Strong Solutions of Stochastic Differential Equations with Coefficients in Mixed-Norm Spaces

Chengcheng Ling¹,² · Longjie Xie³

Received: 12 July 2020 / Accepted: 10 February 2021 / Published online: 4 March 2021
© The Author(s), under exclusive licence to Springer Nature B.V. part of Springer Nature 2021

Abstract
By studying parabolic equations in mixed-norm spaces, we prove the existence and uniqueness of strong solutions to stochastic differential equations driven by Brownian motion with coefficients in spaces with mixed-norm, which extends Krylov and Röckner’s result in Krylov (Probab. Theory Rel. Fields. 131(2)154–196, 2005) and Zhang’s result in Zhang (Electron. J. Probab. 16, 1096–1116, 2011).

Keywords Stochastic differential equations · Zvonkin’s transformation · Krylov’s estimate · Singular drift

Mathematics Subject Classification (2010) 60H10 · 60J60

1 Introduction and main result
Consider the following stochastic differential equation (SDE for short):
\[ dX_t = b(t, X_t)dt + dW_t, \quad X_0 = x \in \mathbb{R}^d, \]
where \( d \geq 1 \), \( b = (b^1, \ldots, b^d) : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \) is a Borel measurable function, and \( (W_t)_{t \geq 0} \) is a standard Brownian motion defined on some filtered probability space \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})\). It is a classical result that if the coefficient \( b \) is global Lipschitz continuous in \( x \) uniformly with respect to \( t \), then there exists a unique strong solution \( (X_t(x))_{t \geq 0} \)

Research of C. Ling is supported by the DFG through the IRTG 2235. L. Xie is supported by the Alexander-von-Humboldt foundation and NNSF of China (No. 12071186, 11931004, 12090011).

Longjie Xie
longjiexie@jsnu.edu.cn

Chengcheng Ling
chengchenglcc@126.com

¹ School of Science, Beijing Jiaotong University, Beijing 100044, China
² Fakultät für Mathematik, Universität Bielefeld, Bielefeld, 33615, Germany
³ School of Mathematics and Statistics, Jiangsu Normal University, Xuzhou, Jiangsu 221000, People’s Republic of China

Published online: 4 March 2021
to SDE (1.1) for every $x \in \mathbb{R}^d$. However, many important applications of this class of SDE show that the Lipschitz continuity imposed on the coefficient is a rather severe restriction. Thus a lot of attentions have been paid to seek a strong solution for (1.1) under weaker assumptions on the drift $b$. A remarkable result due to Zvonkin [23] showed that if $d = 1$ and $b$ is bounded, then SDE (1.1) admits a unique strong solution for each $x \in \mathbb{R}$. Zvonkin’s result was then extended to the multidimensional case by Veretennikov [15]. A further generalization was obtained by Krylov and Röckner [10] where the pathwise uniqueness for SDE (1.1) was shown when

$$b \in L^q_{loc}(\mathbb{R}^d; L^p_{loc}(\mathbb{R}^d)) \quad \text{with} \quad q, p \in (2, \infty) \quad \text{and} \quad d/p + 2/q < 1.$$  

(1.2)

Later, Zhang [20] extends these results to SDE driven by multiplicative noise

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \quad X_0 = x \in \mathbb{R}^d$$

(1.3)

under the assumptions that $\sigma = (\sigma^{ij})_{d \times d} : \mathbb{R}^d \to \mathbb{R}^{d \times d}$ is a bounded, uniformly elliptic matrix-valued function which is uniformly continuous in $x$ locally uniformly with respect to $t$, and

$$b, |\nabla\sigma| \in L^q_{loc}(\mathbb{R}^d; L^p_{loc}(\mathbb{R}^d))$$

with $p, q$ satisfying (1.2). Here and below, $\nabla$ denotes the weak derivative with respect to the $x$ variable. Note that when $\sigma \equiv 0$, SDE (1.3) is just an ordinary differential equation, which is far from being well-posed under the above conditions on the drift coefficient. This reflects that noises can play some regularization effects to the deterministic systems, we refer the readers to [4, 5] for more comprehensive overview. We mention that there are also increasing interests in studying the properties of the unique strong solution to SDE (1.3) with singular coefficients, see e.g. [2, 11, 16, 19, 21] and the references therein.

To the best of our knowledge, the conditions imposed on the coefficients in [10, 20] are known to be the weakest so far in the literature to ensure the strong well-posedness of the SDE (1.1) and (1.3). However, such conditions turn out to be not proper when studying certain complex models involving several phase variables. For example, consider the following skew-product fast-slow system (see [12, Section 10.6]) in $\mathbb{R}^d \times \mathbb{R}^d$:

$$\begin{cases}
    dX^\epsilon_t = -\epsilon^{-1} X^\epsilon_t \, dt + \epsilon^{-1/2} dW^1_t, & X^\epsilon_0 = x, \\
    dY^\epsilon_t = (1 - X^\epsilon_t)^2 F(Y^\epsilon_t) \, dt + dW^2_t, & Y^\epsilon_0 = y,
\end{cases}$$

(1.4)

where $d \geq 1$, $W^1_t, W^2_t$ are independent standard Brownian motions, and the small parameter $0 < \epsilon \ll 1$ represents the separation of time scales between the fast motion $X^\epsilon_t$ (with time order $1/\epsilon$) and the slow component $Y^\epsilon_t$. Such multi-scale model has wide range of applications and has been intensively studied in the past decades. Let $\mu(dx)$ be the unique invariant measure of the Ornstein-Uhlenbeck process

$$dX_t = -X_t \, dt + dW^1_t.$$  

Then the celebrated averaging principle says that the slow component $Y^\epsilon_t$ in system (1.4) will converge as $\epsilon \to 0$ to the solution of the following so-called averaged equation:

$$d\bar{Y}_t = \bar{F}(\bar{Y}_t) \, dt + dW^2_t, \quad \bar{Y}_0 = y,$$

(1.5)

where the new averaged drift coefficient is defined by

$$\bar{F}(y) := \int_{\mathbb{R}^d} (1 - x)^2 \mu(dx) \cdot F(y).$$

Note that according to [10, 20], we need to assume that $F \in L^p_{loc}(\mathbb{R}^d)$ with $p > 2d$ to ensure the strong well-posedness of SDE (1.4). This in turn yields that $\bar{F} \in L^p_{loc}(\mathbb{R}^d)$ with $p > 2d$.  

 Springer
However, according to [10, 20] again, $F \in L^p_{loc}(\mathbb{R}^d)$ with $p > d$ would be enough to ensure the strong well-posedness of SDE (1.5), which implies that the conditions in [10, 20] are not optimal. We point out that such problem will always appear when considering SDEs in multi-dimension, and especially for degenerate noise cases and distribution dependent equations, see e.g. [3, 13, 18, 22].

The main aim of this work is to get rid of the above unreasonableness by studying SDE (1.3) with coefficients in general mixed-norm spaces. To this end, let $p = (p_1, \cdots, p_d) \in [1, \infty)^d$ be a multi-index, we denote by $L^p(\mathbb{R}^d)$ the space of all measurable functions on $\mathbb{R}^d$ with norm
\[
\|f\|_{L^p(\mathbb{R}^d)} := \left( \int_{\mathbb{R}^d} \cdots \left( \int_{\mathbb{R}} |f(x_1, \ldots, x_d)|^{p_1} \, dx_1 \right)^{p_2/p_1} \cdots \left( \int_{\mathbb{R}} |f(x_1, \ldots, x_d)| \, dx_d \right)^{p_d/p_{d-1}} \right)^{1/p_d} < \infty.
\]
When $p_i = \infty$ for some $i = 1, \cdots, d$, the norm is taken as the supreme over $\mathbb{R}$ with respect to the corresponding variable $x_i \in \mathbb{R}$, and by $L^p_{loc}(\mathbb{R}^d)$ we mean the corresponding local space defined as usual. Notice that the order is important when taking the above integrals. If we permute the $p_i$s, then increasing the order of $p_i$ gives the smallest norm, while decreasing the order gives the largest norm.

Our main result in this paper is as follows.

**Theorem 1.1** Assume that for some $p_1, \cdots, p_d, q \in (2, \infty]$ and every $T > 0$,
\[
|b|, |\nabla \sigma| \in L^q([0, T]; L^p_{loc}(\mathbb{R}^d)) \quad \text{with} \quad \frac{2}{q} + \frac{1}{p_1} + \cdots + \frac{1}{p_d} < 1,
\]
and for every $n \in \mathbb{N}$, $\sigma$ is uniformly continuous in $x \in B_n := \{x \in \mathbb{R}^d : |x| \leq n\}$ uniformly with respect to $t \in [0, n]$, and there exist positive constants $\delta_n$ such that for all $(t, x) \in [0, n] \times B_n$,
\[
|\sigma(t, x)\xi| \geq \delta_n|\xi|^2, \quad \forall \xi \in \mathbb{R}^d.
\]
Then, for each $x \in \mathbb{R}^d$ there exists a unique strong solution $X_t(x)$ up to an explosion time $\zeta(x)$ to SDE (1.3) such that
\[
\lim_{t \uparrow \zeta(x)} X_t(x) = +\infty, \quad a.s.
\]

**Remark 1.2** i) The advantage of (1.6) lies in the flexible integrability of the coefficients. More precisely, it allows the integrability of the coefficients to be small in some directions by taking the integrability index large for the other directions (not as functions of the whole space variable). With this condition, the problem of the tricky example mentioned before does not appear since we can take the integrability index with respect to the $x$-variable to be $\infty$, and hence we only need to assume $F \in L^p_{loc}(\mathbb{R}^d)$ with $p > d$ to ensure the strong well-posedness of SDE (1.4). This will imply that $\bar{F} \in L^p_{loc}(\mathbb{R}^d)$ with $p > d$, which coincides with the condition given by [10].

ii) As mentioned in [8], the necessity of mixed-norm spaces arises when the physical processes have different behavior with respect to each component. In view of (1.6), it reflects the classical fact that the integrability of time variable and space variable has the ratio $1:2$. Meanwhile, the integrability of each component of the space variable is the same, which is natural because the noise is non-degenerate. Such kind of mixed-norm spaces will be more important when studying SDEs with degenerate noises. This will be our future works.
Now, let us introduce the proof briefly. The key tool to prove our main result is the \(L^q_p\)-maximal regularity estimate for the following second order parabolic PDEs on \([0, T] \times \mathbb{R}^d\):

\[
\partial_t u(t, x) = \mathcal{L}_2^a u(t, x) + \mathcal{L}_1^b u(t, x) + f(t, x), \quad u(0, x) = 0,
\]

where \(\mathcal{L}_2^a + \mathcal{L}_1^b\) is the infinitesimal generator corresponding to SDE (1.3), i.e.,

\[
\mathcal{L}_2^a u(t, x) := \frac{1}{2} a^{ij}(t, x) \partial_{ij} u(t, x), \quad \mathcal{L}_1^b u(t, x) := b^i(t, x) \partial_i u(t, x)
\]

with \(a(t, x) = (a^{ij}(t, x)) := (\sigma \sigma^T)(t, x)\), and \(\partial_i\) denotes the \(i\)-th partial derivative respect to \(x\). Here we use Einstein’s convention that the repeated indices in a product will be summed automatically. To be more specific, for any \(q \in (1, \infty)\) and \(p = (p_1, \cdots, p_d) \in (1, \infty)^d\), we need to establish the following estimate:

\[
\|\nabla^2 u\|_{L^q_p(T)} \leq C \|f\|_{L^q_p(T)},
\]

see Section 2 for the precise definition of \(L^q_p(T)\). Notice that when \(p_1 = \cdots = p_d = q\), it is a standard procedure to prove Eq. 1.8 by the classical freezing coefficient argument (cf. [21]). However, for general \(q \in (1, \infty)\) and \(p \in (1, \infty)^d\) it seems to be non-trival. When \(a^{ij}\) is independent of \(x\) and \(p_1 = \cdots = p_d\), estimate Eq. 1.8 was first proved by Krylov in [8]. In the spatial dependent diffusion coefficient case, Kim [9] showed Eq. 1.8 only for \(p_1 = \cdots = p_d \leq q\). This was recently generalized to \(p_1 = \cdots = p_d > 1\) and \(q > 1\) in [17] by a duality method. We shall further develop the argument used in [17], and combing with the interpolation technique, to prove that Eq. 1.8 holds for mixed-norms even in the space variable. The main result is provided by Theorem 2.1, which should be of independent interest in the theory of PDEs.

This paper is organized as follows: In Section 2, we study the maximal regularity estimate for second order parabolic equations. In Section 3, we prove our main theorem. Throughout this paper, we use the following convention: \(C\) with or without subscripts will denote a positive constant, whose value may change from one appearance to another, and whose dependence on parameters can be traced from calculations.

## 2 Parabolic Equations in Mixed-Norm Spaces

Fix \(T > 0\) and let \(\mathbb{R}^{d+1}_T := [0, T] \times \mathbb{R}^d\). This section is devoted to study the parabolic equation (1.7) on \(\mathbb{R}^{d+1}_T\) in general mixed-norm spaces. We first introduce some notations. For any \(\alpha \in \mathbb{R}\) and \(p = (p_1, ..., p_d) \in [1, \infty)^d\), let \(H^\alpha_p(\mathbb{R}^d) := (1 - \Delta)^{-\alpha/2}(L^p_\mathbb{R}(\mathbb{R}^d))\) be the usual Bessel potential space with norm

\[
\|f\|_{H^\alpha_p(\mathbb{R}^d)} := \|(1 - \Delta)^{\alpha/2} f\|_{L^p(\mathbb{R}^d)},
\]

where \((1 - \Delta)^{\alpha/2} f\) is defined through Fourier’s transform

\[
(1 - \Delta)^{\alpha/2} f := \mathcal{F}^{-1}((1 + |\cdot|^2)^{\alpha/2} \mathcal{F} f).
\]

Notice that for \(n \in \mathbb{N}\) and \(p = (p_1, ..., p_d) \in [1, \infty)^d\), an equivalent norm in \(H^n_p(\mathbb{R}^d)\) is given by

\[
\|f\|_{H^n_p(\mathbb{R}^d)} = \|f\|_{L^p(\mathbb{R}^d)} + \|\nabla^n f\|_{L^p(\mathbb{R}^d)}.
\]
For \( q \in [1, \infty) \) and any \( S < T \), we denote \( \mathbb{L}^q_T(S, T) := L^q([S, T]; L^p(\mathbb{R}^d)) \). For simplicity, we will write \( \mathbb{L}^q(T) := \mathbb{L}^q(0, T) \), and \( L^\infty(T) \) consists of functions satisfying
\[
\|f\|_{L^\infty(T)} := \sup_{t \in [0, T]} \sup_{x \in \mathbb{R}^d} |f(t, x)| < +\infty.
\]

We also introduce that for \( \alpha \in \mathbb{R} \),
\[
\mathcal{H}_{\alpha}^q(T) := L^q([0, T]; H_x^q(\mathbb{R}^d)),
\]
and the space \( \mathcal{H}_{\alpha, p}^q(T) \) consists of the functions \( u = u(t) \) on \( [0, T] \) with values in the space of distributions on \( \mathbb{R}^d \) such that \( u \in \mathcal{H}_{\alpha}^q(T) \) and \( \partial_t u \in L^q(0, T) \).

Throughout this section, we always assume that
\[\textbf{(Ha)}: a(t, x) = (\sigma \sigma^T)(t, x) \text{ is uniformly continuous in } x \in \mathbb{R}^d \text{ locally uniformly with respect to } t \in \mathbb{R}_+, \text{ and there exists a constant } \delta > 1 \text{ such that for all } x \in \mathbb{R}^d,\]
\[
\delta^{-1}|t|^2 \leq |a(t, x)\xi|^2 \leq \delta|\xi|^2, \quad \forall x \in \mathbb{R}^d.
\]

The main result of this section is as follows.

Theorem 2.1 Let \( T > 0 \) be small enough. Assume that \( \textbf{(Ha)} \) holds, \( p \in (1, \infty)^d \) and \( q \in (1, \infty] \). Let \( b \in \mathbb{L}^{\tilde{q}}_p(T) \) with \( \tilde{p}_i \tilde{q}_i \) satisfying \( \tilde{p}_i \in [p_i, \infty) \) for \( 1 \leq i \leq d \) and \( 2/\tilde{q}_1 + 1/\tilde{p}_1 + \cdots + 1/\tilde{p}_d < 1 \). Then for every \( f \in \mathbb{L}^\tilde{q}_p(T) \), there exists a unique solution \( u \in \mathcal{H}_{\alpha}^q(T) \) to equation \( (1.7) \). Moreover, we have the following estimates:

(i) there is a constant \( C_1 = C(d, p, q, \|b\|_{\mathbb{L}^\tilde{q}_p(T)}), T \geq 0 \) such that
\[
\|\partial_t u\|_{L^q_p(T)} + \|u\|_{H_x^q(\mathbb{R}^d)} \leq C_1 \|f\|_{L^q_p(T)};
\]

(ii) for any \( \alpha \in [0, 2 - \frac{2}{\tilde{q}}) \), there exists a constant \( C_T = C(d, p, q, \|b\|_{\mathbb{L}^\tilde{q}_p(T)}), T \) satisfying
\[
\lim_{T \to 0} C_T = 0 \text{ such that}
\]
\[
\|u\|_{\mathbb{H}^\infty_{\alpha, p}(T)} \leq C_T \|f\|_{L^q_p(T)}.
\]

In particular, we have
\[
\|u\|_{L^\infty(T)} \leq \hat{C}_T \|f\|_{L^q_p(T)}, \quad \text{if} \quad 2/q + 1/p_1 + \cdots + 1/p_d < 2,
\]
and
\[
\|\nabla u\|_{L^\infty(T)} \leq \hat{C}_T \|f\|_{L^q_p(T)}, \quad \text{if} \quad 2/q + 1/p_1 + \cdots + 1/p_d < 1,
\]
where \( \hat{C}_T > 0 \) is a constant satisfying \( \lim_{T \to 0} \hat{C}_T = 0 \).

We shall provide the proof of the above result in the following subsections.

2.1 Smooth Diffusion Coefficients Without Drift

In this subsection, we consider PDE \( (1.7) \) on \( \mathbb{R}^{d+1}_T \) with \( b \equiv 0 \), i.e.,
\[
\partial_t u(t, x) - \mathcal{L}_2^a u(t, x) - f(t, x) = 0, \quad u(0, x) = 0.
\]
We shall focus on the $L^q_p$-maximal regularity a priori estimate for (2.6). To this end, we assume that $a$ is smooth enough, i.e., $a$ satisfies (H$a$) and for all $m \in \mathbb{N}$,
\[ \| \nabla^m a^{ij} (t, \cdot) \|_\infty < \infty. \]
Motivated by [17], we also need to consider the dual equation for (2.6):
\[ \partial_t w(t, x) + \frac{1}{2} \partial_{ij} ((a^{ij} (t, x) w(t, x)) + f(t, x) = 0, \ w(T, x) = 0. \quad (2.7) \]
Our aim in this subsection is to prove the following result.

**Theorem 2.2** For any $p \in (1, \infty)^d$ and $q \in (1, \infty)$, there is a constant $C > 0$ depending only on $d, p, q$, $T$ and the continuity modulus of $a$ such that for every $f \in C^0_0([0, T] \times \mathbb{R}^d)$,
\[ \| \nabla^2 u \|_{L^q_p(T)} \leq C \| f \|_{L^q_p(T)}, \quad \| w \|_{L^q_p(T)} \leq C \| f \|_{L^q_{-2,p}(T)}, \quad (2.8) \]
where $u$ and $w$ are solutions of Eqs. 2.6 and 2.7 respectively. Moreover, for any $\alpha \in [0, 2 - \frac{2}{q})$, we have
\[ \| u \|_{H^{\alpha}_{a,p}(T)} \leq C_T \| f \|_{L^q_p(T)}, \quad \| w \|_{H^{\alpha}_{a-2,p}(T)} \leq C_T \| f \|_{L^q_{-2,p}(T)}. \quad (2.9) \]
where $C_T > 0$ is a constant satisfying $\lim_{T \to 0} C_T = 0$.

Before giving the proof of the above theorem, we first show the following lemma for later use, which generalizes [8, Lemma 2.6] (see also [9, Lemma 3.5]).

**Lemma 2.3** Let $T \in [0, \infty)$, $p \in (1, \infty)$ and $n \in \mathbb{N}$. For $k = 1, \cdots, n$, let $a_k : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ be measurable functions and there exists a constant $\delta \geq 1$ such that for all $t \in [0, T]$,
\[ \delta^{-1} |\xi|^2 \leq a_k^{ij}(t) \xi_i \xi_j \leq \delta |\xi|^2, \quad \forall \xi \in \mathbb{R}^d. \]
Let $\lambda_k \in (0, \infty)$, $\gamma_k \in \mathbb{R}$ and $u^k \in \mathcal{H}^p_{\lambda_k+2,p}(T)$ be the solution to the equation
\[ \partial_t u^k = a_k^{ij} \partial_{ij} u^k + f^k, \quad u^k(0, x) = 0 \]
with $f \in \mathbb{H}^p_{\lambda_k,p}(T)$. Denote by $\Lambda_k = (\lambda_k - \Delta)^{\gamma_k/2}$. Then for any $i = 2, \cdots, d$, we have
\[ \int_0^T \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \prod_{k=1}^n \| \Lambda_k \Delta u^k(t, \cdot, x_i, \cdots, x_d) \|_{L^p(\mathbb{R}^{d-1})} dx_i \cdots dx_d dt \leq C_0 \sum_{k=1}^n \int_0^T \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \| \Lambda_k f^k(t, \cdot, x_i, \cdots, x_d) \|_{L^p(\mathbb{R}^{d-1})} \times \prod_{j \neq k} \| \Delta_j \Delta u^j(t, \cdot, x_i, \cdots, x_d) \|_{L^p(\mathbb{R}^{d-1})} dx_i \cdots dx_d dt, \quad (2.10) \]
where $C_0$ is a positive constant.

**Proof** Without loss of generality we may assume $\gamma_k = 0$. Define $v^k := \Delta u^k$. For fixed $i = 2, \cdots, d$ and $x_i, x_{i+1}, \cdots, x_d \in \mathbb{R}$, take $X = (x^1, \cdots, x^n)$ with $x^i = (x_1^i, \cdots, x_d^i) \in \mathbb{R}^d$.
such that for all $1 \leq j \leq n$, $x_i^j = x_i \in \mathbb{R}$, $x_{i+1}^j = x_{i+1} \in \mathbb{R}$, \ldots, $x_d^j = x_d \in \mathbb{R}$. Hence, $X \in \mathbb{R}^{d + (n-1)(i-1)}$. For such $X$, we define

$$V(t, X) := v^1(t, x^1) \times \cdots \times v^n(t, x^n).$$

Then one can check that

$$\partial_t V(t, X) := \mathbb{P}V(t, X) + F(t, X),$$

where

$$\mathbb{P}V = a^i_j \frac{\partial^2 V}{\partial x_i \partial x_j},$$

$$F(t, X) := \Delta x^j G^j(t, X), \quad G^j(t, X) = f^j(t, x^j) \prod_{j \neq k} v^k(t, x^k).$$

By classical result (cf. [8, Lemma 1.5]) we have

$$\|V\|_{L^p([0,T] \times \mathbb{R}^{d+(n-1)(i-1)})} \leq C_0 \sum_j \|G^j\|_{L^p([0,T] \times \mathbb{R}^{d+(n-1)(i-1)})},$$

which is exactly Eq. 2.10. The lemma is proved.

With the above preparation, we can give:

**Proof of Theorem 2.2** Let $p = (p_1, p_2, \cdots, p_d) \in (1, \infty)^d$ and $q \in (1, \infty)$. We divide the proof into five steps: we first prove estimate Eq. 2.8 in step 1-4, and in the fifth step we show estimate Eq. 2.9.

**Step 1.** [Case $p_1 = \cdots = p_d \in (1, \infty)$ and $q \in (1, \infty)$]. In this case, the estimate (2.8) was proved by [17, Theorem 3.3].

**Step 2.** [Case $p_1 = \cdots = p_{d-1} \in (1, \infty)$ and $p_d = q \in (1, \infty)$]. We only prove the estimate for $w$ since the estimate for $u$ is similar and easier. By duality and the same argument as in the proof of [17, Theorem 3.3], it is sufficient to prove the desired estimate when $q = p_d = np_{d-1} = \cdots = np_1 = : np$ for $n \in \mathbb{N}_+$ and $p \in (1, \infty)$. That is to say, we shall prove:

$$\|w\|_{L^{np}([0,T] \times \mathbb{R}^{d+(n-1)(i-1)})} \leq C \|f\|_{H^{np}_2}^{p} \mathcal{P}, \quad \mathcal{P} = (p, \cdots, p, np).$$

Take a non-negative smooth function $\phi$ supported in the ball $B_r := \{x \in \mathbb{R}^d : |x| < r\}$ and $\int_{\mathbb{R}^d} |\phi|^p dx = 1$, where $r$ is a small constant which will be determined below. For $x, z \in \mathbb{R}^d$, $s \in \mathbb{R}_+$, define $\phi_z(x) := \phi(x-z)$, $w_z(s, x) := w(s, x)\phi_z(x)$, $f_z(s, x) := f(s, x)\phi_z(x)$ and $a_z(s) := a(s, z)$. Then we can write

$$\partial_t w_z + \partial_{ij}(a_{ij}^z w_z) + g_z = 0, \quad w_z(T, x) = 0, \quad (2.11)$$

where

$$g_z = f_z + \partial_i(a_{ij}^z w)\phi_z - \partial_{ij}(a_{ij}^z w\phi_z).$$

Below we drop the time variable for simplicity, and for any $\gamma \in \mathbb{R}$ and fixed $x_d \in \mathbb{R}$, we denote by $\|f(\cdot, x_d)\|_{H^\gamma_{p}^{p}([\mathbb{R}^{d-1})} := \|(1 - \Delta)^{\gamma/2} f(\cdot, x_d)\|_{L^p([\mathbb{R}^{d-1})}$. Notice that

$$g_z = f \phi_z - 2\partial_j(a_{ij}^z w)\partial_i\phi_z - a_{ij}^z w\partial_i \partial_j \phi_z + \partial_i \partial_j((a_{ij}^z - a_{ij}^z)w_z).$$
By the continuity of \(a\), we have

\[
\left( \int_{\mathbb{R}^d} \| g_z(\cdot, x_d) \|^{p}_{H^{-2}d^{-1}} \, dz \right)^{1/p} \leq C_1 \| f(\cdot, x_d) \|^{p}_{H^{-2}d^{-1}} 
+ C_r \sum_{i,j} \| (a^{ij} w)(\cdot, x_d) \|^{p}_{H^{-1}d^{-1}} 
+ C_r \sum_{i,j} \| a^{ij} w(\cdot, x_d) \|^{p}_{H^{-2}d^{-1}} + c_r \| w(\cdot, x_d) \|^{p}_{L^p d^{-1}},
\]

where \(C_r > 0\) and \(\lim_{r \to 0} c_r = 0\). Let \(\rho_n\) be a family of standard mollifiers and \(a_n(t, x) := a(t, \cdot) \ast \rho_n(x)\) be the mollifying approximation of \(a\). For every \(\varepsilon > 0\), we can take \(n\) large enough such that

\[
\sum_{i,j} \| (a^{ij} w)(\cdot, x_d) \|^{p}_{H^{-1}d^{-1}} + \sum_{i,j} \| a^{ij} w(\cdot, x_d) \|^{p}_{H^{-2}d^{-1}} 
\leq C_2 \| (aw)(\cdot, x_d) \|^{p}_{H^{-1}d^{-1}} 
+ C_2 \| (a_n w)(\cdot, x_d) \|^{p}_{H^{-1}d^{-1}} 
+ C_2 \| w(\cdot, x_d) \|^{p}_{H^{-1}d^{-1}} + c_1/n \| w(\cdot, x_d) \|^{p}_{L^p d^{-1}} 
\leq C_2 \| w(\cdot, x_d) \|^{p}_{H^{-2}d^{-1}} + c \| w(\cdot, x_d) \|^{p}_{L^p d^{-1}},
\]

where the last step is due to the interpolation and Young’s inequalities. Hence, we arrive at

\[
\left( \int_{\mathbb{R}^d} \| g_z(\cdot, x_d) \|^{p}_{H^{-2}d^{-1}} \, dz \right)^{1/p} \leq C_3 \| f(\cdot, x_d) \|^{p}_{H^{-2}d^{-1}} 
+ C_r \| w(\cdot, x_d) \|^{p}_{H^{-2}d^{-1}} + c_r \| w(\cdot, x_d) \|^{p}_{L^p d^{-1}}.
\]

(2.12)

Observe that

\[
\| w \|^{np}_{L^{np}(0,T \times \mathbb{R}, L^p d^{-1})} = \int_0^T \int_\mathbb{R} \left( \int_{\mathbb{R}^d} \| w(t, \cdot, x_d) \phi_z \|^{p}_{L^p d^{-1}} \, dz \right)^n \, dx \, dt 
= \int_0^T \int_\mathbb{R} \prod_{k=1}^n \| w_{z_k}(t, \cdot, x_d) \|^{p}_{L^p d^{-1}} dz_1 \cdots dz_n \, dx \, dt.
\]

(2.13)

Using Lemma 2.3, we can deduce that

\[
\int_0^T \int_\mathbb{R} \prod_{k=1}^n \| w_{z_k}(t, \cdot, x_d) \|^{p}_{L^p d^{-1}} \, dx \, dt 
\leq C_4 \sum_{k=1}^n \int_0^T \int_{\mathbb{R}} \| g_{z_k}(t, \cdot, x_d) \|^{p}_{H^{-2}d^{-1}} \prod_{l \neq k} \| w_{z_l}(t, \cdot, x_d) \|^{p}_{L^p d^{-1}} \, dx \, dt,
\]

\(\mathbb{C}\) Springer
which together with Eqs. 2.12 and 2.13 implies
\[
\|w\|_{L^p([0,T] \times \mathbb{R}, L^p(\mathbb{R}^{d-1}))}^{np} \leq C_5 \sum_{k=1}^n \int_0^T \int_{\mathbb{R}^{d}} \|g_{z_k}(t, \cdot, x_d)\|_{H_p^{2}(\mathbb{R}^{d-1})}^p \times \prod_{l \neq k} \|w_{zl}(t, \cdot, x_d)\|_{L^p(\mathbb{R}^{d-1})}^p \, dz_1 \cdots dz_n \, dx_d \, dt
\]
\[
= C_5 n \int_0^T \int_{\mathbb{R}} \left( \int_{\mathbb{R}^{d}} \|g_{z}(t, \cdot, x_d)\|_{H_p^{2}(\mathbb{R}^{d-1})}^p \, dz \right) \times \left( \int_{\mathbb{R}^{d}} \|w_{z}(t, \cdot, x_d)\|_{L^p(\mathbb{R}^{d-1})}^p \, dz \right)^{n-1} \, dx_d \, dt
\]
\[
\leq C_6 \int_0^T \int_{\mathbb{R}} \left( \int_{\mathbb{R}^{d}} \|g_{z}(t, \cdot, x_d)\|_{H_p^{2}(\mathbb{R}^{d-1})}^p \, dz \right) \times \|w(t, \cdot, x_d)\|_{L^p(\mathbb{R}^{d-1})}^{(n-1)p} \, dx_d \, dt
\]
\[
+ Cr \int_0^T \int_{\mathbb{R}} \left( \int_{\mathbb{R}^{d}} \|w(t, \cdot, x_d)\|_{L^p(\mathbb{R}^{d-1})}^p \, dz \right) \times |w(t, \cdot, x_d)|_{L^p(\mathbb{R}^{d-1})} \, dx_d \, dt
\]
\[
\leq C_6 \int_0^T \int_{\mathbb{R}} \int_{\mathbb{R}} \|f(t, \cdot, x_d)\|_{L^p(\mathbb{R}^{d-1})}^{np} \, dx_d \, dt
\]
\[
+ C_r \int_0^T \int_{\mathbb{R}} \|w(t, \cdot, x_d)\|_{L^p(\mathbb{R}^{d-1})}^{np} \, dx_d \, dt
\]
where the last inequality follows from Hölder’s inequality and Young’s inequality for product. Let \( r \) be small enough so that \( c_r < 1 \), we can get that
\[
\|w\|_{L^p([0,T] \times \mathbb{R}, L^p(\mathbb{R}^{d-1}))}^{np} \leq C_7 \left( \int_0^T \int_{\mathbb{R}} \|f(t, \cdot, x_d)\|_{H_p^{2}(\mathbb{R}^{d-1})}^{np} \, dx_d \, dt \right)
\]
\[
+ \int_0^T \int_{\mathbb{R}} \|w(t, \cdot, x_d)\|_{L^p(\mathbb{R}^{d-1})}^{np} \, dx_d \, dt \right) . \quad (2.14)
\]
It remains to control the last term on the right hand side of the above inequality. To this end, let \( \kappa_{s,t}^{z} := \int_t^u dz(u)du \) and
\[
P_{s,t}^{z}(x-y) := \frac{1}{\sqrt{(2\pi)^d \det(\kappa_{s,t}^{z})}} e^{-\frac{((x-y)^{T}(\kappa_{s,t}^{z})^{-1}(x-y))}{2\det(\kappa_{s,t}^{z})}}.
\]
Then the solution of Eq. 2.11 is given by
\[
w_{z}(t, x) = \int_t^T \int_{\mathbb{R}^d} P_{t,u}^{z}(x-y) g_z(u, y) \, dy \, du.
\]
By (H\(a \)) and a standard interpolation technique, we get that for any \( \alpha \in [0, 2) \),
\[
\|w_{z}(t, \cdot, x_d)\|_{H_p^{2}(\mathbb{R}^{d-1})} \leq C_8 \int_t^T (u-t)^{-\frac{\alpha}{2}} \|g_z(u, \cdot, x_d)\|_{H_p^{2}(\mathbb{R}^{d-1})} \, du.
\]
Thus by Minkowski’s inequality we have
\[
\|w(t, \cdot, x_d)\|_{H_p^{2}(\mathbb{R}^{d-1})} \leq \left( \int_{\mathbb{R}^{d}} \|w_{z}(t, \cdot, x_d)\|_{H_p^{2}(\mathbb{R}^{d-1})} \, dz \right)^{1/p}
\]
\[
\leq C_8 \int_t^T (u-t)^{-\frac{\alpha}{2}} \left( \int_{\mathbb{R}^{d}} \|g_z(u, \cdot, x_d)\|_{H_p^{2}(\mathbb{R}^{d-1})} \, dz \right)^{1/p} \, du.
\]
Using Eq. 2.12 and the similar argument as in the proof of Eq. 2.14, we further have
\[
\|w(t, \cdot, x_d)\|_{H^2_p(\mathbb{R}^{d-1})} \leq C \int_t^T (u-t)^{-\frac{q}{2}} \left( \|f(u, \cdot, x_d)\|_{H^2_p(\mathbb{R}^{d-1})} + \|w(u, \cdot, x_d)\|_{H^2_p(\mathbb{R}^{d-1})} \right) du.
\]
Let \(\frac{1}{q} + \frac{1}{np} = 1\), then for any \(\alpha \in (0, 2 - \frac{2}{np})\), we get by Hölder’s inequality that
\[
\|w(t, \cdot, x_d)\|_{H^2_p(\mathbb{R}^{d-1})} \leq C_0 \left( \int_t^T (u-t)^{-\frac{q}{2}} du \right)^{np/q'} \cdot \left( \int_t^T \|f(u, \cdot, x_d)\|_{H^2_p(\mathbb{R}^{d-1})} du + \|w(u, \cdot, x_d)\|_{H^2_p(\mathbb{R}^{d-1})} \right)^{np} du.
\]
(2.15)
where \(C_T > 0\) satisfying \(\lim_{T \to 0} C_T = 0\). Then by taking \(\alpha = 0\) and Gronwall’s inequality, we can obtain
\[
\sup_{s \in [0,T]} \|w(s, \cdot, x_d)\|_{H^2_p(\mathbb{R}^{d-1})} \leq C_T \int_0^T \|f(u, \cdot, x_d)\|_{H^2_p(\mathbb{R}^{d-1})} du,
\]
(2.16)
which in particular implies that
\[
\|w\|_{L^{np}(0,T) \times L^p(\mathbb{R}^{d-1})} \leq C_{10} \|f\|_{L^{np}_{2,p}(T)}, \quad p = (p, \cdots, p, np).
\]

Step 3. [Case \(p_1 = \cdots = p_{d-j} \in (1, \infty)\) and \(p_{d-j+1} = \cdots = p_d = q \in (1, \infty)\) with any \(1 \leq j \leq d - 1\)]. This can be proved by following exactly the same arguments as in the proof of step 2, except that we need to use Eq. 2.10 Lemma 2.3 with \(i = d - j + 1\), we omit the details.

Step 4. [Interpolation] We develop an interpolation scheme to show the following claim:

for every \(1 \leq j \leq d - 1\), (2.8) holds with \(p_1 = \cdots = p_{d-j} \in (1, \infty)\) and \(p_{d-j+1}, p_{d-j+2}, \cdots, p_d, q \in (1, \infty)\). (2.17)

In particular, when \(j = d - 1\), we get the desired result.

Interpolate the results in step 1 and step 2, we can get that Eq. 2.8 holds when \(p_1 = \cdots = p_{d-j} \in (1, \infty)\) and \(p_{d-j+1} = \cdots = p_d, q \in (1, \infty)\). Thus, the assertion (2.17) is true for \(j = 1\).

Assume that (2.17) holds for some \(j = n - 1 \leq d - 2\), we proceed to show that (2.17) is true for \(n\). For this, we first interpolate \(p_1 = \cdots = p_{d-j} \in (1, \infty)\) and \(q \in (1, \infty)\) with \(p_1 = \cdots = p_{d-j} \in (1, \infty)\) and \(p_{d-j+1} = \cdots = p_d, q \in (1, \infty)\) (both of which hold according to step 3) to get that Eq. 2.8 holds for \(p_1 = \cdots = p_{d-j} \in (1, \infty)\) and \(p_{d-j+1} = p_{d-j+2} = \cdots = p_d, q \in (1, \infty)\). Then we interpolate \(p_1 = \cdots = p_{d-j} \in (1, \infty)\) and \(p_{d-j+1} = p_{d-j+2} = \cdots = p_d, q \in (1, \infty)\) with \(p_1 = \cdots = p_{d-1} \in (1, \infty)\) and \(p_{d-1}, p_{d}, q \in (1, \infty)\) (which holds by the induction assumption for \(j = 1\)) to get that Eq. 2.8 holds for \(p_1 = \cdots = p_{d-j} \in (1, \infty)\) and \(p_{d-j+1} = \cdots = p_{d-1} \in (1, \infty)\) and \(p_{d-1}, p_d, q \in (1, \infty)\). We proceed to show that (2.17) holds for \(j = 2\). Keep interpolating with the induction
assumption for \( j = 3, \ldots, n - 1 \), we can get that Eq. 2.8 holds for \( p_1 = \cdots = p_{d-j} \in (1, \infty) \) and \( p_{d-j+1}, p_{d-j+2}, \cdots, p_d, q \in (1, \infty) \).

Step 5. Finally, we proceed to prove estimate (2.9). With the same argument as in the previous 4 steps, it is sufficient to prove the following estimate:

\[
\|w\|_{H^\alpha_{-2, pd}} \leq C_T \|f\|_{H^{\alpha p}_{-2, pd}}, \quad p = (p, \cdots, p, np), \quad \alpha \in \left[ 0, 2 - \frac{2}{np} \right],
\]

where \( \lim_{T \to \infty} C_T = 0 \). In fact, by Eqs. 2.15 and 2.16, we get for any \( \alpha \in [0, 2 - \frac{2}{np}) \),

\[
\sup_{s \in [0, T]} \int_{\mathbb{R}} \|w(s, \cdot, x_d)\|_{H^{\alpha p}_{-2, (R^d - 1)}} dx_d \leq \int_{\mathbb{R}} \sup_{s \in [0, T]} \|w(s, \cdot, x_d)\|_{H^{\alpha p}_{-2, (R^d - 1)}} dx_d \leq C_T \|f\|_{H^{\alpha p}_{-2, pd}}.
\]

The whole proof can be finished. \qed

2.2 Proof of Theorem 2.1

Now, we are in the position to give:

**Proof of Theorem 2.1** By standard continuity method, it suffices to establish the estimates (2.2) and (2.3). Estimates (2.4) and (2.5) then follow by Sobolev embedding theorems, see e.g. [1]. We divide the proof into two steps.

(i) (Case \( b \equiv 0 \)) For \( p \in (1, \infty)^d \) and \( q \in (1, \infty) \), let \( u \in H^q_p(T) \) and \( f \in H^q_p(T) \) satisfy Eq. 2.6, and let \( \rho_n \) be a family of standard mollifiers. Define

\[
u_n(t, x) := u(t, \cdot) * \rho_n(x), \quad a_n(t, x) := a(t, \cdot) * \rho_n(x), \quad f_n(t, x) := f(t, \cdot) * \rho_n(x).
\]

Then, it is easy to see that \( u_n \) satisfies

\[
\partial_t u_n = a_{ij} \partial_{ij} u_n + g_n, \quad u_n(0, x) = 0,
\]

where

\[
g_n := f_n + (a^{ij} \partial_{ij} u) * \rho_n - a^{ij} \partial_{ij} u_n.
\]

As a result of Eqs. 2.8 and 2.9, we have

\[
\|\nabla^2 u_n\|_{L^{q}_p(T)} \leq C_1 \left( \|f_n\|_{L^{q}_p(T)} + \|a^{ij} \partial_{ij} u\| * \rho_n - a^{ij} \partial_{ij} u_n\|_{L^{q}_p(T)} \right),
\]

and there exists a constant \( C_T \) with \( \lim_{T \to \infty} C_T = \infty \) such that

\[
\|u_n\|_{L^{\infty}_q(T)} \leq C_T \left( \|f_n\|_{L^{q}_p(T)} + \|a^{ij} \partial_{ij} u\| * \rho_n - a^{ij} \partial_{ij} u_n\|_{L^{q}_p(T)} \right).
\]

Letting \( n \to \infty \) and by the property of convolution, we can obtain the desired result.

(ii) (Case \( b \neq 0 \)) Let \( \frac{1}{p_1} = \frac{1}{p} + \frac{1}{p_2} \) and \( \frac{1}{q} = \frac{1}{q_1} + \frac{1}{q_2} \), by Hölder’s inequality and Sobolev embedding theorem (see [1]), we get

\[
\|b \cdot \nabla u\|_{L^{q}_p(T)} \leq C_2 \|b\|_{L^{q}_p(T)} \|\nabla u\|_{L^{q}_p(T)} \leq C_2 \|b\|_{L^{q}_p(T)} \|u\|_{L^{1+\theta, q_1}_p(T)} \leq C_T \|b\|_{L^{q}_p(T)} \|u\|_{L^{1+\theta, \infty}_p(T)} \tag{2.18},
\]

where \( \theta \in \left( \frac{1}{p_1} + \cdots + \frac{1}{p_d}, 1 - \frac{2}{q} \right) \subset \left( \frac{1}{p_1} + \cdots + \frac{1}{p_d}, 1 - \frac{2}{q} \right) \). By the result of (i) and Eq. 2.18, we have that

\[
\|u\|_{L^{\infty, \theta, q_1}_p(T)} \leq C_T \left( \|f\|_{L^{q}_p(T)} + \|b \cdot \nabla u\|_{L^{q}_p(T)} \right) \leq C_T \left( \|f\|_{L^{q}_p(T)} + \|b\|_{L^{q}_p(T)} \|u\|_{L^{1+\theta, \infty}_p(T)} \right).
\]
By choosing $T$ small so that $CT \|b\|_{L_p^q(T)} < 1$, we have
\[
\|u\|_{H_\infty^{1+\theta,p}(T)} \leq C_5 \|f\|_{L_p^q(T)}.
\] (2.19)

Then by (i) we get
\[
\|\nabla^2 u\|_{L_p^q(T)} \leq C_4 (\|f\|_{L_p^q(T)} + \|b\cdot \nabla u\|_{L_p^q(T)} + \|u\|_{H_1^{1+\theta,p}(T)})
\]
\[
\leq C_4 \|f\|_{L_p^q(T)}.
\]

Furthermore, Eq. 2.19 shows that for any $\alpha \in \left[0, 2 - \frac{2}{q}\right)$,
\[
\|u\|_{H_\infty^{\alpha,p}(T)} \leq C_T \|f\|_{L_p^q(T)},
\]
where $\lim_{T \to 0} C_T = 0$. The whole proof is finished. \qed

### 3 Well-Posedness of SDEs with Singular Coefficients

We first provide the weak well-posedness of SDE (1.3) and prove the Krylov estimate, which will play an important role below.

**Lemma 3.1** Assume (Ha) holds, $b \in L_p^q(T)$ with $q$, $p_1$, $\cdots$, $p_d \in (2, \infty]$ and $2/q + 1/p_1 + \cdots + 1/p_d < 1$. Then there exists a unique weak solution $(X_t)_{t \geq 0}$ to SDE (1.3). Moreover, for any function $f \in L_p^q(T)$ with $\hat{q}$, $\hat{p}_1$, $\cdots$, $\hat{p}_d \in (1, \infty)$ satisfying $2/\hat{q} + 1/\hat{p}_1 + \cdots + 1/\hat{p}_d < 2$, we have
\[
\mathbb{E} \left( \int_0^T |f(s, X_s)| ds \right) \leq C \|f\|_{L_p^q(T)},
\] (3.1)

where $C = C(d, \hat{p}, \hat{q}, \|b\|_{L_p^q(T)}, T)$ is a positive constant.

**Proof** Firstly, by Eqs. 2.4, 2.5 and following the same argument as in [20, Theorem 2.1], we can show that (3.1) holds when $b \equiv 0$. More precisely, for any $0 < S < T < \infty$ and function $f \in L_p^q(S, T)$ with $2/\hat{q} + 1/\hat{p}_1 + \cdots + 1/\hat{p}_d < 2$, there exists a constant $C(d, \hat{p}, \hat{q}) > 0$ such that
\[
\mathbb{E} \left( \int_S^T |f(t, Y_t)| dt \right) \leq C \|f\|_{L_p^q(S, T)},
\] (3.2)

where $(Y_t)_{t \geq 0}$ solves the following SDE without drift:
\[
dY_t = \sigma(t, Y_t) dW_t, \quad Y_0 = x \in \mathbb{R}^d.
\] (3.3)

Applying Eq. 3.2 to $f = |b|^2$, we can get
\[
\mathbb{E} \left( \int_S^T |b(t, Y_t)|^2 dt \right) \leq C \|b\|^2_{L_p^q(S, T)} = C \|b\|^2_{L_p^q(S, T)}.
\]
It then follows from Khasminskii’s lemma (see [20, Lemma 5.3]) that for any constant \( \kappa > 0 \),
\[
\mathbb{E} \exp \left\{ \kappa \int_0^T |b(s, X_s)|^2 \mathrm{d}s \right\} \leq C(\kappa, d, \mathbf{p}, q, \|b\|_{L^q_p(T)}) < \infty. \tag{3.4}
\]
As a result, we have
\[
\mathbb{E}_{\rho_T} := \mathbb{E} \exp \left\{ \int_0^T [b^T \sigma^{-1}] (s, Y_s) \mathrm{d}W_s - \frac{1}{2} \int_0^T [b^T (\sigma \sigma^T)^{-1} b] (s, Y_s) \mathrm{d}s \right\} = 1.
\]
Thus the existence of a weak solution \((X_t)_{t \geq 0}\) to SDE (1.3) follows by Girsanov’s theorem. Meanwhile, under \((\mathbf{H_a})\) there exists a unique weak solution to SDE (3.3) (see e.g. [14, Theorem 7.2.1]). Thus the weak well-posedness of SDE (1.3) follows by [7, Theorem 4.2]. Furthermore, we can deduce that
\[
\mathbb{E} \left( \int_0^T |f(t, X_t)| \mathrm{d}s \right) = \mathbb{E} \left( \rho_T \int_0^T |f(t, Y_t)| \mathrm{d}t \right) \leq \left( \mathbb{E} \int_0^T \rho_T^\alpha \mathrm{d}t \right)^{1/\alpha} \left( \mathbb{E} \int_0^T |f(t, Y_t)|^\beta \mathrm{d}t \right)^{1/\beta},
\]
where \( \alpha, \beta > 1 \) satisfying \( 1/\alpha + 1/\beta = 1 \). Since
\[
\mathbb{E}_{\rho_T}^\alpha = \mathbb{E} \left[ \left( \exp(-2\alpha \int_0^T [b^T \sigma]^{-1} \sigma \sigma^T \sigma^{-1}) \right)^{1/2} \left( \exp((4\alpha^2 - \alpha) \int_0^T [b^T (\sigma \sigma^T)^{-1} b] \sigma \sigma^T \sigma^{-1}) \right)^{1/2} \right],
\]
by Hölder’s inequality, the fact that exponential martingale is a supermartingale, Eqs. 2.1 and 3.4, we get for every \( \alpha > 1 \), \( \mathbb{E}_{\rho_T}^\alpha \leq C(\alpha, d, \mathbf{p}, q, \|b\|_{L^q_p(T)}) \). Then, it holds that
\[
\mathbb{E} \left( \int_0^T |f(t, X_t)| \mathrm{d}t \right) \leq C(\alpha, d, \mathbf{p}, q, \|b\|_{L^q_p(T)}, T) \left( \mathbb{E} \int_0^T |f(t, Y_t)|^\beta \mathrm{d}t \right)^{1/\beta}.
\]
Choosing \( \beta \) close enough to 1 such that \( 2/\tilde{q} + 1/\tilde{p}_1 + \cdots + 1/\tilde{p}_d < 2/\beta \) and taking \( \tilde{p} = \tilde{p}/\beta \), \( \tilde{q} = \tilde{q}/\beta \) in (3.2), we can get
\[
\mathbb{E} \left( \int_0^T |f(t, X_t)| \mathrm{d}t \right) \leq C \|f\|_{L^\infty_p^p(T)}^{1/\beta} \mathbb{E} \int_0^T |f(t, Y_t)|^\beta \mathrm{d}t^{1/\beta} = C \|f\|_{L^\infty_p^p(T)}.
\]
The proof is finished.

Recall that the Hardy-Littlewood maximal operator \( \mathcal{M} \) is defined by
\[
\mathcal{M} f(x) := \sup_{r \in (0, \infty)^d} \frac{1}{|B_r|} \int_{B_r} f(x + y) \mathrm{d}y, \quad \forall f \in L^1_{loc}(\mathbb{R}^d),
\]
where for \( r = (r_1, r_2, \cdots, r_d) \), \( B_r := \{ x \in \mathbb{R}^d : |x_1| < r_1, |x_2| < r_2, \cdots, |x_d| < r_d \} \). For every \( f \in C^\infty(\mathbb{R}^d) \), it is known that there exists a constant \( C_d > 0 \) such that for all \( x, y \in \mathbb{R}^d \) (see [17, Lemma 2.1]),
\[
|f(x) - f(y)| \leq C_d |x - y| (\mathcal{M} |\nabla f|(x) + \mathcal{M} |\nabla f|(y)), \tag{3.5}
\]
and the following \( L^p(\mathbb{R}^d) \)-boundness of \( \mathcal{M} \) with \( p \in (1, \infty)^d \) holds (see [6, Theorem 4.1]):
\[
\|\mathcal{M} f\|_{L^p(\mathbb{R}^d)} \leq C_d \|f\|_{L^p(\mathbb{R}^d)}. \tag{3.6}
\]
By Theorem 2.1, there exists a function $u \in \mathbb{H}^q_2, p$ satisfying
\[
\partial_t u(t, x) + \mathcal{L}_2 u(t, x) + \mathcal{L}_1 u(t, x) + b(t, x) = 0, \quad u(T, x) = 0.
\]
Define $\Phi(t, x) := x + u(t, x)$. In view of (2.5), we can choose $T$ small such that
\[
1/2 < \|\nabla \Phi^{-1}\|_{L^\infty(T)} \leq 2.
\]  
(3.7)
Assume that SDE (1.3) admits two solutions $X^1_t$ and $X^2_t$. By the Krylov's estimate (3.1), we can use Itô's formula to get that the process $Y^i_t := \Phi(t, X^i_t)$ satisfies
\[
dY^i_t = \nabla \Phi(t, X^i_t) \sigma(t, X^i_t) dW_t =: \Psi(t, X^i_t) dW_t, \quad i = 1, 2.
\]
Let $Z_t := X^1_t - X^2_t$, we have by (3.7) that
\[
\mathbb{E}|Z_t|^2 \leq 4\mathbb{E}|Y^1_t - Y^2_t|^2 \leq 4\mathbb{E} \left( \int_0^t |Z_s|^2 dA_s \right),
\]
where
\[
A_t := \int_0^t \frac{|\Psi(s, X^1_s) - \Psi(s, X^2_s)|^2}{|Z_s|^2} ds.
\]
Let $\rho_n$ be a family of mollifiers on $\mathbb{R}^d$, and define $\Psi^n(t, x) := \Psi(s, \cdot) * \rho^n(x)$. Then we can write
\[
\mathbb{E} A_t \leq \lim_{\epsilon \downarrow 0} \mathbb{E} \left( \int_0^t \frac{|\Psi(s, X^1_s) - \Psi(s, X^2_s)|^2}{|Z_s|^2} \cdot 1_{\{|Z_s| > \epsilon\}} ds \right)
\]
\[
\leq 3 \lim_{\epsilon \downarrow 0} \lim_{n \to \infty} \mathbb{E} \left( \int_0^t \frac{|\Psi^n(s, X^1_s) - \Psi(s, X^2_s)|^2}{|Z_s|^2} \cdot 1_{\{|Z_s| > \epsilon\}} ds \right)
\]
\[
+ \lim_{\epsilon \downarrow 0} \lim_{n \to \infty} \mathbb{E} \left( \int_0^t \frac{|\Psi^n(s, X^2_s) - \Psi(s, X^2_s)|^2}{|Z_s|^2} \cdot 1_{\{|Z_s| > \epsilon\}} ds \right)
\]
\[
+ \sup_{\epsilon \downarrow 0} \sup_{n \in \mathbb{N}} \mathbb{E} \left( \int_0^t \frac{|\Psi^n(s, X^1_s) - \Psi^n(s, X^2_s)|^2}{|Z_s|^2} \cdot 1_{\{|Z_s| > \epsilon\}} ds \right)
\]
\[
=: 3(\mathcal{I}_1(t) + \mathcal{I}_2(t) + \mathcal{I}_3(t)).
\]
By the property of mollification, it is easy to see that
\[
\mathcal{I}_1(t) + \mathcal{I}_2(t) \leq \lim_{\epsilon \downarrow 0} \epsilon^{-2} \lim_{n \to \infty} C \|\Psi^n - \Psi\|_{L^\infty(T)}^2 = 0.
\]
As for the third term, we can use (3.5), the Krylov's estimate Eqs. 3.1 and 3.6 to get that
\[
\mathcal{I}_3(t) \leq C \sup_{n \in \mathbb{N}} \mathbb{E} \left( \int_0^t \left[ |\mathcal{M}| \nabla \Psi^n |(s, X^1_s) + |\mathcal{M}| \nabla \Psi^n |(s, X^2_s) \right]^2 ds \right)
\]
\[
\leq C \sup_{n \in \mathbb{N}} \|\mathcal{M}| \nabla \Psi^n \|_{L^2(T)}^2 \leq C \|\nabla \Psi\|_{L^2(T)}^2 < \infty.
\]
Hence, as a result of the stochastic Gronwall’s inequality [19, Lemma 3.7], we can get $\mathbb{E}[|Z_t|^2] = 0$. The general case can be proved by a standard localization procedure as in [20, Theorem 1.3]. The proof is finished.

**Acknowledgments** The authors would like to thank the referees for their quite useful suggestions.

**References**

1. Burenkov, V., Viktorova, N.: The embedding theorem for Sobolev spaces with mixed norm for limit exponents. Math. Notes 59(1), 45–51 (1996)
2. Fedrizzi, E., Flandoli, F.: Pathwise uniqueness and continuous dependence of SDEs with non-regular drift. Stochastics 83(3), 241–257 (2011)
3. Fedrizzi, E., Flandoli, F., Priola, E., Vovelle, J.: Regularity of stochastic Kinetic equations. Electron. J. Probab. 22, 1–42 (2017)
4. Flandoli, F.: Random Perturbation of PDEs and Fluid Dynamic Models. Lecture Notes in Mathematics. Springer, Heidelberg (2011)
5. Gess, B.: Regularization and well-posedness by noise for ordinary and partial differential equations. Stoch. Partial Diff. Equ. Rel. Fields 229, 43–67 (2016)
6. Ho, K.: Strong maximal operator on mixed-norm spaces. Ann. Univ. Ferrara. 62, 275–291 (2016)
7. Ikeda, M., Watanabe, S.: Stochastic differential equations and diffusion processes. 2nd edn, North-Holland (1989)
8. Krylov, N.V.: The heat equation in $l^q((0, T),l^p)$-spaces with weights. SIAM J. Math. Anal. 32, 1117–1141 (2001)
9. Kyeong-Hun, K.: $L_q(L_p)$-theory of parabolic PDEs with variable coefficients. Bull. Korean. Math. Soc. 45, 169–190 (2008)
10. Krylov, N.V., Röckner, M.: Strong solutions of stochastic equations with singular time dependent drift. Probab. Theory Rel. Fields 131(2), 154–196 (2005)
11. Menoukeu, P.O., Meyer, B.T., Nilssen, T., Proske, F., Zhang, T.: A variational approach to the construction and Malliavin differentiability of strong solutions of SDEs. Math. Ann. 357, 761–799 (2013)
12. Pavliotis, G.A., Stuart, A.M.: Multiscale Methods: Averaging and Homogenization, Volume 53 of Texts in Applied Mathematics. Springer, New York (2008)
13. Röckner, M., Sun, X., Xie, L.: Averaging principle for stochastic differential equations with singular coefficient. preprint
14. Stroock, D.W., Varadhan, S.: Multidimensional Diffusion Processes. Springer (2007)
15. Veretennikov, A.J.U.: On the strong solutions of stochastic differential equations. Theory Probab. Appl. 24, 354–366 (1979)
16. Wang, F.Y.: Integrability conditions for SDEs and semi-linear SPDEs. Annals Prob. 45(5), 3223–3265 (2017)
17. Xia, P., Xie, L., Zhang, X., Zhao, G.: $l^q(l^p)$-theory of stochastic differential equations. Stochastic Process. Appl. 130, 5188–5211 (2020)
18. Xie, L.: Exponential ergodicity for stochastic Langevin equation with partial dissipative drift. Applied Math. L. 79, 34–42 (2018)
19. Xie, L., Zhang, X.: Ergodicity of stochastic differential equations with jumps and singular coefficients. Ann. Inst Henri Poincare-Pr. 56, 175–229 (2020)
20. Zhang, X.: Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients. Electron. J. Probab. 16, 1096–1116 (2011)
21. Zhang, X.: Stochastic differential equations with Sobolev coefficients and applications. Annals Applied Prob. 26(5), 2697–2732 (2016)
22. Zhang, X.: Stochastic Hamiltonian flows with singular coefficients. Science China Math. 61, 1353–1384 (2018)
23. Zvonkin, A.K.: A transformation of the phase space of a diffusion process that removes the drift. Math. Sbornik 93(135), 129–149 (1974)

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.