \in \tilde{L}_{q_1}^{p_1}$ and $\|\nabla Z\|_{T,\infty} < \infty$, the proof of this lemma follows from that of [14, Proposition 2.1] and the approximation argument used in [29, Theorem 1.1 (B), see (iii) in page 20] completely.

Lemma 4.2. *There exists $\tilde{K}_0$ such that for any $f \in \mathscr{B}_b^+(\mathbb{R}^d)$,*

$$\mathcal{T}_T \log f(y) \le \log \mathcal{T}_T f(x) + \frac{\tilde{K}_0|x-y|^2}{\kappa_1 T},\ x,y\in\mathbb{R}^d.$$

Since $P_s f(x) = \mathbb{E}f(X_s^x) = \mathbb{E}f(\Phi_s^{-1}(Y_s^{\Phi_0(x)})) = \mathcal{T}_s\bar{f}(\Phi_0(x))$ with $\bar{f}(\cdot) = f(\Phi_s^{-1}(\cdot))$, Theorem 4.1 follows from this lemma and (4.2) directly.

Next, we shall establish the Harnack inequality with power for (1.1). Before our detailed discussions, we give some remarks on the Harnack inequality with power for SDEs with irregular coefficients.

Remark 4.1. *For $b_0 \in \tilde{L}_{q_1}^{p_1}(t)$ with $p,q\in(1,\infty)$ satisfying $\frac{d}{p_1}+\frac{2}{q_1}<1$ and that σ is a constant matrix, we have by (4.5) that $\|\nabla\Sigma\| \in \tilde{L}_{q_1}^{p_1}$, which does not yield that $\nabla\Sigma(t,\cdot)$ is bounded. Thus (4.4) does not fulfill conditions to derived Harnack inequalities with power in [26]. [20] established Harnack inequalities with an extra constant for SDEs whose drift merely satisfies the L^p-L^q integral condition. Because of the extra constant, one can't derive the strong Feller property from the Harnack inequality established in [20].*

Recently, assuming that the non-regular drift b_0 satisfies $b_0 \in L_q^p(T)$ with $\frac{d}{p}+\frac{2}{q}<1$ and

$$\int_{\mathbb{R}^d} |b_0(t,x+y) - b_0(t,x)|^p \mathrm{d}x \le K^p(t)|y|^p \tag{4.6}$$

with $K \in L_{loc}^q([0,\infty))$, the author in [7] obtained the Harnack inequality. However, the conditions used in [7] are not allowed the drift to be singular in space *variable. In fact, given $t \in [0,T]$, $b_0(t,\cdot) \in L^p(\mathbb{R}^d)$ and (4.6) imply by the definition of the Besov space $B^\gamma_{p,\infty}$, see [22, Section 2.5.12], that $b_0(t,\cdot) \in \cap_{\gamma<1} B^\gamma_{p,\infty}$ and there exists $C > 0$ such that*

$$\|b_0(t,\cdot)\|_{B^\gamma_{p,\infty}} \le C(K(t)+1), \ \text{a.e. } t \in [0,T].$$

Since $\frac{d}{p} + \frac{2}{q} < 1$, there exists $\gamma - \frac{d}{p} > 0$ for γ being closed to 1. By the embedding theorem of Besov space, see [3, Theorem 6.5.1], we have

$$b_0(t,\cdot) \in B^{\gamma-\frac{d}{p}}_{\infty,\infty}(\mathbb{R}^d) = C^{\gamma-\frac{d}{p}}(\mathbb{R}^d), \ \frac{d}{p} < \gamma < 1.$$

Then $b_0(t,\cdot)$ is bounded and $(\gamma - \frac{d}{p})$-Hölder continuous for any $\gamma \in (\frac{d}{p}, 1)$ and there exists $C > 0$ such that

$$\int_0^T \|b_0(t,\cdot)\|^q_{C^{\gamma-\frac{d}{p}}} \mathrm{d}t \le C \int_0^T (K(t)+1)^q \mathrm{d}t,$$

where $\|\cdot\|_{C^{\gamma-\frac{d}{p}}}$ is the $(\gamma - \frac{d}{p})$-Hölder norm.

Our main result on the Harnack inequality with power is the following theorem, which can be applied to SDEs with singular drift without extra regularity assumption and the diffusion coefficient is Hölder continuous with order greater than $\frac{1}{2}$.

Theorem 4.3. *Fix $T > 0$. Assume* **(H1)** *and* **(H3)***, and that there exist $c_T > 0$ and $\beta > 0$ such that*

$$\|\sigma(t,x) - \sigma(t,y)\|_{HS} \le c_T |x-y|^\beta, \ x,y \in \mathbb{R}^d. \tag{4.7}$$

(1) If $\beta \ge \frac{1}{2}$, then for any $\gamma > 1 + 4\left(\sqrt{1+\frac{\lambda_1^2}{8\lambda_2}} - 1\right)^{-1}$ there exist $\tilde{K}_1, \tilde{K}_2 > 0$ such that the following Harnack inequality with extra constant holds

$$(P_t f)^\gamma(y) \le P_t f^\gamma(x) \exp\left\{ t\tilde{K}_2 + \frac{\tilde{K}_1 |x-y|^2}{(1-e^{-\tilde{K}_1 t})} \right\}, \ f \in \mathscr{B}_b^+(\mathbb{R}^d). \tag{4.8}$$

(2) If $\frac{d}{p} + \frac{2}{q} < \frac{1}{2}$ and $\beta > \frac{1}{2}$, then there exists $K_T > 0$ such that for any $\gamma > (1 + \frac{6\sqrt{2}\lambda_2}{\sqrt{\lambda_1}\alpha_0})^2$ with $\alpha_0 = (1 - \frac{d}{p_1} - \frac{2}{q_1}) \wedge \beta$, any $\alpha \in (\frac{1}{2}, 1 - \frac{d}{p_1} - \frac{2}{q_1}) \cap (\frac{1}{2}, \beta]$, we have

$$(P_t f)^\gamma(y) \le P_t f^\gamma(x) \exp\left\{ \frac{\sqrt{\gamma}(\sqrt{\gamma}-1) K_T (|x-y|^2 \vee |x-y|^{2\alpha})}{2\delta_{\gamma,T}(\sqrt{\lambda_T}\alpha(\sqrt{\gamma}-1) - 2\delta_{\gamma,T})(1-e^{-K_T T})} \right\}.$$

where $\delta_{\gamma,T} = \frac{3\lambda_2}{2} \vee \frac{\sqrt{\lambda_1}\alpha(\sqrt{\gamma}-1)}{4\sqrt{2}}$ and $f \in \mathscr{B}_b^+(\mathbb{R}^d)$.

The proof of (1) of Theorem 4.3 just follows that of [20] and Krylov's estimate in Theorem 3.1 directly, and we omit it here. We now focus on establishing Harnack inequality without extra constant. To this end, we first investigate the following equation

$$\mathrm{d}\hat{X}_t = \hat{b}(t, \hat{X}_t)\mathrm{d}t + \hat{\sigma}(t, \hat{X}_t)\mathrm{d}W_t, \tag{4.9}$$

where $\hat{b} : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}^d$ and $\hat{\sigma} : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ satisfying the following conditions

(H4) Fix $T > 0$. Assume there exist $\alpha \in (\frac{1}{2}, 1]$ and positive constants K_T, δ_T, λ_T such that for $x, y \in \mathbb{R}^d$, $t \in [0, T]$

$$2\langle \hat{b}(t, x) - \hat{b}(t, y), x - y\rangle + \|\hat{\sigma}(t, x) - \hat{\sigma}(t, y)\|_{HS}^2 \le K_T|x - y|^2 \vee |x - y|^{2\alpha},$$
$$|(\hat{\sigma}(t, x) - \hat{\sigma}(t, y))^*(x - y)| \le \delta_T|x - y|, \quad \hat{\sigma}(t, x)\hat{\sigma}^*(t, x) \ge \lambda_T.$$

Let $\hat{X}_t$ solve (4.9) with $\hat{X}_0 = x$. Set $\eta(t) = \frac{2\alpha-\theta}{K_T}(1 - e^{K_T(t-T)})$ with $\theta \in (0, 2\alpha)$ and let $\hat{Y}_t$ solve the following equation with $Y_0 = y$

$$\mathrm{d}\hat{Y}_t = \hat{b}(t, \hat{Y}_t)\mathrm{d}t + \hat{\sigma}(t, \hat{Y}_t)\mathrm{d}W_t + \frac{\hat{\sigma}(t, \hat{Y}_t)\hat{\sigma}^{-1}(t, \hat{X}_t)(\hat{X}_t - \hat{Y}_t)}{\eta(t)(|\hat{X}_t - \hat{Y}_t|^{2-2\alpha} \wedge 1)}\mathbb{1}_{[0,T)}(t)\mathrm{d}t, \tag{4.10}$$

The following lemma is crucial to establish the Harnack inequality for (1.1).

Lemma 4.4. *Fix $T > 0$. Assume* **(H4)**. *Suppose that* (4.9) *has a unique strong solution and for any $(x, y) \in \mathbb{R}^{2d}$, and the martingale solution to the system $(\hat{X}_t, \hat{Y}_t)$ with $(\hat{X}_0, \hat{Y}_0) = (x, y)$ is well-posed on $[0, T)$. Let*

$$R_s = \exp\left\{-\int_0^s \left\langle \frac{\hat{\sigma}^{-1}(t, \hat{X}_t)(\hat{X}_t - \hat{Y}_t)}{\eta(t)(|\hat{X}_t - \hat{Y}_t|^{2-2\alpha} \wedge 1)}, \mathrm{d}W_t \right\rangle \right.$$
$$\left. -\frac{1}{2}\int_0^s \left|\frac{\hat{\sigma}^{-1}(t, \hat{X}_t)(\hat{X}_t - \hat{Y}_t)}{\eta(t)(|\hat{X}_t - \hat{Y}_t|^{2-2\alpha} \wedge 1)}\right|^2 \mathrm{d}t\right\}, \quad s \in [0, T]. \tag{4.11}$$

Then we have the following two assertions:

(i) *Let $\gamma_0 = \frac{\lambda_T\theta^2}{8(2\delta_T + \sqrt{\lambda_T}\theta)\delta_T}$, one has*

$$\sup_{s\in[0,T]} \mathbb{E}R_s^{1+\gamma_0} \le \exp\left\{\frac{(4\delta_T + \sqrt{\lambda_T}\theta)\theta K(|x - y|^2 \vee |x - y|^{2\alpha})}{16(2\delta_T + \sqrt{\lambda_T}\theta)(2\alpha - \theta)(1 - e^{-K_T T})\delta_T^2}\right\}. \tag{4.12}$$

(ii) *Let $\hat{P}_t$ be the associated transition semigroup of $\hat{X}_t$. For any $f \in \mathscr{B}_b^+(\mathbb{R}^d)$, $x, y \in \mathbb{R}^d$ and $\gamma > \left(1 + \frac{2\delta_T}{\sqrt{\lambda_T}\alpha}\right)^2$, we have*

$$(\hat{P}_t f)^\gamma(y) \le \hat{P}_t f^\gamma(x) \exp\left\{\frac{\sqrt{\gamma}(\sqrt{\gamma} - 1)K_T(|x - y|^2 \vee |x - y|^{2\alpha})}{4\delta_{\gamma,T}(\sqrt{\lambda_T}\alpha(\sqrt{\gamma} - 1) - 4\delta_{\gamma,T})(1 - e^{-K_T T})}\right\}, \tag{4.13}$$

where $\delta_{\gamma,T} = \delta_T \vee \frac{\sqrt{\lambda_T}\alpha(\sqrt{\gamma}-1)}{4}$.

Proof. Fix $(x,y)\in\mathbb{R}^{2d}$. Since the martingale solution of the system $(\hat X_t,\hat Y_t)$ is well-posed, there exist a system of process $(\hat X_t,\hat Y_t,W_t)_{t\in[0,T)}$ and a probability space with filtration $(\Omega,\mathscr{F},\mathbb{P},\mathscr{F}_t)_{t\in[0,T)}$ such that $\{W\}_{t\in[0,T)}$ is a Brownian motion w.r.t. $(\Omega,\mathscr{F},\mathbb{P},\mathscr{F}_t)_{t\in[0,T)}$ and $(\hat X_t,\hat Y_t,W_t)_{t\in[0,T)}$ satisfies (4.9) and (4.10). Let $\tau_n=\inf\{t\in[0,T)\mid |\hat X_t|+|Y_t|\ge n\}$, and let

$$\tilde W_t=W_t+\int_0^t\frac{\hat\sigma^{-1}(s,\hat X_s)(\hat X_s-\hat Y_s)}{\eta(s)(|\hat X_s-\hat Y_s|^{2-2\alpha}\wedge 1)}\mathrm{d}s,\ t\in[0,T).$$

Then the system (4.9) and (4.10) can be rewritten as

$$\begin{aligned}\mathrm{d}\hat X_t&=\hat b(t,\hat X_t)\mathrm{d}t+\hat\sigma(t,\hat X_t)\mathrm{d}\tilde W_t-\frac{\hat X_t-\hat Y_t}{\eta(t)(|\hat X_t-\hat Y_t|^{2-2\alpha}\wedge 1)}\mathbb{1}_{[0,T)}(t)\mathrm{d}t,\\ \mathrm{d}\hat Y_t&=\hat b(t,\hat Y_t)\mathrm{d}t+\hat\sigma(t,\hat Y_t)\mathrm{d}\tilde W_t,\end{aligned}$$

and it follows from Girsanov's theorem that $\{\tilde W_t\}_{0\le t\le s\wedge\tau_n}$ is a Brownian motion under $R_{s\wedge\tau_n}\mathbb{P}$ for any $s\in[0,T)$. By Itô's formula,

$$\begin{aligned}\mathrm{d}|\hat X_t-\hat Y_t|^2&=2\langle\hat b(t,\hat X_t)-\hat b(t,\hat Y_t),\hat X_t-\hat Y_t\rangle\mathrm{d}t-\frac{2|\hat X_t-\hat Y_t|^{2\alpha}\vee|\hat X_t-\hat Y_t|^2}{\eta(t)}\mathrm{d}t\\ &\quad+2\langle\hat X_t-\hat Y_t,(\hat\sigma(t,\hat X_t)-\hat\sigma(t,\hat Y_t))\mathrm{d}\tilde W_t\rangle\\ &\quad+\|\hat\sigma(t,\hat X_t)-\hat\sigma(t,\hat Y_t)\|_{HS}^2\mathrm{d}t\\ &\le K_T|\hat X_t-\hat Y_t|^2\vee|\hat X_t-\hat Y_t|^{2\alpha}\mathrm{d}t-\frac{2|\hat X_t-\hat Y_t|^{2\alpha}\vee|\hat X_t-\hat Y_t|^2}{\eta(t)}\mathrm{d}t\\ &\quad+2\langle\hat X_t-\hat Y_t,(\hat\sigma(t,\hat X_t)-\hat\sigma(t,\hat Y_t))\mathrm{d}\tilde W_t\rangle,\qquad t<s\wedge\tau_n.\end{aligned}$$

Then

$$\begin{aligned}\mathrm{d}\frac{|\hat X_t-\hat Y_t|^2}{\eta(t)}&\le\frac{K_T\eta(t)-2}{\eta^2(t)}|\hat X_t-\hat Y_t|^{2\alpha}\vee|\hat X_t-\hat Y_t|^2\mathrm{d}t-\frac{\eta'(t)}{\eta^2(t)}|\hat X_t-\hat Y_t|^2\mathrm{d}t\\ &\quad+\frac{2}{\eta(t)}\langle\hat X_t-\hat Y_t,(\hat\sigma(t,\hat X_t)-\hat\sigma(t,\hat Y_t))\mathrm{d}\tilde W_t\rangle\\ &\le-\frac{\eta'(t)+2-K_T\eta(t)}{\eta^2(t)}|\hat X_t-\hat Y_t|^{2\alpha}\vee|\hat X_t-\hat Y_t|^2\mathrm{d}t\\ &\quad+\frac{2}{\eta(t)}\langle\hat X_t-\hat Y_t,(\hat\sigma(t,\hat X_t)-\hat\sigma(t,\hat Y_t))\mathrm{d}\tilde W_t\rangle\\ &\le-\frac{\theta|\hat X_t-\hat Y_t|^{2\alpha}\vee|\hat X_t-\hat Y_t|^2}{\eta^2(t)}\mathrm{d}t\\ &\quad+\frac{2}{\eta(t)}\langle\hat X_t-\hat Y_t,(\hat\sigma(t,\hat X_t)-\hat\sigma(t,\hat Y_t))\mathrm{d}\tilde W_t\rangle,\qquad t<s\wedge\tau_n.\end{aligned}$$

Since $\alpha>\frac12$, it follows from Itô's formula that

$$\mathrm{d}\frac{|\hat X_t-\hat Y_t|^{2\alpha}}{\eta(t)}\le-\frac{\eta'(t)}{\eta^2(t)}|\hat X_t-\hat Y_t|^{2\alpha}\mathrm{d}t+\frac{\alpha K_T}{\eta(t)}|\hat X_t-\hat Y_t|^{4\alpha-2}\vee|\hat X_t-\hat Y_t|^{2\alpha}\mathrm{d}t$$

$$
\begin{aligned}
&- \frac{2\alpha}{\eta^2(t)}|\hat{X}_t - \hat{Y}_t|^{2\alpha} \vee |\hat{X}_t - \hat{Y}_t|^{4\alpha-2}\mathrm{d}t \\
&+ \left\langle \frac{2\alpha(\hat{X}_t - \hat{Y}_t)}{\eta(t)|\hat{X}_t - \hat{Y}_t|^{2-2\alpha}}, (\hat{\sigma}(t,\hat{X}_t) - \hat{\sigma}(t,\hat{Y}_t))\mathrm{d}\tilde{W}_t \right\rangle \\
\le &- \frac{2\alpha + \eta'(t) - \alpha K_T \eta(t)}{\eta^2(t)}|\hat{X}_t - \hat{Y}_t|^{2\alpha} \vee |\hat{X}_t - \hat{Y}_t|^{4\alpha-2} \\
&+ \left\langle \frac{2\alpha(\hat{X}_t - \hat{Y}_t)}{\eta(t)|\hat{X}_t - \hat{Y}_t|^{2-2\alpha}}, (\hat{\sigma}(t,\hat{X}_t) - \hat{\sigma}(t,\hat{Y}_t))\mathrm{d}\tilde{W}_t \right\rangle \\
\le &- \frac{\theta|\hat{X}_t - \hat{Y}_t|^{2\alpha} \vee |\hat{X}_t - \hat{Y}_t|^{4\alpha-2}}{\eta^2(t)} \\
&+ \left\langle \frac{2\alpha(\hat{X}_t - \hat{Y}_t)}{\eta(t)|\hat{X}_t - \hat{Y}_t|^{2-2\alpha}}, (\hat{\sigma}(t,\hat{X}_t) - \hat{\sigma}(t,\hat{Y}_t))\mathrm{d}\tilde{W}_t \right\rangle, \ t < s \wedge \tau_n.
\end{aligned}
$$

Then

$$
\begin{aligned}
&\mathbb{E}R_{s\wedge\tau_n}\left(\int_0^{s\wedge\tau_n} \frac{|\hat{X}_t - \hat{Y}_t|^2 \vee |\hat{X}_t - \hat{Y}_t|^{4\alpha-2}}{\eta^2(t)}\mathrm{d}t\right) \\
&\le \mathbb{E}R_{s\wedge\tau_n}\left(\int_0^{s\wedge\tau_n} \frac{|\hat{X}_t - \hat{Y}_t|^{2\alpha} \vee |\hat{X}_t - \hat{Y}_t|^{4\alpha-2} + |\hat{X}_t - \hat{Y}_t|^{2\alpha} \vee |\hat{X}_t - \hat{Y}_t|^{2}}{\eta^2(t)}\mathrm{d}t\right) \\
&\le \frac{K_T\left(|x-y|^2 + |x-y|^{2\alpha}\right)}{\theta(2\alpha-\theta)(1-e^{-K_T T})},
\end{aligned}
$$

which yields that

$$
\mathbb{E}R_{s\wedge\tau_n}\log R_{s\wedge\tau_n} \le \frac{2K_T\left(|x-y|^{2\alpha} \vee |x-y|^2\right)}{\lambda_T\theta(2\alpha-\theta)(1-e^{-K_T T})}, \ s\in[0,T), \ n\in\mathbb{N}.
$$

Hence, $\{R_{s\wedge\tau_n}\}_{s<T,n\in\mathbb{N}}$ is a uniformly integrable martingale. By martingale convergence theorem and $\tau_n \uparrow T$, $R_{s\wedge\tau_n}$ can be extended to T such that $\{R_s\}_{s\in[0,T]}$ is a martingale. Moreover, it follows from Fatou's lemma that

$$
\sup_{s\in[0,T]} \mathbb{E}R_s\log R_s \le \frac{2K_T(|x-y|^{2\alpha} \vee |x-y|^2)}{\lambda_T\theta(2\alpha-\theta)(1-e^{-K_T T})}, \tag{4.14}
$$

which also implies that $\{R_s\}_{s\in[0,T]}$ is a uniformly integrable martingale. Hence, by Girsanov's theorem, $\{\tilde{W}_t\}_{t\in[0,T]}$ is a Brownian motion under $R_T\mathbb{P}$. Moreover, by the pathwise uniqueness of (4.9), $\{\hat{Y}_t\}_{t\in[0,T)}$ can be extended to T such that $\lim_{t\to T}\hat{Y}_t = \hat{Y}_T$, and $\{\hat{Y}\}_{t\in[0,T]}$ is a weak solution of (4.9) with starting point x replaced by y. Moreover, we have $\hat{X}_T = \hat{Y}_T$ $\mathbb{P}$-a.s. since (4.14) and $\int_0^T \eta^{-2}(t)\mathrm{d}t = \infty$.

Next, we prove (4.12). Since

$$
\int_0^s \frac{|\hat{X}_t - \hat{Y}_t|^2 \vee |\hat{X}_t - \hat{Y}_t|^{4\alpha-2}}{\eta^2(t)}\mathrm{d}t \le \frac{2K_T|x-y|^2 \vee |x-y|^{2\alpha}}{\theta(2\alpha-\theta)(1-e^{-K_T T})}
$$

$$+\int_0^s \left\langle \frac{2(\hat{X}_t-\hat{Y}_t)}{\theta\eta(t)} + \frac{2\alpha(\hat{X}_t-\hat{Y}_t)}{\theta\eta(t)|\hat{X}_t-\hat{Y}_t|^{2-2\alpha}}, (\hat{\sigma}(t,\hat{X}_t)-\hat{\sigma}(t,\hat{Y}_t))\mathrm{d}\tilde{W}_t \right\rangle.$$

Denoting by $\mathbb{E}_{s,n}$ the expectation w.r.t. $R_{s\wedge\tau_n}\mathbb{P}$, then for $r>0$

$$\begin{aligned}
&\mathbb{E}_{s,n}\exp\left\{r\int_0^{s\wedge\tau_n}\frac{|\hat{X}_t-\hat{Y}_t|^2\vee|\hat{X}_t-\hat{Y}_t|^{4\alpha-2}}{\eta^2(t)}\mathrm{d}t\right\}\exp\left\{-\frac{2rK_T(|x-y|^{2\alpha}\vee|x-y|^2)}{\theta(2\alpha-\theta)(1-e^{-K_TT})}\right\}\\
&\le \mathbb{E}_{s,n}\exp\left\{\frac{r}{\theta}\int_0^{s\wedge\tau_n}\left\langle\left(2+\frac{2\alpha}{|\hat{X}_t-\hat{Y}_t|^{2-2\alpha}}\right)\frac{\hat{X}_t-\hat{Y}_t}{\eta(t)}, (\hat{\sigma}(t,\hat{X}_t)-\hat{\sigma}(t,\hat{Y}_t))\mathrm{d}\tilde{W}_t\right\rangle\right\}\\
&\le \left(\mathbb{E}_{s,n}\exp\left\{\frac{2r^2}{\theta^2}\int_0^{s\wedge\tau_n}\left(2+\frac{2\alpha}{|\hat{X}_t-\hat{Y}_t|^{2-2\alpha}}\right)^2\frac{|(\hat{\sigma}(t,\hat{X}_t)-\hat{\sigma}(t,\hat{Y}_t))^*(\hat{X}_t-\hat{Y}_t)|^2}{\eta^2(t)}\mathrm{d}t\right\}\right)^{\frac{1}{2}}\\
&\le \left(\mathbb{E}_{s,n}\exp\left\{\frac{2\delta_T^2r^2}{\theta^2}\int_0^{s\wedge\tau_n}\frac{\left(2|\hat{X}_t-\hat{Y}_t|+2\alpha|\hat{X}_t-\hat{Y}_t|^{2\alpha-1}\right)^2}{\eta^2(t)}\mathrm{d}t\right\}\right)^{\frac{1}{2}}\\
&\le \left(\mathbb{E}_{s,n}\exp\left\{\frac{32\delta_T^2r^2}{\theta^2}\int_0^{s\wedge\tau_n}\frac{|\hat{X}_t-\hat{Y}_t|^2\vee|\hat{X}_t-\hat{Y}_t|^{4\alpha-2}}{\eta^2(t)}\mathrm{d}t\right\}\right)^{\frac{1}{2}}.
\end{aligned}$$

By taking $r=\frac{\theta^2}{32\delta_T^2}$, we have

$$\begin{aligned}
&\mathbb{E}_{s,n}\exp\left\{\frac{\theta^2}{32\delta_T^2}\int_0^{s\wedge\tau_n}\frac{|\hat{X}_t-\hat{Y}_t|^2\vee|\hat{X}_t-\hat{Y}_t|^{4\alpha-2}}{\eta^2(t)}\mathrm{d}t\right\}\\
&\quad\le \exp\left\{\frac{\theta K_T(|x-y|^{2\alpha}\vee|x-y|^2)}{8\delta_T^2(2\alpha-\theta)(1-e^{-K_TT})}\right\}. \qquad (4.15)
\end{aligned}$$

By the Hölder inequality, we have for any $\gamma_1>1$ that

$$\begin{aligned}
\mathbb{E}R_{s\wedge\tau_n}^{1+\gamma_0} &\le \left(\mathbb{E}_{s,n}\exp\left\{\frac{(\gamma_1\gamma_0+1)\gamma_0\gamma_1}{2(\gamma_1-1)}\int_0^{s\wedge\tau_n}\frac{|\hat{\sigma}^{-1}(t,\hat{X}_t)(\hat{X}_t-\hat{Y}_t)|^2}{\eta^2(t)(|\hat{X}_t-\hat{Y}_t|^{4-4\alpha}\wedge 1)}\mathrm{d}t\right\}\right)^{\frac{\gamma_1-1}{\gamma_1}}\\
&\le \left(\mathbb{E}_{s,n}\exp\left\{\frac{(\gamma_1\gamma_0+1)\gamma_0\gamma_1}{2(\gamma_1-1)}\int_0^{s\wedge\tau_n}\frac{|\hat{\sigma}^{-1}(t,\hat{X}_t)(\hat{X}_t-\hat{Y}_t)|^2}{\eta^2(t)(|\hat{X}_t-\hat{Y}_t|^{4-4\alpha}\wedge 1)}\mathrm{d}t\right\}\right)^{\frac{\gamma_1-1}{\gamma_1}}\\
&\le \left(\mathbb{E}_{s,n}\exp\left\{\frac{(\gamma_1\gamma_0+1)\gamma_0\gamma_1}{2(\gamma_1-1)\lambda_T}\int_0^{s\wedge\tau_n}\frac{|\hat{X}_t-\hat{Y}_t|^2\vee|\hat{X}_t-\hat{Y}_t|^{4\alpha-2}}{\eta^2(t)}\mathrm{d}t\right\}\right)^{\frac{\gamma_1-1}{\gamma_1}}.
\end{aligned}$$

Taking $\gamma_1=1+\sqrt{1+\gamma_0^{-1}}$ which minimizes $\frac{\gamma_1(\gamma_1\gamma_0+1)}{\gamma_1-1}$, we have

$$\frac{\gamma_1\gamma_0(\gamma_1\gamma_0+1)}{2(\gamma_1-1)\sqrt{\lambda_T}} = \frac{\gamma_0(\sqrt{\gamma_0}+\sqrt{\gamma_0+1})^2}{2\lambda_T} = \frac{\theta^2}{32\delta_T^2},$$

$$\frac{\gamma_1 - 1}{\gamma_1} = \frac{4\delta_T + \sqrt{\lambda_T}\theta}{4\delta_T + 2\sqrt{\lambda_T}\theta}.$$

Combining this with (4.15), we have

$$\begin{aligned}\mathbb{E}R_{s\wedge\tau_n}^{1+\gamma_0} &\le \exp\left\{\frac{(\gamma_1 - 1)\theta K_T(|x-y|^{2\alpha} \vee |x-y|^2)}{8\gamma_1\delta_T^2(2\alpha-\theta)(1-e^{-K_T T})}\right\} \\ &= \exp\left\{\frac{(4\delta_T + \sqrt{\lambda_T}\theta)\theta K_T(|x-y|^{2\alpha} \vee |x-y|^2)}{16(2\delta_T + \sqrt{\lambda_T}\theta)\delta_T^2(2\alpha-\theta)(1-e^{-K_T T})}\right\}.\end{aligned}$$

Letting $n \to \infty$, we get (4.12).

For any $\gamma > (1 + \frac{2\delta_T}{\sqrt{\lambda_T}\alpha})^2$, we set $\theta = \frac{4\delta_T}{\sqrt{\lambda_T}(\sqrt{\gamma}-1)}$. Then $\theta < 2\alpha$ and $\frac{1}{\gamma-1} = \gamma_0$. Consequently,

$$\begin{aligned}\sup_{s\in[0,T]} \left(\mathbb{E}R_s^{\frac{\gamma}{\gamma-1}}\right)^{\gamma-1} &= \sup_{s\in[0,T]} \left(\mathbb{E}R_s^{1+\gamma_0}\right)^{\gamma-1} \\ &\le \exp\left\{\frac{(\gamma-1)(4\delta_T + \sqrt{\lambda_T}\theta)\theta K_T(|x-y|^{2\alpha} \vee |x-y|^2)}{16(2\delta_T + \sqrt{\lambda_T}\theta)\delta_T^2(2\alpha-\theta)(1-e^{-K_T T})}\right\} \\ &= \exp\left\{\frac{\sqrt{\gamma}(\sqrt{\gamma}-1)K_T(|x-y|^{2\alpha} \vee |x-y|^2)}{2\delta_T[2\alpha(\sqrt{\gamma}-1) - 4\delta_T](1-e^{-K_T T})}\right\}.\end{aligned}$$

Therefore,

$$\begin{aligned}(\hat{P}_T f)^\gamma(y) &= (\mathbb{E}R_T f(\hat{Y}_T))^\gamma \le (\mathbb{E}R_T^{\frac{\gamma}{\gamma-1}})^{\gamma-1}\mathbb{E}f^\gamma(\hat{Y}_T) = (\mathbb{E}R_T^{\frac{\gamma}{\gamma-1}})^{\gamma-1}\mathbb{E}f^\gamma(\hat{X}_T) \\ &\le \hat{P}_T f^\gamma(x)\exp\left\{\frac{\sqrt{\gamma}(\sqrt{\gamma}-1)K_T(|x-y|^{2\alpha} \vee |x-y|^2)}{4\delta_T[\alpha(\sqrt{\gamma}-1) - 2\delta_T](1-e^{-K_T T})}\right\}.\end{aligned}$$

It is clear that this inequality also holds with δ_T replaced by $\delta_{\gamma,T}$.

□

Proof of Theorem 4.3:

Let ϕ be given by (1.2) with $\|\nabla\phi\|_{T,\infty} < \frac{1}{2}$, $\Phi_t(x) = x + \phi(t,x)$ and $Y_t = \Phi_t(X_t)$. By (2.3) and (3) of Lemma 2.3, for any $\delta \in (0, 1 - \frac{d}{p_1} - \frac{2}{q_1})$, we have that

$$\sup_{t\in[0,T]} \|\nabla\phi(t,x) - \nabla\phi(t,y)\| \le C_T|x-y|^\delta.$$

Then

$$\frac{\kappa_1}{4}|h|^2 \le \left|[\sigma^*(I+\nabla\phi)^*](t,\Phi_t^{-1}(x))h\right|^2 \le \frac{9}{4}\kappa_2|h|^2,\ x,h\in\mathbb{R}^d, t\in[0,T],$$

and there exists $C_T > 0$ such that for any $t \in [0,T]$

$$\|(I+\nabla\phi)\sigma(t,\Phi_t^{-1}(x)) - (I+\nabla\phi)\sigma(t,\Phi_t^{-1}(y))\|_{HS} \le C_T|x-y|^\delta \vee |x-y|^\beta.$$

Setting

$$\hat{b}(t,x) = b_1(t, \Phi_t^{-1}(x)) + \lambda\phi(t, \Phi^{-1}(x)),\ \hat{\sigma}(t,x) = (I + \nabla\phi)\sigma(t, \Phi_t^{-1}(x)),$$

it is clear that the conditions for $\hat{b}$ and $\hat{\sigma}$ in Lemma 4.4 holds with $\delta_T = \frac{3}{2}\lambda_2$, $\lambda_T = \frac{\lambda_1}{2}$ and any $\alpha \in (\frac{1}{2}, 1 - \frac{d}{p_1} - \frac{2}{q_1}) \cap (0, \beta]$. Hence, the Harnack inequality with power follows if we prove the well-posedness of the martingale solution to the system $(\hat{X}_t, \hat{Y}_t)$ with $\hat{b}$ and $\hat{\sigma}$ defined as above. To this end, according to [21, Corollary 10.1.2], we only need to investigate the following system $(\hat{X}_t^{(n)}, \hat{Y}_t^{(n)})$ with $n \in \mathbb{N}$:

$$\begin{aligned}
\mathrm{d}\hat{X}_t^{(n)} &= \hat{b}(t, \hat{X}_t^{(n)})\mathrm{d}t + \hat{\sigma}(t, \hat{X}_t^{(n)})\mathrm{d}W_t,\\
\mathrm{d}\hat{Y}_t^{(n)} &= \hat{b}(t, \hat{Y}_t^{(n)})\mathrm{d}t + \hat{\sigma}(t, \hat{Y}_t^{(n)})\mathrm{d}W_t\\
&\quad + \frac{\hat{\sigma}(t, \hat{Y}_t^{(n)})\hat{\sigma}^{-1}(t, \hat{X}_t^{(n)})}{\eta(t)}\pi_n\left(\frac{\hat{X}_t^{(n)} - \hat{Y}_t^{(n)}}{|\hat{X}_t^{(n)} - \hat{Y}_t^{(n)}|^{2-2\alpha} \wedge 1}\right)\mathbb{1}_{[0,T)}(t)\mathrm{d}t,
\end{aligned}$$

where $\pi_n : \mathbb{R}^d \to \mathbb{R}^d$ defined as follows

$$\pi_n(y) = y\mathbb{1}_{[|y|<n]} + n\frac{y}{|y|}\mathbb{1}_{[|y|\geq n]},\ y \in \mathbb{R}^d.$$

Since π_n is a bounded function and $\hat{\sigma}$ is bounded and nondegenerate, the well-posedness of $(\hat{X}_t^{(n)}, \hat{Y}_t^{(n)})$ can be investigated via the following system with any Brownian motion $\{W_t\}_{t\geq 0}$ and Girsanov's theorem:

$$\begin{aligned}
\mathrm{d}\hat{X}_t^{(n)} &= \hat{b}(t, \hat{X}_t^{(n)})\mathrm{d}t + \hat{\sigma}(t, \hat{X}_t^{(n)})\mathrm{d}W_t\\
&\quad - \eta^{-1}(t)\pi_n\left(\frac{\hat{X}_t^{(n)} - \hat{Y}_t^{(n)}}{|\hat{X}_t^{(n)} - \hat{Y}_t^{(n)}|^{2-2\alpha} \wedge 1}\right)\mathbb{1}_{[0,T)}(t)\mathrm{d}t, && (4.16)\\
\mathrm{d}\hat{Y}_t^{(n)} &= \hat{b}(t, \hat{Y}_t^{(n)})\mathrm{d}t + \hat{\sigma}(t, \hat{Y}_t^{(n)})\mathrm{d}W_t. && (4.17)
\end{aligned}$$

By **(H1)** and **(H3)**, (4.17) has a unique strong solution. Next, we prove (4.16) has a pathwise unique solution. For any two solutions of (4.16), say $\hat{X}_t^{(n),1}$ and $\hat{X}_t^{(n),2}$ with the same initial value and $\hat{Y}_t^{(n)}$, there exists $C > 0$ such that

$$\begin{aligned}
\mathrm{d}|\hat{X}_t^{(n),1} - \hat{X}_t^{(n),2}|^2 \leq C|\hat{X}_t^{(n),1} - \hat{X}_t^{(n),2}|^2\mathrm{d}t + \|\sigma(t, \hat{X}_t^{(n),1}) - \sigma(t, \hat{X}_t^{(n),2})\|_{HS}^2\mathrm{d}t\\
+ 2\langle \hat{X}_t^{(n),1} - \hat{X}_t^{(n),2}, (\hat{\sigma}(t, \hat{X}_t^{(n),1}) - \hat{\sigma}(\hat{t}, X_t^{(n),2}))\mathrm{d}W_t\rangle,
\end{aligned}$$

where we have used the following inequality

$$\left\langle \pi_n\left(\frac{y_1}{|y_1|^{2-2\alpha} \wedge 1}\right) - \pi_n\left(\frac{y_2}{|y_2|^{2-2\alpha} \wedge 1}\right), y_1 - y_2\right\rangle \geq 0.$$

Since **(H3)** and $\phi \in \tilde{W}_{q_1}^{2,p_1}(T)$, we have by (4.5) that $\hat{\sigma}$ is bounded and $\nabla\hat{\sigma}(t, \cdot) \in L_{loc}^1$. Hence, by [29, Lemma 2.1],

$$\|\hat{\sigma}(t, \hat{X}_t^{(n),1}) - \hat{\sigma}(t, \hat{X}_t^{(n),2})\|_{HS}$$

$$\leq \left(|\mathcal{M}_1(\nabla\hat{\sigma}(t,\cdot))(\hat{X}_t^{(n),1})| + |\mathcal{M}_1(\nabla\hat{\sigma}(t,\cdot))(\hat{X}_t^{(n),2})| + \|\hat{\sigma}\|_\infty\right) |\hat{X}_t^{(n),1} - \hat{X}_t^{(n),2}|,$$

where $\mathcal{M}$ is the Hardy-Littlewood maximal function defined by

$$(\mathcal{M}_1 f)(x) = \frac{1}{|B_1(0)|} \int_{B_1(0)} f(x+y)\mathrm{d}y,$$

where $|B_1(0)|$ is the Lebesgue measure of $B_1(0)$. Consequently,

$$\begin{aligned} \mathrm{d}|\hat{X}_t^{(n),1} - \hat{X}_t^{(n),2}|^2 \leq C|\hat{X}_t^{(n),1} - \hat{X}_t^{(n),2}|^2 & \Big(1 + |\mathcal{M}_1(\nabla\hat{\sigma}(t,\cdot))(\hat{X}_t^{(n),1})|^2 \\ & + |\mathcal{M}_1(\nabla\hat{\sigma}(t,\cdot))(\hat{X}_t^{(n),2})|^2\Big)\,\mathrm{d}t \\ & + 2\langle \hat{X}_t^{(n),1} - \hat{X}_t^{(n),2}, (\hat{\sigma}(t,\hat{X}_t^{(n),1}) - \hat{\sigma}(t,\hat{X}_t^{(n),2}))\mathrm{d}W_t\rangle. \end{aligned} \tag{4.18}$$

Since π_n is a bounded function, for any $s < T$ and $r \in [0,s]$,

$$\tilde{W}_r = W_r - \int_0^r \eta^{-1}(t)\hat{\sigma}^{-1}(t,\hat{X}_t^{(n),1})\pi_n\left(\frac{\hat{X}_t^{(n),1} - \hat{Y}_t^{(n)}}{|\hat{X}_t^{(n),1} - \hat{Y}_t^{(n)}|^{2-2\alpha} \wedge 1}\right) \mathbb{1}_{[0,T)}(t)\mathrm{d}t,$$

is a Brownian motion under $\hat{R}_s\mathbb{P}$ with

$$\begin{aligned} \hat{R}_s = \exp\Bigg\{ & \int_0^r \left\langle \eta^{-1}(t)\hat{\sigma}^{-1}(t,\hat{X}_t^{(n),1})\pi_n\left(\frac{\hat{X}_t^{(n),1} - \hat{Y}_t^{(n)}}{|\hat{X}_t^{(n),1} - \hat{Y}_t^{(n)}|^{2-2\alpha} \wedge 1}\right), \mathrm{d}W_t\right\rangle \\ & -\frac{1}{2}\int_0^r \left|\eta^{-1}(t)\hat{\sigma}^{-1}(t,\hat{X}_t^{(n),1})\pi_n\left(\frac{\hat{X}_t^{(n),1} - \hat{Y}_t^{(n)}}{|\hat{X}_t^{(n),1} - \hat{Y}_t^{(n)}|^{2-2\alpha} \wedge 1}\right)\right|^2 \mathrm{d}t \Bigg\} \end{aligned}$$

and $\hat{X}_t^{(n),1}$ is a weak solution of (4.17) under $\hat{R}_s\mathbb{P}$. Due to (4.5), Theorem 3.1 and [29, Lemma 2.1], for any $0 \leq s_0 < s_1 \leq s$

$$\begin{aligned} & \mathbb{E}\hat{R}_s\left[\int_{r_0}^{r_1} |\mathcal{M}_1(\nabla\hat{\sigma}(t,\cdot))(\hat{X}_t^{(n),1})|^2\mathrm{d}t \Big| \mathscr{F}_{s_0}\right] \\ & \quad \leq C\left(\||\mathcal{M}_1(\nabla\sigma)|^2\|_{\tilde{L}^{p_2/2}_{q_2/2}(s_0,s_1)} + \||\mathcal{M}_1(\nabla^2\phi)|^2\|_{\tilde{L}^{p_1/2}_{q_1/2}(s_0,s_1)}\right) \\ & \leq C\left(\|\nabla\sigma\|_{\tilde{L}^{p_2}_{\tilde{q}_2}(s_0,s_1)} + \|\nabla^2\phi\|_{\tilde{L}^{p_1}_{\tilde{q}_1}(s_0,s_1)}\right). \end{aligned}$$

Then, it follows from [31, Lemma 3.5] that for any $c \in \mathbb{R}$

$$\mathbb{E}\hat{R}_s \exp\left\{c\int_0^s |\mathcal{M}_1(\nabla\hat{\sigma}(t,\cdot))(\hat{X}_t^{(n),1})|^2\mathrm{d}t\right\} < \infty,\ \ s \in [0,T).$$

Since π_n is bounded, $\mathbb{E}\hat{R}_s^{-m} < \infty$ for any $m \in \mathbb{N}$ and $s \in [0,T)$. Then the Hölder inequality yields that for any c

$$\mathbb{E}\exp\left\{c\int_0^s |\mathcal{M}_1(\nabla\hat{\sigma}(t,\cdot))(\hat{X}_t^{(n),1})|^2\mathrm{d}t\right\} < \infty,\ \ s \in [0,T).$$

A similar inequality can be established for $\hat{X}_t^{(n),2}$. Combining these with (4.18) and stochastic Gronwall's inequality, see [31, Lemma 3.8], the pathwise uniqueness follows. Therefore, the proof is completed.

Acknowledgements

The authors would like to thank the editor and referee who kindly provided valuable suggestions and comments which helped to improve the manuscript. The authors also would like to thank Professor Feng-Yu Wang and Dr. Xing Huang for their helpful suggestions. The first author was supported by the National Natural Science Foundation of China (Grant No. 11901604, 11771326).

References

[1] H. Amann, Linear and Quasilinear Parabolic Problems, Vol. 1, Abstract Linear Theory, Birkhäuser, Basel, 1995.

[2] V. Barbu, Nonlinear Differential Equations of Monotone Types in Banach Spaces, Springer, 2010.

[3] J. Bergh, J. Löfström, Interpolation Spaces, An Introduction, Springer-Verlag, 1976.

[4] G. Cupini, S. Fornaro, Maximal regularity in $L^p(\mathbb{R}^N)$ for a class of elliptic operators with unbounded coefficients, *Differential Integral Equations*, **17** (3-4) (2004), 259–296.

[5] C. Gallarati, M. Veraar, Maximal regularity for non-autonomous equations with measurable dependence on time, *Potent. Anal.* (2017) **46**, 527–567.

[6] I. Gyöngy, T. Martinez, On stochastic differential equations with locally unbounded drift, *Czechoslovak Math. J.*, **51** (4) (2001), 763–783.

[7] X. Huang, Harnack and shift Harnack inequalities for SDEs with integrable drifts, Stochastics and Dynamics (2018): 1950034.

[8] X. Huang, F.-Y. Wang, Distribution dependent SDEs with singular coefficients, to appear at *Stoch. Proc. Appl.*

[9] X. Huang, S.-Q. Zhang, Mild solutions and Harnack inequality for functional stochastic partial differential equations with Dini drift, *J. Theor. Probab.* (2019) **32**: 303–329

[10] N. V. Krylov, Estimates of the maximum of the solution of a parabolic equation and estimates of the distribution of a semimartingale, *Mat. Sb.* (N.S.) **130** (172) (1986), 207–221.

[11] N. V. Krylov, The heat equation in $L_q((0,T),L_p)$-spaces with weights, *SIAM J. Math. Anal.*, **31** (5) (2001), 1117–1141.

[12] N. V. Krylov, Some properties of traces for stochastic and deterministic parabolic weighted Sobolev spaces, *J. Funct. Anal.*, **183** (2001), 1–41.

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

[14] H. Li, D. Luo and J. Wang, Harnack inequalities for SDEs with multiplicative noise and non-regular drift, *Stochastics and Dynamics*, **15** (3) (2015), 1550015.

[15] A. Lunardi, V. Vespri, Generation of strongly continuous semigroups by elliptic operators with unbounded coefficients in $L^p(\mathbb{R}^N)$, *Rend. Istit. Mat. Univ. Trieste* **28** (1997), 251–279.

[16] A. Lunardi, V. Vespri, Optimal L^∞ and Schauder estimates for elliptic and parabolic operators with unbounded coefficients, in: Proc. Conf. "Reaction-diffusion systems", G. Caristi, E. Mitidieri eds., Lecture notes in pure and applied mathematics 194, M. Dekker (1998), 217-239.

[17] G. Metafune, D. Pallara, V. Vespri, L^p-estimates for a class of elliptic operators with unbounded coefficients in $\mathbb{R}^N$, *Houston J. Math*, **31** (2005), 605–620.

[18] S. Monniaux, Maximal regularity and applications to PDEs, *Analytical and Numerical Aspects of Partial Differential Equations*, Editors E. Emmrich, P. Wittbold, Publisher: Walter de Gruyter, Berlin, 247–287 (2009).

[19] M. Röckner, F.-Y. Wang, Log-Harnack inequality for stochastic differential equations in Hilbert spaces and its consequences, *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **13** (2010), 27–37.

[20] J. Shao, Harnack inequalities and heat kernel estimates for SDEs with singular drifts, *Bull. Sci. math.* **137** (2013) 589–601.

[21] D. W. Stroock and S.R. S. Varadhan, Multidimensional Diffusion Processes, Springer, 2005.

[22] H. Triebel, Theory of Function Spaces, Birkhäuser Verlag, 1983.

[23] A. J. Veretennikov, On Strong solutions and explicit formulars for solutions of stochastic integral equations, *Mathematics of the USSR-Sbornik*, **39** (1981) 387–403.

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

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

[26] F.-Y. Wang, Harnack inequality for SDE with multiplicative noise and extension to Neumann semigroup on non-convex manifolds, *Ann. Probab.* **39** (2011), 1449–1467.

[27] F.-Y. Wang, Harnack inequalities for stochastic partial differential equations, Springer Verlag, 2013.

[28] 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.

[29] 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.

[30] L. Xie, X. Zhang, Sobolev differentiable flows of SDEs with local Sobolev and super-linear growth coefficients, *Ann. Probab.* **44** (2016) 3661–3687.

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

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

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

[34] X. Zhang, Stochastic differential equations with Sobolev diffusion and singular drift and applications, *Ann. Appl. Probab.*, **26** (2016), 2697–2732.

[35] X. Zhang, G. Zhao, Heat kernel and ergodicity of SDEs with distribution drifts, arXiv: 1710.10537v2.

[36] A. K. Zvonkin, A transformation of the phase space of a diffusion process that will remove the drift, *Mat. Sb.*, **93** (135) (1974), 129–149.