Degenerate SDEs with Singular Drift and Applications to Heisenberg Groups *

Xing Huang[a)] **and Feng-Yu Wang**[a),b)]

[a)]Center for Applied Mathematics, Tianjin University, Tianjin 300072, China

[b)]Department of Mathematics, Swansea University, Singleton Park, SA2 8PP, United Kingdom

May 4, 2018

Abstract

By using the ultracontractivity of a reference diffusion semigroup, Krylov's estimate is established for a class of degenerate SDEs with singular drifts, which leads to existence and pathwise uniqueness by means of Zvonkin's transformation. The main result is applied to singular SDEs on generalized Heisenberg groups.

AMS subject Classification: 60H15, 35R60.
Keywords: Degenerate SDEs, Krylov estimate, Zvonkin transformation, Heisenberg group.

1 Introduction

Since 1974 when Zvonkin [29] proved the well-posedness of the Brownian motion with bounded drifts, his argument (known as Zvonkin's transformation) has been developed for more general models with singular drifts, see [19, 13, 26, 24] and references within for non-degenerate SDEs, and [4]-[7] and [11, 20] for non-degenerate semilinear SPDEs. In these references only Gaussian noise is considered, see also [17, 25] for extensions to the case with jump.

In recent years, Zvonkin's transformation has been applied in [2, 16, 22, 23, 27] to a class of degenerate SDEs/SPDEs with singular drifts. This type degenerate stochastic systems are called stochastic Hamiltonian systems in probability theory. Consider, for instance, the following SDE for (X_t, Y_t) on $\mathbb{R}^{2d}$ $(d \geqslant 1)$:

$$(1.1)\qquad \begin{cases} \mathrm{d}X_t = Y_t \mathrm{d}t, \\ \mathrm{d}Y_t = b_t(X_t, Y_t)\mathrm{d}t + \sigma_t(X_t, Y_t)\mathrm{d}W_t, \end{cases}$$

where W_t is the d-dimensional Brownian motion, and

$$b : [0, \infty) \times \mathbb{R}^{2d} \to \mathbb{R}^d, \quad \sigma : [0, \infty) \times \mathbb{R}^{2d} \to \mathbb{R}^d \otimes \mathbb{R}^d$$

*Supported in part by NNSFC (11771326, 11431014).

are measurable. According to [27, Theorem 1.1], if there exists a constant $K > 1$ such that

$$K^{-1}|v| \leqslant |\sigma_t v| \leqslant K|v|, \quad t \geqslant 0, v \in \mathbb{R}^d,$$

and for some constant $p > 2(1 + 2d)$,

$$\sup_{t\geqslant 0} \|\nabla\sigma_t\|_{L^p(\mathbb{R}^{2d})} + \int_0^\infty \|(1 - \Delta_x)^{\frac{1}{3}} b_t\|_{L^p(\mathbb{R}^{2d})}^p \mathrm{d}t < \infty,$$

then the SDE (1.1) has a unique strong solution for any initial points. By a standard truncation argument, the existence and pathwise uniqueness up to the life time hold under the corresponding local conditions.

In this paper, we aim to extend this result to general degenerate SDEs, in particular, for singular diffusions on generalized Heisenberg groups. As typical models of hypoelliptic systems, smooth SDEs on Heisenberg groups have been intensively investigated, see for instance [1, 8, 9, 10, 14, 21] and references within for the study of functional inequalities, gradient estimates, Harnack inequalities, and Riesz transforms. We will use these results to establish Krylov's estimates for singular SDEs and to prove the existence and uniqueness of strong solutions using Zvonkin's transformation.

In Section 2, we present a general result (see Theorem 2.1(3)) for the existence and uniqueness of degenerate SDEs with singular drifts, and then apply this result in Section 3 to singular diffusions on generalized Heisenberg groups.

2 General results

For fixed constant $T > 0$, consider the following SDE on $\mathbb{R}^N$:

$$\mathrm{d}X_t = Z_t(X_t)\mathrm{d}t + \sigma_t(X_t)\mathrm{d}B_t, \quad t \in [0, T], \tag{2.1}$$

where B_t is the m-dimensional Brownian motion with respect to a complete filtered probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\in[0,T]}, \mathbb{P})$, and

$$Z : [0, T] \times \mathbb{R}^N \to \mathbb{R}^N, \quad \sigma : [0, T] \times \mathbb{R}^N \to \mathbb{R}^N \otimes \mathbb{R}^m$$

are measurable and locally bounded. We are in particular interested in the case that $m < N$ such that this SDE is degenerate.

Throughout the paper, we assume that for any $x \in \mathbb{R}^N$ and $s \in [0, T)$, this SDE has a unique solution $(X_{s,t}^x)_{t\in[s,T]}$ with $X_{s,s}^x = x$; i.e. it is a continuous adapted process such that

$$X_{s,t}^x = x + \int_s^t Z_r(X_{s,r}^x)\mathrm{d}r + \int_s^t \sigma_r(X_{s,r}^x)\mathrm{d}B_r, \quad t \in [s, T].$$

Let $(P_{s,t})_{0\leqslant s\leqslant t\leqslant T}$ be the associated Markov semigroup. We have

$$P_{s,t}f(x) = \mathbb{E}f(X_{s,t}^x), \quad f \in \mathscr{B}_b(\mathbb{R}^N) \cup \mathscr{B}^+(\mathbb{R}^N), \quad x \in \mathbb{R}^N, 0 \leqslant s \leqslant t \leqslant T,$$

where $\mathscr{B}_b$ (resp. $\mathscr{B}^+$) denotes the set of bounded (resp. nonnegative) measurable functions. The infinitesimal generator of the solution is

$$\mathscr{L}_s := \frac{1}{2}\sum_{i,j=1}^N (\sigma_s\sigma_s^*)_{ij}\,\partial_i\partial_j + \sum_{i=1}^N (Z_s)_i\,\partial_i.$$

Now, let $\mathbf{b} : [0,T]\times\mathbb{R}^N \to \mathbb{R}^m$ be measurable. We intend to find reasonable conditions on $\mathbf{b}$ such that the following perturbed SDE is well-posed:

$$\mathrm{d}\tilde{X}_t = \{Z_t + \sigma_t\mathbf{b}_t\}(\tilde{X}_t)\mathrm{d}t + \sigma_t(\tilde{X}_t)\mathrm{d}B_t, \quad t\in[0,T]. \tag{2.2}$$

To state the main result, we introduce two spaces $L^q_{p,loc}([0,T]\times\mathbb{R}^N)$ and $W^q_{p,loc}([0,T]\times\mathbb{R}^N)$, for $p,q\geqslant 1$. A real measurable function f defined on $[0,T]\times\mathbb{R}^N$ is said in $L^q_p([0,T]\times\mathbb{R}^N)$, if

$$\|f\|_{L^q_p} := \left(\int_0^T \|f_t\|^q_{L^p(\mathbb{R}^N)}\mathrm{d}t\right)^{\frac{1}{q}} < \infty.$$

Next, if $f\in L^q_p([0,T]\times\mathbb{R}^N)$ such that ∇f_s exists in weak sense for a.e. $s\in[0,T]$ and $|\nabla f|\in L^q_p([0,T]\times\mathbb{R}^N)$, where ∇ is the gradient operator on $\mathbb{R}^N$, we write $f\in W^q_p([0,T]\times\mathbb{R}^N)$. Consequently, we write

$$f\in L^q_{p,loc}([0,T]\times\mathbb{R}^N)$$

if $hf\in L^q_p([0,T]\times\mathbb{R}^N)$ for any $h\in C_0^\infty(\mathbb{R}^N)$, and

$$f\in W^q_{p,loc}([0,T]\times\mathbb{R}^N)$$

if $hf\in W^q_p([0,T]\times\mathbb{R}^N)$ for any $h\in C_0^\infty(\mathbb{R}^N)$. Moreover, a vector-valued function is in one of these spaces if so are its components.

Finally, a real function f on $\mathbb{R}^N$ is called σ_t-differentiable, if for any $v\in\mathbb{R}^m$ it is differentiable along the direction $\sigma_t v$; i.e.

$$\nabla_{\sigma_t v}f(x) := \frac{\mathrm{d}}{\mathrm{d}r}f(x+r\sigma_t(x)v)\Big|_{r=0}$$

exists for any $x\in\mathbb{R}^N$. A real function f on $[0,T]\times\mathbb{R}^N$ is called σ-differentiable if f_t is σ_t-differentiable for every $t\in[0,T]$. In this case, $\nabla_\sigma f : [0,T]\times\mathbb{R}^N\to\mathbb{R}^m$ is defined by

$$\langle(\nabla_\sigma f)_t(x), v\rangle := \nabla_{\sigma_t v}f_t(x), \quad v\in\mathbb{R}^m, t\in[0,T], x\in\mathbb{R}^N.$$

When ∇f exists, we have $\nabla_\sigma f = \sigma^*\nabla f$. Let

$$\mathbb{B} = \big\{f\in C_b([0,T]\times\mathbb{R}^N) : \nabla_\sigma f\in C_b([0,T]\times\mathbb{R}^N;\mathbb{R}^m)\big\}.$$

Then $\mathbb{B}$ is a Banach space with

$$\|f\|_{\mathbb{B}} := \|f\|_\infty + \|\nabla_\sigma f\|_\infty,$$

where $\|\cdot\|_\infty$ is the uniform norm. An $\mathbb{R}^m$-valued function g is said in the space $\mathbb{B}^m$, if its components belong to $\mathbb{B}$. Let

$$\|g\|_{\mathbb{B}^m} = \|g\|_\infty + \|\nabla_\sigma g\|_\infty, \quad g\in\mathbb{B}^m.$$

We make the following assumptions.

(A_1) $\sigma_t(x)$ is locally bounded in $(t,x) \in [0,T] \times \mathbb{R}^N$, and for any $R > 0$, there exists a constant $c > 0$ such that

$$|\sigma_t(x)\mathbf{b}_t(x)| \geqslant c|\mathbf{b}_t(x)|, \quad t \in [0,T], |x| \leqslant R.$$

(A_2) For any $f \in C_0^\infty([0,T] \times \mathbb{R}^N)$ and $\lambda \geqslant 0$, the function

$$(Q_\lambda f)_s(x) := \int_s^T \mathrm{e}^{-\lambda(t-s)} P_{s,t} f_t(x) \mathrm{d}t, \quad s \in [0,T], x \in \mathbb{R}^N \tag{2.3}$$

satisfies that $\partial_s (Q_\lambda f)_s, \mathscr{L}_s (Q_\lambda f)_s$ exist and are locally bounded on $[0,T] \times \mathbb{R}^N$ with

$$(\partial_s + \mathscr{L}_s - \lambda)(Q_\lambda f)_s + f_s = 0. \tag{2.4}$$

Assumption (A_1) holds provided σ is continuous and has rank m. Assumption (A_2) holds if $Z_t(x)$ and $\sigma_t(x)$ are regular enough in x, for instance, C^2-smooth in x uniformly in $t \in [0,T]$.

For any $p, q \geqslant 1$, let $\|\cdot\|_{p \to q}$ denote the operator norm from $L^p(\mathbb{R}^N; \mathrm{d}x)$ to $L^q(\mathbb{R}^N; \mathrm{d}x)$. To introduce the integrability conditions for the drift $\mathbf{b}$, we need the following two classes of pairs $(p,q) \in (1,\infty]^2$:

$$\begin{aligned}
\mathscr{K}_1 := \big\{(p,q) \in (1,\infty]^2 : &\text{ there exists } \gamma \in L^{\frac{q}{q-1}}([0,T]) \text{ such that} \\
&\|P_{s,t}\|_{p \to \infty} \leqslant \gamma(t-s), \quad 0 \leqslant s < t \leqslant T\big\}, \\
\mathscr{K}_2 := \big\{(p,q) \in (1,\infty]^2 : &\text{ there exists } \gamma \in L^{\frac{q}{q-1}}([0,T]) \text{ such that} \\
&\|\sigma_t^* \nabla P_{s,t}\|_{p \to \infty} \leqslant \gamma(t-s), \quad 0 \leqslant s < t \leqslant T\big\}.
\end{aligned}$$

Obviously, both $\mathscr{K}_1$ and $\mathscr{K}_2$ are increasing sets; that is, if $(p,q) \in \mathscr{K}_i$ then $(p',q') \in \mathscr{K}_i$ for $p' \geqslant p$ and $q' \geqslant q, i = 1,2$. We will also use the following class

$$2\mathscr{K}_1 := \{(2p, 2q) : (p,q) \in \mathscr{K}_1\}.$$

Clearly, $2\mathscr{K}_1 \subset \mathscr{K}_1$. When $Z = 0$ and $\sigma\sigma^* = I_{N \times N}$, the $N \times N$-identity matrix, we have $P_{s,t} = P_{t-s}$ for the standard heat semigroup P_t, so that

$$\|P_{s,t}\|_{p \to \infty} \leqslant C(t-s)^{-d/(2p)}, \quad \|\sigma^* \nabla P_{s,t}\|_{p \to \infty} \leqslant C(t-s)^{-\frac{1}{2} - d/(2p)}, \quad 0 \leqslant s < t \leqslant T \tag{2.5}$$

for some constant $C > 0$, then

$$\mathscr{K}_1 \supset \Big\{(p,q) \in (1,\infty]^2 : \ \frac{1}{q} + \frac{d}{2p} < 1\Big\}, \quad \mathscr{K}_2 \supset \Big\{(p,q) \in (1,\infty]^2 : \ \frac{2}{q} + \frac{d}{p} < 1\Big\}. \tag{2.6}$$

These formulas also hold for elliptic diffusions satisfying (2.5). But for degenerate diffusions the dimension d in this display will be enlarged, see for instance the proof of Theorem 3.1 below.

We are now ready to state the main result in this section. In particular, the first assertion implies that if $\mathscr{K}_1 \cap \mathscr{K}_2 \neq \emptyset$, then for any $\mathbf{b} \in C_0^\infty([0,T] \times \mathbb{R}^N)$ and large enough $\lambda > 0$, the equation

$$\mathbf{u}_s = \int_s^T \mathrm{e}^{-\lambda(t-s)} P_{s,t} \big\{\nabla_{\sigma_t \mathbf{b}_t} \mathbf{u}_t + \sigma_t \mathbf{b}_t\big\} \mathrm{d}t, \quad s \in [0,T] \tag{2.7}$$

has a unique solution $\mathbf{u} =: \Xi_\lambda \mathbf{b} \in \mathbb{B}^m$. We write $f \in C^{1,2}([0,T] \times \mathbb{R}^N)$ if f is a function on $[0,T] \times \mathbb{R}^N$ such that $\partial_t f_t(x)$ and $\nabla^2 f_t(x)$ exist and continuous in (t,x).

Theorem 2.1. *Assume (A_1) and (A_2). Then the the following assertions hold.*

(1) *For any $(p,q)\in\mathscr{K}_1\cap\mathscr{K}_2$ and $L>0$, there exists $\tilde\lambda>0$ such that for any $\lambda\geqslant\tilde\lambda$, and $\mathbf{b}:[0,T]\times\mathbb{R}^N\to\mathbb{R}^m$ with $|\sigma\mathbf{b}|+|\mathbf{b}|\in L_p^q([0,T]\times\mathbb{R}^N)$ and $\||\sigma\mathbf{b}|\|_{L_p^q}\vee\||\mathbf{b}|\|_{L_p^q}\leqslant L$, the equation (2.7) has a unique solution $\mathbf{u}=:\Xi_\lambda\mathbf{b}$ in $\mathbb{B}^m$. Moreover, there exists a decreasing function $\psi:[\tilde\lambda,\infty)\to(0,\infty)$ with $\psi(\infty):=\lim_{\lambda\to\infty}\psi(\lambda)=0$ such that for any $\mathbf{b},\tilde{\mathbf{b}}:[0,T]\times\mathbb{R}^N\to\mathbb{R}^m$ with $\||\sigma\mathbf{b}|\|_{L_p^q},\||\sigma\tilde{\mathbf{b}}|\|_{L_p^q},\||\mathbf{b}|\|_{L_p^q},\||\tilde{\mathbf{b}}|\|_{L_p^q}\leqslant L$,*

$$\|\Xi_\lambda\mathbf{b}-\Xi_\lambda\tilde{\mathbf{b}}\|_{\mathbb{B}^m}\leqslant\psi(\lambda)\||\mathbf{b}-\tilde{\mathbf{b}}|+|\sigma(\mathbf{b}-\tilde{\mathbf{b}})|\|_{L_p^q},\quad\lambda\geqslant\tilde\lambda.\tag{2.8}$$

(2) *If $|\mathbf{b}|\in L_p^q([0,T]\times\mathbb{R}^N)$ for some $p,q\geqslant1$ with $(p,q)\in2\mathscr{K}_1$, then for any $x\in\mathbb{R}^N$ the SDE (2.2) has a weak solution $(\tilde X_t)_{t\in[0,T]}$ starting at x with respect to a probability $\mathbb{Q}$ such that $\mathbb{E}_{\mathbb{Q}}\mathrm{e}^{\lambda\int_0^T|\mathbf{b}_t(\tilde X_t)|^2\mathrm{d}t}<\infty$ holds for all $\lambda>0$.*

(3) *Assume further that*

   (i) for large enough $\lambda>0$, $\Xi_\lambda f\in C^{1,2}([0,T]\times\mathbb{R}^N)$ holds for any $f\in C_0^\infty([0,T]\times\mathbb{R}^N)$;

   (ii) there exists $(p,q)\in2\mathscr{K}_1\cap\mathscr{K}_2$ such that $|\mathbf{b}|+|\nabla\sigma|\in L^q_{p,loc}([0,T]\times\mathbb{R}^N)$, $Z\in W^q_{p,loc}([0,T]\times\mathbb{R}^N)$, and for large enough $\lambda>0$ there hold

$$\nabla_\sigma\Xi_\lambda(h\mathbf{b})\in W^q_{p,loc}([0,T]\times\mathbb{R}^N),\quad h\in C_0^\infty(\mathbb{R}^N),\tag{2.9}$$

$$\lim_{\lambda\to\infty}\|h\nabla\Xi_\lambda(h\mathbf{b})\|_\infty=0,\quad h\in C_0^\infty(\mathbb{R}^N).\tag{2.10}$$

Then for any $x\in\mathbb{R}^N$ the SDE (2.2) has a unique strong solution $\tilde X_t$ starting at x up to the life time $\zeta:=\lim_{n\to\infty}T\wedge\zeta_n:=\inf\{t\in[0,T]:|\tilde X_t|\geqslant n\}$ with $\int_0^{T\wedge\zeta_n}|\mathbf{b}|_t^2(\tilde X_t)\mathrm{d}t<\infty$ for any $n\geqslant1$.

By (2.7) and the definition of Q_λ in (A_2), we have

$$\Xi_\lambda\mathbf{b}=Q_\lambda\big\{\nabla_{\sigma\mathbf{b}}\Xi_\lambda\mathbf{b}+\sigma\mathbf{b}\big\}.\tag{2.11}$$

If (2.6) holds, Theorem 2.1 (3) ensures the strong well-posedness when $|\mathbf{b}|\in L_p^q$ for some $p,q\geqslant1$ with $\frac{d}{p}+\frac{2}{q}<1$, which coincides known optimal result in the elliptic setting. In the elliptic case there exists much weaker sufficient conditions for the well-posedness, for instance, in a recent paper by Xicheng Zhang and Guohua Zhao [28], the drift is allowed to be distributions (not necessarily functionals).

To prove Theorem 2.1, we first investigate the Krylov estimate and the weak existence for (2.2).

2.1 Krylov's estimate and weak existence

Theorem 2.2. *Assume (A_1) and (A_2). Let $p,q\geqslant1$ such that $(p,q)\in\mathscr{K}_1$ and $|\mathbf{b}|^2\in L_p^q([0,T]\times\mathbb{R}^N)$.*

(1) *For any* $(p',q') \in \mathscr{K}_1$ *there exists a constant* $\kappa > 0$ *such that for any* $s \in [0,T)$ *and solution* $(\tilde{X}_{s,t})_{t\in[s,T]}$ *of* (2.2) *from time* s *with* $\int_s^T |\mathbf{b}_t(\tilde{X}_{s,t})|^2 \mathrm{d}t < \infty$,

$$\mathbb{E}\bigg(\int_s^T |f_t(\tilde{X}_{s,t})|\mathrm{d}t \bigg| \mathscr{F}_s \bigg) \leqslant \kappa \|f\|_{L_{p'}^{q'}}, \quad f \in L_{p'}^{q'}([0,T]\times\mathbb{R}^N), s \in [0,T]. \tag{2.12}$$

Consequently, for any $f \in L_{p'}^{q'}([0,T]\times\mathbb{R}^N)$ *with* $(p',q') \in \mathscr{K}_1$ *and any* $\lambda > 0$, *there exists a constant* $c(f,\lambda) \in (0,\infty)$ *such that*

$$\mathbb{E}\Big(\mathrm{e}^{\lambda\int_s^T |f_t(\tilde{X}_{s,t})|\mathrm{d}t} \Big| \mathscr{F}_s \Big) \leqslant c(f,\lambda), \quad s \in [0,T]. \tag{2.13}$$

(2) *The assertion in Theorem 2.1(2) holds.*

To prove this result, we need the following lemma.

Lemma 2.3. *Assume* (A_1) *and* (A_2).

(1) *For any* $(p,q) \in \mathscr{K}_1$, *there exists a decreasing function* $\psi : [0,\infty) \to (0,\infty)$ *with* $\psi(\infty) := \lim_{\lambda\uparrow\infty}\psi(\lambda) = 0$ *such that for any* $\lambda \geqslant 0$, $(Q_\lambda, C_0^\infty([0,T]\times\mathbb{R}^N))$ *in* (A_2) *extends uniquely to a bounded linear operator* $Q_\lambda : L_p^q([0,T]\times\mathbb{R}^N) \to L^\infty([0,T]\times\mathbb{R}^N)$ *with*

$$\|Q_\lambda f\|_\infty \leqslant \psi(\lambda)\|f\|_{L_p^q}, \quad f \in L_p^q([0,T]\times\mathbb{R}^N), \ \lambda \geqslant 0.$$

(2) *For any* $(p,q) \in \mathscr{K}_1 \cap \mathscr{K}_2$, $(Q_\lambda, C_0^\infty([0,T]\times\mathbb{R}^N))$ *extends to a unique bounded linear operator* $Q_\lambda : L_p^q([0,T]\times\mathbb{R}^N) \to \mathbb{B}$ *such that*

$$\|Q_\lambda f\|_{\mathbb{B}} := \|Q_\lambda f\|_\infty + \|\nabla_\sigma Q_\lambda f\|_\infty \leqslant \psi(\lambda)\|f\|_{L_p^q}, \quad \lambda \geqslant 0, f \in L_p^q([0,T]\times\mathbb{R}^N)$$

holds for some decreasing $\psi : [0,\infty) \to (0,\infty)$ *with* $\psi(\infty) = 0$.

Proof. We only prove (1) since that of (2) is completely similar. For any $(p,q) \in \mathscr{K}_1$, there exists $\gamma \in L^{\frac{q}{q-1}}([0,T])$ such that for any $\lambda \geqslant 0$ and $f \in C_0^\infty([0,T]\times\mathbb{R}^N)$,

$$\begin{aligned}\|Q_\lambda f\|_\infty &\leqslant \sup_{s\in[0,T]} \int_s^T \mathrm{e}^{-\lambda(t-s)}\gamma(t-s)\|f_t\|_{L^p(\mathbb{R}^N)}\mathrm{d}t \\ &\leqslant \sup_{s\in[0,T]} \bigg(\int_s^T \big|\mathrm{e}^{-\lambda(t-s)}\gamma(t-s)\big|^{\frac{q}{q-1}}\mathrm{d}t \bigg)^{\frac{q-1}{q}} \|f\|_{L_p^q}, \quad \lambda \geqslant 0.\end{aligned}$$

So, assertion (1) holds with

$$\psi(\lambda) := \bigg(\int_0^T \big|\mathrm{e}^{-\lambda t}\gamma(t)\big|^{\frac{q}{q-1}}\mathrm{d}t \bigg)^{\frac{q-1}{q}}.$$

□

We will also need the following lemma which reduces the desired Krylov's estimate to $f \in C_0^\infty([0,T]\times\mathbb{R}^N)$. It can be proved using a standard approximation argument.

Lemma 2.4. *Let $s \in [0,T]$ and $p,q \geqslant 1$ For any two stopping times $\tau_1 \leqslant \tau_2$, measurable process $(\xi_t)_{t\in[s,T]}$ on $\mathbb{R}^N$, and random variable $\eta \geqslant 0$, if the inequality*

$$\mathbb{E}\bigg(\int_{s\wedge\tau_1}^{T\wedge\tau_2} f_t(\xi_t)\mathrm{d}t\bigg|\mathscr{F}_s\bigg) \leqslant \|f\|_{L_p^q}\mathbb{E}(\eta|\mathscr{F}_s) \tag{2.14}$$

holds for all nonnegative $f \in C_0^\infty([0,T]\times\mathbb{R}^N)$, it holds for all nonnegative $f \in L_p^q([0,T]\times\mathbb{R}^N)$.

Proof of Theorem 2.2. (1) According to the Khasminskii estimate, see [26, Lemma 5.3], (2.13) follows from (2.12). For simplicity, we only prove for $s = 0$. To prove (2.12), we first consider $\mathbf{b} = 0$. Let $(X_t)_{t\in[0,T]}$ solve (2.1) and let

$$\tau_n = \inf\{t \in [0,T] : \ |X_t| \geqslant n\}, \quad n \geqslant 1.$$

For $0 \leqslant f \in C_0^\infty([0,T]\times\mathbb{R}^N)$, take $u^{(\lambda)} = Q_\lambda f$ for $\lambda \geqslant 0$. By (A_2) and Itô's formula, we obtain

$$\begin{aligned}
0 &\leqslant \mathbb{E}\big(u^{(\lambda)}_{T\wedge\tau_n}(X_{T\wedge\tau_n})\big|\mathscr{F}_s\big)\\
&= u_0^{(\lambda)}(X_{s\wedge\tau_n}) + \mathbb{E}\bigg(\int_0^{T\wedge\tau_n}(\partial_t+\mathscr{L}_t)u_t^{(\lambda)}(X_t)\mathrm{d}t\bigg|\mathscr{F}_s\bigg)\\
&\leqslant (1+\lambda T)\|u^{(\lambda)}\|_\infty - \mathbb{E}\bigg(\int_0^{T\wedge\tau_n} -f_t(X_t)\mathrm{d}t\bigg|\mathscr{F}_s\bigg).
\end{aligned}$$

Noting that $u^{(\lambda)} = Q_\lambda f$, combining this with Lemma 2.3 and Lemma 2.4, for any $(p',q') \in \mathscr{K}_1$ there exists decreasing $\psi : [0,\infty) \to (0,\infty)$ with $\psi(\infty) = 0$ such that

$$\mathbb{E}\bigg(\int_0^{T\wedge\tau_n} f_t(X_t)\mathrm{d}t\bigg|\mathscr{F}_0\bigg) \leqslant \psi(\lambda)\|f\|_{L_{p'}^{q'}}(1+\lambda T), \quad 0 \leqslant f \in L_{p'}^{q'}([0,T]\times\mathbb{R}^N). \tag{2.15}$$

Letting $n \to \infty$ we prove (2.12) for $\mathbf{b} = 0$. In general, let $(\tilde{X}_t)_{t\in[0,T]}$ solve (2.2) with $\int_0^T |\mathbf{b}_t(\tilde{X}_t)|^2\mathrm{d}t < \infty$. Define

$$T_n = \inf\bigg\{t \in [0,T] : \int_0^t |\mathbf{b}_r(\tilde{X}_r)|^2\mathrm{d}r \geqslant n\bigg\}, \quad n \geqslant 0,$$

where $\inf\emptyset := \infty$ by convention. Let

$$\begin{aligned}
R_n &= \exp\bigg[-\int_0^{T\wedge T_n}\big\langle \mathbf{b}_r(\tilde{X}_r), \mathrm{d}B_r\big\rangle - \frac{1}{2}\int_0^{T\wedge T_n}|\mathbf{b}_r(\tilde{X}_r)|^2\mathrm{d}r\bigg],\\
\tilde{B}_t &= B_t + \int_0^{t\wedge T_n}\mathbf{b}_r(\tilde{X}_r)\mathrm{d}r, \quad t \in [s,T].
\end{aligned}$$

Then under the probability $R_n\mathbb{P}$, $(\tilde{X}_r, \tilde{B}_r)_{r\in[0,T\wedge T_n]}$ is a weak solution to the SDE (2.2) for $\mathbf{b} = 0$. So, by the assertion for $\mathbf{b} = 0$, there exists a constat $c > 0$ such that

$$\mathbb{E}\bigg[R_n\bigg(\int_0^{T\wedge T_n} f(r,\tilde{X}_r)\mathrm{d}r\bigg)^2\bigg|\mathscr{F}_0\bigg] \leqslant c\|f\|^2_{L_{p'}^{q'}(T)}. \tag{2.16}$$

By Hölder inequality and (2.12) for $\mathbf{b}=0$, there exists a constant $c'>0$ such that

$$\begin{aligned}\mathbb{E}\big(R_n^{-1}\big|\mathscr{F}_0\big)&=\mathbb{E}\Big(R_n\mathrm{e}^{2\int_0^{T\wedge T_n}\langle\mathbf{b}_r(\tilde X_r),\mathrm{d}B_r\rangle+\int_0^{T\wedge T_n}|\mathbf{b}_r(\tilde X_r)|^2\mathrm{d}r}\Big|\mathscr{F}_0\Big)\\&\leqslant\sqrt{\mathbb{E}\Big(R_n\mathrm{e}^{4\int_0^{T\wedge T_n}\langle\mathbf{b}_r(\tilde X_r),\mathrm{d}\tilde B_r\rangle-8\int_0^{T\wedge T_n}|\mathbf{b}_r(\tilde X_r)|^2\mathrm{d}r}\Big|\mathscr{F}_0\Big)}\\&\quad\times\sqrt{\mathbb{E}\Big(R_n\mathrm{e}^{6\int_0^{T\wedge T_n}|\mathbf{b}_r(\tilde X_r)|^2\mathrm{d}r}\Big|\mathscr{F}_0\Big)}\\&=\sqrt{\mathbb{E}\Big(R_n\mathrm{e}^{6\int_0^{T\wedge T_n}|\mathbf{b}_r(\tilde X_r)|^2\mathrm{d}r}\Big|\mathscr{F}_0\Big)}\leqslant c',\end{aligned}$$

where the last step follows from $|\mathbf{b}|^2\in L^{q/2}_{p/2}([0,T]\times\mathbb{R}^N)$ for some $(p,q)\in 2\mathscr{K}_1$ and (2.13) for $\mathbf{b}=0$. Then there exists a constant $C>0$ such that

$$\begin{aligned}&\Big\{\mathbb{E}\Big(\int_0^{T\wedge T_n}f_r(\tilde X_r)\mathrm{d}r\Big|\mathscr{F}_0\Big)\Big\}^2\\&\leqslant\mathbb{E}\Big[R_n\Big(\int_0^{T\wedge T_n}f_r(\tilde X_r)\mathrm{d}r\Big)^2\Big|\mathscr{F}_0\Big]\cdot\mathbb{E}\big(R_n^{-1}\big|\mathscr{F}_0\big)\leqslant C\|f\|^2_{L^{q'}_{p'}(T)}.\end{aligned}$$

By letting $n\to\infty$ we prove (2.12).

(2) Assume that $|\mathbf{b}|\in L^q_p([0,T]\times\mathbb{R}^N)$ for some $(p,q)\in 2\mathscr{K}_1$. Let $(X_t)_{t\in[0,T]}$ solve (2.1) with $X_0=\tilde X_0$. We intend to show that it is a weak solution of (2.2) under a weighted probability $\mathbb{Q}:=R\mathbb{P}$, where $R\geqslant 0$ is a probability density, and thus finish the proof. Since by (2.12) for $\mathbf{b}=0$ we have

$$\mathbb{E}\int_0^T|\mathbf{b}_t(X_t)|^2\mathrm{d}t\leqslant\kappa\||\mathbf{b}|^2\|_{L^{q/2}_{p/2}}=\kappa\|\mathbf{b}\|^2_{L^q_p}<\infty \tag{2.17}$$

for some constant $\kappa>0$. Then

$$T_n:=\inf\Big\{s\in[0,T]:\int_0^s|\mathbf{b}_t|^2(X_t)\mathrm{d}t\geqslant n\Big\}\uparrow\infty\text{ as }n\uparrow\infty,$$

where we set $\inf\emptyset=\infty$ by convention. For any $n\geqslant 1$, let

$$R_n=\exp\Big[\int_0^{T\wedge T_n}\langle\mathbf{b}_t(X_t),\mathrm{d}B_t\rangle-\frac12\int_0^{T\wedge T_n}|\mathbf{b}_t(X_t)|^2\mathrm{d}t\Big].$$

By Girsanov's theorem, $\{R_n\}_{n\geqslant1}$ is a martingale and $\mathbb{Q}_n:=R_n\mathbb{P}$ is a probability measure such that

$$\tilde B_t:=B_t-\int_0^{t\wedge T_n}\mathbf{b}_s(X_s)\mathrm{d}s,\quad t\in[0,T]$$

is an m-dimensional Brownian motion. Rewriting (2.1) by

$$\mathrm{d}X_t=(Z_t+\sigma_t\mathbf{b}_t)(X_t)\mathrm{d}t+\sigma_t(X_t)\mathrm{d}\tilde B_t,\quad t\in[0,T\wedge T_n], \tag{2.18}$$

we see that $(X_t,\tilde B_t)_{t\in[0,T\wedge T_n]}$ is a weak solution of (2.2) up to time $T\wedge T_n$. To extend this solution to time T, it suffices to show that the martingale $(R_n)_{n\geqslant1}$ is uniformly integrable, so that $R := \lim_{n\to\infty} R_n$ is a probability density, and $(X_t, \tilde{B}_t)_{t\in[0,T]}$ is a weak solution of (2.2) under the probability $\mathbb{Q} := R\mathbb{P}$. Therefore, it remains to prove

$$\sup_{n\geqslant 1} \mathbb{E}[R_n \log R_n] < \infty, \quad n \geqslant 1. \tag{2.19}$$

Since $(\tilde{B}_t)_{t\in[0,T]}$ is an m-dimensional Brownian motion under probability $\mathbb{Q}_n := R_n\mathbb{P}$, by (2.18) and Theorem 2.1(1), for any $(p', q') \in \mathscr{K}_1$ there exists a constant $\kappa > 0$ such that

$$\mathbb{E}_{\mathbb{Q}_n} \int_0^{T\wedge T_n} f_t(X_t)\mathrm{d}t \leqslant \kappa \|f\|_{L^{q'}_{p'}}, \quad 0 \leqslant f \in L^{q'}_{p'}([0,T]\times\mathbb{R}^N), n \geqslant 1.$$

Applying this estimate to $f = |\mathbf{b}|^2$, we arrive at

$$2\mathbb{E}[R_n \log R_n] = \mathbb{E}_{\mathbb{Q}_n} \int_0^{T\wedge T_n} |\mathbf{b}_t|^2(X_t)\mathrm{d}t \leqslant \kappa \||\mathbf{b}|^2\|_{L^{q/2}_{p/2}}, \quad n \geqslant 1.$$

This implies (2.19) and $\mathbb{E}_{\mathbb{Q}} \int_0^T |\mathbf{b}_t|^2(X_t)\mathrm{d}t < \infty$. Then the proof is finished since by (2.13) for $f = |\mathbf{b}|^2$ in Theorem 2.2 (1), $\||\mathbf{b}|^2\|_{L^{q/2}_{p/2}} < \infty$ implies $\mathbb{E}_{\mathbb{Q}} \mathrm{e}^{\lambda \int_0^T |\mathbf{b}_t|^2(X_t)\mathrm{d}t} < \infty$ for all $\lambda > 0$. $\square$

2.2 Proof of Theorem 2.1

Since Theorem 2.1(2) follows from Theorem 2.2(2), we only prove Theorem 2.1(1),(3).

Proof of Theorem 2.1(1). Let $\||\mathbf{b}|\|_{L^q_p} \leqslant L$ for some $(p,q) \in \mathscr{K}_1 \cap \mathscr{K}_2$. We first prove the existence and uniqueness of $\Xi_\lambda \mathbf{b}$ for large enough $\lambda > 0$. Consider the operator $\mathscr{K}_\lambda$ on $\mathbb{B}^m$:

$$\mathscr{K}_\lambda \mathbf{u} := Q_\lambda \big\{ \nabla_{\sigma\mathbf{b}} \mathbf{u} + \sigma\mathbf{b} \big\}, \quad \mathbf{u} \in \mathbb{B}^m.$$

By the fixed-point theorem, it suffices to show that $\mathscr{K}_\lambda$ is contractive in $\mathbb{B}^m$ for large enough $\lambda > 0$.

By Lemma 2.3, for any $\mathbf{u}, \tilde{\mathbf{u}} \in \mathbb{B}^m$ we have

$$\begin{aligned}
&\|\mathscr{K}_\lambda \mathbf{u} - \mathscr{K}_\lambda \tilde{\mathbf{u}}\|_{\mathbb{B}^m} = \|Q_\lambda \nabla_{\sigma\mathbf{b}}(\mathbf{u} - \tilde{\mathbf{u}})\|_{\mathbb{B}^m} \\
&\leqslant \psi(\lambda) \|\nabla_\sigma(\mathbf{u} - \tilde{\mathbf{u}})\|_\infty \||\mathbf{b}|\|_{L^q_p} \leqslant \psi(\lambda) L \|\mathbf{u} - \tilde{\mathbf{u}}\|_{\mathbb{B}^m}.
\end{aligned}$$

Since $\psi(\lambda) \to 0$ as $\lambda \to \infty$, there exists $\lambda_L > 0$ such that $\psi(\lambda_L) \leqslant \frac{1}{2L}$. So, when $\lambda \geqslant \lambda_L$, the map $\mathscr{K}_\lambda$ is contractive in $\mathbb{B}^m$. By the fixed point theorem, there exists a unique $\mathbf{u} \in \mathbb{B}^m$ such that $\mathbf{u} = \mathscr{K}_\lambda \mathbf{u}$, which is denoted by $\Xi_\lambda \mathbf{b}$.

Next, let $\||\mathbf{b}|\|_{L^q_p}, \||\tilde{\mathbf{b}}|\|_{L^q_p}, \||\sigma\mathbf{b}|\|_{L^q_p}, \||\sigma\tilde{\mathbf{b}}|\|_{L^q_p} \leqslant L$. By (2.11), Lemma 2.3 and $\psi(\lambda) \leqslant \frac{1}{2L}$ for $\lambda \geqslant \lambda_L$, we have

$$\begin{aligned}
&\|\Xi_\lambda \mathbf{b} - \Xi_\lambda \tilde{\mathbf{b}}\|_{\mathbb{B}^m} \leqslant \psi(\lambda) \big(\|\nabla_{\sigma\mathbf{b}} \Xi_\lambda \mathbf{b} - \nabla_{\sigma\tilde{\mathbf{b}}} \Xi_\lambda \tilde{\mathbf{b}}\|_{L^q_p} + \|\sigma(\mathbf{b} - \tilde{\mathbf{b}})\|_{L^q_p} \big) \\
&\leqslant \psi(\lambda) \big(\|\nabla_{\sigma\mathbf{b}}(\Xi_\lambda \mathbf{b} - \Xi_\lambda \tilde{\mathbf{b}}) + \nabla_{\sigma(\mathbf{b}-\tilde{\mathbf{b}})} \Xi_\lambda \tilde{\mathbf{b}}\|_{L^q_p} + \|\sigma(\mathbf{b} - \tilde{\mathbf{b}})\|_{L^q_p} \big) \\
&\leqslant \psi(\lambda) \big(\|\sigma(\mathbf{b} - \tilde{\mathbf{b}})\|_{L^q_p} + \|\Xi_\lambda \tilde{\mathbf{b}}\|_{\mathbb{B}^m} \|\mathbf{b} - \tilde{\mathbf{b}}\|_{L^q_p} \big) + \psi(\lambda) \||\mathbf{b}|\|_{L^q_p} \|\Xi_\lambda \mathbf{b} - \Xi_\lambda \tilde{\mathbf{b}}\|_{\mathbb{B}^m} \\
&\leqslant \psi(\lambda) \big(\|\sigma(\mathbf{b} - \tilde{\mathbf{b}})\|_{L^q_p} + \|\Xi_\lambda \tilde{\mathbf{b}}\|_{\mathbb{B}^m} \|\mathbf{b} - \tilde{\mathbf{b}}\|_{L^q_p} \big) + \frac{1}{2} \|\Xi_\lambda \mathbf{b} - \Xi_\lambda \tilde{\mathbf{b}}\|_{\mathbb{B}^m}, \quad \lambda \geqslant \lambda_L.
\end{aligned}$$

Thus,

$$\|\Xi_\lambda \mathbf{b} - \Xi_\lambda \tilde{\mathbf{b}}\|_{\mathbb{B}^m} \leqslant 2\psi(\lambda)\big(\|\sigma(\mathbf{b}-\tilde{\mathbf{b}})\|_{L^q_p} + \|\Xi_\lambda \tilde{\mathbf{b}}\|_{\mathbb{B}^m}\|\mathbf{b}-\tilde{\mathbf{b}}\|_{L^q_p}\big), \quad \lambda \geqslant \lambda_L. \tag{2.20}$$

Applying this inequality to $\mathbf{b} = 0$ we obtain $\|\Xi_\lambda \tilde{\mathbf{b}}\|_{\mathbb{B}^m} \leqslant 2\psi(\lambda)\|\sigma\tilde{\mathbf{b}}\|_{L^q_p} \leqslant 1$, so that (2.20) gives

$$\|\Xi_\lambda \mathbf{b} - \Xi_\lambda \tilde{\mathbf{b}}\|_{\mathbb{B}^m} \leqslant 2\psi(\lambda)\big(\||\mathbf{b}-\tilde{\mathbf{b}}|\|_{L^q_p} + \|\sigma(\mathbf{b}-\tilde{\mathbf{b}})\|_{L^q_p}\big), \quad \lambda \geqslant \lambda_L.$$

Then the proof is finished. □

To prove Theorem 2.1(3), we consider Zvonkin's transformation

$$\theta_t^{(\lambda)}(x) := x + (\Xi_\lambda \mathbf{b})_t(x), \quad x \in \mathbb{R}^N, t \in [0,T] \tag{2.21}$$

for large enough $\lambda > 0$. We have the following result.

Lemma 2.5. *Assume* (A_1)-(A_2) *and Theorem 2.1(3)(i). If* $|\mathbf{b}| \in L^q_{p,loc}([0,T]\times\mathbb{R}^N)$ *for some* $(p,q) \in \mathscr{K}_1 \cap \mathscr{K}_2$, *then for large enough* $\lambda > 0$, *any solution* $(\tilde{X}_t)_{t\in[0,T]}$ *of the SDE* (2.2) *with* $\int_0^T |\mathbf{b}|_t^2(\tilde{X}_t)\mathrm{d}t < \infty$, *any* $k \geqslant 1$, *and* $h_k \in C_0^\infty(\mathbb{R}^N)$ *such that* $h_k|_{B(0,k)} = 1$,

$$\mathrm{d}\theta_t^{(\lambda,k)}(\tilde{X}_t) = \big\{Z_t(\tilde{X}_t) + \lambda(\Xi_\lambda h_k \mathbf{b})_t(\tilde{X}_t)\big\}\mathrm{d}t + \nabla_{\sigma_t(\tilde{X}_t)\mathrm{d}B_t}\theta_t^{(\lambda,k)}(\tilde{X}_t), \quad t \in [0, T\wedge\tilde{\tau}_k], \tag{2.22}$$

where

$$\tilde{\tau}_k := \inf\{t \in [0,T] : |\tilde{X}_t| \geqslant k\},$$

and

$$\theta_t^{(\lambda,k)}(x) := x + (\Xi_\lambda h_k \mathbf{b})_t(x), \quad x \in \mathbb{R}^N, t \in [0,T].$$

Proof. When $\theta_t^{(\lambda,k)}$ is second-order differentiable with bounded derivatives, the desired formula follows from (2.4),(2.11) and Itô's formula. In general, we use the following approximation argument as in [24]. Let $\{\mathbf{b}^{(n)}\}_{n\geqslant 1} \subset C_0^\infty([0,T]\times\mathbb{R}^N)$ such that

$$\lim_{n\to\infty} \|h_k\mathbf{b} - h_k\mathbf{b}^{(n)}\|_{L^q_p} = 0. \tag{2.23}$$

Since σ is locally bounded, we have

$$\lim_{n\to\infty} \|\sigma h_k\mathbf{b} - \sigma h_k\mathbf{b}^{(n)}\|_{L^q_p} = 0. \tag{2.24}$$

Let $\theta^{(\lambda,n,k)}$ be defined in (2.21) for $h_k\mathbf{b}^{(n)}$ replacing $\mathbf{b}$ respectively, i.e.

$$\theta_t^{(\lambda,n,k)}(x) := x + (\Xi_\lambda h_k\mathbf{b}^{(n)})_t(x), \quad x \in \mathbb{R}^N, t \in [0,T], \lambda \geqslant 0. \tag{2.25}$$

By (A_3), (2.11) and (2.25), we have

$$\begin{aligned}
&(\partial_s + \mathscr{L}_s + \nabla_{\sigma_s\mathbf{b}_s})\theta_s^{(\lambda,n,k)} \\
&= Z_s + \sigma_s\mathbf{b}_s + \nabla_{\sigma_s\mathbf{b}_s}(\Xi_\lambda h_k\mathbf{b}^{(n)})_s + \lambda(\Xi_\lambda h_k\mathbf{b}^{(n)})_s - \big\{\nabla_{\sigma_s h_k\mathbf{b}_s^{(n)}}(\Xi_\lambda h_k\mathbf{b}^{(n)})_s + \sigma_s h_k\mathbf{b}_s^{(n)}\big\} \\
&= Z_s + \lambda(\Xi_\lambda h_k\mathbf{b}^{(n)})_s + \sigma_s(\mathbf{b}_s - h_k\mathbf{b}_s^{(n)}) + \nabla_{\sigma_s(\mathbf{b}_s - h_k\mathbf{b}_s^{(n)})}(\Xi_\lambda h_k\mathbf{b}^{(n)})_s.
\end{aligned}$$

So, by (2.2) and Itô's formula, we have

$$
\begin{aligned}
&\theta^{(\lambda,n,k)}_{t\wedge\tilde\tau_k}(\tilde X_{t\wedge\tilde\tau_k})-\theta^{(\lambda,n,k)}_0(\tilde X_0)\\
&=\int_0^{t\wedge\tilde\tau_k}\Big(Z_s+\lambda(\Xi_\lambda h_k\mathbf{b}^{(n)})_s+\sigma_s(\mathbf{b}_s-h_k\mathbf{b}^{(n)}_s)+\nabla_{\sigma_s(\mathbf{b}_s-h_k\mathbf{b}^{(n)}_s)}(\Xi_\lambda h_k\mathbf{b}^{(n)})_s\Big)(\tilde X_s)\mathrm{d}s \qquad (2.26)\\
&\quad+\int_0^{t\wedge\tilde\tau_k}\sigma_s(\tilde X_s)\mathrm{d}B_s+\int_0^{t\wedge\tilde\tau_k}\nabla_{\sigma_s(\tilde X_s)\mathrm{d}B_s}(\Xi_\lambda h_k\mathbf{b}^{(n)})_s(\tilde X_s),\quad k\geqslant 1,t\in[0,T].
\end{aligned}
$$

By Theorem 2.1(1) and (2.23), for large enough $\lambda>0$,

$$\lim_{n\to\infty}\|\theta^{(\lambda,n,k)}-\theta^{(\lambda,k)}\|_{\mathbb{B}^m}=\lim_{n\to\infty}\|\Xi_\lambda h_k\mathbf{b}-\Xi_\lambda h_k\mathbf{b}^{(n)}\|_{\mathbb{B}^m}=0.$$

Then

$$
\begin{aligned}
&\lim_{n\to\infty}\big\{\theta^{(\lambda,n,k)}_{t\wedge\tau_k}(\tilde X_{t\wedge\tau_k})-\theta^{(\lambda,n,k)}_0(\tilde X_0)\big\}=\theta^{(\lambda,k)}_{t\wedge\tau_k}(\tilde X_{t\wedge\tau_k})-\theta^{(\lambda,k)}_0(\tilde X_0),\\
&\lim_{n\to\infty}\int_0^{T\wedge\tilde\tau_k}|(\Xi_\lambda h_k\mathbf{b}^{(n)}-\Xi_\lambda h_k\mathbf{b})_s|(\tilde X_s)\mathrm{d}s=0.
\end{aligned}
$$

Since $h_k|_{B(0,k)}=1$, combining these with (2.12) and the local boundedness of σ, we may find out a constant $C>0$ such that

$$
\begin{aligned}
&\lim_{n\to\infty}\mathbb{E}\int_0^{T\wedge\tilde\tau_k}\Big(|\nabla_{\sigma_s}(\Xi_\lambda h_k\mathbf{b}_s-\Xi_\lambda h_k\mathbf{b}^{(n)}_s)|^2\\
&\qquad\qquad+|\nabla_{\sigma_s(\mathbf{b}_s-h_k\mathbf{b}^{(n)}_s)}(\Xi_\lambda h_k\mathbf{b}^{(n)})_s|+|\sigma_s(\mathbf{b}_s-h_k\mathbf{b}^{(n)}_s)|\Big)(\tilde X_s)\mathrm{d}s\\
&\leqslant C\lim_{n\to\infty}\Big(\|\Xi_\lambda h_k\mathbf{b}-\Xi_\lambda h_k\mathbf{b}^{(n)})\|^2_{\mathbb{B}^m}+\|h_k\mathbf{b}-h_k\mathbf{b}^{(n)}\|_{L^q_p}\Big)=0.
\end{aligned}
$$

Therefore, letting $n\to\infty$ in (2.26), we obtain

$$
\begin{aligned}
\theta^{(\lambda,k)}_{t\wedge\tilde\tau_k}(\tilde X_{t\wedge\tilde\tau_k})-\theta^{(\lambda,k)}_0(\tilde X_0)&=\int_0^{t\wedge\tilde\tau_k}\Big(Z_s+\lambda(\Xi_\lambda h_k\mathbf{b})_s\Big)(\tilde X_s)\mathrm{d}s\\
+\int_0^{t\wedge\tilde\tau_k}\sigma_s(\tilde X_s)\mathrm{d}B_s&+\int_0^{t\wedge\tilde\tau_k}\nabla_{\sigma_s(\tilde X_s)\mathrm{d}B_s}(\Xi_\lambda h_k\mathbf{b})_s(\tilde X_s),\quad k\geqslant 1,t\in[0,T].
\end{aligned}
$$

This means that (2.22) holds for $t\leqslant T\wedge\tilde\tau_k$. □

By Lemma 2.5, the uniqueness of the SDE (2.2) follows from that of (2.22). As in [13, 26], to prove the uniqueness of (2.22) we will use the following result for the maximal operator: for any $N\geqslant 1$,

$$\mathscr{M}h(x):=\sup_{r>0}\frac{1}{|B(x,r)|}\int_{B(x,r)}h(y)\mathrm{d}y,\quad h\in L^1_{loc}(\mathbb{R}^N),x\in\mathbb{R}^N,$$

where $B(x,r):=\{y:|x-y|<r\}$, see [3, Appendix A].

Lemma 2.6. *There exists a constant $C_N > 0$ such that for any continuous and weak differentiable function f,*

$$|f(x) - f(y)| \leqslant C_N|x-y|(\mathscr{M}|\nabla f|(x) + \mathscr{M}|\nabla f|(y)), \quad \text{a.e. } x, y \in \mathbb{R}^N. \tag{2.27}$$

Moreover, for any $p > 1$, there exists a constant $C_{N,p} > 0$ such that

$$\|\mathscr{M} f\|_{L^p} \leqslant C_{N,p}\|f\|_{L^p}, \quad f \in L^p(\mathbb{R}^N). \tag{2.28}$$

Proof of Theorem 2.1(3). It suffices to prove that for any $h \in C_0^\infty(\mathbb{R}^N)$, (2.2) with hb replacing b has a unique solution. So, without loss of generality, we may and do assume that b has a compact support. Then $h_n b = b$ with $h_n \in C_0^\infty(\mathbb{R}^N)$ such that $h_n|_{B(0,n)} = 1$ for large $n \geqslant 1$. By Theorem 2.1(2) and the Yamada-Watanabe principle, it suffices to prove the pathwise uniqueness. Let $\tilde{X}_t, \tilde{Y}_t$ be two solutions of (2.2) with $\tilde{X}_0 = \tilde{Y}_0$, life times ξ, η respectively, and $\int_0^{T\wedge\xi_n} |\mathbf{b}|_t^2(\tilde{X}_t)\mathrm{d}t + \int_0^{T\wedge\eta_n} |\mathbf{b}|_t^2(\tilde{Y}_t)\mathrm{d}t < \infty$, where

$$\xi_n := \inf\{t \in [0,T] : |\tilde{X}_t| \geqslant n\}, \quad \eta_n := \inf\{t \in [0,T] : |\tilde{Y}_t| \geqslant n\}, \quad n \geqslant 1.$$

Let $T_n = \xi_n \wedge \eta_n$. It remains to prove $\mathbb{P}$-a.s.

$$|\tilde{X}_{t\wedge T_n} - \tilde{Y}_{t\wedge T_n}| = 0, \quad n \geqslant 1, t \in [0,T]. \tag{2.29}$$

Let $h_n \in C_0^\infty(\mathbb{R}^N)$ such that $h_n|_{B(0,n)} = 1$. Then, up to time $T \wedge T_n$, $\tilde{X}_t$ and $\tilde{Y}_t$ solve the SDE (2.2) for $h_n\mathbf{b}$ replacing $\mathbf{b}$.

By (2.10), we take large enough $\lambda > 0$ such that

$$\sup_{t\in[0,T]} \|h_n \nabla \Xi_\lambda(h_n\mathbf{b})_t\|_\infty \leqslant \frac{1}{2}.$$

Simply denote $\mathbf{u} = \Xi_\lambda(h_n\mathbf{b})$ and $\theta_s(x) = x + \mathbf{u}_s(x)$. Then

$$\frac{1}{2}|\theta_t(x) - \theta_t(y)| \leqslant |x-y| \leqslant 2|\theta_t(x) - \theta_t(y)|, \quad t \in [0,T], x, y \in B(0,n). \tag{2.30}$$

By Lemma 2.5 and Itô's formula, we have

$$\begin{aligned}
&|\theta_{t\wedge T_n}(\tilde{X}_{t\wedge T_n}) - \theta_{t\wedge T_n}(\tilde{Y}_{t\wedge T_n})|^2 \\
&= 2\int_0^{t\wedge T_n} \big\langle Z_s(\tilde{X}_s) - Z_s(\tilde{Y}_s) + \lambda(\mathbf{u}_s(\tilde{X}_s) - \mathbf{u}_s(\tilde{Y}_s)), \theta_s(\tilde{X}_s) - \theta_s(\tilde{Y}_s)\big\rangle \mathrm{d}s \\
&\quad + \int_0^{t\wedge T_n} \Big\|[\nabla_{\sigma_s(\tilde{X}_s)}\theta_s(\tilde{X}_s) - \nabla_{\sigma_s(\tilde{Y}_s)}\theta_s(\tilde{Y}_s)]\Big\|_{\mathrm{HS}}^2 \mathrm{d}s \\
&\quad + 2\int_0^{t\wedge T_n} \big\langle \nabla_{\sigma_s(\tilde{X}_s)(\theta_s(\tilde{X}_s)-\theta_s(\tilde{Y}_s))}\theta_s(\tilde{X}_s) - \nabla_{\sigma_s(\tilde{Y}_s)(\theta_s(\tilde{X}_s)-\theta_s(\tilde{Y}_s))}\theta_s(\tilde{Y}_s), \mathrm{d}B_s\big\rangle \\
&= \int_0^t \beta_n(s)|\theta_s(\tilde{X}_s) - \theta_s(\tilde{Y}_s)|^2\mathrm{d}s + \int_0^t \langle \alpha_n(s)|\theta_s(\tilde{X}_s) - \theta_s(\tilde{Y}_s)|^2, \mathrm{d}B_s\rangle, \quad t \in [0,T],
\end{aligned} \tag{2.31}$$

where

$$\beta_n(s) := \frac{1_{\{s<T_n\}}1_{\{\tilde{X}_s\neq\tilde{Y}_s\}}}{|\theta_s(\tilde{X}_s) - \theta_s(\tilde{Y}_s)|^2}\Big(\Big\|\nabla_{\sigma_s(\tilde{X}_s)}\theta_s(\tilde{X}_s) - \nabla_{\sigma_s(\tilde{Y}_s)}\theta_s(\tilde{Y}_s)\Big\|_{\mathrm{HS}}^2$$

$$+ 2\Big\langle Z_s(\tilde X_s) - Z_s(\tilde Y_s) + \lambda(\mathbf{u}_s(\tilde X_s) - \mathbf{u}_s(\tilde Y_s)), \theta_s(\tilde X_s) - \theta_s(\tilde Y_s)\Big\rangle\Big),$$

$$\alpha_n(s) := \frac{2 1_{\{s<T_n\}} 1_{\{\tilde X_s \neq \tilde Y_s\}}}{|\theta_s(\tilde X_s) - \theta_s(\tilde Y_s)|^2}\Big(\nabla_{\sigma_s(\tilde X_s)(\theta_s(\tilde X_s)-\theta_s(\tilde Y_s))}\theta_s(\tilde X_s) - \nabla_{\sigma_s(\tilde Y_s)(\theta_s(\tilde X_s)-\theta_s(\tilde Y_s))}\theta_s(\tilde Y_s)\Big).$$

Since $h_n|_{B(0,n)} = 1$, β_n and α_n do not change if $Z, \nabla_\sigma\theta$, and $\mathbf{u}$ are replaced by $h_n Z, h_n\nabla_\sigma\theta$ and $h_n\mathbf{u}$ respectively. So, letting

$$\Phi_s = \|\nabla(h_n Z)_s\| + \|\nabla(h_n\mathbf{u})_s\| + \|\nabla(h_n\nabla_{\sigma_s}\theta_s)\|^2,$$

by Lemma 2.6 we may find a constant $C_1 > 0$ such that

$$|\alpha_n(s)|^2 + |\beta_n(s)| \leqslant C_1 1_{\{s<T_n\}}(\mathscr{M}\Phi_s(\tilde X_s) + \mathscr{M}\Phi_s(\tilde Y_s)), \quad s\in[0,T]. \tag{2.32}$$

Applying Theorem 2.2(1) for $h_n\mathbf{b}$ replacing $\mathbf{b}$ and using (2.32), we obtain

$$\mathbb{E}\bigg(\int_s^T (|\alpha_n(s)|^2 + |\beta_n(s)|)\mathrm{d}s\bigg|\mathscr{F}_s\bigg) \leqslant \kappa\|\mathscr{M}\Phi\|_{L^q_p}, \quad s\in[0,T]$$

for some constant $\kappa > 0$. Since Lemma 2.6 and our conditions in Theorem 2.1(3) imply

$$\|\mathscr{M}\Phi\|_{L^q_p} \leqslant \kappa'\|\Phi\|_{L^q_p} < \infty$$

for some constant $\kappa' > 0$, using the Khasminskii estimate as in (2.13) we conclude that

$$\mathbb{E}\exp\bigg[c\int_0^T (|\alpha_n(s)|^2 + |\beta_n(s)|)\mathrm{d}s\bigg] < \infty, \quad c>0.$$

So, by Dolèans-Dade's exponential formula, (2.31) implies

$$|\theta_{t\wedge T_n}(\tilde X_{t\wedge T_n}) - \theta_{t\wedge T_n}(\tilde Y_{t\wedge T_n})|^2 = |\theta_0(\tilde X_0) - \theta_0(\tilde Y_0)|^2 \mathrm{e}^{2\int_0^t \langle\alpha_n(s),\mathrm{d}B_s\rangle + \int_0^t \left(\beta_n(s) - 2|\alpha_n(s)|^2\right)\mathrm{d}s}, \quad t\in[0,T].$$

Since $\tilde X_0 = \tilde Y_0$, we have proved (2.29).

□

3 Singular SDEs on generalized Heisenberg groups

3.1 Framework and main result

Consider the following vector fields on $\mathbb{R}^{m+d}$, where $m\geqslant 2, d\geqslant 1$:

$$U_i(x,y) = \sum_{k=1}^m \theta_{ki}\partial_{x_k} + \sum_{l=1}^d (A_l x)_i \partial_{y_l}, \quad 1\leqslant i\leqslant m, \tag{3.1}$$

where $(x,y) = (x_1,\cdots,x_m,y_1,\cdots,y_d)\in\mathbb{R}^{m+d}$, $\Theta := (\theta_{ij})$ and $A_l (1\leqslant l\leqslant d)$ are $m\times m$-matrices satisfying the following assumption:

(H) α is invertible, $G_l := A_l\alpha - \alpha^* A_l^* \neq 0 (1 \leqslant l \leqslant d)$, and there exists $\varepsilon \in [0,1)$ such that

$$\varepsilon \sum_{l=1}^d a_l^2 |G_l u|^2 \geqslant \sum_{1\leqslant l\neq k\leqslant d} |a_l a_k \langle G_l u, G_k u\rangle|, \quad a \in \mathbb{R}^d, u \in \mathbb{R}^m.$$

As showing in the beginning of [21, §1], this assumption implies

$$\sum_{i,j=1}^m \bigg|\sum_{l=1}^d (G_l)_{ij} a_l\bigg|^2 \geqslant (1-\varepsilon)\Big(\inf_{1\leqslant l\leqslant d} \|G_l\|_{HS}^2\Big)|a|^2, \quad a \in \mathbb{R}^d. \tag{3.2}$$

Consequently, $\{U_i, [U_i, U_j]\}_{1\leqslant i,j\leqslant m}$ spans the tangent space of $\mathbb{R}^{m+d}$. Since $\mathrm{div} U_i = 0$, the operator

$$\mathscr{L} := \frac{1}{2}\sum_{i=1}^m U_i^2$$

is subelliptic and symmetric in $L^2(\mathbb{R}^{m+d})$, and the associated diffusion process solves the SDE for $(X_t, Y_t) \in \mathbb{R}^{m+d}$:

$$\mathrm{d}(X_t, Y_t) = \sum_{i=1}^m U_i(X_t) \circ \mathrm{d}B_t^i = Z\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t, \tag{3.3}$$

where $B_t := (B_t^i)_{1\leqslant i\leqslant m}$ is the m-dimensional Brownian motion, and

$$\sigma(x) := (\Theta, A_1 x, \cdots, A_d x), \quad Z := \sum_{i=1}^m \nabla_{U_i} U_i = \sum_{l=1}^d \mathrm{tr}(\Theta A_l)\partial_{y_l}.$$

We now consider the following SDE with a singular drift $\mathbf{b} : [0,T] \times \mathbb{R}^{m+d} \to \mathbb{R}^m$:

$$\mathrm{d}(\tilde{X}_t, \tilde{Y}_t) = \Big\{\sigma(\tilde{X}_t)\mathbf{b}_t(\tilde{X}_t, \tilde{Y}_t) + Z\Big\}\mathrm{d}t + \sigma(\tilde{X}_t)\mathrm{d}B_t. \tag{3.4}$$

Remark 3.1. Take $d = m-1, \Theta = I_{m\times m}$ and for some constants $a_l \neq \beta_l$,

$$(A_l)_{ij} = \begin{cases} a_l, & \text{if } i = 1, j = l+1, \\ \beta_l, & \text{if } i = l+1, j = 1, \\ 0, & \text{otherwise.} \end{cases}$$

Then $G_l^* G_k = 0$ for $l \neq k$, so that (H) holds with $\varepsilon = 0$. In particular, for $a_l = -\beta_l = \frac{1}{2}$, $\mathscr{L}$ is the Kohn-Laplacian operator on the $(2m-1)$-dimensional Heisenberg group. In general, $\mathbb{R}^{m+d}$ is a group under the action

$$(x,y) \bullet (x',y') := (x+x', y+y' + \langle(\Theta^*)^{-1} A.x, x'\rangle), \quad (x,y),(x',y') \in \mathbb{R}^{m+d}, \tag{3.5}$$

and $U_i, 1 \leqslant i \leqslant m$ are left-invariant vector fields. So, we call (3.4) a singular SDE on the generalized Heisenberg group.

For two nonnegative functions F_1, F_2, we write $F_1 \preceq F_2$ if there exists a constant $C > 0$ such that $F_1 \leqslant C F_2$, and write $F_1 \asymp F_2$ if $F_1 \preceq F_2$ and $F_2 \preceq F_1$.

Let $\Delta_y = \sum_{l=1}^d \partial_{y_l}^2$. Then $(\Delta_y, W^{2,2}(\mathbb{R}^d))$ is a negative definite operator in $L^2(\mathbb{R}^d)$. For any $\alpha > 0$ and $\lambda \geqslant 0$, we consider the operator $(\lambda - \Delta_y)^\alpha$ defined on domain $\mathscr{D}((-\Delta_y)^\alpha) := W^{2\alpha,2}(\mathbb{R}^d)$. This operator extends naturally to a measurable function f on the produce space $\mathbb{R}^{m+d}$ such that $f(x,\cdot) \in \mathscr{D}((-\Delta_y)^\alpha)$ for $x \in \mathbb{R}^m$:

$$(\lambda - \Delta_y)^\alpha f(x,y) := (\lambda - \Delta_y)^\alpha f(x,\cdot)(y).$$

For any $\beta > 0, p \geqslant 1$, let $\mathbb{H}_y^{\alpha,p}$ be the space of measurable functions on $\mathbb{R}^{m+d}$ such that

$$\|f\|_{\mathbb{H}_y^{\beta,p}} := \|(1-\Delta_y)^{\frac{\beta}{2}} f\|_p \asymp \|f\|_p + \|(-\Delta_y)^{\frac{\beta}{2}} f\|_p < \infty.$$

Recall that for $\beta \in (0,2)$, we have

$$-(-\Delta_y)^{\frac{\beta}{2}} f(z) := \int_{\mathbb{R}^d} (f(z+(0,y')) - f(z))|y'|^{-(m+\beta)} \mathrm{d}y', \quad z \in \mathbb{R}^{m+d}. \tag{3.6}$$

For any $\beta > 0, p, q \geqslant 1$, let $\mathbb{H}_y^{\beta,p,q}$ be the completion of $C_0^\infty([0,T]\times\mathbb{R}^{m+d})$ with respect to the norm

$$\|f\|_{\mathbb{H}_y^{\beta,p,q}} := \|(1-\Delta_y)^{\frac{\beta}{2}} f\|_{L_p^q} \asymp \|f\|_{L_p^q} + \|(-\Delta_y)^{\frac{\beta}{2}} f\|_{L_p^q}.$$

Applying Theorem 2.1 to the present model, we will prove the following result.

Theorem 3.1. *Assume* (H) *and let* $p, q \geqslant 1$ *satisfy*

$$\frac{2}{q} + \frac{m+2d}{p} < 1. \tag{3.7}$$

1. *If* $|\mathbf{b}| \in L_p^q([0,\infty]\times\mathbb{R}^N)$, *then for any initial value* $x \in \mathbb{R}^{m+d}$, *the SDE* (3.4) *has a weak solution* $(X_t)_{t\in[0,T]}$ *starting at* x *with* $\mathbb{E}\mathrm{e}^{\lambda\int_0^T |\mathbf{b}_t(X_t)|^2 \mathrm{d}t} < \infty$ *for all* $\lambda > 0$.
2. *If* $(h\mathbf{b}) \in \mathbb{H}_y^{\frac{1}{2},p,q}$ *holds for any* $h \in C_0^\infty(\mathbb{R}^{m+d})$, *then for any initial value* $x \in \mathbb{R}^{m+d}$, *the SDE* (3.4) *has a unique strong solution* $\tilde{X}_t$ *starting at* x *up to the life time* $\zeta := \lim_{n\to\infty} T \wedge \zeta_n := \inf\{t \in [0,T] : |\tilde{X}_t| \geqslant n\}$ *with* $\int_0^{T\wedge\zeta_n} |\mathbf{b}|_t^2(\tilde{X}_t)\mathrm{d}t < \infty$ *for any* $n \geqslant 1$.

3.2 Proof of Theorem 3.1

To apply Theorem 2.1, we first collect some known assertions about $\mathscr{L}$ and the associated Markov semigroup P_t. Let $\|\cdot\|_{p\to q}$ denote the operator norm from $L^p(\mathbb{R}^{m+d})$ to $L^q(\mathbb{R}^{m+d})$, and let $\|\cdot\|_p = \|\cdot\|_{p\to p}$. For any $\alpha > 0, p \geqslant 1$, let $\mathbb{H}_\sigma^{\alpha,p}$ be the completion of $C_0^\infty(\mathbb{R}^{m+d})$ with respect to the norm

$$\|f\|_{\mathbb{H}_\sigma^{\alpha,p}} := \|(1-\mathscr{L})^{\frac{\alpha}{2}} f\|_p \asymp \|f\|_p + \|(-\mathscr{L})^{\frac{\alpha}{2}} f\|_p.$$

It is classical that

$$\|(-\Delta_y)^{\frac{1}{2}} f\|_p \asymp \|\nabla_y f\|_p, \tag{3.8}$$

and for any $\beta > 0$,

$$\|f\|_{\mathbb{H}_y^{\beta,p}} \asymp \|f\|_p + \|(-\Delta_y)^{\frac{\beta-[\beta]}{2}} \nabla_y^{[\beta]} f\|_p, \tag{3.9}$$

where $[\beta] := \sup\{k \in \mathbb{Z}_+ : k \leqslant \beta\}$ is the integer part of β.

Moreover, by the interpolation inequality, for any $0 \leqslant \alpha < \beta < \infty$ we have

$$\|f\|_{\mathbb{H}_y^{\alpha,p}} \preceq \|f\|_p^{\frac{\beta-\alpha}{\beta}} \|f\|_{\mathbb{H}_y^{\beta,p}}^{\frac{\alpha}{\beta}}. \tag{3.10}$$

Lemma 3.2. *Assume (H).*

(1) *There exists a constant $C > 0$ such that*

$$\|P_t\|_{L^1 \to L^\infty} \leqslant C t^{-\frac{m+2d}{2}}, \quad t > 0. \tag{3.11}$$

Moreover, for any $p > 1$ there exists a constant $c_p > 0$ such that

$$|\nabla_\sigma P_t f| \leqslant \frac{c_p}{\sqrt{t}} (P_t |f|^p)^{\frac{1}{p}}, \quad f \in \mathscr{B}_b(\mathbb{R}^{m+d}), t > 0. \tag{3.12}$$

(2) *For any $r \geqslant 0, p \in (1, \infty)$,*

$$\|(1-\mathscr{L})^{r+\frac{1}{2}} f\|_p \asymp \|(1-\mathscr{L})^r f\|_p + \|(1-\mathscr{L})^r \nabla_\sigma f\|_p, \quad f \in \mathbb{H}_\sigma^{1+2r,p},$$

and

$$\|(-\mathscr{L})^{r+\frac{1}{2}} f\|_p \asymp \|(-\mathscr{L})^r \nabla_\sigma f\|_p, \quad f \in \mathbb{H}_\sigma^{1+2r,p}.$$

(3) *For any $r > 0$ and $p \in (1, \infty)$,*

$$\|(1-\mathscr{L})^r f\|_p \asymp \|f\|_p + \|(-\mathscr{L})^r f\|_p, \quad f \in \mathbb{H}_\sigma^{2r,p}.$$

(4) *For any $r \in (0, 1)$ and $p > \frac{m+2d}{2r}$,*

$$\|f\|_\infty \preceq \|(1-\mathscr{L})^r f\|_p, \quad f \in \mathbb{H}_\sigma^{2r,p}.$$

(5) *For any $p \in (1, \infty), \alpha_1, \alpha_2 \geqslant 0, \theta \in (0, 1)$, and $f \in \mathbb{H}_\sigma^{2\alpha_2,p} \cap \mathbb{H}_y^{2\alpha_1,p}$,*

$$\|(1-\Delta_y)^{\theta\alpha_1} (1-\mathscr{L})^{(1-\theta)\alpha_2} f\|_p \preceq \|(1-\Delta_y)^{\alpha_1} f\|_p^\theta \|(1-\mathscr{L})^{\alpha_2} f\|_p^{1-\theta},$$
$$\|(-\Delta_y)^{\theta\alpha_1} (-\mathscr{L})^{(1-\theta)\alpha_2} f\|_p \preceq \|(-\Delta_y)^{\alpha_1} f\|_p^\theta \|(-\mathscr{L})^{\alpha_2} f\|_p^{1-\theta}.$$

Proof. The inequalities in (1) follow from Lemma 2.4 and Corollary 1.2 in [21] respectively. Assertion (2) is due to [12, Theorem 4.10]. Since P_t is contractive in $L^p(\mathbb{R}^{m+d})$ and

$$(1-\mathscr{L})^{-\alpha} = c \int_0^\infty e^{-t} t^{\alpha-1} P_t \mathrm{d}t \tag{3.13}$$

for some constant $c > 0$, $(1-\mathscr{L})^{-\alpha}$ is bounded in $L^p(\mathbb{R}^{m+d})$ for all $p \geqslant 1$. Combining this with the closed graph theorem that

$$\|f\|_p + \|(-\mathscr{L})^\alpha f\|_p \asymp \|f\|_p + \|(1-\mathscr{L})^\alpha f\|_p,$$

we prove assertion (3). By the first inequality in assertion (1) and using (3.13), we have

$$\|(1-\mathscr{L})^{-\frac{r}{2}} f\|_\infty \preceq \|f\|_p \int_0^\infty \mathrm{e}^{-t} \frac{\|P_t\|_{p\to\infty}}{t^{1-r}} \mathrm{d}t \leqslant C\|f\|_p$$

for some constant $C > 0$. So, assertion (4) holds. Finally, let

$$\mathscr{A} = (1-\Delta_y)^{\alpha_1}(1-\mathscr{L})^{-\alpha_2}.$$

By the interpolation theorem (see [15, Theorem 6.10]), we have

$$\|\mathscr{A}^\theta g\|_p \preceq \|g\|_p^{1-\theta} \|\mathscr{A} g\|_p^\theta.$$

Applying this inequality to $g = (1-\mathscr{L})^{\alpha_2} f$, we obtain

$$\|(1-\Delta_y)^{\theta\alpha_1}(1-\mathscr{L})^{(1-\theta)\alpha_2} f\|_p = \|\mathscr{A}^\theta g\|_p \preceq \|(1-\mathscr{L})^{\alpha_2}\|_p^{1-\theta} \|(1-\Delta_y)^{\alpha_1}\|_p^\theta.$$

□

Proof of Theorem 3.1. We first estimate $\mathscr{K}_1$ and $\mathscr{K}_2$. Let $P_{s,t} = P_{t-s}$. By (3.11) and using the interpolation theorem, we have

$$\|P_{s,t} f\|_\infty \preceq (t-s)^{-\frac{m+2d}{2p}} \|f\|_p, t > s \geqslant 0, p \geqslant 1. \tag{3.14}$$

So,

$$\mathscr{K}_1 \supset \Big\{(p,q) \in (1,\infty]^2 : \ \frac{1}{q} + \frac{m+2d}{2p} < 1\Big\}. \tag{3.15}$$

Combining (3.14) with (3.12), we see that for any $\varepsilon \in (0, p-1)$,

$$\|\nabla_\sigma P_{s,t} f\|_\infty \preceq (t-s)^{-\frac{1}{2}} \|P_{s,t}|f|^{1+\varepsilon}\|_\infty^{\frac{1}{1+\varepsilon}} \preceq (t-s)^{-\frac{1}{2}-\frac{(m+2d)(1+\varepsilon)}{2p}} \|f\|_p, \quad t > s \geqslant 0. \tag{3.16}$$

So,

$$\mathscr{K}_2 \supset \Big\{(p,q) \in (1,\infty]^2 : \ \frac{2}{q} + \frac{m+2d}{p} < 1\Big\}. \tag{3.17}$$

Therefore, the first assertion follows from Theorem 2.1(2).

Next, we verify (A_1), (A_2) and the assumption in Theorem 2.1(3). Since Θ is invertible, there exists a constant $\lambda > 0$ such that

$$|\sigma v| \geqslant |\Theta v| \geqslant \lambda |v|, \quad v \in \mathbb{R}^m.$$

So, (A_1) holds. Next, since U_i are smooth vector fields with constant or linear coefficients, $\partial_t P_t f = \mathscr{L} P_t f$ for $f \in C_0^\infty(\mathbb{R}^N)$ and

$$\|\nabla P_t f\|_\infty \leqslant C\|\nabla f\|_\infty, \quad t \in [0, T], f \in C_b^1(\mathbb{R}^N)$$

for some constant $C > 0$. So, (A_2) and the assumption in Theorem 2.1(3) (i) hold. So, for (p, q) satisfy (3.7), by (3.15) and (3.17) we have $(p, q) \in \mathscr{K}_1 \cap \mathscr{K}_2$. According to Theorem 2.1(3), it remains to prove that for $h \in C_0^\infty(\mathbb{R}^{m+d})$,

$$\lim_{\lambda\to\infty} \|\nabla\{h\Xi_\lambda(h\mathbf{b})\}\|_\infty = 0, \tag{3.18}$$

$$\limsup_{\lambda\to\infty} \|\nabla\{h\nabla_\sigma\Xi_\lambda(h\mathbf{b})\}\|_{L_p^q} < \infty. \tag{3.19}$$

We leave the proofs to the following subsection. □

3.3 Proofs of (3.18) and (3.19)

We first investigate the regularity of the solution to the following PDE:

$$\partial_t u_t = (\lambda - \mathscr{L})u_t - f_t, \quad u_T = 0. \tag{3.20}$$

For this, we need some preparations.

The following interpolation theorem comes from [12, 18].

Lemma 3.3. *Let $p \in (1, \infty)$, $0 < \alpha < \beta$ and $f \in \mathbb{H}_\sigma^{2\alpha,p} \cap \mathbb{H}_\sigma^{2\beta,p}$. For any $\theta \in (0, 1)$, let $\gamma = \theta\alpha + (1-\theta)\beta$. Then $f \in \mathbb{H}_\sigma^{2\gamma,p}$ and*

$$\|(-\mathscr{L})^\gamma f\|_p \leqslant C\|(-\mathscr{L})^\alpha f\|_p^\theta \|(-\mathscr{L})^\beta f\|_p^{1-\theta},$$
$$\|(1-\mathscr{L})^\gamma f\|_p \leqslant C\|(1-\mathscr{L})^\alpha f\|_p^\theta \|(1-\mathscr{L})^\beta f\|_p^{1-\theta},$$

where C only depends on α, β, γ.

Next, let P_t be the diffusion semigroup associated with the SDE (3.3). We estimate derivatives of P_t by following the line of [21].

Lemma 3.4. *Let $p > 1$, $t > 0$. Then the following assertions hold.*

(1) *There exists a constant $c_p > 0$ such that for any $f \in \mathscr{B}_b(\mathbb{R}^{m+d})$,*

$$|\nabla_y P_t f| \leqslant \frac{c_p}{t}(P_t|f|^p)^{\frac{1}{p}}, \tag{3.21}$$

and

$$|\nabla_\sigma\nabla_\sigma P_t f| \leqslant \frac{c_p}{t}(P_t|f|^p)^{\frac{1}{p}}. \tag{3.22}$$

(2) *For any $\alpha \in (0, 1)$, there exists a constant $C = C(p, \alpha)$ such that for all $f \in L^p(\mathbb{R}^{m+d})$,*

$$\|\nabla_\sigma P_t f\|_{\mathbb{H}_\sigma^{\alpha,p}} + \|(-\Delta_y)^{\frac{1}{4}} P_t f\|_{\mathbb{H}_\sigma^{\alpha,p}} \leqslant Ct^{-\frac{\alpha}{2}-\frac{1}{2}}\|f\|_p. \tag{3.23}$$

Proof. (3.21) follows from [21, Theorem 1.1] for $u = 0$. Moreover, combining (3.21) with (3.8) and (3.10), we obtain

$$\|(-\Delta_y)^{\frac{\beta}{2}} P_t f\|_p \preceq \frac{1}{t^\beta}(P_t|f|^p)^{\frac{1}{p}}, \quad \beta \in (0,1), p > 1, t > 0. \tag{3.24}$$

Then we can claim that that it suffices to prove (3.22) holds. Indeed, by (3.22), (3.12) and Lemma 3.3 we obtain

$$\|\nabla_\sigma P_t f\|_{\mathbb{H}^{\alpha,p}_\sigma} \preceq t^{-\frac{\alpha}{2}-\frac{1}{2}}\|f\|_p$$

for $\alpha \in (0,1)$. On the other hand, by Lemma 3.2(5), (3.22) and (3.24), we have

$$\begin{aligned}\|(-\Delta_y)^{\frac{1}{4}}(-\mathscr{L})^{\frac{\alpha}{2}} P_t f\|_p &= \|(-\Delta_y)^{\frac{1}{4(1-\frac{\alpha}{2})}(1-\frac{\alpha}{2})}(-\mathscr{L})^{\frac{\alpha}{2}} P_t f\|_p \\ &\preceq \|(-\Delta_y)^{\frac{1}{4(1-\frac{\alpha}{2})}} P_t f\|_p^{1-\frac{\alpha}{2}} \|(-\mathscr{L}) P_t f\|_p^{\frac{\alpha}{2}} \\ &\preceq t^{-\frac{1}{2}} t^{-\frac{\alpha}{2}}\|f\|_p.\end{aligned}$$

Therefore, (3.23) holds.

We now prove (3.22) by using the derivative formula in [21, Theorem 1.1]. Let $\mathbf{Q}_t = (q_{kl}(t))_{1\leqslant k,l\leqslant d}$ with

$$q_{lk}(t) := \int_0^t \left\langle G_l^* G_k \left(B_s - \frac{1}{t}\int_0^t B_s \mathrm{d}s\right), \left(B_s - \frac{1}{t}\int_0^t B_s \mathrm{d}s\right)\right\rangle \mathrm{d}s.$$

Then $\mathbf{Q}_t$ is invertible for $t > 0$. Next, for $x, w \in \mathbb{R}^m$ and $v \in \mathbb{R}^d$, let

$$(\tilde{\alpha}_{t,w,v,x})_l := v_l - \langle \Theta^{-1} w, A_l x\rangle - \frac{1}{t}\int_0^t \langle G_l^* \Theta^{-1} w, B_s\rangle \mathrm{d}s, \quad 1 \leqslant l \leqslant d. \tag{3.25}$$

Then for the functional $(x, y) \mapsto \tilde{\alpha}_{t,w,v,x}$ we have

$$\nabla_{(w',v')}(\tilde{\alpha}_{t,w,v,x})_l = -\langle \Theta^{-1} w, A_l w'\rangle, \quad (w', v') \in \mathbb{R}^{m+d}, 1 \leqslant l \leqslant d. \tag{3.26}$$

Next, the solution of (3.3) starting at (x, y) is given by

$$X_t = x + \alpha B_t, \quad (Y_t)_l = y_l + \langle A_l x, B_t\rangle + \int_0^t \langle A_l \alpha B_s, \mathrm{d}B_s\rangle, \quad 1 \leqslant l \leqslant d.$$

Then

$$\nabla_{(w',v')}(X_t, (Y_t)_l) = (w', (v')'_l + \langle A_l w', B_t\rangle), \quad 1 \leqslant l \leqslant d. \tag{3.27}$$

According to [21, Theorem 1.1(3)], we have the Bismut derivative formula

$$\nabla_{w,v} P_t f = \mathbb{E}[f(X_t, Y_t) M_t], \tag{3.28}$$

where by the formulation of $\tilde{h}'$ given in [21, Theorem 1.1],

$$\begin{aligned} M_t := D^*\tilde{h} &= \frac{1}{t}\langle \Theta^{-1} w, B_t\rangle + \sum_{k=1}^d (\mathbf{Q}_t^{-1}\tilde{\alpha}_{t,w,v,x})_k \int_0^t \langle G_k B_s, \mathrm{d}B_s\rangle \\ &\quad - \sum_{k=1}^d D_{\beta_k}(\mathbf{Q}_t^{-1}\tilde{\alpha}_{t,w,v,x})_k - \sum_{k=1}^d \frac{(\mathbf{Q}_t^{-1}\tilde{\alpha}_{t,w,v,x})_k}{t}\left\langle \int_0^t G_k B_s \mathrm{d}s, B_t\right\rangle \\ &\quad + \sum_{k=1}^d \sum_{i=1}^m \frac{D_{h_i}(\mathbf{Q}_t^{-1}\tilde{\alpha}_{t,w,v,x})_k}{t} \int_0^t (G_k B_s)_i \mathrm{d}s + \sum_{k=1}^d \sum_{i=1}^m \frac{t}{2}(\mathbf{Q}_t^{-1}\tilde{\alpha}_{t,w,v,x})_k (G_k)_{ii}, \end{aligned} \tag{3.29}$$

for $h_i(s) := se_i, \beta_k(s) := \int_0^s G_k B_r \mathrm{d}r, s \in [0,t]$, and $\{e_i\}_{i=1,\cdots,m}$ being the orthonormal basis of $\mathbb{R}^m$. According to step (1) in the proof of [21, Theorem 1.1], for any $p > 1$, we have

$$\{\mathbb{E}|M_t|^p\}^{1/p} \preceq \frac{(|v| + |w|(|x| + \sqrt{t}))}{t}. \tag{3.30}$$

Moreover, by (3.29) we have

$$\begin{aligned}
\nabla_{(w',v')} M_t &= \sum_{k=1}^d (\mathbf{Q}_t^{-1} \nabla_{(w',v')} \tilde{\alpha}_{t,w,v,x})_k \int_0^t \langle G_k B_s, \mathrm{d}B_s \rangle \\
&\quad - \sum_{k=1}^d D_{\beta_k} (\mathbf{Q}_t^{-1} \nabla_{(w',v')} \tilde{\alpha}_{t,w,v,x})_k - \sum_{k=1}^d \frac{(\mathbf{Q}_t^{-1} \nabla_{(w',v')} \tilde{\alpha}_{t,w,v,x})_k}{t} \left\langle \int_0^t G_k B_s \mathrm{d}s, B_t \right\rangle \\
&\quad + \sum_{i=1}^m \sum_{k=1}^d \frac{D_{h_i} (\mathbf{Q}_t^{-1} \nabla_{(w',v')} \tilde{\alpha}_{t,w,v,x})_k}{t} \int_0^t (G_k B_s)_i \mathrm{d}s \\
&\quad + \sum_{k=1}^d \sum_{i=1}^m \frac{t}{2} (\mathbf{Q}_t^{-1} \nabla_{(w',v')} \tilde{\alpha}_{t,w,v,x})_k (G_k)_{ii}.
\end{aligned}$$

Combining this with (3.26) we prove

$$\{\mathbb{E}|\nabla_{(w',v')} M_t|^p\}^{1/p} \preceq \frac{|w||w'|}{t}, \quad (w',v') \in \mathbb{R}^{m+d}. \tag{3.31}$$

By the Markov property and (3.28), we derive

$$\nabla_{(w,v)} P_t f = \nabla_{(w,v)} P_{\frac{t}{2}} (P_{\frac{t}{2}} f) = \mathbb{E}[(P_{\frac{t}{2}} f)(X_{\frac{t}{2}}, Y_{\frac{t}{2}}) M_{\frac{t}{2}}],$$

and by the chain rule,

$$\begin{aligned}
&\nabla_{(w',v')} \nabla_{(w,v)} P_t f = \nabla_{(w',v')} \mathbb{E}[(P_{\frac{t}{2}} f)(X_{\frac{t}{2}}, Y_{\frac{t}{2}}) M_{\frac{t}{2}}] \\
&= \mathbb{E}\left[\left(\nabla_{\nabla_{(w',v')}(X_{\frac{t}{2}}, Y_{\frac{t}{2}})} P_{\frac{t}{2}} f\right)(X_{\frac{t}{2}}, Y_{\frac{t}{2}}) M_{\frac{t}{2}}\right] + \mathbb{E}\left[(P_{\frac{t}{2}} f)(X_{\frac{t}{2}}, Y_{\frac{t}{2}}) \nabla_{(w',v')} M_{\frac{t}{2}}\right].
\end{aligned} \tag{3.32}$$

By (3.27), (3.30) and using Hölder's inequality, we obtain

$$\begin{aligned}
&\mathbb{E}\left|\left(\nabla_{\nabla_{(w',v')}(X_{\frac{t}{2}}, Y_{\frac{t}{2}})} P_{\frac{t}{2}} f\right)(X_{\frac{t}{2}}, Y_{\frac{t}{2}}) M_{\frac{t}{2}}\right| \\
&\preceq (P_t|f|^p)^{1/p} \frac{(|v'|w'(|x| + \sqrt{t}))(|v| + w(|x| + \sqrt{t}))}{t^2},
\end{aligned}$$

while by (3.31) and Hölder's inequality,

$$\mathbb{E}\left|(P_{\frac{t}{2}} f)(X_{\frac{t}{2}}, Y_{\frac{t}{2}}) \nabla_{(w',v')} M_{\frac{t}{2}}\right| \preceq (P_t|f|^p)^{1/p} \frac{|u||u'|}{t}.$$

Therefore, it follows from (3.26) that

$$\begin{aligned}
&|\nabla_{(w',v')} \nabla_{(w,v)} P_t f|(x,y) \\
&\preceq (P_t|f|^p)^{1/p} \left(\frac{(|w'| + v'(|x| + \sqrt{t}))(|v| + w(|x| + \sqrt{t}))}{t^2} + \frac{|w| \cdot |v'|}{t}\right).
\end{aligned} \tag{3.33}$$

Finally, by the definition of U_i, we have

$$
\begin{aligned}
U_iU_j &= \left(\sum_{k=1}^m \theta_{ki}\partial_{x_k} + \sum_{l=1}^d (A_l x)_i \partial_{y_l}\right)\left(\sum_{k=1}^m \theta_{kj}\partial_{x_k} + \sum_{l=1}^d (A_l x)_j \partial_{y_l}\right)\\
&= \sum_{k=1}^m\sum_{l=1}^m \theta_{ki}\theta_{lj}\partial_{x_k}\partial_{x_l} + \sum_{k=1}^m\sum_{l=1}^d \theta_{ki}(A_l)_{jk}\partial_{y_l} + \sum_{k=1}^m\sum_{l=1}^d \theta_{ki}(A_l x)_j\partial_{x_k}\partial_{y_l}\\
&+ \sum_{l=1}^d\sum_{k=1}^m (A_l x)_i\theta_{kj}\partial_{y_l}\partial_{x_k} + \sum_{l=1}^d\sum_{k=1}^d (A_l x)_i (A_k x)_j \partial_{y_l}\partial_{y_k}.
\end{aligned}
$$

Combining this with (3.21) and (3.33) with $(x,y)=(0,0)$, we arrive at

$$|U_iU_jP_tf(0,0)| \preceq \frac{1}{t}(P_t|f|^p(0,0))^{1/p}, \quad 1 \leqslant i,j \leqslant m, p>1, t>0.$$

As explained in the proof of [21, Proof of Corollary 1.2], by the left-invariant property of U_i and ∂_{y_l} under the group action in (3.5), this is equivalent to (3.22). □

The next lemma due to [12, Theorem 5.15] generalizes the classical Sobolev embedding theorem.

Lemma 3.5. *Suppose $p \in (1,\infty)$ and $\alpha > \frac{m+2d}{p}$, then*

$$\|f\|_{\Gamma_\gamma} \leqslant C(p, m+2d, \alpha)\|f\|_{\mathbb{H}^{\alpha,p}_\sigma}, \quad \gamma \in [0, \alpha-(m+2d)/p], \tag{3.34}$$

where

$$\|f\|_{\Gamma_\gamma} := \|f\|_\infty + |f|_\gamma, \quad |f|_\gamma := \sup_{x\in\mathbb{R}^{d+m}, y\neq 0} \frac{|f(x\bullet y)-f(x)|}{|y|^\gamma}.$$

Finally, we introduce the following lemma.

Lemma 3.6. *Let $p > m+2d$. For any $\beta \in (0,1]$ and $\alpha \in (\frac{m+2d}{p}, 1]$, there exists a constant $C = C(\alpha,\beta,m+2d,p) > 0$ such that for $\mathbb{R}^m$-valued function $\mathbf{b} \in \mathbb{H}^{\beta,p}_y$ and real function $u \in L^p(\mathbb{R}^{m+d})$ with $(-\mathscr{L})^{\frac{1}{2}+\frac{\alpha}{2}}u \in \mathbb{H}^{\beta,p}_y$,*

$$\|\nabla_{\sigma\mathbf{b}}u\|_{\mathbb{H}^{\beta,p}_y} \preceq \|\mathbf{b}\|_{\mathbb{H}^{\beta,p}_y}\left(\|(-\Delta_y)^{\frac{\beta}{2}}\nabla_\sigma u\|_{\mathbb{H}^{\alpha,p}_\sigma} + \|\nabla_\sigma u\|_{\mathbb{H}^{\alpha,p}_\sigma}\right).$$

Proof. By the definition of $\|\cdot\|_{\mathbb{H}^{\beta,p}_y}$ and noting that $\nabla_{\sigma\mathbf{b}}u = \langle\nabla_\sigma u, \mathbf{b}\rangle$, we have

$$
\begin{aligned}
\|\nabla_{\sigma\mathbf{b}}u\|_{\mathbb{H}^{\beta,p}_y} &\preceq \|\langle\nabla_\sigma u,\mathbf{b}\rangle\|_p + \|(-\Delta_y)^{\frac{\beta}{2}}\langle\nabla_\sigma u,\mathbf{b}\rangle\|_p\\
&\preceq \|\nabla_\sigma u\|_\infty\|\mathbf{b}\|_p + \|(-\Delta_y)^{\frac{\beta}{2}}\langle\nabla_\sigma u,\mathbf{b}\rangle\|_p.
\end{aligned}
\tag{3.35}
$$

According to [27, (2.5)],

$$\int_{\mathbb{R}^d} |f(x,y+y')-f(x,y)|^p \mathrm{d}y \leqslant \|(1-\Delta_y)^{\frac{\beta}{2}}f(x,\cdot)\|_p^p(|y'|^{p\beta}\wedge 1), \quad f \in \mathbb{H}^{\beta,p}_y.$$

Then

$$
\begin{aligned}
\|(f(\cdot+(0,y'))-f(\cdot))\|_p^p &= \int_{\mathbb{R}^{m+d}} |f(x,y+y')-f(x,y)|^p \mathrm{d}x\mathrm{d}y \\
&\leqslant \|f\|^p_{\mathbb{H}^{\beta,p}_y}(|y'|^{p\beta}\wedge 1).
\end{aligned}
$$

Combining this with Lemma 3.5, for any $\gamma\in(0,\alpha-\frac{m+2d}{p})$, we obtain

$$
\begin{aligned}
&\|\langle(\nabla_\sigma u)(\cdot+(0,y'))-(\nabla_\sigma u)(\cdot),\mathbf{b}(\cdot+(0,y'))-\mathbf{b}(\cdot)\rangle\|_p^p \\
=&\int_{\mathbb{R}^{m+d}} |\langle(\nabla_\sigma u)(x,y+y')-(\nabla_\sigma u)(x,y),\mathbf{b}(x,y+y')-\mathbf{b}(x,y)\rangle|^p\,\mathrm{d}x\mathrm{d}y \\
\leqslant&\int_{\mathbb{R}^m}\left(\sup_{y\in\mathbb{R}^d}|(\nabla_\sigma u)(x,y+y')-(\nabla_\sigma u)(x,y)|^p\int_{\mathbb{R}^d}|\mathbf{b}(x,y+y')-\mathbf{b}(x,y)|^p\,\mathrm{d}y\right)\mathrm{d}x \\
\preceq&\|\nabla_\sigma u\|^p_{\mathbb{H}^{\alpha,p}_\sigma}(|y'|^{\gamma p}\wedge 1)\int_{\mathbb{R}^{m+d}}|\mathbf{b}(x,y+y')-\mathbf{b}(x,y)|^p\,\mathrm{d}x\mathrm{d}y \\
\preceq&\|\nabla_\sigma u\|^p_{\mathbb{H}^{\alpha,p}_\sigma}(|y'|^{\gamma p}\wedge 1)\|\mathbf{b}\|^p_{\mathbb{H}^{\beta,p}_y}(|y'|^{p\beta}\wedge 1).
\end{aligned}
$$

By (3.6), Minkovskii inequality and Lemma 3.5, we have

$$
\begin{aligned}
&\|(-\Delta_y)^{\frac{\beta}{2}}\langle\nabla_\sigma u,\mathbf{b}\rangle\|_p \\
&\preceq\|\langle(-\Delta_y)^{\frac{\beta}{2}}\nabla_\sigma u,\mathbf{b}\rangle\|_p+\|\langle\nabla_\sigma u,(-\Delta_y)^{\frac{\beta}{2}}\mathbf{b}\rangle\|_p \\
&+\int_{\mathbb{R}^d}\|\langle(\nabla_\sigma u)(\cdot+(0,y'))-(\nabla_\sigma u)(\cdot),\mathbf{b}(\cdot+(0,y'))-\mathbf{b}(\cdot)\rangle\|_p\,|y'|^{-\beta-d}\mathrm{d}y' \\
&\preceq\|\mathbf{b}\|_p\|(-\Delta_y)^{\frac{\beta}{2}}\nabla_\sigma u\|_\infty+\|(-\Delta_y)^{\frac{\beta}{2}}\mathbf{b}\|_p\|\nabla_\sigma u\|_\infty+\|\nabla_\sigma u\|_{\mathbb{H}^{\alpha,p}_\sigma}\|\mathbf{b}\|_{\mathbb{H}^{\beta,p}_y} \\
&\preceq\|\mathbf{b}\|_p\|(-\Delta_y)^{\frac{\beta}{2}}\nabla_\sigma u\|_{\mathbb{H}^{\alpha,p}_\sigma}+\|\mathbf{b}\|_{\mathbb{H}^{\beta,p}_y}\|\nabla_\sigma u\|_{\mathbb{H}^{\alpha,p}_\sigma}.
\end{aligned}
$$

Substituting this into (3.35), we finish the proof. □

It is now ready to prove the following regularity estimates for solutions of (3.20).

Theorem 3.7. *Let $p,q\geqslant 1$ satisfy*

$$
\frac{2}{q}+\frac{m+2d}{p}<1. \tag{3.36}
$$

For any $f\in C_0^\infty([0,T]\times\mathbb{R}^{m+d})$ and $\lambda\geqslant 0$, (3.20) has a unique solution $u^\lambda=Q_\lambda f$, where $Q_\lambda f$ is in (2.3) for $P_{s,t}=P_{t-s}$. Moreover:

(1) *There exists a constant $C>0$ such that*

$$
\begin{aligned}
&\|\nabla_\sigma\nabla_\sigma u^\lambda\|_{L^q_p}+\|\nabla_y u^\lambda\|_{L^q_p}+\|(-\Delta_y)^{\frac{1}{4}}\nabla_\sigma u^\lambda\|_{L^q_p} \\
&\leqslant C\|f\|_{L^q_p},\ \ f\in C_0^\infty([0,T]\times\mathbb{R}^{m+d}).
\end{aligned} \tag{3.37}
$$

For any $\alpha\in(0,1)$ with $\alpha<1-\frac{2}{q}$,

$$
\begin{aligned}
&\|(-\Delta_y)^{\frac{1}{4}}u_t^\lambda\|_{\mathbb{H}^{\alpha,p}_\sigma}+\|\nabla_\sigma u_t^\lambda\|_{\mathbb{H}^{\alpha,p}_\sigma} \\
&\leqslant\phi(\lambda)\|f\|_{L^q_p},\quad t\in[0,T],f\in C_0^\infty([0,T]\times\mathbb{R}^{m+d})
\end{aligned} \tag{3.38}
$$

holds for some decreasing function $\phi:(0,\infty)\to(0,\infty)$ with $\lim_{\lambda\to\infty}\phi(\lambda)=0$.

(2) *There exists a constant $C > 0$ such that*

$$\|\nabla_y \nabla_\sigma u^\lambda\|_{L^q_p} \leqslant C\|f\|_{\mathbb{H}_y^{\frac{1}{2},p,q}}, \quad f \in C_0^\infty([0,T] \times \mathbb{R}^{m+d}). \tag{3.39}$$

For any $\alpha \in (0,1)$ with $\alpha < 1 - \frac{2}{q}$,

$$\begin{aligned} &\|\nabla_y u_t^\lambda\|_{\mathbb{H}_\sigma^{\alpha,p}} + \|(-\Delta_y)^{\frac{1}{4}} \nabla_\sigma u_t^\lambda\|_{\mathbb{H}_\sigma^{\alpha,p}} \\ &\leqslant \phi(\lambda)\|f\|_{\mathbb{H}_y^{\frac{1}{2},p,q}}, \quad t \in [0,T], f \in C_0^\infty([0,T] \times \mathbb{R}^{m+d}) \end{aligned} \tag{3.40}$$

holds for some decreasing function $\phi : (0,\infty) \to (0,\infty)$ with $\lim_{\lambda\to\infty} \phi(\lambda) = 0$.

Proof. (a) By (A_2) and Lemma 2.3, (3.20) has a unique solution $u^\lambda = Q_\lambda f$ such that

$$\|u^\lambda\|_{\mathbb{B}} := \|u^\lambda\|_\infty + \|\nabla_\sigma u^\lambda\|_\infty \leqslant \psi(\lambda)\|f\|_{L^q_p}, \quad \lambda \geqslant 0,$$

holds for some decreasing function $\psi : (0,\infty) \to (0,\infty)$ with $\lim_{\lambda\to\infty} \psi(\lambda) = 0$. Let $g \in C_0^\infty(\mathbb{R}^{m+d})$. By (A_2), the heat equation $\partial_t P_t g = \mathscr{L} P_t g$, and the contraction of P_t in $L^p(\mathbb{R}^{m+d})$, we have

$$\|\mathscr{L} u^\lambda\|_{L^q_p} \preceq \|f\|_{L^q_p}. \tag{3.41}$$

Since $(\nabla_\sigma g)_i = U_i g$, $(\nabla_\sigma \nabla_\sigma g)_{ij} = U_i U_j g$, $1 \leqslant i,j \leqslant m$, Lemma 3.2(2) gives

$$\|\nabla_\sigma \nabla_\sigma g\|_p \preceq \|(-\mathscr{L})^{\frac{1}{2}} \nabla_\sigma g\|_p \preceq \|(-\mathscr{L})g\|_p. \tag{3.42}$$

Combining this with (3.41), we obtain

$$\|U_i U_j u^\lambda\|_{L^q_p} \preceq \|f\|_{L^q_p}, \quad 1 \leqslant i,j \leqslant m. \tag{3.43}$$

Since (3.1) implies $U_i U_j = \sum_{l=1}^d (G_l)_{ij} \partial_{y_l}, i \neq j$, it follows from (3.2) that

$$\sum_{l=1}^d |\partial_{y_l} g|^2 \preceq \sum_{i,j=1}^m |U_i U_j g|^2 \tag{3.44}$$

This together with (3.43) leads to

$$\|\nabla_y u^\lambda\|_{L^q_p} \preceq \|f\|_{L^q_p}. \tag{3.45}$$

On the other hand, (3.8) implies

$$\|(-\Delta_y)^{\frac{1}{2}} u^\lambda\|_{L^q_p} \preceq \|f\|_{L^q_p}. \tag{3.46}$$

Applying Lemma 3.2(5) with $\theta = \frac{1}{2}$ and Young's inequality, we have

$$\|(-\Delta_y)^{\frac{1}{4}} \nabla_\sigma u^\lambda\|_{L^q_p} \preceq \|(-\Delta_y)^{\frac{1}{2}} u^\lambda\|_{L^q_p} + \|(-\mathscr{L}) u^\lambda\|_{L^q_p} \preceq \|f\|_{L^q_p}.$$

Combining this with (3.43) and (3.45), we prove (3.37).

Next, recall that

$$u_s^\lambda = (Q_\lambda f)_s := \int_s^T \mathrm{e}^{-\lambda(t-s)} P_{t-s} f_t \mathrm{d}t.$$

By (3.23), Hölder's inequality, and noticing that $\alpha < 1 - \frac{2}{q}$ implies $\frac{q}{q-1}(-\frac{\alpha}{2} - \frac{1}{2}) > -1$, we obtain

$$\begin{aligned}
&\|\nabla_\sigma u_s^\lambda\|_{\mathbb{H}_\sigma^{\alpha,p}} + \|(-\Delta_y)^{\frac{1}{4}} u_s^\lambda\|_{\mathbb{H}_\sigma^{\alpha,p}} \\
&\preceq \int_s^T \mathrm{e}^{-\lambda(t-s)} (t-s)^{-\frac{\alpha}{2}-\frac{1}{2}} \|f_t\|_p \mathrm{d}t \\
&\preceq \left(\int_0^T \mathrm{e}^{-\lambda \frac{q}{q-1}(t-s)} (t-s)^{\frac{q}{q-1}(-\frac{\alpha}{2}-\frac{1}{2})} \mathrm{d}t \right)^{\frac{q-1}{q}} \|f\|_{L_p^q} \\
&=: \phi(\lambda) \|f\|_{L_p^q},
\end{aligned}$$

where ϕ is decreasing with $\lim_{\lambda\to\infty} \phi(\lambda) = 0$. Therefore, assertion (1) is proved.

(b) Let $w^\lambda = (-\Delta_y)^{\frac{1}{4}} u^\lambda$, where $u^\lambda := Q_\lambda f$ is the unique solution of (3.20). We have

$$\partial_t w_t^\lambda = (\lambda - \mathscr{L}) w_t^\lambda - (-\Delta_y)^{\frac{1}{4}} f_t, \quad w_T^\lambda = 0. \tag{3.47}$$

Applying (3.8) and (3.37) for $(-\Delta_y)^{\frac{1}{4}} f$ replacing f, we obtain

$$\|\nabla_y \nabla_\sigma u^\lambda\|_{L_p^q} \asymp \|(-\Delta_y)^{\frac{1}{2}} \nabla_\sigma u^\lambda\|_{L_p^q} = \|(-\Delta_y)^{\frac{1}{4}} \nabla_\sigma w^\lambda\|_{L_p^q} \preceq \|(-\Delta_y)^{\frac{1}{4}} f\|_{L_p^q} \preceq \|f\|_{\mathbb{H}_y^{\frac{1}{2},p,q}}.$$

So, (3.39) holds.

Finally, applying (3.39) to $(w^\lambda, (-\Delta_y)^{\frac{1}{4}} f)$ replacing (u^λ, f), we prove (3.40). □

We now investigate the regularity of the solution to the following singular equation for $\mathbb{R}^{m+d}$-valued $u_t = (u_t^1, \cdots, u_t^{m+d})$:

$$\partial_t u_t = (\lambda - \mathscr{L}) u_t - \nabla_{\sigma \mathbf{b}_t} u_t - \sigma \mathbf{b}_t, \quad u_T = 0. \tag{3.48}$$

Theorem 3.8. *Let $p, q \geqslant 1$ satisfy* (3.36).

(1) *Assume $\mathbf{b} \in C_0^\infty([0,T] \times \mathbb{R}^{m+d}; \mathbb{R}^m)$. Then there exists a constant $\lambda_0 > 0$ such that for any $\lambda \geqslant \lambda_0$, the equation* (3.48) *has a unique solution* (*denoted by $\Xi_\lambda \mathbf{b}$*) *satisfying*

$$\|\nabla_\sigma \nabla_\sigma \Xi_\lambda \mathbf{b}\|_{L_p^q} \preceq \|\mathbf{b}\|_{L_p^q}^2 + \|\mathbf{b}\|_{L_p^q} \|\sigma \mathbf{b}\|_{L_p^q} + \|\sigma \mathbf{b}\|_{L_p^q}. \tag{3.49}$$

(2) *There exists a constant $\lambda_1 \geqslant \lambda_0$ such that for any $\lambda \geqslant \lambda_1$,*

$$\|\nabla_y \nabla_\sigma \Xi_\lambda \mathbf{b}\|_{L_p^q} \preceq \|\sigma \mathbf{b}\|_{\mathbb{H}_y^{\frac{1}{2},p,q}} + \|\sigma \mathbf{b}\|_{\mathbb{H}_y^{\frac{1}{2},p,q}} \|\mathbf{b}\|_{\mathbb{H}_y^{\frac{1}{2},p,q}}. \tag{3.50}$$

Moreover, for any $\alpha \in (\frac{m+2d}{p}, 1 - \frac{2}{q})$,

$$\sup_{t\in[0,T]} \|\nabla_y (\Xi_\lambda \mathbf{b})_t\|_{\mathbb{H}_\sigma^{\alpha,p}} \preceq \phi(\lambda) \left(\|\sigma \mathbf{b}\|_{\mathbb{H}_y^{\frac{1}{2},p,q}} + \|\sigma \mathbf{b}\|_{\mathbb{H}_y^{\frac{1}{2},p,q}} \|\mathbf{b}\|_{\mathbb{H}_y^{\frac{1}{2},p,q}} \right) \tag{3.51}$$

holds for some decreasing function $\phi : (0,\infty) \to (0,\infty)$ with $\lim_{\lambda\to\infty} \phi(\lambda) = 0$.

Proof. (1) By Lemma 2.3 and Theorem 2.1, there exists a constant $\lambda_0 > 0$ such that for any $\lambda \geqslant \lambda_0$, the equation (3.48) has a unique solution $u^\lambda (=: \Xi_\lambda \mathbf{b})$. By (3.37), we have

$$\begin{aligned}\|\nabla_\sigma \nabla_\sigma \Xi_\lambda \mathbf{b}\|_{L^q_p} &\preceq \|\nabla_{\sigma\mathbf{b}} \Xi_\lambda \mathbf{b} + \sigma \mathbf{b}\|_{L^q_p} \\ &\preceq \|\mathbf{b}\|_{L^q_p} \|\nabla_\sigma \Xi_\lambda \mathbf{b}\|_\infty + \|\sigma \mathbf{b}\|_{L^q_p} \\ &\preceq \|\mathbf{b}\|^2_{L^q_p} + \|\mathbf{b}\|_{L^q_p} \|\sigma \mathbf{b}\|_{L^q_p} + \|\sigma \mathbf{b}\|_{L^q_p},\end{aligned}$$

(2) Let H be the space of measurable functions $u : [0, T] \times \mathbb{R}^{m+d} \to \mathbb{R}^{m+d}$ such that

$$\|u\|_H := \sup_{t\in[0,T]} \left(\|\nabla_\sigma u_t\|_{\mathbb{H}^{\alpha,p}_\sigma} + \|(-\Delta_y)^{\frac14} \nabla_\sigma u_t\|_{\mathbb{H}^{\alpha,p}_\sigma} \right) < \infty.$$

Then H is a Banach space with the norm $\|\cdot\|_H$ defined above. For any $u \in H$, let Φu be the solution to the following equation:

$$\partial_t(\Phi u)_t = (\lambda - \mathscr{L})(\Phi u)_t - \nabla_{\sigma\mathbf{b}_t} u_t - \sigma \mathbf{b}_t, \quad u_T = 0. \tag{3.52}$$

By (2.4), we have $\Phi u = Q_\lambda(\nabla_{\sigma\mathbf{b}} u - \sigma \mathbf{b})$. By Lemma 3.6 with $\beta = \frac12$, we have

$$\|\nabla_{\sigma\mathbf{b}} u\|_{\mathbb{H}^{\frac12,p,q}_y} \preceq \|\mathbf{b}\|_{\mathbb{H}^{\frac12,p,q}_y} \left(\sup_{t\in[0,T]} \|(-\Delta_y)^{\frac14} \nabla_\sigma u_t\|_{\mathbb{H}^{\alpha,p}_\sigma} + \sup_{t\in[0,T]} \|\nabla_\sigma u_t\|_{\mathbb{H}^{\alpha,p}_\sigma} \right). \tag{3.53}$$

Combining this with (3.38) and (3.40), we obtain

$$\|\Phi u\|_H \preceq \phi(\lambda) \|\nabla_{\sigma\mathbf{b}} u + \sigma \mathbf{b}\|_{\mathbb{H}^{\frac12,p,q}_y} \preceq \phi(\lambda) \left(\|\sigma \mathbf{b}\|_{\mathbb{H}^{\frac12,p,q}_y} + \|\mathbf{b}\|_{\mathbb{H}^{\frac12,p,q}_y} \|u\|_H \right) < \infty. \tag{3.54}$$

So, $\Phi u \in H$ for $u \in H$. Moreover, for any u, $\tilde{u} \in H$, (3.52) and (3.54) imply

$$\|\Phi u - \Phi \tilde{u}\|_H \preceq \phi(\lambda) \|\nabla_{\sigma\mathbf{b}}(u - \tilde{u})\|_{\mathbb{H}^{\frac12,p,q}_y} \preceq \phi(\lambda) \|\mathbf{b}\|_{\mathbb{H}^{\frac12,p,q}_y} \|u - \tilde{u}\|_H.$$

Since $\phi(\lambda) \to 0$ as $\lambda \to \infty$, there exists a constant $\lambda_1 > 0$ such that $\phi(\lambda)\|\mathbf{b}\|_{\mathbb{H}^{\frac12,p,q}_y} < \frac12$ for $\lambda \geqslant \lambda_1$. Then by the fixed point theorem, for any $\lambda \geqslant \lambda_1$, the equation (3.48) has a unique solution $\Xi_\lambda \mathbf{b} \in H$. Furthermore, (3.54) implies

$$\|\Xi_\lambda \mathbf{b}\|_H \preceq \phi(\lambda) \|\sigma \mathbf{b}\|_{\mathbb{H}^{\frac12,p,q}_y}, \quad \lambda \geqslant \lambda_1. \tag{3.55}$$

This together with (3.53) gives

$$\|\nabla_{\sigma\mathbf{b}} \Xi_\lambda \mathbf{b}\|_{\mathbb{H}^{\frac12,p,q}_y} \preceq \|\sigma \mathbf{b}\|_{\mathbb{H}^{\frac12,p,q}_y} \|\mathbf{b}\|_{\mathbb{H}^{\frac12,p,q}_y}, \quad \lambda \geqslant \lambda_1. \tag{3.56}$$

Then (3.39) and (3.56) imply

$$\|\nabla_y \nabla_\sigma \Xi_\lambda \mathbf{b}\|_{L^q_p} \preceq \|\nabla_{\sigma\mathbf{b}} u^\lambda + \sigma \mathbf{b}\|_{\mathbb{H}^{\frac12,p,q}_y} \preceq \|\sigma \mathbf{b}\|_{\mathbb{H}^{\frac12,p,q}_y} + \|\sigma \mathbf{b}\|_{\mathbb{H}^{\frac12,p,q}_y} \|\mathbf{b}\|_{\mathbb{H}^{\frac12,p,q}_y}.$$

Similarly, (3.40) and (3.56) yield

$$\begin{aligned}\sup_{t\in[0,T]} \|\nabla_y (\Xi_\lambda \mathbf{b})_t\|_{\mathbb{H}^{\alpha,p}_\sigma} &\preceq \phi(\lambda) \|\nabla_{\sigma\mathbf{b}} u^\lambda + \sigma \mathbf{b}\|_{\mathbb{H}^{\frac12,p,q}_y} \\ &\preceq \phi(\lambda) \left(\|\sigma \mathbf{b}\|_{\mathbb{H}^{\frac12,p,q}_y} + \|\sigma \mathbf{b}\|_{\mathbb{H}^{\frac12,p,q}_y} \|\mathbf{b}\|_{\mathbb{H}^{\frac12,p,q}_y} \right).\end{aligned}$$

Then the proof is finished. $\square$

We are now ready to prove (3.18) and (3.19).

Proof of (3.18) *and* (3.19). We first consider smooth $\mathbf{b}$ then extend to the situation of Theorem 3.1.

(a) Let $\mathbf{b} \in C^\infty([0,T] \times \mathbb{R}^{m+d}, \mathbb{R}^m)$. Then for any $h \in C_0^\infty(\mathbb{R}^{m+d})$, $h\mathbf{b} \in C^\infty([0,T] \times \mathbb{R}^{m+d}, \mathbb{R}^m)$. Applying Theorem 3.8, we obtain that

$$\|\nabla_\sigma \nabla_\sigma \Xi_\lambda h\mathbf{b}\|_{L_p^q} \preceq \|h\mathbf{b}\|_{L_p^q}^2 + \|h\mathbf{b}\|_{L_p^q} \|\sigma h\mathbf{b}\|_{L_p^q} + \|\sigma h\mathbf{b}\|_{L_p^q} \tag{3.57}$$

and

$$\|\nabla_y \nabla_\sigma \Xi_\lambda h\mathbf{b}\|_{L_p^q} \preceq \|\sigma h\mathbf{b}\|_{\mathbb{H}_y^{\frac{1}{2},p,q}} + \|\sigma h\mathbf{b}\|_{\mathbb{H}_y^{\frac{1}{2},p,q}} \|h\mathbf{b}\|_{\mathbb{H}_y^{\frac{1}{2},p,q}} \tag{3.58}$$

for any $\lambda \geqslant \lambda_1$. Next, by Lemma 3.2(4), (3.51), and Theorem 2.1(1),

$$\|\nabla_y(\Xi_\lambda h\mathbf{b})\|_\infty \preceq \phi(\lambda) \left(\|\sigma h\mathbf{b}\|_{\mathbb{H}_y^{\frac{1}{2},p,q}} + \|\sigma h\mathbf{b}\|_{\mathbb{H}_y^{\frac{1}{2},p,q}} \|h\mathbf{b}\|_{\mathbb{H}_y^{\frac{1}{2},p,q}} \right) \tag{3.59}$$

and

$$\|\nabla_\sigma(\Xi_\lambda h\mathbf{b})\|_\infty \leqslant \phi(\lambda)(\|\sigma h\mathbf{b}\|_{L_p^q} + \|h\mathbf{b}\|_{L_p^q}) \tag{3.60}$$

hold for large $\lambda > 0$ and some decreasing function $\phi : (0,\infty) \to (0,\infty)$ with $\lim_{\lambda\to\infty} \phi(\lambda) = 0$.

Moreover, for any $R > 0$, there exists a constant $c(R) > 0$ such that

$$\begin{aligned} |\nabla f|^2(x) &\leqslant c(R) \left(\sum_{i=1}^m |U_i f|^2(x) + \sum_{i=1}^d |\partial_{y_i} f|^2(x) \right) \\ &= c(R) \left(|\nabla_\sigma f|^2(x) + |\nabla_y f|^2(x) \right), \quad |x| \leqslant R, f \in C^1(\mathbb{R}^{m+d}). \end{aligned} \tag{3.61}$$

Combining (3.59)-(3.61), we conclude that for large $\lambda > 0$,

$$\|h\nabla(\Xi_\lambda h\mathbf{b})\|_\infty \leqslant C_{\sigma,h} \|h\|_\infty \phi(\lambda)(\|h\mathbf{b}\|_{\mathbb{H}_y^{\frac{1}{2},p,q}} + \|h\mathbf{b}\|^2_{\mathbb{H}_y^{\frac{1}{2},p,q}}), \tag{3.62}$$

where $C_{\sigma,h} > 0$ is a constant depending on $\mathrm{supp} h$ and $\|\sigma 1_{\mathrm{supp} h}\|_\infty$. Similarly, (3.57), (3.58) and (3.61) imply

$$\|\nabla h \nabla_\sigma(\Xi_\lambda h\mathbf{b})\|_{L_p^q} \leqslant C_{\sigma,h}\Big(\|h\|_\infty + \|h'\|_\infty)(\|h\mathbf{b}\|_{\mathbb{H}_y^{\frac{1}{2},p,q}} + \|h\mathbf{b}\|^2_{\mathbb{H}_y^{\frac{1}{2},p,q}} \Big) \tag{3.63}$$

for large $\lambda > 0$ and some constant $C_{\sigma,h} > 0$ depending on $\mathrm{supp} h$ and $\|\sigma 1_{\mathrm{supp} h}\|_\infty$. Therefore, (3.18) and (3.19) are proved.

(b) Now, assume that for any $h \in C_0^\infty(\mathbb{R}^{m+d})$ we have

$$\|h\mathbf{b}\|_{\mathbb{H}_y^{\frac{1}{2},p,q}} = \|(1-\Delta_y)^{\frac{1}{4}}(h\mathbf{b})\|_{L_p^q} < \infty. \tag{3.64}$$

Let ρ be a non-negative smooth function with compact support in $\mathbb{R}^{m+d}$ and $\int_{\mathbb{R}^{m+d}} \rho(z)\mathrm{d}z = 1$. For any $n \in \mathbb{N}$, let

$$\rho_n(z) = n^{m+d}\rho(nz), \quad \mathbf{b}^n = \rho_n * (h\mathbf{b}), \quad z \in \mathbb{R}^{m+d}. \tag{3.65}$$

Then

$$\lim_{n\to\infty} \|\mathbf{b}^n - h\mathbf{b}\|_{\mathbb{H}_y^{\frac{1}{2},p,q}} = 0.$$

Combining this with (3.62) and (3.63) for $\mathbf{b}^n$ replacing $h\mathbf{b}$, and by an approximation method, we may find out a constant $\lambda_1 > 0$ not depending on n, such that for any $\lambda \geqslant \lambda_1$, the unique solution $u^\lambda (=: \Xi_\lambda h\mathbf{b})$ of (3.48) satisfies

$$\|h\nabla(\Xi_\lambda h\mathbf{b})\|_\infty \preceq (1 + C_{\sigma,h})\|h\|_\infty \phi(\lambda)\big(\|h\mathbf{b}\|_{\mathbb{H}_y^{\frac{1}{2},p,q}} + \|h\mathbf{b}\|^2_{\mathbb{H}_y^{\frac{1}{2},p,q}}\big), \tag{3.66}$$

and

$$\|\nabla h\nabla_\sigma(\Xi_\lambda h\mathbf{b})\|_{L_p^q} \preceq (1 + C_{\sigma,h})(\|h\|_\infty + \|h'\|_\infty)\big(\|h\mathbf{b}\|_{\mathbb{H}_y^{\frac{1}{2},p,q}} + \|h\mathbf{b}\|^2_{\mathbb{H}_y^{\frac{1}{2},p,q}}\big), \tag{3.67}$$

here, $C_{\sigma,h}$, ϕ are in (3.62) and (3.63). Combining these with (3.64), we finish the proof. $\square$

Acknowledgement. The authors would like to thank the referee for corrections and helpful comments.

References

[1] F. Baudoin, N. Garofalo, *A note on the boundedness of Riesz transform for some subelliptic operators,* arxiv:1105,0467.

[2] P. E. Chaudru de Raynal, *Strong existence and uniqueness for stochastic differential equation with Hörlder drift and degenerate noise,* arxiv:1205.6688v4.

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

[4] G. Da Prato, F. Flandoli, *Pathwise uniqueness for a class of SDE in Hilbert spaces and applications,* J. Funct. Anal. 259(2010), 243-267.

[5] G. Da Prato, F. Flandoli, E. Priola, M. Röckner, *Strong uniqueness for stochastic evolution equations in Hilbert spaces perturbed by a bounded measurable drift,* Ann. Probab. 41(2013), 3306-3344.

[6] G. Da Prato, F. Flandoli, E. Priola, M. Röckner, *Strong uniqueness for stochastic evolution equations with unbounded measurable drift term,* J. Theor. Probab. 28(2015), 1571-1600.

[7] G. Da Prato, F. Flandoli, M. Röckner, A. Yu. Veretennikov, *Strong uniqueness for SDEs in Hilbert spaces with non-regular drift,* Ann. Probab. 44(2016), 1985–2023.

[8] D. Bakry, F. Baudoin, M. Bonnefont, D. Chafai, *On gradient bounds for the heat kernel on the Heisenberg group,* J. Funct. Anal. 255(2008), 1905–1938.

[9] F. Baudoin, M. Bonnefont, *Log-Sobolev inequalities for subelliptic operators satisfying a generalized curvature dimension inequality,* J. Funct. Anal. 262(2012), 2646–2676.

[10] F. Baudoin, M. Gordina, T. Melcher, *Quasi-invariance for heat kernel measures on sub-Riemannian infinite-dimensional Heisenberg groups,* Trans. Amer. Math. Soc. 365(2013), 4313–4350.

[11] E. Fedrizzi, F. Flandoli, *Pathwise uniqueness and continuous dependence for SDEs with nonirregular drift,* Stochastics 83(2011), 241-257.

[12] G. B. Folland, *Subelliptic estimates and function spaces on nilpotent Lie groups,* Arkiv För Matematik 13(1975), 161–207.

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

[14] H. Q., Li, *Estimation optimale du gradient du semi-groupe de la chaleur sur le groupe de Heisenberg,* J. Funct. Anal. 236(2006), 369–394.

[15] A. Pazy *Semigroups of linear operators and applications to partial differential equations,* Springer-Verlag, New York, (1983).

[16] E. Priola, *On weak uniqueness for some degenerate SDEs by global L^p estimates,* Potential Anal. 42(2015), 247-281.

[17] E. Priola, *Davies type uniqueness for a class of SDEs with jumps,* arxiv:1509.07448.

[18] K. Saka, *Besov Spaces and Sobolv Spaces on a Nilpotent Lie Group,* Tôhoku Math. Journ. 31 (1979), 383-437.

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

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

[21] F.-Y. Wang, *Derivative formulas and Poincaré inequality for Kohn-Laplacian type semigroups,* Sci. China Math. 59(2016), 261–280.

[22] F.-Y. Wang, X. Zhang, *Degenerate SDEs in Hilbert spaces with rough drifts,* Infin. Dimens. Anal. Quantum Probab. Relat. Top. 18(2015), 1550026, 25 pp.

[23] F.-Y. Wang, X. Zhang, *Degenerate SDE with Hölder-Dini drift and non-Lipschitz noise coefficient,* SIAM J. Math. Anal. 48(2016), 2189-2226.

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

[25] L. Xie, X. Zhang *Ergodicity of stochastic differential equations with jumps and singular coefficients,* arxiv:1705.0742.

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

[27] X. Zhang, *Stochastic Hamiltonian flows with singular coefficients,* Sci. China Math. (2018). https://doi.org/10.1007/s11425-017-9127-0.

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

[29] A.K.Zvonkin, *A transformation of the phase space of a diffusion process that removes the drift,* Mat. Sbornik 93(1974), 129–149.