STOCHASTIC DIFFERENTIAL EQUATIONS WITH SOBOLEV COEFFICIENTS
AND APPLICATIONS

XICHENG ZHANG

School of Mathematics and Statistics, Wuhan University, Wuhan, Hubei 430072, P.R.China
Email: XichengZhang@gmail.com

Abstract. In this work we study the properties of solutions to stochastic differential equations
with Sobolev diffusion coefficients and singular drifts such as: the stability with respect to the
coefficients, weak differentiability with respect to the starting point, and the Malliavin differenti-
ability with respect to the sample path. We also establish Bismut-Elworthy-Li’s formula. As
applications, we use a stochastic Lagrangian representation to Navier-Stokes equations given by
Constantin-Iyer [1] to prove the local well-posedness of NSEs in $\mathbb{R}^d$ with initial values in the
first order Sobolev space $W^1_{\text{loc}}(\mathbb{R}^d)$ provided $p > d$.

1. Introduction and Main Results

Consider the following stochastic differential equation (abbreviated as SDE) in $\mathbb{R}^d$:

$$dX_t = b_t(X_t)dt + dW_t, \quad t \geq 0, \quad X_0 = x.$$ (1.1)

It is a classical result due to Veretennikov [23] that when $b$ is bounded and measurable, the
above SDE admits a unique strong solution, and more strongly, for almost all $\omega$, the uniqueness
holds for the following random ODE (cf. Davies [3]):

$$dX_t(\omega) = b_t(X_t(\omega) + W_t(\omega))dt, \quad t \geq 0, \quad X_0 = x.$$ (1.2)

Recently, in [16] and [17], the authors studied the Malliavin and Sobolev differentiabilities of
$X_t(x, \omega)$ with respect to the sample path $\omega$ and the starting point $x$ respectively, and also applied
to stochastic transport equations. Moreover, when

$$b \in L^q(\mathbb{R}_+; L^p(\mathbb{R}^d)) \quad \text{with} \quad p, q \in (1, \infty) \quad \text{and} \quad \frac{d}{p} + \frac{2}{q} < 1,$$

in a remarkable paper [13], Krylov and Röckner proved the existence and uniqueness of strong
solutions for SDE (1.1) by using Girsanov’s transformation and analytic estimates from PDE
theory. Later, their results are extended to the case of multiplicative noises in [26] (see also
[8] [24] for related results). Moreover, the Sobolev differentiability of solutions are also obtained
in [6] and [21]. The interesting of studying the Sobolev differentiability for SDEs (1.1) with
singular drifts is partly due to the discovery of Flandoli, Gubinelli and Priola [7] that noises can
prevent the singularity in linear transport equations (see also [5]).

In this work we consider the following SDE: for given $T < S$ with $T - S \leq 1$,

$$dX_{t,s} = b_s(X_{t,s})ds + \sigma_s(X_{t,s})dW_s, \quad X_{t,t} = x, \quad T \leq t \leq s \leq S,$$ (1.2)

where $b : [T, S] \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : [T, S] \times \mathbb{R}^d \to \mathbb{M}^d$ are two measurable functions,
and $\{W_t\}_{t \in [T,S]}$ is a standard $d$-dimensional Brownian motion on the classical Wiener space
$(\Omega, \mathcal{F}, P, \mathbb{H})$. Here, $\mathbb{M}^d$ denotes the set of all $d \times d$-matrices, $\Omega$ is the space of all continuous
functions from $[T, S]$ to $\mathbb{R}^d$, $\mathcal{F}$ is the Borel-$\sigma$ field, $P$ is the Wiener measure, and $\mathbb{H} \subset \Omega$ is the
Cameron-Martin space. We make the following assumption on $\sigma$:

Key words and phrases. Weak differentiability, Malliavin differentiability, Stability, Krylov’s estimates,
Zvonkin’s transformation.
(H$_K^p$) there exists a positive constant $K$ such that for all $(t, x) \in [T, S] \times \mathbb{R}^d$ and $\xi \in \mathbb{R}^d$,

$$K^{-1}|\xi|^2 \leq \sum_{i,k} |\sigma_i^k(x)\xi^k|^2 \leq K|\xi|^2,$$  

(1.3)

and for some $\alpha \in (0, 1),$

$$[\sigma_i]_\alpha := \sup_{x \neq y} \frac{|\sigma_i(x) - \sigma_i(x')|}{|x - y|^\alpha} \leq K.$$

Our main result of this paper is:

**Theorem 1.1.** Assume (H$_K^p$) and one of the following two conditions hold:

(i) $\sigma_i(x) = \sigma_i$ is independent of $x$ and for some $p, q \in (2, \infty)$ with $\frac{d}{p} + \frac{2}{q} < 1,$

$$b \in L^q([T, S]; L^p(\mathbb{R}^d)).$$

(ii) $\nabla \sigma, b \in L^q([T, S]; L^p(\mathbb{R}^d))$ for some $q = p > d + 2.$

Then we have the following conclusions:

(A) For any $(t, x) \in [T, S] \times \mathbb{R}^d,$ there is a unique strong solution denoted by $X_{t,s}(x)$ or $X_{t,s}^{b,\sigma}(x)$ to SDE (1.2), which has a bicontinuous version with respect to $t, x.$

(B) For each $s \geq t$ and almost all $\omega,$ the random field $x \mapsto X_{t,s}(x, \omega)$ is weakly differentiable and the Jacobian matrix $\nabla X_{t,s}(x)$ satisfies that for any $p' \geq 1,$

$$\text{ess. sup}_{x \in \mathbb{R}^d} \mathbb{E} \left( \sup_{x \in [t,s]} \|\nabla X_{t,s}(x)\|^{p'} \right) \leq C = C(d, p, q, K, \alpha, p', \|b\|_{L_q^p(t,s)}, \|\nabla \sigma\|_{L_q^p(t,s)}),$$

(1.4)

where the constant $C$ is increasing with respect to $\|b\|_{L_q^p(t,s)}$ and $\|\nabla \sigma\|_{L_q^p(t,s)}$.

(C) For each $s \geq t$ and $x \in \mathbb{R}^d,$ the random variable $\omega \mapsto X_{t,s}(x, \omega)$ is Malliavin differentiable, and for any $p' \geq 1,$

$$\sup_{x \in \mathbb{R}^d} \mathbb{E} \left( \sup_{x \in [t,s]} \|D X_{t,s}(x)\|^{p'} \right) < +\infty,$$

(1.5)

where $D$ is the Malliavin derivative (cf. [13]).

(D) For any $f \in C_b^1(\mathbb{R}^d),$ we have the following formula: for Lebesgue-almost all $x \in \mathbb{R}^d$,

$$\nabla \mathbb{E}f(X_{t,s}(x)) = \frac{1}{s-t} \mathbb{E} \left( f(X_{t,s}(x)) \int_t^s \sigma^{-1}(X_{t,r}(x))\nabla X_{t,r}(x) dW_r \right).$$

(1.6)

(E) Assume that $b' \in L_q^p(T, S)$ with the same $p, q$ as in the assumption, then

$$\sup_{x \in \mathbb{R}^d} \mathbb{E} \left( \sup_{x \in [t,s]} |X_{t,s}^{b,\sigma}(x) - X_{t,s}^{b',\sigma}(x)|^2 \right) \leq C\|b - b'\|_{L_q^p(t,s)}^2,$$

(1.7)

where $C = C(d, p, q, K, \alpha, \|b\|_{L_q^p(t,s)}, \|b'\|_{L_q^p(t,s)}, \|\nabla \sigma\|_{L_q^p(t,s)}).$

**Remark 1.2.** In the case of (i), conclusions (A) and (B) are not new, and has been contained in [13, 6, 26, 21], while, (C), (D) and (E) seem to be new. Our proofs are based on Zvonkin’s transformation (cf. [28]) and some results from the theory of PDEs. In particular, the global $L^p$-integrability are crucial. In a forthcoming paper, we shall study the Sobolev differentiability of solutions under some local $L^p$-integrability assumptions on $b$ and $\sigma.$

**Remark 1.3.** The stability estimate (1.7) has potential application for numerical solutions of SDEs with singular drifts. For example, consider the following SDE:

$$dX_t = 1_{A}(X_t)dt + dW_t, \quad X_0 = x.$$
where $A$ is a bounded open set of $\mathbb{R}^d$. Let $b_n(x) = 1_A * \varphi_n(x)$ be the mollifying approximation. By (1.7), the solution $X^n_{t}$ of SDE (1.7) corresponding to $b_n$ converges to $X_t$ in $L^2$. Next, we can approximate $X^n_{t}$ by Euler’s scheme. Thus, we obtain a numerical approximation.

As an application of the above theorem (one of our motivation), let us consider the following classical Navier-Stokes equation in $\mathbb{R}^d$:

$$\partial_t u = \nu \Delta u - u \cdot \nabla u + \nabla p, \quad \text{div} u = 0, \quad u_0 = \varphi,$$

where $u$ is the velocity field, $\nu$ is the viscosity constant and $p$ is the pressure of the fluid, $\varphi$ is the initial velocity with vanishing divergence. In [1], Constantin and Iyer provided a probabilistic representation to the above NSE as follows:

$$\begin{align*}
X_t(x) &= x + \int_0^t u_s(X_s(x))ds + \sqrt{2\nu}W_t, \\
u_t(x) &= \mathcal{P}[\nabla X_t^{-1} \cdot \varphi(X_t^{-1})](x),
\end{align*}$$

(1.8)

where $X_t^{-1}(x)$ denotes the inverse flow of $x \mapsto X_t(x)$, $\nabla X_t^{-1}$ is the transpose of Jacobian matrix, and $\mathcal{P} = \mathbb{I} - \nabla(-\Delta)^{-1}\text{div}$ is the Leray’s projection onto the space of all divergence free vector fields. When $d = 3$, let $\omega = \text{curl}(u) = \nabla \times u$ be the vorticity. Then the above second equation can be written as

$$\omega_t(x) = \mathbb{E}[\nabla^{-1}X_t^{-1}(x) \cdot \omega_0(X_t^{-1}(x))], \quad \omega_0 = \nabla \times \varphi. \quad (1.9)$$

In this case, the velocity $u$ can be recovered from $\omega$ by the Biot-Savart law (cf. [15]):

$$u_t(x) = \int_{\mathbb{R}^3} K_3(x - y)\omega_t(y)dy =: K_3 \omega_t(x), \quad (1.10)$$

where

$$K_3(x) = \frac{1}{4\pi} \frac{x \times h}{|x|^3}, \quad x, h \in \mathbb{R}^3.$$

In other words, we have the following stochastic representation to vorticity:

$$\begin{align*}
X_t(x) &= x + \int_0^t K_3(X_s(x) - y)\omega_t(y)dy + \sqrt{2\nu}W_t, \\
\omega_t(x) &= \mathbb{E}[\nabla^{-1}X_t^{-1}(x) \cdot \omega_0(X_t^{-1}(x))].
\end{align*}$$

(1.11)

Now if we substitute (1.9) and (1.10) into (1.11), then we obtain a closed equation:

$$X_t(x) = x + \mathbb{E} \int_0^t \int_{\mathbb{R}^3} [K_3(X_s(x) - y)\nabla^{-1}\tilde{X}_t^{-1}(y) \cdot \omega_0(\tilde{X}_t^{-1}(y))]dyds + \sqrt{2\nu}W_t,$$

where the random field $\{\tilde{X}_t(y)\}_{y \in \mathbb{R}^d}$ is an independent copy of $\{X_t(x)\}_{x \in \mathbb{R}^d}$, and $\mathbb{E}$ denotes the expectation with respect to $\tilde{X}$. By the change of variable $\tilde{X}_t^{-1}(y) = x'$ and noting that

$$\det \nabla \tilde{X}_t(x') = 1, \quad \nabla^{-1}\tilde{X}_t^{-1}(\tilde{X}_t(x')) = \nabla \tilde{X}_t(x'),$$

we further have

$$X_t(x) = x + \mathbb{E} \int_0^t \int_{\mathbb{R}^3} [K_3(X_s(x) - \tilde{X}_t(x'))\nabla \tilde{X}_t(x') \cdot \omega_0(x')]dx'ds + \sqrt{2\nu}W_t.$$

This is just the random vortex method for Navier-Stokes equations studied in [15] Chapter 6.

Recently, in [25] and [27], we studied a backward analogue for stochastic representation (1.8): for $t \leq s \leq 0$,

$$\begin{align*}
X_{t,s}(x) &= x + \int_s^t u_r(X_{r,s}(x))dr + \sqrt{2\nu}(W_s - W_t), \\
u_t(x) &= \mathcal{P}[\nabla^{\dagger}X_{t,0} \cdot \varphi(X_{t,0})](x),
\end{align*}$$

(1.12)
The advantage of this representation is that the inverse of stochastic flow \( x \mapsto X_{t,0}(x) \) does not appear. In particular, \( u_t(x) \) solves the following backward Navier-Stokes equation:

\[
\partial_t u + \nu \Delta u - u \cdot \nabla u + \nabla p = 0, \quad \text{div} u = 0, \quad u(0) = \varphi,
\]

Using Theorem 1.1 we have the following local well-posedness to stochastic system (1.12).

**Theorem 1.4.** For any \( p > d \) and \( \varphi \in \mathcal{W}_1^1(\mathbb{R}^d; \mathbb{R}^d) \), where \( \mathcal{W}_1^1 \) is the first order Sobolev space, there exist a time \( T = T(p,d,\nu,\|\varphi\|_{\mathcal{W}_1^1}) \) and a unique pair of \((u, X)\) with \( u \in L^\infty([T,0]; \mathcal{W}_1^1) \) solving stochastic system (1.12).

This paper is organised as follows: In Section 2, we recall some well-known results and give some preliminaries about the Sobolev differentiability of random vector fields. In Section 3, we study a class of parabolic partial differential equations with time dependent coefficients and give some necessary estimates. In Section 4, we prove Krylov’s and Khasminskii’s type estimates. In Section 5, we prove our main theorem 1.1 for SDEs (1.2) with \( b = 0 \). In Section 6, we prove Theorem 1.1. In Section 7, we prove Theorem 1.4 by Theorem 1.1 and a fixed point argument.

Throughout this paper, we use the following convention: \( C \) with or without subscripts will denote a positive constant, whose value may change in different places, and whose dependence on the parameters can be traced from the calculations.

2. Preliminaries

We first introduce some spaces and notations. For \( p, q \in [1, \infty] \) and \( T < S \), we denote by \( L^q_p(T,S) \) the space of all real Borel measurable functions on \([T,S] \times \mathbb{R}^d\) with the norm

\[
\|f\|_{L^q_p(T,S)} := \left( \int_T^S \left( \int_{\mathbb{R}^d} |f(t,x)|^p \, dx \right) \frac{dt}{T} \right)^{1/p} < +\infty.
\]

For \( m \in \mathbb{N} \) and \( p \geq 1 \), let \( \mathcal{W}_p^m = \mathcal{W}_p^m(\mathbb{R}^d) \) be the usual Sobolev space over \( \mathbb{R}^d \) with the norm

\[
\|f\|_{\mathcal{W}_p^m} := \sum_{k=0}^m \|\nabla^k f\|_p < +\infty,
\]

where \( \nabla^k \) denotes the \( k \)-order gradient operator, and \( \| \cdot \|_p \) is the usual \( L^p \)-norm. We also write

\[
\mathcal{W}_p^{2,q} = \mathcal{W}_p^2(T,S) = L^q(T,S; \mathcal{W}_p^2),
\]

and for a function \( f \) on \( \mathbb{R}^d \) and \( \alpha \in (0,1) \),

\[
[f]_{\alpha} := \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|^\alpha}.
\]

Let \( f \) be a locally integrable function on \( \mathbb{R}^d \). The Hardy-Littlewood maximal function is defined by

\[
\mathcal{M}f(x) := \sup_{0 < r < \infty} \frac{1}{|B_r|} \int_{B_r} f(x + y) \, dy,
\]

where \( B_r := \{ x \in \mathbb{R}^d : |x| < r \} \). We recall the following result (cf. [2] Appendix B).

**Lemma 2.1.** (i) There exists a constant \( C_d > 0 \) such that for all \( f \in \mathcal{W}_1^1 \) and Lebesgue-almost all \( x, y \in \mathbb{R}^d \),

\[
|f(x) - f(y)| \leq C_d |x - y| (\mathcal{M}|\nabla f|(x) + \mathcal{M}|\nabla f|(y)).
\]

(ii) For any \( p > 1 \), there exists a constant \( C_{d,p} \) such that for all \( f \in L^p(\mathbb{R}^d) \),

\[
\|\mathcal{M}f\|_p \leq C_{d,p} \|f\|_p.
\]
For $p > 1$, let $\mathcal{V}_p$ be the set of all continuous random fields $X : \mathbb{R}^d \times \Omega \rightarrow \mathbb{R}^d$ with

$$||X||_{\mathcal{V}_p} := ||X(0)||_{L_p^\infty} + ||\nabla X||_{L_p^\infty(V)} < \infty,$$

where $\nabla$ denotes the generalized gradient, and

$$L_{x}^\infty(L_{\omega}^p) := L_{x}^\infty(\mathbb{R}^d; L_{\omega}^p(\Omega)).$$

Let $\mathcal{V}_p^0 \subset \mathcal{V}_p$ be the set of random fields satisfying the additional condition

$$\int_{\mathbb{R}^d} \mathbb{E}f(X(x))dx = \int_{\mathbb{R}^d} f(x)dx. \quad \text{(2.3)}$$

**Remark 2.2.** The continuous assumption of $x \mapsto X(x)$ in the definition of $\mathcal{V}_p$ is purely technical. In fact, if $X \in \mathcal{V}_p$ for $p > d$, then by Sobolev’s embedding theorem, $x \mapsto X(x)$ always has a continuous version. Moreover, condition (2.3) means that $x \mapsto X(x)$ preserves the volume in the sense of mean values. Below, we also write

$$\mathcal{V}_{\infty-} := \bigcap_{p>1} \mathcal{V}_p, \quad \mathcal{V}_{\infty-}^0 := \bigcap_{p>1} \mathcal{V}_p^0, \quad L_x^{\infty-}(L_\omega^p) := \bigcap_{p>1} L_x^{\infty-}(L_\omega^p).$$

Let $\varrho : \mathbb{R}^d \rightarrow [0, 1]$ be a smooth function with support in $B_1$ and $\int \varrho dx = 1$. For $n \in \mathbb{N}$, define a family of mollifiers $\varrho_n(x)$ as follows:

$$\varrho_n(x) := n^{-d} \varrho(n^{-1} x). \quad \text{(2.4)}$$

For $X \in \mathcal{V}_p$, define

$$X_n(x) := \varrho_n \ast X(x) = \int_{\mathbb{R}^d} X(x-y)\varrho_n(y)dy. \quad \text{(2.5)}$$

Clearly, by Jensen’s inequality we have

$$\sup_{x \in \mathbb{R}^d} \mathbb{E}|\nabla X_n(x)|^p \leq \sup_{x \in \mathbb{R}^d} \mathbb{E}|\nabla X(x)|^p = ||\nabla X||_{L_x^\infty(V)}^p. \quad \text{(2.6)}$$

**Lemma 2.3.** For any $X \in \mathcal{V}_p$, we have

$$\mathbb{E}|X(x) - X(y)|^p \leq |x-y|^p ||\nabla X||_{L_x^\infty(V)}^p, \quad \forall x, y \in \mathbb{R}^d. \quad \text{(2.7)}$$

**Proof.** Let $X_n$ be defined by (2.5). By Fatou’s lemma and (2.6), we have for all $x, y \in \mathbb{R}^d$,

$$\mathbb{E}|X(x) - X(y)|^p \leq \lim_{n \rightarrow \infty} \mathbb{E}|X_n(x) - X_n(y)|^p$$

$$\leq |x-y|^p \lim_{n \rightarrow \infty} \int_0^1 \mathbb{E}||\nabla X_n(x + \theta(y-x))||^p d\theta$$

$$\leq |x-y|^p \sup_{x \in \mathbb{R}^d} \mathbb{E}||\nabla X_n(x)||^p \leq |x-y|^p ||\nabla X||_{L_x^\infty(V)}^p,$$

where in the first inequality, we have used the continuity of $x \mapsto X(x)$. \hfill \Box

**Lemma 2.4.** Let $\{X_n, n \in \mathbb{N}\} \subset \mathcal{V}_p$ be a bounded sequence and $X(x)$ a continuous random field. We assume that for each $x \in \mathbb{R}^d$, $X_n(x)$ converges to $X(x)$ in probability. Then $X \in \mathcal{V}_p$ and

$$||\nabla X||_{L_x^\infty(V)} \leq \sup_n ||\nabla X_n||_{L_x^\infty(V)}.$$

Moreover, for some subsequence $n_k$, $\nabla X_{n_k}$ weakly converges to $\nabla X$ as random variables in $L^p(\Omega \times B_R; \mathbb{M}^d)$ for any $R \in \mathbb{N}$. 

\[5\]
Proof. Since \( \sup_n \|X_n(0)\|_{L_p^p} < \infty \), by (2.6) and (2.7), we have for any \( R > 0 \),
\[
\sup_n \int_{B_R} (\mathbb{E}|X_n(x)|^p + \mathbb{E}|
abla X_n(x)|^p) dx < \infty,
\]
where \( B_R = \{ x : |x| < R \} \). This means that \( \{X_n(\cdot), n \in \mathbb{N}\} \) is bounded in \( L^p(\Omega; \mathcal{W}^{1,p}_p(B_R)) \), where \( \mathcal{W}^{1,p}_p(B_R) \) is the first-order Sobolev space over \( B_R \). Since \( L^p(\Omega; \mathcal{W}^{1,p}_p(B_R)) \) is weakly compact, by a diagonal argument, there are a subsequence \( n_k \) and a random field \( \tilde{X} \in \cap_{R \in \mathbb{N}} L^p(\Omega; \mathcal{W}^{1,p}_p(B_R)) \) such that for any \( R \in \mathbb{N} \),
\[
X_{n_k}(x) \to \tilde{X}(x) \text{ weakly in } L^p(\Omega; \mathcal{W}^{1,p}_p(B_R)).
\]
In particular, for any \( Z \in C_0^\infty(\mathbb{R}^d, \mathbb{R}^d) \) and \( \xi \in L^\infty(\Omega) \), we have
\[
\lim_{k \to \infty} \mathbb{E} \int_{\mathbb{R}^d} \langle X_{n_k}(x), Z(x)\xi \rangle dx = \mathbb{E} \int_{\mathbb{R}^d} \langle \tilde{X}(x), Z(x)\xi \rangle dx.
\]
Since for each \( x \in \mathbb{R}^d \), \( X_{n_k}(x) \) converges to \( X(x) \) in probability, by (2.8) and the dominated convergence theorem, we also have
\[
\lim_{k \to \infty} \mathbb{E} \int_{\mathbb{R}^d} \langle X_{n_k}(x), Z(x)\xi \rangle dx = \mathbb{E} \int_{\mathbb{R}^d} \langle X(x), Z(x)\xi \rangle dx.
\]
Thus, for all \( Z \in C_0^\infty(\mathbb{R}^d, \mathbb{R}^d) \) and \( \xi \in L^\infty(\Omega) \),
\[
\mathbb{E} \int_{\mathbb{R}^d} \langle X(x), Z(x)\xi \rangle dx = \mathbb{E} \int_{\mathbb{R}^d} \langle \tilde{X}(x), Z(x)\xi \rangle dx,
\]
which implies that \( X(x, \omega) = \tilde{X}(x, \omega) \) for \( dx \times P(d\omega) \)-almost all \( (x, \omega) \). In particular, for almost all \( \omega, x \mapsto X(x, \omega) \) is Sobolev differentiable, and by (2.9), \( \nabla X_{n_k} \) weakly converges to \( \nabla X \) as random variables in \( L^p(\Omega \times B_R; \mathbb{M}^d) \) for each \( R \in \mathbb{N} \).

Now, let \( \mathcal{Y}^\infty_c \) be the set of all \( \mathbb{M}^d \)-valued smooth random fields with compact supports and bounded derivatives. Let \( p_* = p/(p-1) \). Since the dual space of \( L^1(\mathbb{R}^d; L^p(\Omega)) \) is \( L^\infty(\mathbb{R}^d; L^p(\Omega)) \) and \( \mathcal{Y}^\infty_c \) is dense in \( L^1(\mathbb{R}^d; L^p(\Omega)) \), we have
\[
\|\nabla X\|_{L^p_p(L^p_{\mathcal{Y}_c^\infty}(\mathbb{R}^d))} = \sup_{U \in \mathcal{Y}^\infty_c \|U\|_{L^1(L^p_p)}, 1 < 1} \left| \mathbb{E} \int_{\mathbb{R}^d} \langle \nabla X(x), U(x) \rangle_{\mathbb{M}^d} dx \right|
\]
\[
= \sup_{U \in \mathcal{Y}^\infty_c \|U\|_{L^1(L^p_p)}, 1 < 1} \left| \mathbb{E} \left( \mathcal{E} \int_{\mathbb{R}^d} \langle X(x), \text{div} U(x) \rangle_{\mathbb{R}^d} dx \right) \right|
\]
\[
= \sup_{U \in \mathcal{Y}^\infty_c \|U\|_{L^1(L^p_p)}, 1 < 1} \left| \lim_{n \to \infty} \mathbb{E} \left( \mathcal{E} \int_{\mathbb{R}^d} \langle X_n(x), \text{div} U(x) \rangle_{\mathbb{R}^d} dx \right) \right|
\]
\[
= \sup_{U \in \mathcal{Y}^\infty_c \|U\|_{L^1(L^p_p)}, 1 < 1} \left| \lim_{n \to \infty} \mathbb{E} \left( \mathcal{E} \int_{\mathbb{R}^d} \langle \nabla X_n(x), U(x) \rangle_{\mathbb{M}^d} dx \right) \right|
\]
\[
\leq \sup_{n \in \mathbb{N}} \sup_{U \in \mathcal{Y}^\infty_c \|U\|_{L^1(L^p_p)}, 1 < 1} \left| \mathbb{E} \left( \mathcal{E} \int_{\mathbb{R}^d} \langle \nabla X_n(x), U(x) \rangle_{\mathbb{M}^d} dx \right) \right| = \sup_{n \in \mathbb{N}} \|\nabla X_n\|_{L^p_p(L^p_{\mathcal{Y}_c^\infty}(\mathbb{R}^d))}.
\]
The proof is complete. \( \square \)

**Proposition 2.5.** Let \( p_1, p_2, p_3 \in (1, \infty) \) with \( \frac{1}{p_3} = \frac{1}{p_1} + \frac{1}{p_2} \). Let \( X \in \mathcal{Y}_{p_1}, Y \in \mathcal{Y}_{p_2} \) be two independent random fields. Then we have \( X \circ Y \in \mathcal{Y}_{p_3} \) and
\[
\|\nabla (X \circ Y)\|_{L^p_{\mathcal{Y}_c^\infty}(\mathbb{R}^d)} \leq \|\nabla X\|_{L^p_{\mathcal{Y}_c^\infty}(\mathbb{R}^d)}\|\nabla Y\|_{L^p_{\mathcal{Y}_c^\infty}(\mathbb{R}^d)},
\]
(2.10)
Moreover, if for each \( x \in \mathbb{R}^d \), \( \omega \mapsto X(x, \omega), Y(x, \omega) \) are Malliavin differentiable and
\[
\sup_{x \in \mathbb{R}^d} \mathbb{E}\|DX(x)\|_{L^p}^{p_1} < \infty, \sup_{x \in \mathbb{R}^d} \mathbb{E}\|DY(x)\|_{L^p}^{p_2} < \infty,
\]
then \( X \circ Y(x) \) is also Malliavin differentiable and
\[
\sup_{x \in \mathbb{R}^d} \mathbb{E}\|DX \circ Y(x)\|_{L^p}^{p_3} < \infty.
\tag{2.11}
\]

**Proof.** Let \( X_n \) be defined by (2.5). By (2.7), we have
\[
\sup_{x \in \mathbb{R}^d} \mathbb{E}|X_n(x) - X(x)|^{p_1} \leq \sup_{x \in \mathbb{R}^d} \mathbb{E} \int_{\mathbb{R}^d} |X(x - y) - X(x)|^{p_1} g_n(y)dy \\
\leq \|\nabla X\|_{L^p(L^p)}^{p_1} \int_{\mathbb{R}^d} |y|^{p_1} \rho_n(y)dy \leq \|\nabla X\|_{L^p(L^p)}^{p_1}/n^{p_1}.
\]
In view of \( (X_n(x), X(x))_{x \in \mathbb{R}^d} \) and \( (Y_n(x), Y(x))_{x \in \mathbb{R}^d} \) are independent, we have for each \( x \in \mathbb{R}^d \),
\[
\mathbb{E}|X_n \circ Y(x) - X \circ Y(x)|^{p_1} = \mathbb{E}\left(\mathbb{E}|X_n(y) - X(y)|^{p_1}_{|y=Y(x)}\right) \\
\leq \sup_{y} \mathbb{E}|X_n(y) - X(y)|^{p_1} \leq \|\nabla X\|_{L^p(L^p)}^{p_1}/n^{p_1},
\]
and
\[
\|X_n \circ Y(x) - X_n \circ Y(x)\|_{L^p(L^p)}^{p_3} \leq \left\|\nabla Y_n(Y(x) + \theta(Y(x) - Y_n(x)))\right\|_{L^p(L^p)}^{p_1} \\
\leq \|\nabla Y_n\|_{L^p(L^p)}^{p_1} \sup_x \|\nabla X_n(x)\|_{L^p(L^p)}^{p_1} \leq \|\nabla X\|_{L^p(L^p)}^{p_1}\|\nabla Y\|_{L^p(L^p)}^{p_1}/n.
\]
Hence, in view of \( p_3 \leq p_1 \),
\[
\lim_{n \to \infty} \sup_{x \in \mathbb{R}^d} \mathbb{E}|X_n \circ Y(x) - X \circ Y(x)|^{p_3} = 0.
\tag{2.12}
\]
On the other hand, by the chain rule and Hölder’s inequality, we have
\[
\|\nabla (X_n \circ Y_n)\|_{L^p(L^p)} \leq \sup_{x \in \mathbb{R}^d} \left[\mathbb{E}(|\nabla X_n) \circ Y_n(x)|^{p_1} (\mathbb{E}|\nabla Y_n(x)|^{p_2})^{1/p_2}\right]^{1/p_1} \\
\leq \|\nabla X_n\|_{L^p(L^p)}\|\nabla Y_n\|_{L^p(L^p)} \leq \|\nabla X\|_{L^p(L^p)}\|\nabla Y\|_{L^p(L^p)}.
\]
which, together with (2.12) and by Lemma [2.4], yields (2.10).

Similarly, since by the chain rule,
\[
D(X_n \circ Y_n(x)) = (DX_n) \circ Y_n(x) + \nabla X_n \circ Y_n(x) \cdot DY_n(x),
\]
and \( (DX_n(x), \nabla X_n(x))_{x \in \mathbb{R}^d} \) and \( (Y_n(x))_{x \in \mathbb{R}^d} \) are independent, as above, we have
\[
\|D(X_n \circ Y_n)\|_{L^p(L^p)} \leq \|DX_n \circ Y_n\|_{L^p(L^p)} + \|\nabla X_n \circ Y_n \cdot DY_n\|_{L^p(L^p)} \\
\leq \|DX_n\|_{L^p(L^p)} + \|\nabla X_n \circ Y_n\|_{L^p(L^p)} \|DY_n\|_{L^p(L^p)} \\
\leq \|DX_n\|_{L^p(L^p)} + \|\nabla X_n\|_{L^p(L^p)} \|DY_n\|_{L^p(L^p)} \\
\leq \|DX\|_{L^p(L^p)} + \|\nabla X\|_{L^p(L^p)} \|DY\|_{L^p(L^p)},
\]
which, together with (2.12) and by [13, p.79, Lemma 1.5.3], yields (2.11). \( \square \)

Let \( \mathbf{P} = I - \nabla(-\Delta)^{-1} \) be the Leray’s projection on the space of divergence free vector fields. It is well-known that the singular integral operator \( \mathbf{P} \) is a bounded linear operator from \( L^p \) to \( L^p \). We also need the following result (cf. [1] and [25]).

**Lemma 2.6.** Let \( \varphi \in L^p(\mathbb{R}^d; \mathbb{R}^d) \) for some \( p > 1 \). We have the following conclusions:
(i) For any \( X \in L^p_\chi(L^{\infty}_\chi \cap \mathcal{V}_{\infty} \) and \( Y \in \mathcal{V}_{\infty} \), we have
\[
\mathbf{P}\mathbb{E}[\nabla X \cdot \varphi(Y)] = -\mathbf{P}\mathbb{E}[\nabla Y \cdot \nabla \varphi(Y) \cdot X].
\tag{2.13}
\]
(ii) For any \( X \in \mathcal{V}_{\infty}^0 \), we have
\[
\nabla \mathbb{P}[\nabla^i X \cdot \varphi(X)] = \mathbb{P}[\nabla^i X \cdot (\nabla^i \varphi - \nabla \varphi)(X) \cdot \nabla X].
\] (2.14)

Proof. Let \( X_n, Y_n, \varphi_n \) be the mollifying approximations of \( X, Y, \varphi \) defined by (2.5).

(i) Notice that
\[
\mathbb{P}[\nabla^i X_n \cdot \varphi_n(Y_m)] + \mathbb{P}[\nabla^i Y_m \cdot \nabla^i \varphi_n(Y_m) \cdot X_n] = \mathbb{P}[\nabla^i X_n \cdot \varphi_n(Y_m)] = 0.
\]
By (2.6), the dominated convergence theorem and Hölder’s inequality, it is easy to see that for each \( n \),
\[
\mathbb{E}[\nabla^i X_n \cdot \varphi_n(Y_m)] \rightarrow \mathbb{E}[\nabla^i X_n \cdot \varphi_n(Y)] \text{ in } L^p \text{ as } m \rightarrow \infty,
\]
and
\[
\mathbb{E}[\nabla^i Y_m \cdot \nabla^i \varphi_n(Y_m) \cdot X_n] \rightarrow \mathbb{E}[\nabla^i Y \cdot \nabla^i \varphi_n(Y) \cdot X_n] \text{ in } L^p \text{ as } m \rightarrow \infty.
\]
Hence,
\[
\mathbb{P}[\nabla^i X_n \cdot \varphi_n(Y)] = -\mathbb{P}[\nabla^i Y \cdot \nabla^i \varphi_n(Y) \cdot X_n].
\]
By letting \( n \rightarrow \infty \), we obtain (2.13).

(ii) As above calculations, we have
\[
\nabla \mathbb{P}[\nabla^i X_m \cdot \varphi_n(X_m)] = \mathbb{P}[\nabla^i X_m \cdot (\nabla^i \varphi_n - \nabla \varphi_n)(X_m) \cdot \nabla X_m].
\]
By Hölder’s inequality, we have
\[
\sup_{n,m} \| \nabla \mathbb{P}[\nabla^i X_m \cdot \varphi_n(X_m)] \|_p < \infty.
\]
Firstly letting \( m \rightarrow \infty \) and then \( n \rightarrow \infty \), we find that
\[
\mathbb{E}[\nabla^i X_m \cdot (\nabla^i \varphi_n - \nabla \varphi_n)(X_m) \cdot \nabla X_m] \rightarrow \mathbb{E}[\nabla^i X \cdot (\nabla^i \varphi - \nabla \varphi)(X) \cdot \nabla X] \text{ in } L^p,
\]
and
\[
\mathbb{E}[\nabla^i X_m \cdot \varphi_n(X_m)] \rightarrow \mathbb{E}[\nabla^i X \cdot \varphi(X)] \text{ in } L^p.
\]
Combining the above calculations, we obtain (2.14). \( \square \)

3. A study of PDE \( \partial_t u + L_t^\sigma u + b \cdot \nabla u + f = 0 \)

In the remainder of this paper, we shall fix \( T < S \) with \( S - T \leq 1 \). Let \( \sigma : [T, S] \times \mathbb{R}^d \rightarrow \mathbb{M}^d \) be a Borel measurable function, where \( \mathbb{M}^d \) is the set of all \( d \times d \)-matrices. Consider the following second order differentiable operator:
\[
L_t^\sigma u(x) := \frac{1}{2} \sigma(x) \sigma(x) \partial_t \partial_j u(x).
\] (3.1)

Here and below, we use the convention that the repeated indices in a product will be summed automatically. Under \( (H^d_x) \), it is a classical fact that operator \( \partial_t + L_t^\sigma \) has a fundamental solution \( \rho(t, x; s, y) \) (cf. [14]), i.e., for any \( f \in C_b(\mathbb{R}^d) \), the function
\[
\mathcal{T}_{t,s} f(x) := \int_{\mathbb{R}^d} \rho(t, x; s, y) f(y) dy
\]
satisfies that for all \( (t, x) \in [T, s] \times \mathbb{R}^d \),
\[
\mathcal{T}_{t,s} f(x) = \int_t^s L_r^\sigma \mathcal{T}_{r,s} f(x) dr, \quad \lim_{t \uparrow s} \mathcal{T}_{t,s} f(x) = f(x).
\] (3.2)
Moreover, we have the following estimates: for all \( x, y \in \mathbb{R}^d \) and \( T \leq t < s \leq S \) (cf. [14], p.376, (13.1)),
\[
\nabla_i \rho(t, x; s, y) \leq C_j (s - t)^{-\frac{j}{2}} \varphi_j(s - t, x - y), \quad j = 0, 1, 2.
\] (3.3)
where \( C_j, \kappa_j > 0 \) only depend on \( \alpha, K \) and \( d \), and
\[
\varphi_k(t, x) := t^{-\frac{d}{2}} e^{-\frac{|x|^2}{4t}},
\]

**Remark 3.1.** Although in [14 Chapter IV], the Hölder’s continuity of \( \sigma \) in \( t \) is also assumed, obviously, if we only freeze the spatial variable and do not care about the \( C^1 \)-smoothness with respect to the time variable, then the proof still works.

We first prove the following easy corollary of estimates \((3.3)\).

**Lemma 3.2.** For any \( p > d \) and \( \gamma \in (0, 1) \), there exist positive constants \( C_j = C_j(d, p, \alpha, K) \), \( j = 0, 1, 2 \) and \( C_2 = C_2(d, p, \gamma, \alpha, K) \) such that for all \( f \in L^p(\mathbb{R}^d) \) and \( T < t < S \),
\[
\|\nabla^j T_{t,s} f\|_{L^p} \leq C_j (s-t)^{-\frac{1}{2}-\frac{j}{p}} \|f\|_p, \quad j = 0, 1, 2, \tag{3.4}
\]
and
\[
[\nabla T_{t,s} f]_y \leq C_2 (s-t)^{-\frac{1}{2}-\frac{1}{p}} \|f\|_p. \tag{3.5}
\]

**Proof.** Without loss of generality, we assume \( f \in C_0^\infty(\mathbb{R}^d) \) is nonnegative. By the heat kernel estimate \((3.3)\), we have for all \( p \in [1, \infty] \),
\[
\|\nabla^j T_{t,s} f\|_p \leq C (s-t)^{-\frac{j}{p}} \|P^t s f\|_p, \quad j = 0, 1, 2, \tag{3.6}
\]
where
\[
P^t s f(x) := \int_{\mathbb{R}^d} \varphi_k(s-t, x-y)f(y)dy.
\]
Notice that the Gaussian heat kernel also has the property: for \( p \in [1, \infty] \),
\[
\|\nabla^j P^t s f\|_p \leq C (s-t)^{-\frac{j}{p}} \|f\|_p, \quad j = 0, 1. \tag{3.7}
\]
By Gagliardo-Nirenberg’s interpolation inequality, \((3.6)\) and \((3.7)\), we have for \( p > d \),
\[
\|T_{t,s} f\|_{L^p} \leq C \|T_{t,s} f\|_p^{1-p} \|\nabla T_{t,s} f\|_p^p \leq C (s-t)^{-\frac{1}{p}} \|f\|_p,
\]
\[
\|\nabla T_{t,s} f\|_{L^p} \leq C \|T_{t,s} f\|_p^{1-p} \|\nabla^2 T_{t,s} f\|_p^p \leq C (s-t)^{-\frac{1}{p}} \|f\|_p,
\]
which gives \((3.4)\). Similarly,
\[
\|P^t s f\|_{L^p} \leq C \|P^t s f\|_p^{1-p} \|\nabla P^t s f\|_p^p \leq C (s-t)^{-\frac{1}{p}} \|f\|_p. \tag{3.8}
\]
On the other hand, for \( \gamma \in (0, 1) \), we have
\[
|\nabla T_{t,s} f(x) - \nabla T_{t,s} f(x')| \leq C |x-x'|^\gamma \|\nabla^2 T_{t,s} f\|_p \|\nabla T_{t,s} f\|_{L^p}^{1-\gamma}
\]
\[
\leq C |x-x'|^\gamma (s-t)^{-\frac{1}{p}} \|P^t s f\|_p \|\nabla T_{t,s} f\|_{L^p}^{1-\gamma}, \tag{3.9}
\]
which gives \((3.5)\) by combining \((3.4)\) and \((3.8)\).

Below, for \( p > 1 \) and \( f \in L^p_T(\mathbb{R}^d) \), let us consider the following backward PDE:
\[
\partial_t u + L^p u + f = 0, \quad u(S) = 0. \tag{3.9}
\]

We have the following result.

**Theorem 3.3.** Assume \((H^\sigma_p)\) and \( p \in (1, \infty) \). For any \( f \in L^p_T(\mathbb{R}^d) \), there exists a unique solution \( u \in W^{2,p}_p(T, S) \) to PDE \((3.9)\) with
\[
\|u\|_{L^p_T(\mathbb{R}^d)} + \|\nabla^2 u\|_{L^p_T(\mathbb{R}^d)} \leq C \|f\|_{L^p_T(\mathbb{R}^d)}, \tag{3.10}
\]
where \( C = C(K, \alpha, p, d) > 0 \), and \( u(t, x) \) is explicitly given by
\[
u(t, x) = \int_t^S T_{t,s} f(s, x)ds. \tag{3.11}
\]
Moreover, we have the following estimates: for \( p, q \in (1, \infty) \) and \( \gamma \in (0, 1) \),

\[
\begin{align*}
\|u(0)\|_\infty & \leq C(S - t)^{1 - \frac{d}{2p} - \frac{1}{q}} \|f\|_{L^p_b(t, S)}^{\frac{d}{p} + \frac{2}{q}} < 2, \\
\|\nabla u(t)\|_\infty & \leq C(S - t)^{\frac{d}{2p} - \frac{1}{q}} \|f\|_{L^p_b(t, S)}^{\frac{d}{p} + \frac{2}{q} < 1}, \\
\|\nabla u(t)\|_\gamma & \leq C(S - t)^{\frac{1}{\gamma} \cdot \frac{d}{2p} - \frac{1}{q}} \|f\|_{L^p_b(t, S)}^{\frac{d}{p} + \frac{2}{q} + \gamma < 1},
\end{align*}
\]

where \( C = C(d, \alpha, K, p, q, \gamma) \) is independent of \( t \in [T, S] \).

**Proof.** (1) First of all, we assume \( \sigma(x) = \sigma_t \) does not depend on \( x \). In this case, by [12] Theorem 1.1, for any \( p, q \in (1, \infty) \), there exists a constant \( C = C(d, K, p, q) > 0 \) such that

\[
\int_T^S \left\| \nabla^2 \int_T^s T_{t,s} f(s, \cdot) ds \right\|^q_p dt \leq C \int_T^S \left( \int_T^S \|f(s)\|_p ds \right)^q dt \leq C \|f\|_{L^p_b(t, S)}^q.
\]

Moreover, by (3.6), (3.7) and Hölder’s inequality, it is easy to see that

\[
\int_T^S \left\| \int_T^s T_{t,s} f(s, \cdot) ds \right\|^q_p dt \leq C \int_T^S \left( \int_T^S \|f(s)\|_p ds \right)^q dt \leq C \|f\|_{L^p_b(t, S)}^q.
\]

In particular, the estimate (3.10) holds. When \( \sigma(x) \) depends on \( x \), and is uniformly continuous in \( x \), for \( p = q \), the apriori estimate (3.10) follows by a standard freezing coefficients argument (cf. [11]). As for the existence and stability, it follows by a standard continuity argument.

(2) Let us now prove (3.11). Let \( \varrho \) be a nonnegative smooth function in \( \mathbb{R}^{d+1} \) with support in \( \{x \in \mathbb{R}^{d+1} : |x| \leq 1\} \) and \( \int_{\mathbb{R}^{d+1}} \varrho(t, x) dx = 1 \). Set \( \varrho_n(t, x) := n^{d+1} \varrho(nt, nx) \) and extend \( u(s) \) to \( \mathbb{R} \) by setting \( u(s, \cdot) = u(S, \cdot) \) for \( s \in [T, S] \). Define

\[
u_n(t, x) := u_n(t, x) := \int_{\mathbb{R}^{d+1}} u(s, y) \varrho_n(t - s, x - y) ds dy
\]

and

\[
f_n(t, x) := -[\partial_t u_n(t, x) + L^r u_n(t, x)].
\]

Clearly, by (3.2) and the uniqueness of solutions to equation (3.9) in the class of \( C^{1,2}_b \) (cf. [22] Theorem 3.1.1), we have

\[
u_n(t, x) = \int_t^S T_{t,s} f_n(s, x) ds,
\]

and by (3.9),

\[
\|f_n - f\|_{L^p_b(t, S)} \leq \|\partial_t u_n - u\|_{L^p_b(t, S)} + K\|\nabla^2 u_n - u\|_{L^p_b(t, S)}
\]

\[
\leq \|\partial_t u \ast \varrho_n - \partial_t u\|_{L^p_b(t, S)} + K\|\nabla^2 u \ast \varrho_n - \nabla^2 u\|_{L^p_b(t, S)}
\]

\[
\leq \|f \ast \varrho_n - f\|_{L^p_b(t, S)} + 2K\|\nabla^2 u \ast \varrho_n - \nabla^2 u\|_{L^p_b(t, S)},
\]

which converges to zero as \( n \to \infty \) by the property of convolutions. Now taking limits for (3.17), we obtain (3.11).

(3) Next we prove (3.12). Let \( q^* := \frac{d}{q-1} \). By (3.11) and Hölder’s inequality, we have

\[
\|u(t)\|_\infty \leq \int_t^S \|T_{t,s} f(s)\|_\infty ds \leq C \int_t^S (s-t)^{\frac{d}{p} + \frac{2}{q}} \|f(s)\|_p ds
\]

\[
\leq C \left( \int_t^S (s-t)^{\frac{d}{q} + \frac{2}{q}} ds \right)^{\frac{1}{q}} \|f\|_{L^p_b(t, S)} \leq C(S - T)^{\frac{1}{q} \cdot \frac{d}{p} + \frac{2}{q}} \|f\|_{L^p_b(t, S)},
\]
which gives the first estimate in (3.12). Similarly, by (3.11) we also have
\[
\|\nabla u(t)\|_\infty \leq C \left( \int_t^S (s-t)^{-\frac{d}{2p} - \frac{d}{2q}} \, ds \right)^{\frac{1}{p}} \|f\|_{L_p^{\beta}(T,S)}
\]
and
\[
[\nabla u(t)]_T \leq C \left( \int_t^S (s-t)^{-\frac{d}{2p} - \frac{d}{2q}} \, ds \right)^{\frac{1}{p}} \|f\|_{L_p^{\beta}(T,S)},
\]
which gives the other estimates in (3.12). The proof is complete. \(\square\)

**Remark 3.4.** Although Krylov has proved (3.13) for \(p \neq q\) in [12], we do not know whether estimate (3.10) holds for \(p \neq q\) if \(\sigma\) depends on \(x\). The problem is that when we use the freezing coefficients argument, it seems that we have to require \(p = q\). This is also the reason why we have to consider two cases in Theorem 3.7.

Now we consider the following backward PDE:
\[
\partial_t u + L_p^\beta u + b \cdot \nabla u + f = 0, \quad u(S) = 0. \tag{3.19}
\]

**Theorem 3.5.** Assume \((H^r_K)\) and one of the following two conditions hold:

(i) \(\sigma_i(x) = \sigma_i\) is independent of \(x\) and for some \(p, q \in (1, \infty)\) with \(\frac{d}{p} + \frac{2}{q} < 1\), \(b \in L_p^q(T, S)\).

(ii) \(\nabla \sigma, b \in L_p^q(T, S)\) for some \(q = p > d + 2\).

For any \(f \in L_p^q(T, S)\) with the same \(p, q\) as above, there exists a unique solution \(u = u_f^b \in \mathcal{W}_p^{2,q}(T, S)\) to PDE (3.19) with
\[
\|u\|_{L_p^q(T,S)} + \|\nabla^2 u\|_{L_p^q(T,S)} \leq C_1 \exp \left\{ C_1 \|b\|_{L_p^q(T,S)}^q \right\} \|f\|_{L_p^q(T,S)},
\tag{3.20}
\]
and for all \(t \in [T, S]\),
\[
\|\nabla u(t)\|_\infty + \|\nabla^2 u(t)\|_{T_p} \leq C_1 (S - T)^{\beta/2} \exp \left\{ C_1 (S - T)^{\beta} \|b\|_{L_p^q(T,S)}^q \right\} \|f\|_{L_p^q(T,S)}, \tag{3.21}
\]
where \(\beta := \frac{1}{2} - \frac{d}{2p} - \frac{1}{q} > 0\) and \(C_1 = C_1(K, \alpha, p, q, d) > 0\). Moreover, we have
\[
\|u_f^b(t) - u_f^{b'}(t)\|_\infty + \|\nabla u_f^b(t) - \nabla u_f^{b'}(t)\|_{T_p} \leq C_2 \left( \|f - f'\|_{L_p^q(T,S)} + \|b - b'\|_{L_p^q(T,S)} \right), \tag{3.22}
\]
\[
\|u_f^b - u_f^{b'}\|_{L_p^q(T,S)} + \|\nabla^2 u_f^b - \nabla^2 u_f^{b'}\|_{L_p^q(T,S)} \leq C_2 \left( \|f - f'\|_{L_p^q(T,S)} + \|b - b'\|_{L_p^q(T,S)} \right). \tag{3.23}
\]
where \(b', f' \in L_p^q(T, S)\) and \(C_2 = C_2(K, \alpha, p, q, d, \|b\|_{L_p^q(T,S)}, \|b'\|_{L_p^q(T,S)}, \|f\|_{L_p^q(T,S)}, \|f'\|_{L_p^q(T,S)})\).

**Proof.** It suffices to prove the apriori estimates (3.20), (3.21) and (3.22), (3.23). First of all, by Duhamel’s formula (see (3.11)), the solution of equation (3.19) also satisfies the following integral equation:
\[
u(t) = \int_t^S \mathcal{T}_{t,s}(b \cdot \nabla u(s) + f(s)) \, ds. \tag{3.24}
\]
Let \(\beta := \frac{1}{2} - \frac{d}{2p} - \frac{1}{q}\). By (3.12), we have
\[
\|\nabla u(t)\|_{T_p}^q \leq C(S - T)^{\beta} \int_t^S \|b \cdot \nabla u(s) + f(s)\|_p^q \, ds
\]
\[
\leq C(S - T)^{\beta} \int_t^S \|b(s)\|_p^q \|\nabla u(s)\|_\infty^q + \|f(s)\|_p^q \, ds,
\]
which, by Gronwall’s inequality, yields that
\[
\|\nabla u(t)\|_\infty \leq C(S - T)^{\beta} \exp \left\{ C(S - T)^{\beta} \|b\|_{L_p^q(T,S)}^q \right\} \|f\|_{L_p^q(T,S)}^q.
\]
Moreover, by (3.12) again, we have
\[
[\nabla u(t)]^q_{p/2} \leq C(S - T)^{\beta q/2} \int_t^S \|b \cdot \nabla u(s) + f(s)\|_p^q \, ds \\
\leq C(S - T)^{\beta q/2} \int_t^S \|b(s)\|_p^q \|\nabla u(s)\|_{\infty}^q + \|f(s)\|_p^q \, ds \\
\leq C(S - T)^{\beta q/2} \exp \left( C(S - T)^{\beta q} \|b\|_{L^q_p(T,S)}^q \right) \|f\|_{L^q_p(T,S)}.
\]
Thus, we obtain (3.21).

On the other hand, in the case of (i), by (3.13) and (3.14), we have
\[
\|u\|_{L^q_p(T,S)} + \|\nabla^2 u\|_{L^q_p(T,S)} \leq C\|(b \cdot \nabla u) + f\|_{L^q_p(T,S)} \leq C\|b\|_{L^q_p(T,S)} \|\nabla u\|_{\infty} + C\|f\|_{L^q_p(T,S)} \\
\leq C\left( \|b\|_{L^q_p(T,S)} \exp \left( C\|b\|_{L^q_p(T,S)}^q \right) + 1 \right) \|f\|_{L^q_p(T,S)},
\]
which in turn gives (3.20). In the case of (ii), by (3.12) we still have (3.20).

By (3.24), we can write
\[
u(t) - u(t) = \int_t^S T_{t,s}(b \cdot \nabla(u(s) - u(t))) + ((b' - b) \cdot \nabla u(t))(s) + f(s) - f(t) \, ds.
\]
As above, using (3.12) and (3.21), by Gronwall’s inequality, we have
\[
\|\nabla u(t) - \nabla u(t)\|_{\infty} \leq C_1 \exp \left( C\left( \|b\|_{L^q_p(T,S)}^q + \|b'\|_{L^q_p(T,S)}^q \right) \right) \|f\|_{L^q_p(T,S)} + 1 \\
\times \left( \|f\|_{L^q_p(T,S)} + \|b - b'\|_{L^q_p(T,S)} \right).
\]
The desired estimates (3.22) and (3.23) follow by (3.12) and (3.10). 

4. Krylov and Khasminskii’s estimates

The following Krylov’s estimate was proved in [26, Theorem 2.1]. For the reader’s convenience, we reproduce the proof here.

**Theorem 4.1.** Assume \((H_\infty^K)\) and \(q, p \in (1, \infty)\) with \(\frac{2}{p} + \frac{2}{q} < 2\). Let \(0 < S - T \leq 1\). For any \(s \in [T, S]\) and \(x \in \mathbb{R}^d\), let \(X_T(x)\) solve SDE (1.2) with \(b = 0\). Then there exists a positive constant \(C = C(K, \alpha, d, p, q)\) such that for all \(f \in L^q_p(T, S)\), \(T \leq t \leq s \leq S\) and \(x \in \mathbb{R}^d\),
\[
\mathbb{E}\left( \int_t^s f(r, X_T(x)) \, dr \right) \leq C(s - t)^{1 - \frac{2}{q} + \frac{1}{2} - \frac{1}{2}} \|f\|_{L^q_p(T,S)}.
\]
where \(\mathcal{F}_t := \sigma(W_s : s \leq t)\).

**Proof.** Let \(p' = d + 1\). Since \(L^q_{p'}(T, S) \cap L^q_p(T, S)\) is dense in \(L^q_p(T, S)\), it suffices to prove (4.1) for \(f \in L^q_{p'}(T, S) \cap L^q_p(T, S)\).

Fix \(s \in [T, S]\). By Theorem 3.3, there exists a unique solution \(u \in W^{2, p'}(T, S)\) to the following backward PDE:
\[
\partial_t u + L^q_p u + f = 0, \quad t \in [T, s], \quad u(s, x) = 0.
\]
Moreover, by (3.10), we have for some constant \(C = C(K, \alpha, d, p, q) > 0\),
\[
\|u\|_{L^q_{p'}(t, s)} + \|\nabla^2 u\|_{L^q_{p'}(t, s)} \leq C\|f\|_p, \quad \forall t \in [T, s],
\]
and by (3.12),
\[
\sup_{t \in [T, s]} \|u(t)\|_{\infty} \leq C(s - t)^{1 - \frac{2}{q} + \frac{1}{2} - \frac{1}{2}} \|f\|_{L^q_p(T,S)}, \quad \forall t \in [T, s].
\]

12
Let \( u_n \) and \( f_n \) be defined as in (3.15) and (3.16). As in the calculations in (3.18), we have

\[
\lim_{n \to \infty} \| f_n - f \|_{L_p^p(t, s)} = 0.
\]

So, by the classical Krylov’s estimate (cf. [10] Lemma 5.1) or [8] Lemma 3.1], we have

\[
\lim_{n \to \infty} \mathbb{E} \left( \int_t^s |f_n(r, X_{r,t}) - f(r, X_{r,t})|^p dr \right) \leq C \lim_{n \to \infty} \| f_n - f \|_{L_p^p(t, s)} = 0. \tag{4.4}
\]

Now using Itô’s formula for \( u_n(t, x) \) and by (3.16), we have for any \( T \leq t \leq s \leq S \),

\[
u_n(s, X_{T,s}) = u_n(t, X_{T,t}) - \int_t^s f_n(r, X_{r,t})dr + \int_t^s \partial_t u_n(r, X_{r,t})\sigma_r^i(X_{r,t})dW_r^i.
\]

In view of

\[
\sup_{s,x} |\partial_t u_n(s, x)| \leq C_n,
\]

by Doob’s optional theorem, we have

\[
\mathbb{E} \left[ \left| \int_t^s \partial_t u_n(r, X_{r,t})\sigma_r^i(X_{r,t})dW_r^i \right|_{\mathcal{F}_t} \right] = 0.
\]

Hence,

\[
\mathbb{E} \left( \int_t^s f_n(r, X_{r,t})dr \right)_{\mathcal{F}_t} = \mathbb{E} \left( u_n(t, X_{T,t}) - u_n(s, X_{T,s}) \right)_{\mathcal{F}_t} \leq 2 \sup_{(r,x) \in [t,s] \times \mathbb{R}^d} |u_n(r, x)| \leq 2 \sup_{r \in [t,s]} \| u(r) \|_{\infty} \leq C(s-t)^{-\frac{1}{2} + \frac{1}{1+1}} \| f \|_{L_p^p(T,S)}.
\tag{4.3}
\]

The proof is thus completed by (4.4) and letting \( n \to \infty \).

We also need the following Khasminskii’s type estimate (cf. [19] Lemma 1.1]).

**Lemma 4.2.** Given \( T < S \), let \( \xi(t), \zeta(t), \beta(t), t \in [T, S] \) be three measurable \( \mathcal{F}_t \)-adapted processes and \( \eta(t), \alpha(t) \) two \( \mathbb{R}^d \)-valued measurable \( \mathcal{F}_t \)-adapted processes. Suppose that for any \( T \leq t \leq s \leq S \),

\[
\mathbb{E} \left( \int_t^s [\beta(r)] + |\alpha(r)|^2 dr \right)_{\mathcal{F}_t} \leq c_0 (s-t)\beta,
\tag{4.5}
\]

where \( c_0 > 0 \) and \( \beta \in (0, 1) \), and

\[
\xi(s) = \xi(T) + \int_T^s \zeta(r) dr + \int_T^s \eta(r) dW_r + \int_T^s \xi(r) \beta(r) dr + \int_T^s \xi(r) \alpha(r) dW_r.
\]

Then for any \( p > 0 \) and \( \gamma_1, \gamma_2, \gamma_3 > 1 \), we have

\[
\mathbb{E} \left( \sup_{s \in [T,S]} \xi^+(s)^p \right) \leq C \left[ \| \xi^+(T)^p \|_{\gamma_1} + \left\| \left( \int_T^S \xi^+(r) dr \right)^{\gamma_1} \right\|_{\gamma_2} + \left\| \left( \int_T^S \eta(r)^2 dr \right)^{\gamma_1} \right\|_{\gamma_2} \right],
\tag{4.6}
\]

where \( a^+ = \max[0, a], \quad C = C(c_0, \beta, \beta, \gamma_1, \gamma_3) > 0 \) and \( \| \cdot \|_\gamma \) denotes the norm in \( L^\gamma(\Omega) \).

**Proof.** Write

\[
M(s) := \exp \left\{ \int_T^s \alpha(r) dW_r - \frac{1}{2} \int_T^s |\alpha(r)|^2 dr + \int_T^s \beta(r) dr \right\}.
\]

By Itô’s formula, one sees that

\[
\xi(s) = M(s) \left\{ \xi(T) + \int_T^s M^{-1}(r)(\eta(r) dW_r + [\zeta(r) - \langle \alpha(r), \eta(r) \rangle] dr) \right\}.
\tag{4.7}
\]
By (4.5) and Khasminskii’s estimate (cf. [19, Lemma 1.1]), we have for any $p \geq 1$,
\[
\mathbb{E} \exp \left\{ p \int_T^{\infty} |\alpha(r)|^2 \, dr + p \int_T^{\infty} |\beta(r)| \, dr \right\} \leq C = C(c_0, \beta, p) < \infty,
\]
which implies that for any $p \in \mathbb{R}$,
\[
s \mapsto \exp \left\{ p \int_T^{\infty} \alpha(r) \, dW_r - \frac{p^2}{2} \int_T^{\infty} |\alpha(r)|^2 \, dr \right\}
\]
is an exponential martingale. Thus, by Hölder’s inequality and Doob’s maximal inequality, we have for any $p \in \mathbb{R}$,
\[
\mathbb{E} \left( \sup_{s \in [T, S]} |M(s)|^p \right) \leq C = C(c_0, \beta, p) < \infty.
\]
The desired estimate now follows by (4.7), Hölder and Burkholder’s inequalities.

5. SDEs without Drifts

In this section, we consider the following SDE:
\[
dX_{t,s} = \sigma_s(X_{t,s}) \, dW_s, \quad X_{t,t} = x, \quad s \geq t,
\]
where $\sigma : [T, S] \times \mathbb{R}^d \to \mathbb{M}^d_\sigma$ satisfies (H$_K^\sigma$). Under (H$_K^\sigma$), it is well-known that SDE (5.1) is well-posed in the sense of Stroock-Varadhan’s martingale solutions (cf. [22, p187, Theorem 7.2.1]). Indeed, Hölder’s continuity can be replaced with weaker conditions that $\sigma$ is uniformly continuous in $x$ with respect to $t$. Moreover, $\{X_{t,s}(x)\}$ defines a family of time non-homogeneous Markov processes. The aim of this section is to prove Theorem 1.1 for SDE (5.1). More precisely, we want to prove

**Theorem 5.1.** Assume (H$_K^\sigma$) and for some $q, p \in (1, \infty)$ with $\frac{d}{p} + \frac{2}{q} < 1$,
\[
\nabla \sigma_t \in L_{p, q}^\sigma(T, S).
\]

Then we have the following conclusions:

(a) For any $(t, x) \in [T, S] \times \mathbb{R}^d$, there is a unique strong solution denoted by $X_{t,s}(x)$ or $X_{t,s}^\sigma(x)$ to SDE (5.1), which has a bicontinuous version with respect to $t, x$.

(b) For each $s \geq t$ and almost all $\omega$, the random field $x \mapsto X_{t,s}(x, \omega)$ is weakly differentiable. Let $\nabla X_{t,s}(x)$ be the Jacobian matrix. If we let $J_{t,s}(x)$ solve the following linear matrix SDE:
\[
J_{t,s}(x) = 1 + \int_t^s \nabla \sigma_s(X_{t,s}(x)) J_{t,s}(x) \, dW_r,
\]
then $J_{t,s}(x) = \nabla X_{t,s}(x)$ a.s. for Lebesgue almost all $x \in \mathbb{R}^d$ and for any $p' \geq 1$,
\[
\sup_{x \in \mathbb{R}^d} \mathbb{E} \left( \sup_{x \in [T, S]} |J_{t,s}(x)|_{p'} \right) \leq C = C(p, q, d, K, \alpha, p', \|\nabla \sigma\|_{L^p_q(T, S)}),
\]
where the constant $C$ is increasing with respect to $\|\nabla \sigma\|_{L^p_q(T, S)}$.

(c) For each $s \geq t$ and $x \in \mathbb{R}^d$, the random variable $\omega \mapsto X_{t,s}(x, \omega)$ is Malliavin differentiable, and for any $p' \geq 1$,
\[
\sup_{x \in \mathbb{R}^d} \mathbb{E} \left( \sup_{x \in [T, S]} |DX_{t,s}(x)|_{p'} \right) \leq C = C(p, q, d, K, \alpha, p', \|\nabla \sigma\|_{L^p_q(T, S)}).
\]

Moreover, for any adapted vector field $h$ with $\mathbb{E} \int_T^{\infty} |h(r)|^2 \, dr < \infty$, the Malliavin derivative $D_h X_{t,s}(x)$ along $h$ solves the following linear SDE:
\[
D_h X_{t,s}(x) = \int_t^s \nabla \sigma_r(X_{t,r}(x)) D_h X_{t,r}(x) \, dW_r + \int_t^s \sigma_r(X_{t,r}(x)) h(r) \, dr.
\]
(d) For any \( f \in C^1_b(\mathbb{R}^d) \), we have the following formula: for Lebesgue almost all \( x \in \mathbb{R}^d \),
\[
\nabla \mathbb{E} f(X_{t,s}(x)) = \frac{1}{s-t} \mathbb{E} \left( f(X_{t,s}(x)) \int_t^s \sigma_r^{-1}(X_{t,r}(x)) \nabla X_{t,r}(x) \, dW_r \right).
\] (5.6)

(e) Assume that \( \sigma' \) satisfies the assumptions of the theorem with the same \( K, \alpha \) and \( p, q \), then
\[
\sup_{x \in \mathbb{R}^d} \mathbb{E} \left( \sup_{s \in [t,S]} |X_{t,s}^{\sigma'}(x) - X_{t,s}^{\sigma''}(x)|^2 \right) \leq C(S-t)^\theta \| \sigma - \sigma' \|^2_{L_p^q(t,S)},
\]
provided \( \| \sigma - \sigma' \|^2_{L_p^q(t,S)} < \infty \), where \( \theta \in (0, 1) \) only depends on \( p, q, d \).

5.1. Some apriori estimates. In this section, we assume that \( \sigma \) satisfies \( (H_0^\alpha) \) and
\[
\sup_{t,x} |\nabla^j \sigma_j(x)| < \infty, \ \forall j \in \mathbb{N}.
\]
In this case, it is well-known that the unique solution of SDE (5.1) denoted by \( X_{t,s}^{\sigma'}(x) \) or simply \( X_{t,s} \) forms a \( C^\infty \)-diffeomorphism flow (cf. [20, p.312, Theorem 39]). Let \( J_{t,s} := \nabla X_{t,s} \) be the Jacobian matrix, and \( DX_{t,s} \) the Malliavin derivative of \( X_{t,s} \) with respect to the sample path. Then we have (cf. [20, p.312, Theorem 39])
\[
J_{t,s} = \mathbb{I} + \int_t^s \nabla \sigma_r(X_{t,r}) J_{t,r} \, dW_r,
\] (5.7)
and for any \( h \in \mathbb{H} \),
\[
D_h X_{t,s} = \int_t^s \nabla \sigma_r(X_{t,r}) D_h X_{t,r} \, dW_r + \int_t^s \sigma_r(X_{t,r}) h_r \, dr.
\] (5.8)
We have the following apriori estimates.

**Proposition 5.2.** For any \( p' \geq 1 \), there exists a constant \( C = C(K, \alpha, p, q, d, p', \| \nabla \sigma \|_{L_p^q(T,S)}) \) such that
\[
\sup_{x \in \mathbb{R}^d} \mathbb{E} \left( \sup_{s \in [t,S]} |X_{t,s}(x)|^p \right) + \sup_{x \in \mathbb{R}^d} \mathbb{E} \left( \sup_{s \in [t,S]} \| DX_{t,s}(x) \|^p_{L_p^q} \right) \leq C,
\] (5.9)
where the constant \( C \) is increasing with respect to \( \| \nabla \sigma \|_{L_p^q(T,S)} \).

**Proof.** Without loss of generality, we assume \( t = T \) and write \( X_t := X_{T,s} \) and \( J_t := J_{T,t} \).

1. Below, we use the convention \( \frac{0}{0} := 0 \). If we let
\[
\beta(r) := |\nabla \sigma_t(X_r) J_r|^2 / |J_r|^2, \ \alpha(r) := 2 \langle J_r, \nabla \sigma_t(X_r) J_r \rangle / |J_r|^2,
\]
then by (5.7) and Itô’s formula, we have
\[
|J_s|^2 = |J_T|^2 + \int_T^s |J_r|^2 \beta(r) \, dr + \int_T^s |J_r|^2 \alpha(r) \, dW_r.
\]
By (4.1), we have for any \( T \leq t \leq s \leq S \),
\[
\mathbb{E} \left( \int_T^s \left( |\alpha(r)|^2 + |\beta(r)| |dr| \right) \right) \leq 5 \mathbb{E} \left( \int_T^s |\nabla \sigma_t(X_r)^2 | \, dr \right) \leq C(s-t)^{1-\frac{d}{2}-\frac{d}{p}} \| \nabla \sigma_t \|^2_{L_p^q(T,S)} = C(s-t)^{1-\frac{d}{2}-\frac{d}{p}} \| \nabla \sigma_t \|^2_{L_p^q(T,S)},
\]
which in turn gives the first estimate in (5.9) by (4.6).

2. For \( T \leq r \leq s \leq S \), let \( J_{r,s} \) solve the following linear SDE:
\[
J_{r,s} = \mathbb{I} + \int_r^s \nabla \sigma_r(X_r) J_{r,r}(x) \, dW_r.
\]
By (5.8) and the variation formula of constants, we have

\[ D_{n}X_s = \int_{T}^{s} J_{r,s} \sigma_r(X_r) h_r dr. \]  

(5.10)

Let \( \Sigma_{ij} := \langle DX^i_j, DX^j_i \rangle \) be the Malliavin covariance matrix. Then by (5.10), we have

\[ \Sigma_s = \int_{T}^{s} J_{r,s} \sigma_r(X_r) (J_{r,s} \sigma_r(X_r))' dr. \]  

(5.11)

As in step (1), one can show that for any \( p' \geq 1 \),

\[ \sup_{r \in [T,S]} \mathbb{E} \left( \sup_{x \in [T,S]} |J_{r,s}|^{p'} \right) \leq C. \]  

(5.12)

Thus, by (5.11) and (5.12) we have

\[ \mathbb{E} \left( \sup_{r \in [T,S]} |\Sigma_r|^{p'} \right) \leq C \mathbb{E} \left( \sup_{r \in [T,S]} \int_{T}^{s} |J_{r,s}|^{2p'} dr \right) \leq C \mathbb{E} \left( \int_{T}^{S} \sup_{s \in [T,S]} |J_{r,s}|^{2p'} dr \right) \leq C. \]

The proof is complete. \( \square \)

**Lemma 5.3.** Let \( \sigma, \sigma' : [T, S] \times \mathbb{R}^d \to \mathbb{M}^d \) satisfy (H\(^{\alpha}_K\)) with the same \( K, \alpha \) and for some \( p, q \in (2, \infty) \) with \( \frac{2}{p} + \frac{2}{q} < 1 \),

\[ \nabla \sigma, \nabla \sigma' \in L_{p}^{\alpha}(T, S). \]

Then there exists a constant \( C = C(K, \alpha, p, d, q, \| \nabla \sigma \|_{L_{p}^{\alpha}(T, S)}, \| \nabla \sigma' \|_{L_{p}^{\alpha}(T, S)}) > 0 \) such that

\[ \sup_{x \in \mathbb{R}^d} \mathbb{E} \left( \sup_{s \in [T,S]} |X_{t,s}^{\sigma}(x) - X_{t,s}^{\sigma'}(x)|^2 \right) \leq C(S - t)^{\theta} \| \sigma - \sigma' \|_{L_{p}^{\alpha}(T, S)}^2, \]  

(5.13)

where \( \theta \in (0, 1) \) only depends on \( p, q, d \). Moreover, for any \( \gamma > 1 \) and \( x \in \mathbb{R}^d \),

\[ \mathbb{E} \left( \sup_{s \in [T,S]} |\nabla X_{t,s}^{\sigma}(x) - \nabla X_{t,s}^{\sigma'}(x)|^2 \right) \leq C \left\| \int_{T}^{S} |\nabla \sigma_r(X_{t,s}^{\sigma'}(x)) - \nabla \sigma_r'(X_{t,s}^{\sigma'}(x))|^2 dr \right\|_{L_{\gamma}(\Omega)}. \]  

(5.14)

**Proof.** Without loss of generality, we assume \( t = T \) and write \( X_{T,s}^{\sigma} := X_{t,s}^{\sigma} \).

(1) Set \( Z_s := X_{s}^{\sigma} - X_{s}^{\sigma'}, \) then

\[ Z_s = \int_{T}^{s} [\sigma_r(X_{t,s}^{\sigma}) - \sigma'_r(X_{t,s}^{\sigma'}))] dW_r. \]

By Itô's formula, we have

\[ |Z_s|^2 = \int_{T}^{s} |\sigma(r, X_{t,s}^{\sigma}) - \sigma'_r(X_{t,s}^{\sigma'})|^2 dr + 2 \int_{T}^{s} [\sigma(r, X_{t,s}^{\sigma}) - \sigma'_r(X_{t,s}^{\sigma'})] Z_r dW_r \]

\[ = \int_{T}^{s} \zeta(r) dr + \int_{T}^{s} \eta(r) dW_r + \int_{T}^{s} Z_r^2 \beta(r) dr + \int_{T}^{s} Z_r^2 \alpha(r) dW_r, \]

where

\[ \zeta(r) := |\sigma_r(X_{t,s}^{\sigma}) - \sigma'_r(X_{t,s}^{\sigma'})|^2 - 2|\sigma_r(X_{t,s}^{\sigma}) - \sigma_r(X_{t,s}^{\sigma'})|^2, \]

\[ \eta(r) := 2[\sigma(r, X_{t,s}^{\sigma'}) - \sigma'_r(X_{t,s}^{\sigma'}))] Z_r, \]

\[ \beta(r) := 2|\sigma_r(X_{t,s}^{\sigma}) - \sigma_r(X_{t,s}^{\sigma'})|^2 / |Z_r|^2, \]

\[ \alpha(r) := 2|\sigma_r(X_{t,s}^{\sigma}) - \sigma_r(X_{t,s}^{\sigma'})| Z_r / |Z_r|^2. \]

Here we have used the convention \( \frac{1}{0} := 0. \)
By Lemma 2.1, we have for any $T \leq t < s \leq S$,
\[
\mathbb{E}\left(\int_t^s |\beta(r) + \alpha(r)|^2 dr \right) \leq C\mathbb{E}\left(\int_t^s \left[ M|\nabla\sigma_r|^2(X_r^\gamma) + M|\nabla\sigma_r|^2(X_r^\sigma) \right] dr \right)
\]
\[
\leq C(s-t)^{1-\frac{d}{p} - \frac{d}{q}} \|M|\nabla\sigma|^2\|_{L_{\infty}(T,S)}^{\gamma},
\]
and for any $\gamma \in (1, 1/(2/q + d/p))$,
\[
\mathbb{E}\left(\int_T^S |\sigma_r(X_r^\sigma) - \sigma_r^\prime(X_r^\sigma)|^2 dr \right) \leq C(S-T)^{1-\frac{d}{p} - \frac{d}{q}} \|\sigma - \sigma'\|_{L_{\infty}(T,S)}^{2\gamma},
\]
(5.15)
Since
\[
\zeta^\prime(r) \leq 2|\sigma_r(X_r^\sigma) - \sigma_r^\prime(X_r^\sigma)|^2,
\]
using (4.6) with $p = 1$, $\gamma_2 = \gamma$ and $\gamma_3 = \frac{2\gamma}{\gamma+1}$ and by Hölder’s inequality, we obtain
\[
\mathbb{E}\left(\sup_{s \in [T,S]} |Z_s|^2 \right) \leq C \left\| \left(\int_T^S |\sigma_r(X_r^\sigma) - \sigma_r^\prime(X_r^\sigma)|^2 dr \right)^{\frac{1}{2}} \right\|_{L^2(\Omega)}
\]
\[
+ C \left\| \int_T^S |\sigma_r(X_r^\sigma) - \sigma_r^\prime(X_r^\sigma)|^2 dr \right\|_{L^2(\Omega)}
\]
\[
\leq C \left\| \sup_{r \in [T,S]} |Z_r| \right\|_{L^2(\Omega)} \left\| \int_T^S |\sigma_r(X_r^\sigma) - \sigma_r^\prime(X_r^\sigma)|^2 dr \right\|_{L^2(\Omega)}
\]
\[
+ C \left\| \int_T^S |\sigma_r(X_r^\sigma) - \sigma_r^\prime(X_r^\sigma)|^2 dr \right\|_{L^2(\Omega)}
\]
\[
\leq \frac{1}{2} \left\| \sup_{r \in [T,S]} |Z_r| \right\|_{L^2(\Omega)}^2 + C \left\| \int_T^S |\sigma_r(X_r^\sigma) - \sigma_r^\prime(X_r^\sigma)|^2 dr \right\|_{L^2(\Omega)},
\]
which together with (5.15) yields (5.13).

(2) Set $U_s := J_s^\sigma - J_s^\sigma^\prime$. Then by (5.7), we have
\[
U_s = \int_T^s \left[ \nabla\sigma_r(X_r^\sigma) J_r^\sigma - \nabla\sigma_r^\prime(X_r^\sigma) J_r^\sigma \right] dW_r.
\]
By Itô’s formula, we have
\[
|U_s|^2 = 2 \int_T^s \left\langle U_r, \left[ \nabla\sigma_r(X_r^\sigma) J_r^\sigma - \nabla\sigma_r^\prime(X_r^\sigma) J_r^\sigma \right] \right\rangle dW_r
\]
\[
+ \int_T^s |\nabla\sigma_r(X_r^\sigma) J_r^\sigma - \nabla\sigma_r^\prime(X_r^\sigma) J_r^\sigma|^2 dr.
\]
As in the proof of (5.16), we have for any $\gamma > 1$,
\[
\mathbb{E}\left(\sup_{s \in [T,S]} |U_s|^2 \right) \leq \frac{1}{2} \left\| \sup_{r \in [T,S]} |U_r| \right\|_{L^2(\Omega)}^2 + C \left\| \int_T^S |\nabla\sigma_r(X_r^\sigma) - \nabla\sigma_r^\prime(X_r^\sigma)|^2 J_r^\sigma dr \right\|_{L^2(\Omega)}^2.
\]
Using (5.9) and by Hölder’s inequality, we obtain that for \( \gamma' > \gamma > 1 \),
\[
\mathbb{E} \left( \sup_{s \in [0,T]} |U_s|^2 \right) \leq C \left\| \int_T^S |\nabla \sigma^r_r(X^{\gamma'}_r) - \nabla \sigma^r_r(X^{\gamma'}_r)| J^r_r|^2 \, dr \right\|_{L^\gamma(\Omega)}
\leq C \left\| \sup_{r \in [T,S]} |J^r_r|^2 \int_T^S |\nabla \sigma^r_r(X^{\gamma'}_r) - \nabla \sigma^r_r(X^{\gamma'}_r)|^2 \, dr \right\|_{L^\gamma(\Omega)}
\leq C \left\| \int_T^S |\nabla \sigma^r_r(X^{\gamma'}_r) - \nabla \sigma^r_r(X^{\gamma'}_r)|^2 \, dr \right\|_{L^\gamma(\Omega)},
\]
which gives (5.14) by changing \( \gamma' \) to \( \gamma \).

5.2. **Proof of Theorem 5.1** (a) Under the assumptions, the pathwise uniqueness was proved in [26, Theorem 3.1] (see also the proof of (5.13)).

(b) Define \( \sigma^n_r(x) := \sigma_r * \varrho_n(x) \), where \( \varrho_n \) is a mollifier in \( \mathbb{R}^d \). Consider the following SDE:
\[
X^{\gamma,n}_{t,s}(x) = x + \int_t^s \sigma^n_r(X^{\gamma,n}_{t,r}(x)) \, dW_r, \quad s \geq t.
\]
Since \( \sigma^n \) is uniformly bounded, it is easy to see that for any \( p' > 1 \),
\[
\sup_n \mathbb{E} \left( \sup_{x \in [t,S]} |X^{\gamma,n}_{t,s}(x)|^{p'} \right) \leq C (1 + |x|^{p'}).
\]
Moreover, by (5.9) we have
\[
\sup_n \sup_{x \in \mathbb{R}^d} \mathbb{E} \left( \sup_{s \in [t,S]} |\nabla X^{\gamma,n}_{t,s}(x)|^{p'} \right) < \infty,
\]
and by (5.13),
\[
\lim_{n \to \infty} \sup_{x \in \mathbb{R}^d} \mathbb{E} \left( \sup_{s \in [t,S]} |X^{\gamma,n}_{t,s}(x) - X_{t,s}(x)|^2 \right) \leq C \lim_{n \to \infty} \left\| \sigma^n - \sigma \right\|^2_{L^p(t,S)} = 0. \tag{5.17}
\]
Thus, by Lemma 2.4, the random field \( x \mapsto X_{t,s}(x, \omega) \) is weakly differentiable almost surely, and for some subsequence \( n_k \) and any \( R \in \mathbb{N} \),
\[
\nabla X^{\gamma,n_k}_{t,s} \text{ weakly converges to } \nabla X_{t,s} \text{ as random variables in } L^{p'}(\Omega \times B_R; \mathbb{M}^d). \tag{5.18}
\]
Let \( J^{\gamma,n}_{t,s}(x) \) be the solution of SDE (5.2). We need to show that \( \nabla X_{t,s}(x) = J^{\gamma,n}_{t,s}(x) \). As in the proof of (5.9), we have
\[
\sup_{x \in \mathbb{R}^d} \mathbb{E} \left( \sup_{s \in [t,S]} |J^{\gamma,n}_{t,s}(x)|^{p'} \right) \leq C. \tag{5.19}
\]
Moreover, letting \( J^{\gamma,n}_{t,s}(x) := \nabla X^{\gamma,n}_{t,s}(x) \), by (5.14) we have
\[
\mathbb{E} \left( \sup_{s \in [t,S]} |J^{\gamma,n}_{t,s}(x) - J_{t,s}(x)|^2 \right) \leq C \left\| \int_t^S |\nabla \sigma^r_r(X^{\gamma,n}_{t,r}(x)) - \nabla \sigma_r(X_{t,r}(x))|^2 \, dr \right\|_{L^\gamma(\Omega)} \tag{5.20}.
\]
By (5.15), we have for \( \gamma' \in (1, 1/(2/(q + d)/p)) \),
\[
\sup_{x \in \mathbb{R}^d} \left\| \int_t^S |\nabla \sigma^r_r(X^{\gamma,n}_{t,r}(x)) - \nabla \sigma^r_r(X_{t,r}(x))|^2 \, dr \right\|_{L^\gamma(\Omega)} \leq C \left\| \nabla \sigma^m - \nabla \sigma \right\|^2_{L^p(t,S)}, \tag{5.21}
\]
where \( C \) is independent of \( n \). For fixed \( m \), by (5.17) we have
\[
\lim_{n \to \infty} \sup_{x \in \mathbb{R}^d} \left\| \int_t^S |\nabla \sigma^r_r(X^{\gamma,n}_{t,r}(x)) - \nabla \sigma^r_r(X_{t,r}(x))|^2 \, dr \right\|_{L^\gamma(\Omega)} = 0. \tag{5.22}
\]
Combining (5.20)–(5.22), we obtain
\[
\lim_{n \to \infty} \sup_{x \in \mathbb{R}^d} \left( \sup_{s \in [t, S]} |J^n_{t,s}(x) - J_{t,s}(x)|^2 \right) = 0,
\]
which together with (5.18) implies \(\nabla X_{t,s} = J_{t,s}\) a.e.

(c) By (5.9) again, we have for any \(p' \geq 1\),
\[
\sup_n \sup_{x \in \mathbb{R}^d} \mathbb{E} \left( \sup_{s \in [t, S]} ||DX^n_{t,s}(x)||^p_{\mathbb{E}} \right) \leq C,
\]
which together with (5.17) and by [18, p.79, Lemma 1.5.3] yields that \(X_{t,s}(x)\) is Malliavin differentiable and (5.4) holds. Let \(h\) be an adapted vector field with \(\mathbb{E} \int_t^S |h(r)|^2 dr < \infty\). Then we have
\[
D_{h}X^n_{t,s} = \int_t^s \nabla \sigma_r(X^n_{t,r}) D_{h}X^n_{t,r} dW_r + \int_t^s \sigma^n_r(X^n_{t,r}) h_r dr.
\]
Let \(Z^h_{t,s}\) solve
\[
Z^h_{t,s} = \int_t^s \nabla \sigma_r(X_{t,r}) Z^h_{t,r} dW_r + \int_t^s \sigma_r(X_{t,r}) h_r dr.
\]
As above, one can show that \(D_{h}X^n_{t,s} \to Z^h_{t,s}\) in \(L^2(\Omega)\). Moreover, for some subsequence \(n_k\), \(D_{h}X^n_{t,s}\) also weakly converges to \(D_{h}X_{t,s}\) in \(L^2(\Omega)\). Thus, \(Z^h_{t,s} = D_{h}X_{t,s}\) and equation (5.5) is obtained.

(d) By the classical Bismut-Elworthy-Li’s formula (cf. [4]), we have for any \(f \in C^1_b(\mathbb{R}^d)\),
\[
\nabla \mathbb{E} f(X^n_{t,s}(x)) = \frac{1}{s-t} \mathbb{E} \left[ f(X^n_{t,s}(x)) \int_t^s [\sigma^n_r(X^n_{t,r}(x))]^{-1} \nabla X^n_r(x) dW_r \right].
\]
Using limits (5.17) and (5.23), by taking limits for both sides of the above formula, we obtain (5.6). A more direct way for proving (5.6) is to use (b) and (c). We sketch it as follows: For any \(v \in \mathbb{R}^d\), define
\[
h_v(s') := \frac{1}{s-t} \int_t^{s'} [\sigma_r(X_{t,r})]^{-1} \nabla X_{t,r} dr, \quad s' \in [t, s].
\]
By (5.3), we have
\[
\mathbb{E} \int_t^s |h_v(r)|^2 dr = \frac{1}{(s-t)^2} \mathbb{E} \int_t^s [\sigma_r(X_{t,r})]^{-1} \nabla X_{t,r} dr < \infty.
\]
Since \(D_{h_v}X_{t,s'}\) satisfies
\[
D_{h_v}X_{t,s'} = \int_t^{s'} \nabla \sigma_r(X_{t,r}) D_{h_v}X_{t,r} dW_r + \frac{1}{s-t} \int_t^{s'} \nabla X_{t,r} dr.
\]
By (5.2) and the variation of constants formula, we have
\[
D_{h_v}X_{t,s} = \nabla X_{t,s}.
\]
Hence, by the integration by parts formula in the Malliavin calculus, we obtain
\[
\nabla_v \mathbb{E} \phi(X_{t,s}) = \mathbb{E} [\nabla \phi(X_{t,s}) \nabla X_{t,s}] = \mathbb{E} [\nabla \phi(X_{t,s}) D_{h_v}X_{t,s}] = \mathbb{E} [D_{h_v}(\phi(X_{t,s}))] = \frac{1}{s-t} \mathbb{E} \left( \phi(X_{t,s}) \int_t^s [\sigma_r(X_{t,r})]^{-1} \nabla X_{t,r} dW_r \right).
\]
(e) It follows by taking limits for
\[
\mathbb{E} \left( \sup_{x \in [t, S]} |X^n_{t,s}(x) - X^{\sigma^n}_{t,s}(x)|^2 \right) \leq C(s-t)^\theta \sup_{x \in [t, S]} \|\sigma_n - \sigma^n\|^2_{L^2(t,S)},
\]
6. Proof of Theorem 1.1

In this section we assume \((H^1_K)\) and one of the following two conditions hold:

(i) \(\sigma_t(x) = \sigma_t\) is independent of \(x\) and for some \(p, q \in (2, \infty)\) with \(\frac{2}{p} + \frac{2}{q} < 1\), \(b \in L^q_p(T, S)\).

(ii) \(\nabla \sigma, b \in L^q_p(T, S)\) for some \(q = p > d + 2\).

Let \([t_0, s_0] \subset [T, S]\) be any subinterval. For \(\ell = 1, \ldots, d\), by Theorem 3.5 let \(u^\ell\) solve the following PDE:

\[
\partial_t u^\ell + L_t^\ell u^\ell + b \cdot \nabla u^\ell + b^\ell = 0, \quad u^\ell(s_0, x) = 0.
\]

Let us set

\[
\Phi_t(x) := \Phi_t^b(x) := x + \sum_{\ell=1}^d u^\ell_t(x).
\]

We now prove the following Zvonkin’s transformation.

**Lemma 6.1.** Under (i) or (ii), for any \(U > 0\), there is a positive constant \(\beta = \delta(K, \alpha, d, p, q, U)\) being small enough such that if \(s_0 - t_0 \leq \delta\) and \(\|b\|_{L^q_p(t_0, s_0)} \leq U\), then for each \(t \in [t_0, s_0]\), \(x \mapsto \Phi_t(x)\) is a \(C^1\)-diffeomorphism with

\[
\frac{1}{2}|x - y| \leq |\Phi_t(x) - \Phi_t(y)| \leq \frac{1}{2}|x - y|.
\]

Moreover, letting \(\beta := \frac{1}{2} - \frac{d}{2p} - \frac{1}{q} > 0\), we have the following conclusions:

1. \(\|\nabla \Phi_t\|_{\infty} + \|\nabla \Phi_t^{-1}\|_{\infty} \leq K\), where \(K\) is a universal constant.

2. \(\|\nabla^2 \Phi\|_{L^q_p(t_0, s_0)} + \|\nabla \Phi\|_{L^q_p(t_0, s_0)} \leq C\), where \(C\) only depends on \(K, \alpha, p, q, d, \beta, U\).

3. Let \(b' \in L^q_p(t_0, s_0)\) be another function with \(\|b'\|_{L^q_p(t_0, s_0)} \leq U\). We have

\[
\|\Phi^b - \Phi^{b'}\|_{L^q_p(t_0, s_0)} + \|\nabla \Phi^b - \nabla \Phi^{b'}\|_{L^q_p(t_0, s_0)} \leq C\|b - b'\|_{L^q_p(t_0, s_0)}.
\]

4. \(X_{t_0,s}\) solves SDE (6.2) on \([t_0, s_0]\) if and only if \(Y_{t_0,s} := \Phi_s(X_{t_0,s})\) solves the following SDE:

\[
dY_{t_0,s} = \Sigma_s(Y_{t_0,s})dW_s, \quad s \in [t_0, s_0], \quad Y_{t_0,s_0} = \Phi_{t_0}(x),
\]

where \(\Sigma_s(y) := [\nabla \Phi_s, \sigma_s] \circ (\Phi^{-1}_s(y))\) satisfies \((H^1_K)\) with \(\alpha' = \alpha \wedge (\beta/2)\) and \(K' = K\).

5. Let \(\Sigma^b\) be defined as above through \(\Phi_t\). In the case of (3), we also have

\[
\|\Sigma^b - \Sigma^{b'}\|_{L^q_p(t_0, s_0)} \leq C\|b - b'\|_{L^q_p(t_0, s_0)},
\]

where \(C = C(K, \alpha, p, d, \beta, U)\).

**Proof.** Let \(\beta := \frac{1}{2} - \frac{d}{2p} - \frac{1}{q} > 0\). By (3.2.1), there is a \(C_0 = C_0(K, \alpha, p, d) > 0\) such that for all \([t_0, s_0] \subset [T, S]\),

\[
\|\nabla u\|_{\infty} + \|\nabla u\|_{L^q_p} \leq C_0(s_0 - t_0)^{\beta/2} \exp \left\{ C_0(s_0 - t_0)^{\beta/2} \|\nabla^q u\|_{L^q_p} \right\} \|b\|_{L^q_p(t_0, s_0)}.
\]

For given \(U > 0\), let us choose \(\beta = \delta(\beta, q, C_0, U)\) being small enough so that for all \(s_0 - t_0 \leq \delta\) and \(\|b\|_{L^q_p(t_0, s_0)} \leq U\),

\[
\sup_{t \in [t_0, s_0]} \|\nabla u_t\|_{\infty} + \|\nabla u_t\|_{L^q_p} < 1/2.
\]

In particular, we have

\[
|u_t(x) - u_t(y)| \leq |x - y|/2, \quad t \in [t_0, s_0],
\]

which then gives (6.2) by definition (6.1).

(1) It is obvious from (6.2).
(2) It follows from definition (6.1) and (3.20), (3.21).

(3) It follows from definition (6.1) and (3.22), (3.23).

(4) It follows by generalized Itô’s formula (see [10] or [26] for more details).

(5) By definition, let us write

\[ \Sigma^b_s(y) - \Sigma^b_s'(y) = [\nabla \Phi^b_s \cdot \sigma_s] \circ \Phi^b_{s-}(y) - [\nabla \Phi^b_s \cdot \sigma_s] \circ \Phi^b_{s-}'(y) + [(\nabla \Phi^b_s - \nabla \Phi^b_{s-}) \cdot \sigma_s] \circ \Phi^b_{s-}'(y) =: I_1(s, y) + I_2(s, y). \]

For \( I_1(s, y) \), by (2.1) we have

\[ |I_1(s, y)| \leq C(Mg_s(\Phi^b_{s-}(y)) + Mg_s(\Phi^b_{s-}'(y)))|\Phi^b_{s-}(y) - \Phi^b_{s-}'(y)|, \]

where \( g_s(x) := |\nabla [\nabla \Phi^b_s \cdot \sigma_s](x)| \in L^q_\mu(t_0, s_0) \) by (2), and \( Mg_s \) is the Hardy-Littlewood maximal function. Noticing that

\[ \sup_y |\Phi^b_{s-}(y) - \Phi^b_{s-}'(y)| = \sup_y |y - \Phi^b_{s-}'(y)||\nabla \Phi^b_{s-}'(y)| \leq \|\nabla \Phi^b_{s-}'(y)\|_{\infty} |\Phi^b_{s-}'(y) - \Phi^b_{s-}'(y)|, \]

by the change of variables and (3), we obtain

\[ \|I_1\|_{L^q_\mu(t_0, s_0)} \leq C\|Mg(\Phi^b_{s-}) + Mg(\Phi^b_{s-}')\|_{L^p_{\mu}(t_0, s_0)} \|\Phi^b_{s-}(y) - \Phi^b_{s-}'(y)|_\infty \]

\[ = C\|Mg\|_{L^p_{\mu}(t_0, s_0)} \|b - b'\|_{L^p_{\mu}(t_0, s_0)} \leq C\|g\|_{L^p_{\mu}(t_0, s_0)} \|b - b'\|_{L^p_{\mu}(t_0, s_0)}. \]

For \( I_2(s, y) \), by the change of variables and (3) again, we have

\[ \|I_2\|_{L^q_\mu(t_0, s_0)} \leq C\|\nabla \Phi^b - \nabla \Phi^b\|_{L^p_{\mu}(t_0, s_0)} \leq C\|b - b'\|_{L^p_{\mu}(t_0, s_0)}. \]

Combining the above calculations, we obtain (6.4).

We are now in a position to give

**Proof of Theorem 5.1** Let \( \delta \) be as in Lemma 6.1. Fix \( t_0 \in [T, S) \) and \( s_0 \in (t_0, S) \) with

\[ s_0 - t_0 \leq \delta. \]

Let us first prove the theorem on the time interval \([t_0, s_0]\). By Lemma 6.1 and Theorem 5.1, it is easy to see that (A), (B) and (C) hold. Let us look at (D). By (d) of Theorem 5.1, we have

\[ \nabla \mathbb{E}f(Y_{t_0, s}(y)) = \frac{1}{s - t_0} \mathbb{E} \left( f(Y_{t_0, s}(y)) \int_{t_0}^{s} \Sigma^{-1}_r(Y_{t_0, r}(y)) \nabla Y_{t_0, r}(y) dW_r \right). \]

(6.5)

Since \( Y_{t_0, s}(y) = \Phi_s \circ X_{t_0, s} \circ \Phi_{s-}^{-1}(y) \), by replacing \( f \) with \( f \circ \Phi_{s-}^{-1} \) and the change of variables \( y \rightarrow \Phi_s(x) \), we obtain (1.6). As for (E), it follows by (e) of Theorem 5.1 and (6.4).

Let us now consider the time interval \([t_1, s_1]\) where \( t_1 := \frac{t_0 + s_0}{2} \) and \( s_1 := \frac{3s_0 - t_0}{2} \). By the uniqueness of solutions, we have for all \( s \in [t_1, s_1] \),

\[ X_{t_0, s}(x) = X_{t_0, t_1} \circ X_{t_1, s}(x), \]

where \( X_{t_0, t_1}(\cdot) \) and \( X_{t_1, s}(\cdot) \) are independent. Thus, we can patch up the solutions and conclude the proof by Proposition 2.5.
7. Proof of Theorem 1.4

Given \( p > d \) and \( T \in [-1, 0] \), let \( b \in \mathbb{L}_p^\infty(T, 0) \) be divergence free, and let \( X_{t,s}(x) \) solve

\[
X_{t,s}(x) = x + \int_t^s b_s(X_{t,r}(x))dr + \sqrt{2}v(W_r - W_t).
\]

For given \( \varphi \in L^p(\mathbb{R}^d) \), define

\[
\mathcal{T}(b)_t(x) := u_t(x) := \mathbb{P}\mathbb{E}[\nabla^i X_{t,0} \cdot \varphi(X_{t,0})](x).
\]

**Lemma 7.1.** For any \( f \in L^1(\mathbb{R}^d) \), we have

\[
\mathbb{E} \int_{\mathbb{R}^d} f(X_{t,s}(x))dx = \int_{\mathbb{R}^d} f(x)dx. \tag{7.1}
\]

**Proof.** By a density and monotonic class argument, it suffices to prove it for \( f \in C_0^\infty(\mathbb{R}^d) \). Let \( b^n_t(x) = \varphi_n \ast b_t(x) \), where \( \varphi_n \) is a mollifier. Then \( \|\nabla b^n\|_\infty < \infty \) and \( \text{div} b^n = 0 \). By the classical fact, one has

\[
\mathbb{E} \int_{\mathbb{R}^d} f(X^n_{t,s}(x))dx = \int_{\mathbb{R}^d} f(x)dx.
\]

On the other hand, by (1.7) we have

\[
\lim_{n \to \infty} \mathbb{E} \left( \sup_{x \in [t,0]} |X^n_{t,s}(x) - X_{t,s}(x)|^2 \right) = 0.
\]

By taking limits, we obtain the desired equality. \( \square \)

Below we fix

\[
p > d \text{ and } q > \frac{2p}{p-d}.
\]

**Lemma 7.2.** For any given \( \varphi \in L^p(\mathbb{R}^d) \), there exist a constant \( C_0 = C_0(d, p, q, \nu) > 0 \) and a time \( T_0 = T_0(C_0, \|\varphi\|_p) < 0 \) such that if \( \|b\|_{L_p^\infty(T_0, 0)} \leq 2C_0\|\varphi\|_p \) and \( \text{div} b = 0 \), then

\[
\|\mathcal{T}(b)_t\|_{L^p_\nu} \leq 2C_0\|\varphi\|_p, \quad t \in [T_0, 0].
\]

**Proof.** By definition, we have

\[
\|\mathcal{T}(b)_t\|_{L^p_\nu} \leq C_{d,p}\|\mathbb{E}[\nabla^i X_{t,0} \cdot \varphi(X_{t,0})]\|_{L^p_\nu} \\
\leq C_{d,p} \text{ess. sup.} \|\nabla^i X_{t,0}(x)\|_{L^p_\nu} \|\varphi(X_{t,0})\|_{L^\infty_{\mathbb{R}^d}} \\
\leq C_{d,p} \text{ess. sup.} \|\nabla^i X_{t,0}(x)\|_{L^q_{\mathbb{R}^d}} \|\varphi\|_{L^p_\nu} \\
\leq C(d, q, p, \nu, \|b\|_{L^q_{\mathbb{R}^d}})\|\varphi\|_{L^p_\nu}.
\]

Since the constant \( C \) is increasing as a function of \( \|b\|_{L^q_{\mathbb{R}^d}} \) and goes to some \( C_0 = C_0(d, p, q, \nu) \) as \( \|b\|_{L^q_{\mathbb{R}^d}} \to 0 \) and

\[
\|b\|_{L^q_{\mathbb{R}^d}} \leq \|b\|_{L^q_{\mathbb{R}^d}}|t|^{1/q} \leq 2C_0|t|^{1/q}\|\varphi\|_p,
\]

one can choose \( T_0 < 0 \) being close to zero so that

\[
C(d, q, p, \nu, 2C_0|T_0|^{1/q}\|\varphi\|_p) \leq 2C_0.
\]

The proof is complete. \( \square \)
Lemma 7.3. For given \( \varphi \in \mathcal{W}^1_p(\mathbb{R}^d, \mathbb{R}^d) \), let \( C_0 \) and \( T_0 \) be as in Lemma 7.2 and \( U := 2C_0\|\varphi\|_{\mathcal{W}^1_p} \), there exists a time \( T_1 = T_1(d, \nu, p, q, U) \in (T_0, 0) \) such that for all \( b, b' \in L^\infty_p(T_1, 0) \) with
\[
\|b\|_{L^\infty_p(T_1, 0)} \|b'\|_{L^\infty_p(T_1, 0)} \leq U, \quad \text{div} b = \text{div} b' = 0,
\]
it holds that for all \( t \in [T_1, 0) \),
\[
\|\mathbb{T}(b)_t - \mathbb{T}(b')_t\|_{L^p} \leq \frac{1}{2}\|b - b'\|_{L^\infty_p(T_1, 0)}.
\]

**Proof.** We have
\[
\|\mathbb{T}(b)_t - \mathbb{T}(b')_t\|_p \leq \|\mathbb{P}\mathbb{E}(\nabla^i X^b_{t,0} \cdot \varphi(X^b_{t,0})) - \mathbb{P}\mathbb{E}(\nabla^i X^{b'}_{t,0} \cdot \varphi(X^{b'}_{t,0}))\|_p
\]
\[
\leq \|\mathbb{P}\mathbb{E}(\nabla^i X^b_{t,0} \cdot (\varphi(X^b_{t,0}) - \varphi(X^{b'}_{t,0}))\|_p
\]
\[
+ \|\mathbb{E}(\nabla^i (X^b_{t,0} - X^{b'}_{t,0}) \cdot \varphi(X^{b'}_{t,0}))\|_p =: I_1 + I_2.
\]

For \( I_1 \), by the boundedness of \( \mathbb{P} \) in \( L^p \) and Hölder’s inequality, we have
\[
I_1 \leq C\|\nabla^i X^{b'}_{t,0} - \varphi(X^b_{t,0})\|_p
\]
\[
\leq C\|\nabla X^{b'}_{t,0}\|_{L^p} \|\varphi(X^b_{t,0}) - \varphi(X^{b'}_{t,0})\|_{L^p} \|b - b'\|_{L^\infty_p(T_1, 0)}, \quad (7.2)
\]
where \( \frac{1}{p} + \frac{1}{p_2} = 1 \) with \( p_2 \in (1, \frac{2p}{p+1}) \). By (2.1) and (E) of Theorem 1.1 we have
\[
\mathbb{E}(\varphi(X^b_{t,0}) - \varphi(X^{b'}_{t,0}))^{p_2} \leq C\mathbb{E}\left(\mathbb{M}|\nabla\varphi(X^b_{t,0})|^p + \mathbb{M}|\nabla\varphi(X^{b'}_{t,0})|^{p_2}|X^b_{t,0} - X^{b'}_{t,0}|^{p_2}\right)
\]
\[
\leq C\left(\mathbb{E}(\mathbb{M}|\nabla\varphi(X^b_{t,0})|^p + \mathbb{M}|\nabla\varphi(X^{b'}_{t,0})|^{2p_2})\right)^{\frac{p_2}{2}} \leq C\left(\mathbb{E}(\mathbb{M}|\nabla\varphi(X^b_{t,0})|^p + \mathbb{M}|\nabla\varphi(X^{b'}_{t,0})|^{2p_2})\right)^{\frac{p_2}{2}} \|b - b'\|_{L^\infty_p(T_1, 0)}.
\]

Substituting this into (7.2), and by (B) of Theorem 1.1 and (7.1), we obtain
\[
I_1 \leq C \left(\int_{\mathbb{R}^d} \mathbb{E}(\mathbb{M}|\nabla\varphi(X^b_{t,0})|^p + \mathbb{M}|\nabla\varphi(X^{b'}_{t,0})|^{p_2}) dx\right)^{\frac{1}{p}} \|b - b'\|_{L^\infty_p(T_1, 0)} \leq C\|\nabla\varphi\|_{L^p} \|b - b'\|_{L^\infty_p(T_1, 0)}, \quad (7.3)
\]

As for \( I_2 \), letting \( p' = \frac{2p_2}{p+1} \), by (2.13) and Hölder’s inequality again, we have
\[
I_2 = \|\mathbb{P}\mathbb{E}(\nabla^i X^b_{t,0} \cdot \nabla\varphi(X^b_{t,0}) \cdot (X^b_{t,0} - X^{b'}_{t,0}))\|_p
\]
\[
\leq C\|X^b_{t,0} - X^{b'}_{t,0}\|_{L^p} \|\nabla\varphi(X^b_{t,0})\|_{L^p} \|\nabla X^b_{t,0}\|_{L^p} \|X^b_{t,0} - X^{b'}_{t,0}\|_{L^p} \|\nabla X^{b'}_{t,0} - \nabla X^b_{t,0}\|_{L^p} \|b - b'\|_{L^\infty_p(T_1, 0)} \|b - b'\|_{L^\infty_p(T_1, 0)}.
\]

which, together with (7.3), and letting \( T_1 \in [T_0, 0) \) be small enough, yields the estimate. \( \Box \)

We are now in a position to give

**Proof of Theorem 1.4** By Lemmas 7.2 and 7.3 the nonlinear operator \( \mathbb{T} \) is a contraction operator in the ball with radius \( U = 2C_0\|\varphi\|_{\mathcal{W}^1_p} \) of \( L^\infty_p(T_1, 0) \). Hence, by the fixed point theorem, there is a unique point \( u \in L^\infty_p(T_1, 0) \) such that for each \( t \in [T_1, 0) \),
\[
\mathbb{T}(u)_t = u_t.
\]

On the other hand, by (2.14), Hölder’s inequality and (1.4), (7.1), we also have
\[
\|\nabla\mathbb{T}(u)_t\|_p \leq C\mathbb{E}|\nabla X^b_{t,0}|^2 \|\nabla\varphi - \nabla\varphi(X^b_{t,0})\|_p < +\infty.
\]

The proof is finished. \( \Box \)
**Acknowledgements:**
This work is supported by NNSFs of China (Nos. 11271294, 11325105).

**References**

[1] Constantin, P. and Iyer, G.: A stochastic Lagrangian representation of the three-dimensional incompressible Navier-Stokes equations. Comm. Pure Appl. Math. LXI, 330-345 (2008).

[2] Crippa G. and De Lellis C.: Estimates and regularity results for the DiPerna-Lions flow. J. Reine Angew. Math. 616 (2008), 15-46.

[3] Davie A.M.: Uniqueness of solutions of stochastic differential equations. Int. Math. Res. Not. IMRN, no.24, 26pp, 2007.

[4] Elworthy K.D. and Li X.M.: Formulae for the derivatives of heat semigroups. J.Funct. Anal. 125, 252-286(1994).

[5] Fedrizzi E., Flandoli F.: Noise Prevents Singularities in Linear Transport Equations. J. Func. Anal., 2012.

[6] Fedrizzi E., Flandoli F.: Hölder Flow and Differentiability for SDEs with Nonregular Drift. Stochastic Analysis and Applications (2013).

[7] Flandoli F., Gubinelli M., Priola E.: Well-posedness of the transport equation by stochastic perturbation, Invent. Math. 180 (1) (2010), 1-53.

[8] Gyöngy I., Martinez, T.: On stochastic differential equations with locally unbounded drift. Czechoslovak Math. J. 51(126) (2001), no. 4, 763–783.

[9] Ikeda N., Watanabe, S.: Stochastic differential equations and diffusion processes, 2nd ed., North-Holland/Kodanska, Amsterdam/Tokyo, 1989.

[10] Krylov N.V.: Controlled diffusion processes. Translated from the Russian by A.B. Aries. Applications of Mathematics, 14. Springer-Verlag, New York-Berlin, 1980.

[11] Krylov N.V.: Lectures on elliptic and parabolic equations in Sobolev spaces. Graduate Studies in Math., Vol.96, AMS, 2008.

[12] Krylov N.V.: The heat equation in $L^p((0,T),L^p)$-spaces with weights. SIAM J. Math. Ana, Vol. 32, No. 5, pp. 1117-1141(2001).

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

[14] Ladyzenskaja O.A., Solonnikov V.A. and Urалceva N.N.: Linear and quasi-linear equations of parabolic type. Vol. 23, Trans. of Math. Mono. from Russian, 1968.

[15] Majda A.J. and Bertozzi A.L.: Vorticity and Incompressible Flow. Cambridge University Press, 2002.

[16] Menoukeu-Pamen O, Meyer-Brandis T., Nilssen T., Proske F. and Zhang T.: A variational approach to the construction and Malliavin differentiability of strong solutions of SDEs. Mathematische Annalen, Volume 357, Issue 2, pp 761-799(2013).

[17] Mohammed S.E.A., Nilsen T. and Proske F.: Sobolev Differentiable Stochastic Flows of SDE’s with Measurable Drift and Applications. to appear in Annals of Probability, http://arxiv.org/abs/1204.3867

[18] Nualart D.: The Malliavin calculus and related topics. Springer-Verlag, Berlin, 2007.

[19] Portenko N.I.: Generalized diffusion processes. Nauka, Moscow, 1982 In Russian; English translation: Amer. Math. Soc. Providence, Rhode Island, 1990.

[20] Protter P.: Stochastic integration and differential equations, 2nd ed., Springer-Verlag, Berlin, 2004.

[21] Rezakhanlou F.: Regular flows for diffusions with rough drifts. arXiv:1405.5856

[22] Stroock D. and Varadhan S.R.S.: Multidimensional diffusion processes. Springer-Verlag, Berlin, 1997.

[23] Veretennikov, A. Ju.: On the strong solutions of stochastic differential equations. Theory Probab. Appl., 24(1979), 354-366.

[24] Zhang X.: Strong solutions of SDEs with singular drift and Sobolev diffusion coefficients. Stoch. Proc. and Appl., 115/11 pp. 1805-1818(2005).

[25] Zhang X.: A stochastic representation for backward incompressible Navier-Stokes equations. Prob. Theory and Rela. Fields, Volume 148, Numbers 1-2, 305-332(2010).

[26] Zhang X.: Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients. Elec. Jour. of Prob., Vol. 16, no. 38,1096-1116(2011).

[27] Zhang X.: Stochastic Lagrangian particle approach to fractal Navier-Stokes equations. Comm. in Math Phys., Volume 311, Issue 1, Page 133-155(2012).

[28] Zvonkin, A.K.: A transformation of the phase space of a diffusion process that removes the drift. Mat. Sbornik, No.1. 93(135),129-149(1974).