Abstract. In this paper we develop a new method based on Littlewood-Paley’s decomposition and heat kernel estimates of integral form, to establish Schauder’s estimate for the following degenerate nonlocal equation in $\mathbb{R}^d$ with H"older coefficients:

$$\partial_t u = L^{(\alpha)}_{\kappa,v} u + b \cdot \nabla u + f, \ u_0 = 0,$$

where $u = u(t,x,v)$ and $L^{(\alpha)}_{\kappa,v}$ is a nonlocal $\alpha$-stable-like operator with $\alpha \in (1,2)$ and kernel function $\kappa$, which acts on the variable $v$. As an application, we show the strong well-posedness to the following degenerate stochastic differential equation with H"older drift $b$:

$$dZ_t = b(t,Z_t)dt + (0,\sigma(t,Z_t)dL^{(\alpha)}_t), \ Z_0 = (x,v) \in \mathbb{R}^d,$$

where $L^{(\alpha)}_t$ is a $d$-dimensional rotationally invariant and symmetric $\alpha$-stable process with $\alpha \in (1,2)$, and $b : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ is a $(\gamma,\beta)$-H"older continuous function in $(x,v)$ with $\gamma \in \left(\frac{2\alpha}{2(1+\alpha)}, 1\right)$ and $\beta \in (1 - \frac{\alpha}{2}, 1)$, $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ is a Lipschitz function. Moreover, we also show that for almost all $\omega$, the following random transport equation has a unique $C^1_b$-solution:

$$\partial_t u(t,x,\omega) + (b(t,x) + L^{(\alpha)}_t(\omega)) \cdot \nabla_x u(t,x,\omega) = 0, \ u(0,x) = \varphi(x),$$

where $\varphi \in C^1_b(\mathbb{R}^d)$ and $b : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ is a bounded continuous function in $(t,x)$ and $\gamma$-order H"older continuous in $x$ uniformly in $t$ with $\gamma \in \left(\frac{2\alpha}{2(1+\alpha)}, 1\right)$.

Contents

1. Introduction 2
2. Anisotropic Besov and H"older-Zygmund spaces 5
3. Schauder’s estimates for heat equations 7
4. Commutator estimates 10
5. Heat kernel estimates of nonlocal kinetic operators 15
6. Schauder’s estimates for non-local degenerate equations 17
7. Degenerate SDEs with H"older drifts 26
7.1. Pathwise uniqueness of SDEs with multiplicative Lévy noises 26
7.2. $C^1$-stochastic diffeomorphism flows for SDEs with additive Lévy noise 29
7.3. Application to random transport equations with H"older coefficients 31
8. Appendix 33
References 35

Keywords: Schauder’s estimate, Littlewood-Paley’s decomposition, Heat kernel, Nonlocal kinetic equation, Random transport equation.

Research of X. Zhang is partially supported by NNSFC grant of China (No. 11731009) and the DFG through the CRC 1283 “Taming uncertainty and profiting from randomness and low regularity in analysis, stochastics and their applications”.

1
1. Introduction

Let $T > 0$. Consider the following backward transport equation (a first order PDE):

$$\partial_s u + b \cdot \nabla_x u + f = 0, \quad u(T, x) = \varphi(x), \quad (1.1)$$

where $b(s, x), f(s, x) : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $\varphi(x) : \mathbb{R}^d \rightarrow \mathbb{R}$ are measurable functions. It is a classical fact that if $b, f, \varphi$ are $C^1_b$-functions in $x$ uniformly with respect to $s$, then the above equation is well-posed, and the unique solution is in fact given by

$$u(s, x) = \varphi(X_{s,T}(x)) + \int_s^T f(t, X_{s,t}(x))dt,$$

where for each $x \in \mathbb{R}^d$, $X_{s,t}(x)$ solves the following ordinary differential equation (abbreviated as ODE):

$$\dot{X}_{s,t}(x) = b(t, X_{s,t}(x)), \quad X_{s,s}(x) = x, \quad t \geq s. \quad (1.2)$$

Basing on this representation, DiPerna and Lions [16] developed a well-posed theory for ODE (1.2) when $b \in \mathcal{W}^{1,1}_{loc}$ (the first order Sobolev space) has bounded divergence through studying the transport equation (1.1) (see also [1] for the investigation of ODE (1.2) with BV-vector field $b$). The corresponding results for SDEs are referred to [19], [46], [18] and [48]. It should be noticed that if $b$ is only Hölder continuous, PDE (1.1) would be ill-posed (see [22] for counterexamples). On the other hand, when $b$ is Hölder continuous, under some random perturbations, it was shown in [22] that the following transport equation (an stochastic PDE) is well-posed:

$$du + (b \cdot \nabla_x u)ds + u \circ dW_s = 0, \quad u(T, x) = \varphi(x), \quad (1.3)$$

where $\circ$ stands for the Stratonovich integral, and $W$ is a standard $d$-dimensional Brownian motion on some probability space.

In the same spirit, we consider the following backward heat equation:

$$\partial_s u + \Delta u + b \cdot \nabla_x u + f = 0, \quad u(T, x) = \varphi(x), \quad (1.4)$$

where $\Delta$ is the Laplacian in $\mathbb{R}^d$. When $b, f, \varphi$ are $C^2_b$-functions, the unique solution $u$ still has the following representation:

$$u(s, x) = \mathbb{E}\varphi(X_{s,T}(x)) + \int_s^T \mathbb{E}f(t, X_{s,t}(x))dt,$$

where for each $x \in \mathbb{R}^d$, $X_{s,t}(x)$ is the stochastic flow defined by stochastic differential equation (abbreviated as SDE)

$$dX_{s,t}(x) = b(s, X_{s,t}(x))ds + \sqrt{2}dW_s, \quad X_{s,s}(x) = x, \quad t \geq s. \quad (1.5)$$

Basing on the $L^p$-theory to PDE (1.4), Krylov and Röckner [32] (see also [45], [47], [20]) developed a well-posedness theory for SDE (1.5) with very singular drift $b$, which reveals the regularization effect of Brownian noises. In particular, when $b$ is Hölder continuous, it can be shown that $\{X_{s,t}(x), 0 \leq s < t < \infty, x \in \mathbb{R}^d\}$ defines a $C^1$-stochastic dihfeomorphism flow so that $u(s, x) := \varphi(X_{s,T}(x))$ solves SPDE (1.3) in generalized sense (cf. [33], [22]).

In this work we are concerning with the following degenerate nonlocal equation in $\mathbb{R}^{2d}$:

$$\partial_t u = \mathcal{L}^{(\alpha)}_{k,V} u + b \cdot \nabla u + f, \quad u_0 = 0, \quad (1.6)$$
where $u = u(t, x, v)$ and $\mathcal{L}^{(\alpha)}_{\kappa, v}$ is an $\alpha$-stable-like operator acting on the variable $v$ with the form:

$$
\mathcal{L}^{(\alpha)}_{\kappa, v}u(x, v) := \int_{\mathbb{R}^d} (u(x, v + w) + u(x, v - w) - 2u(x, v)) \frac{\kappa(t, x, v, w)}{|w|^{d+\alpha}} \, dw,
$$

where $\alpha \in (0, 2)$ and $\kappa(t, x, v, w)$ is symmetric in $w$, and $b(t, x, v)$ takes the form

$$
b(t, x, v) = (b^{(1)}(t, x, v), b^{(2)}(t, x, v)) : \mathbb{R}_+ \times \mathbb{R}^{2d} \to \mathbb{R}^{2d}.
$$

Notice that the typical example of equation (1.6) is the following nonlocal kinetic equation:

$$
\partial_t u = \mathcal{K} u + f \text{ with } \mathcal{K} u := \mathcal{L}^{(\alpha)}_{\kappa, v} u + v \cdot \nabla u,
$$

which naturally occurs in the study of spatial inhomogeneous Boltzmann equations (cf. [42]).

The first goal of this paper is to establish the following Schauder’s type estimate for (1.6):

$$
\|u\|_{L^{\infty}([0,T]; C^{(\gamma+\alpha)(1+\alpha);\gamma}_{\kappa,v})} \leq C \|f\|_{L^{\infty}([0,T]; C^{(\gamma+\alpha)(1+\alpha);\gamma}_{\kappa,v})},
$$

where $\alpha \in (1, 2)$ and $\beta \in (0, 1)$, $\gamma \in [\beta, 1 + \alpha)$. Here $b$ and $\kappa$ satisfy some Hölder assumptions (see $H_{\rho, \gamma}^{\alpha, \beta}$ below for precise statement). In PDE’s theory, Schauder’s estimate plays a basic role in constructing the classical solution for quasilinear PDEs. Nowadays, there are many ways to prove such an estimate for heat equations (see [24], [30], [31]). In recent years, Schauder’s estimate for nonlocal equations are also drawn great interests (see [4], [2], [17], [28], [50], etc.). However, most of the works are concentrated on the non-degenerate case. In the degenerate case, Lunardi [35] showed Schauder’s estimate in anisotropic Hölder spaces for linear degenerate Kolmogorov’s equations. Here it is natural to use the anisotropic Hölder spaces due to the feature of multiple scales in different directions. Later, in [34] and [36], the authors established Schauder’s estimate for hypoelliptic Kolmogorov equations with partial nonlinear smooth drifts (corresponding to $b^{(1)}(t, x, v) = v$ in (1.8)). For general variable coefficient $b$ case, to the best of our knowledge, the authors in [6] first establish the sharp Schauder estimate for degenerate nonlinear Kolmogorov equations under some weak Hörmander’s conditions, which in our case corresponds to (1.10) with $\alpha = 2$ and $\gamma = \beta$. Their method is based on complex forward parametrix expansions. We mention that the $L^p$-maximal regularity for degenerate nonlocal Kolmogorov’s equations with constant coefficients was also obtained in [11, 12], [26].

To establish Schauder’s estimate (1.10), we develop a completely new method, that is based on Littlewood-Paley’s decomposition and heat kernel estimates of some integral forms. Roughly speaking, when we consider the usual heat equation, due to Besov’s characterization of Hölder spaces, the key point is the following integral form estimate of the heat kernel (see Lemma 3.1 below): for any $\beta \geq 0$ and some constant $C = C(d, \beta) > 0$,

$$
\int_0^\infty \left( \int_{\mathbb{R}^d} |x|^{\beta} |R_j p_s(x)| \, dx \right) dt \leq C 2^{-2j - \beta j}, \quad \forall t \geq 0, \quad j \in \mathbb{N},
$$

where $R_j$ is the usual block operator in Littlewood-Paley’s decomposition, and $p_s(x)$ is the Gaussian heat kernel. Unlike the usual method by firstly showing Schauder’s estimate for constant coefficient equations, then freezing it for variable coefficient equations, we directly do it for variable coefficient equations by Duhamel’s formula (see Theorem 3.2 below), which looks simpler. Moreover, the advantage of our method is that it provides more flexibility to borrow the spatial regularity of coefficients to compensate the time singularity when we use it to treat the degenerate equation, which allows us to obtain the sharp Schauder estimate (1.10).
Another goal of this paper is to use (1.10) to show the strong well-posedness as well as the $C^1$-stochastic diffeomorphism flow property to degenerate SDEs driven by $\alpha$-stable processes with Hölder drifts. In particular, we shall prove the well-posedness to the following random transport equation with Hölder coefficient:

$$
\partial_t u(t, x, \omega) + (b(t, x) + L^{(\alpha)}_t(\omega)) \cdot \nabla_x u(t, x, \omega) = 0, \quad u(0, x) = \varphi(x),
$$

where $b : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ is a bounded continuous function in $(t, x)$ and $\gamma$-order Hölder continuous function in $x$ uniformly in $t$ with $\gamma \in \left(\frac{2\alpha}{2+\alpha}, 1\right)$ and $\varphi \in C_0^1(\mathbb{R}^d)$. Here $L^{(\alpha)}_t$ is a symmetric and rotationally invariant $\alpha$-stable process with $\alpha \in (1, 2)$. Compared with Flandoli, Gubinelli and Priola’s work [22] for stochastic PDE (1.3), it is a little surprise that as a deterministic equation, PDE (1.11) would be ill-posed for each fixed $\omega$, while in the pathwise sense, random PDE (1.11) could be well-posed for almost all $\omega$ (see Theorem 7.8 below).

In the nondegenerate Brownian diffusion case, as mentioned above, there are numerous works devoted to the studies of strong and weak well-posedness for the SDEs with singular and even distributional drifts (see [32], [45], [20], [47], [49] and references therein). While in the non-degenerate and $\alpha$-stable noise case, recently there are also several works (see [37], [10], [13], [8]) to study this problem, especially for the supercritical case $\alpha \in (0, 1)$, because in this case, from the view point of PDEs, the drift term plays a dominant role. On the other hand, in the degenerate Brownian diffusion case, Chaudru [5] showed a strong uniqueness result for SDEs with Hölder drifts (see also [43], [44]). More recently, Chaudru, Honoré and Menozii [7] applied their Schauder’s estimate [6] to establish the strong uniqueness for a chain of oscillators driven by Brownian noises. However, it seems that there are few works to study the degenerate SDEs with jumps and Hölder drifts.

This paper is organized as follows: In Section 2, we recall the well-known anisotropic Besov and Hölder-Zygmund spaces for later use. In Section 3, we introduce the basic idea of using Littlewood-Paley’s decomposition to establish Schauder’s estimate for heat equations with variable coefficients. In Section 4, we prove several commutator estimates, which plays a crucial role in showing the Schauder estimate (1.10). In Section 5, we give the heat kernel estimate of integral form for nonlocal kinetic operators, which is the basic tool for proving Schauder’s estimate. In Section 6, we prove Schauder’s estimate (1.10) under some natural Hölder’s assumptions on $\kappa$ and $b$ (see Theorem 6.3). In Section 7, we apply the Schauder estimate to the well-posedness of degenerate SDEs with Hölder drifts and also show the well-posedness of a random transport equation with Hölder drift. The key point is to establish the $C^1$-stochastic diffeomorphism flow property to the degenerate SDEs. Finally, in Section 8 we show the existence of smooth solutions for degenerate nonlocal equations with unbounded coefficients by a purely probabilistic argument, which has independent interest. Throughout this paper we shall use the following conventions:

- We use $A \lesssim B$ to denote $A \leq CB$ or some unimportant constant $C > 0$.
- We use $A \approx B$ to denote $C^{-1}B \lesssim A \lesssim CB$ for some unimportant constant $C \geq 1$.
- For any $\varepsilon \in (0, 1)$, we use $A \lesssim \varepsilon B + D$ to denote $A \leq \varepsilon B + C_\varepsilon D$ for some constant $C_\varepsilon > 0$.
- For two operators $\mathcal{A}_1, \mathcal{A}_2$, we use $[\mathcal{A}_1, \mathcal{A}_2] := \mathcal{A}_1\mathcal{A}_2 - \mathcal{A}_2\mathcal{A}_1$ to denote their commutator.
- For a Banach space $B$ and $T > 0$, we denote

$$
L^\infty_T(B) := L^\infty([0, T]; B), \quad L^\infty_{10}(B) := \cap_{T > 0} L^\infty_T(B), \quad L^\infty_T := L^\infty([0, T] \times \mathbb{R}^d).
$$
\[ \bullet \mathbb{N}_0 := \mathbb{N} \cup \{0\}, \mathbb{R}_+ := [0, \infty), a \vee b := \max(a, b), a \wedge b := \min(a, b). \]

2. Anisotropic Besov and Hölder-Zygmund spaces

We first introduce the Hölder (and Hölder-Zygmund) spaces. For \( h \in \mathbb{R}^d \) and \( f : \mathbb{R}^d \to \mathbb{R} \), the first order difference operator is defined by
\[ \delta_h f(x) := f(x + h) - f(x). \]
For \( \beta > 0 \), let \( C^\beta \) be the usual \( \beta \)-order Hölder space consisting of all functions \( f : \mathbb{R}^d \to \mathbb{R} \) with
\[ \| f \|_{C^\beta} := \| f \|_\infty + \cdots + \| \nabla^{[\beta]} f \|_\infty + [\nabla^{[\beta]} f]_{C^{|\beta|-\beta}} < \infty, \]
where \([\beta]\) denotes the greatest integer less than \( \beta \), and \( \nabla^j \) stands for the \( j \)-order gradient, and
\[ [f]_{C^\gamma} := \sup_h \| \delta_h f \|_\infty / |h|^\gamma, \quad \gamma \in [0, 1). \]
The \( \beta \)-order Hölder-Zygmund space \( C^\beta \) is defined by
\[ \| f \|_{C^\beta} := \| f \|_\infty + [f]_{C^\beta} < \infty, \quad [f]_{C^\beta} := \sup_h \| \delta_h^{[\beta]+1} f \|_\infty / |h|^\beta, \]
where for an integer \( m \), \( \delta_h^m := \delta_h \cdots \delta_h \) denotes the \( m \)-order difference operator. Notice that for \( 0 < \beta \in \mathbb{N} \) and \( m \in \mathbb{N} \) (cf. [40]),
\[ \| f \|_{C^\beta} \approx \| f \|_{C^\gamma}, \quad \| f \|_{C^\gamma} \lesssim \| f \|_{C^m}. \quad (2.1) \]
Let
\[ \langle f, g \rangle := \int_{\mathbb{R}^d} f(x)g(x)dx. \]
The adjoint operator of \( \delta_h \) with respect to the above \( \langle \cdot, \cdot \rangle \) is given by
\[ \delta_h^* := -\delta_{-h} \iff \langle \delta_h f, g \rangle = \langle f, \delta_h^* g \rangle. \]
In particular, we have
\[ \delta_h \delta_h^* f(x) = f(x + h) + f(x - h) - 2f(x), \quad (2.2) \]
and for any \( f \in C^2 \),
\[ \| \delta_h^* \delta_h f \|_\infty \leq (2\| \nabla^2 f \|_\infty |h|^2) \wedge (4\| f \|_\infty). \quad (2.3) \]
Let \( \mathcal{S}(\mathbb{R}^d) \) be the Schwartz space of all rapidly decreasing functions on \( \mathbb{R}^d \), and \( \mathcal{S}'(\mathbb{R}^d) \) the dual space of \( \mathcal{S}(\mathbb{R}^d) \) called Schwartz generalized function (or tempered distribution) space. Given \( f \in \mathcal{S}(\mathbb{R}^d) \), the Fourier transform \( \hat{f} \) and inverse Fourier transform \( \hat{f} \) are defined by
\[ \hat{f}(\xi) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}^d} e^{-i\xi \cdot x} f(x)dx, \quad \xi \in \mathbb{R}^d, \]
\[ \hat{f}(x) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}^d} e^{i\xi \cdot x} f(\xi)d\xi, \quad x \in \mathbb{R}^d. \]
Let \( m = (m_1, \cdots, m_n) \in \mathbb{N}^n \) with \( m_1 + \cdots + m_n = d \) and \( a = (a_1, \cdots, a_n) \in [1, \infty)^n \) be fixed. We introduce the following distance in \( \mathbb{R}^d \) by
\[ |x - y|_d := \sum_{i=1}^n |x_i - y_i|^{1/a_i}, \quad x, y \in \mathbb{R}^m. \]
For \( x = (x_1, \cdots, x_n) \), \( t > 0 \) and \( s \in \mathbb{R} \), we denote
\[ t^{sa} x := (t^{s_1} x_1, \cdots, t^{s_n} x_n) \in \mathbb{R}^d, \quad B^a_t := \{ x \in \mathbb{R}^d : |x|_a \leq t \}. \]
Clearly we have
\[ |t^a x_{|a} = t|x_{|a}, \; t > 0. \]
Let \( \phi_0^a \) be a radial \( C^\infty \)-function on \( \mathbb{R}^d \) with
\[ \phi_0^a(\xi) = 1 \; \text{for} \; \xi \in B_1^d \; \text{and} \; \phi_0^a(\xi) = 0 \; \text{for} \; \xi \notin B_2^d. \]
For \( \xi = (\xi_1, \cdots, \xi_n) \in \mathbb{R}^d \) and \( j \in \mathbb{N} \), define
\[ \phi_j^a(\xi) := \phi_0^a(2^{-aj}\xi) - \phi_0^a(2^{-aj-1}\xi). \]
It is easy to see that for \( j \in \mathbb{N} \), \( \phi_j^a(\xi) = \phi_j^a(2^{-aj-1}\xi) \geq 0 \) and
\[ \text{supp} \phi_j^a \subset B_{2^{j+1}} - B_{2^j}, \quad \sum_{j=0}^{k} \phi_j^a(\xi) = \phi_0^a(2^{-ak}\xi) \rightarrow 1, \; k \rightarrow \infty. \]

**Definition 2.1** (Anisotropic Besov and Hölder-Zygmund spaces). For given \( j \in \mathbb{N}_0 \), the block operator \( \mathcal{R}_j^a \) is defined on \( \mathcal{S}'(\mathbb{R}^d) \) by
\[ \mathcal{R}_j^a f(x) := (\phi_j^a \hat{f})(x) = \hat{\phi}_j^a \ast f(x) = 2^{aj} \int_{\mathbb{R}^d} \phi_j^a(2^a(y-x))f(y)dy, \] (2.4)
where \( a \cdot m = a_1m_1 + \cdots + a_nm_n. \) For any \( s \in \mathbb{R} \), the anisotropic Besov space \( \mathcal{B}_{s,\infty}^a \) is defined by
\[ \mathcal{B}_{s,\infty}^a := \left\{ f \in \mathcal{S}'(\mathbb{R}_s^d) : \|f\|_{\mathcal{B}_{s,\infty}^a} := \sup_{j \geq 0} \left( 2^j \|\mathcal{R}_j^a f\|_\infty \right) < \infty \right\}, \]
and for \( s \geq 0 \), the anisotropic Hölder-Zygmund space \( \mathcal{C}_s^a \) is defined by
\[ \mathcal{C}_s^a := \left\{ f \in \mathbb{R}_s^d \rightarrow \mathbb{R} : \|f\|_{\mathcal{C}_s^a} := \|f\|_\infty + [f]_{\mathcal{C}_s^a} < \infty \right\}, \]
where
\[ [f]_{\mathcal{C}_s^a} := \sup_{h} \|\delta^{[s]+1}_h f\|_\infty / |h|_a^s. \]
In particular, if \( a = (1, \cdots, 1) \), we shall drop the index \( a \) in \( \mathcal{B}_{s,\infty}^a, \mathcal{R}_j \) and \( \mathcal{C}_s^a \).

For \( j \in \mathbb{N}_0 \), by definition it is easy to see that
\[ \mathcal{R}_j^a = \mathcal{R}_j^a \mathcal{R}_j^a, \quad \text{where} \; \mathcal{R}_j^a := \mathcal{R}_{j-1}^a + \mathcal{R}_j^a + \mathcal{R}_{j+1}^a \; \text{with} \; \mathcal{R}_{-1}^a \equiv 0, \] (2.5)
and \( \mathcal{R}_j^a \) is symmetric in the sense that
\[ \langle \mathcal{R}_j^a f, g \rangle = \langle f, \mathcal{R}_j^a g \rangle. \]
The cut-off low frequency operator \( S_k \) is defined by
\[ S_k f := \sum_{j=0}^{k-1} \mathcal{R}_j^a f = 2^{a-k} \sum_{j=0}^{k-1} \phi_j^a(2^{ka}(x-y))f(y)dy \rightarrow f. \] (2.6)
For \( f, g \in \mathcal{S}'(\mathbb{R}^d) \), define
\[ T_{fg} = \sum_{k \geq 2} S_{k-1} f \mathcal{R}_k^a g, \quad R(f, g) := \sum_{k \in \mathbb{N}} \sum_{|l| < 1} \mathcal{R}_l^a f \mathcal{R}_k^a g. \]
The Bony decomposition of \( fg \) is formally given by (cf. [3])
\[ fg = T_{fg} + T_{fg} + R(f, g). \] (2.7)
The key point of Bony’s decomposition is
\[ \mathcal{R}_j^a (S_{k-1} f \mathcal{R}_k^a g) = 0 \; \text{for} \; |k - j| > 4. \] (2.8)
Indeed, by Fourier’s transform, we have

\[
(R^a_k(S_{k-1}fR^a_kg)) = \phi_j^a \cdot (\tilde{(\phi_0(2^{(2-k)} \cdot f))} \ast (\phi_j^a g)).
\]

Since the support of \((\phi_0(2^{(2-k)} \cdot f)) \ast (\phi_j^a g)\) is contained in \(B_{2^{k+2}}^a \setminus B_{2^{k-2}}^a\), if \(|k - j| > 4\), then

\[
\phi_j^a \cdot (\tilde{(\phi_0(2^{(2-k)} \cdot f))} \ast (\phi_j^a g)) = 0.
\]

The following result gives the equivalence between \(B_{a,\infty}^s\) and \(C_a^s\) (cf. [41], [14]).

**Theorem 2.2.** For any \(s > 0\), it holds that

\[
\|f\|_{B_{a,\infty}^s} \leq \|f\|_{C_a^s} \leq \|f\|_{C_{a,\infty}^s} + \cdots + \|f\|_{C_{a,\infty}^{s+n}},
\]

where \(\|f\|_{C_{a,j}^{s+n}} := \sup_{A_{i,j} \in \mathbb{R}^d} \|f(x_i, \ldots, x_{i-1}, \cdot, x_{i+1}, \ldots, x_{A_{i,j}})\|_{C_{a,j}^{s+n}}\). By convention we denote \(C_a^s := B_{a,\infty}^s, s < 0\).

We have the following interpolation inequality.

**Corollary 2.3.** For any \(s < r < t\), there is a constant \(C > 0\) such that for any \(\varepsilon \in (0, 1)\),

\[
\|f\|_{C_a^s} \leq C\|f\|_{C_a^r}^{(s-r)/(s-t)}\|f\|_{C_a^t}^{(t-s)/(s-t)} \leq \varepsilon\|f\|_{C_a^r} + C\varepsilon^{(s-r)/(t-r)}\|f\|_{C_a^t}.
\]

**Proof.** By (2.9) and the definition of \(B_{a,\infty}^s\), we have

\[
\|f\|_{C_a^s} \leq \|f\|_{B_{a,\infty}^s} = \sup_{j \geq 0} 2^j \|R^a_j f\|_{\infty} \leq \left( \sup_{j > 0} 2^j \|R^a_j f\|_{\infty} \right)^{(t-r)/(s-t)} \left( \sup_{j > 0} 2^j \|R^a_j f\|_{\infty} \right)^{(r-s)/(t-s)}.
\]

The desired interpolation inequality follows. \(\square\)

3. **Schauder’s estimates for heat equations**

In this section we present the basic idea of proving Schauder’s estimate for heat equations by Littlewood-Paley’s theory. Let \((a^{ij}(t))\) be a measurable symmetric matrix-valued function on \(\mathbb{R}_+\) and satisfy that for some \(c_0 \geq 1\),

\[
c_0^{1} |\xi|^2 \leq a^{ij}(t)\xi_i\xi_j \leq c_0 |\xi|^2, \quad \xi \in \mathbb{R}^d, t \geq 0.
\]

Define for \(0 \leq s < t < \infty\) and \(x \in \mathbb{R}^d\),

\[
p_{s,t}(x) := \frac{e^{-\langle (\overline{A}^d_{s,t})^{-1} \rangle (x,s)/(2(t-s))}}{\sqrt{2\pi \det(A_{s,t})}} = \frac{e^{-\langle (\overline{A}^d_{s,t})^{-1} \rangle (x,s)/(2(t-s))}}{\sqrt{2\pi(t-s)^d \det(A_{s,t})}}.
\]

where \(A_{s,t} := \int_s^t a(r)dr = (t-s) \int_0^1 a(s + (t-s)r)dr =: (t-s)\overline{A}_{s,t}\).

The following lemma is the key observation for Schauder’s estimate of heat equation.

**Lemma 3.1.** Under (3.1), for any \(\beta \geq 0\), there is a constant \(C = C(c_0, \beta, d) > 0\) such that for all \(t \geq 0\) and \(j \in \mathbb{N}\),

\[
\int_0^t \left( \int_{\mathbb{R}^d} |x|^\beta |R_j p_{s,t}(x)|dx \right) ds \leq C2^{-2j-\beta j}.
\]
Proof. We first show that for any $m \in \mathbb{N}_0$ and $\beta \geq 0$, there is a constant $C = C(c_0, \beta, m, d) > 0$ such that for all $0 \leq s < t < \infty$ and $j \in \mathbb{N}$,

$$
\int_{\mathbb{R}^d} |x|^{\beta} |\mathcal{R}_j p_{s,t}(x)| dx \leq C 2^{-2jm}(t-s)^{-m}(2^{-j} + (t-s)^{1/2})^\beta. \tag{3.4}
$$

Recalling (2.4) and by the change of variables, we have

$$
\int_{\mathbb{R}^d} |x|^{\beta} |\mathcal{R}_j p_{s,t}(x)| dx = 2^{-jd-j\beta} \int_{\mathbb{R}^d} |x|^{\beta} \left| \int_{\mathbb{R}^d} p_{s,t}(2^{-j}(x-y)) \phi_1(y) dy \right| dx.
$$

Since the support of $\phi_1$ is contained in the annulus, by Fourier’s transform we have

$$
\int_{\mathbb{R}^d} p_{s,t}(2^{-j}(x-y)) \phi_1(y) dy = \int_{\mathbb{R}^d} \Delta^m p_{s,t}(2^{-j}(x-y)) \cdot \Delta^m \phi_1(y) dy, \quad m \in \mathbb{N}_0,
$$

where $\Delta^m \phi_1 := \langle |\xi|^m \phi_1(\xi) \rangle$. Moreover, by (3.2) and elementary calculations, we have

$$
2^{-j\beta-jd} \int_{\mathbb{R}^d} |x|^{\beta} |\Delta^m p_{s,t}(2^{-j})(x)| dx \leq C 2^{-2jm}(t-s)^{\beta/2-m}.
$$

Hence,

$$
\int_{\mathbb{R}^d} |x|^{\beta} |\mathcal{R}_j p_{s,t}(x)| dx \leq 2^{-jd-j\beta} \int_{\mathbb{R}^d} |x|^{\beta} |\Delta^m p_{s,t}(2^{-j}x)| dx \int_{\mathbb{R}^d} |\Delta^m \phi_1(y)| dy
$$

$$
+ 2^{-jd-j\beta} \int_{\mathbb{R}^d} |\Delta^m p_{s,t}(2^{-j}x)| dx \int_{\mathbb{R}^d} |y|^{\beta} |\Delta^m \phi_1(y)| dy
$$

$$
\leq 2^{-2jm}(t-s)^{\beta/2-m} + 2^{-j\beta-2jm}(t-s)^{-m},
$$

which in turn gives (3.4).

Let $\mathcal{J}$ be the left hand side of (3.3). We make the following decomposition:

$$
\mathcal{J} = \left( \int_{t^{-1/2} - 1/2} \int_0 \right) \left( \int_{\mathbb{R}^d} |x|^{\beta} |\mathcal{R}_j p_{s,t}(x)| dx \right) ds =: \mathcal{J}_1 + \mathcal{J}_2.
$$

For $\mathcal{J}_1$, by (3.4) with $m = 0$, we have

$$
\mathcal{J}_1 \leq \int_{t^{-1/2} - 1/2} \left( (2^{-j} + (t-s)^{-1/2})^\beta \right) ds = \int_0^{t^{-1/2} - 1/2} \left( 2^{-j} + s^{1/2} \right)^\beta ds \leq 2^{-j\beta j}.
$$

For $\mathcal{J}_2$, by (3.4) with $m = 2$, we have

$$
\mathcal{J}_2 \leq \int_0^{t^{-1/2} - 1/2} 2^{-4j}(t-s)^{-2} \left( 2^{-j} + (t-s)^{-1/2} \right)^\beta ds = 2^{-4j} \int_0^{t^{-1/2} - 1/2} s^{-2\beta/2} ds \leq 2^{-2j\beta j}.
$$

Combining the above two estimates, we obtain (3.3). \hfill \Box

Now we consider the following heat equation with variable coefficients:

$$
\partial_t u = a^{ij}(x) \partial_{ij} u + f, \quad u(0) = 0, \tag{3.5}
$$

where $a : \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}^d \otimes \mathbb{R}^d$ is a measurable symmetric matrix-valued function and satisfies (H$^0$) For some $c_0 \geq 1$ and $\beta \in (0, 1)$, it holds that for all $t \geq 0$ and $x, y, \xi \in \mathbb{R}^d$,

$$
c_0^{-1} |\xi|^2 \leq a^{ij}(t,x) \xi_i \xi_j \leq c_0 |\xi|^2, \quad |a(t,x) - a(t,y)| \leq c_0 |x-y|^\beta.
$$

Below we use Lemma 3.1 to establish Schauder’s estimate for heat equation (3.5).
Theorem 3.2. Let $\beta \in (0, 1)$. Under $\mathbf{(H^\beta)}$, there is a constant $C = C(c_0, \beta, d) > 0$ such that for any $T > 0$, and $u \in L^\infty_T(B^\beta_{\infty\infty})$ with $\partial_t u \in L^\infty_T(B^\beta_{\infty\infty})$ solving PDE (3.5),

$$
\|u\|_{L^\infty_T(C^{2,\beta})} \leq C(\|f\|_{L^\infty_T(C^0)} + \|u\|_{L^\infty_T}).
$$

Proof. Fix $x_0 \in \mathbb{R}^d$ and define

$$
u_{x_0}(t, x) := u(t, x + x_0), \quad \tilde{\nu}_{x_0}(t, x) := a(t, x + x_0) - a(t, x_0).$$

It is easy to see that

$$
\partial_t \nu_{x_0} = \partial^{i}(t, x_0)\partial_j \nu_{x_0} + \tilde{\nu}_{x_0} \partial_j \nu_{x_0} + f_{x_0}, \quad \nu_{x_0}(0) = 0.
$$

Let $p^\nu_{s,t}$ be defined by (3.2) in terms of $a(t, x_0)$. For a space-time function $f$, define

$$
P^\nu_{s,t}f(s, x) := \int_{\mathbb{R}^d} p^\nu_{s,t}(x - y)f(s, y)dy.
$$

By Duhamel’s formula we have

$$
u_{x_0}(t, x) = \int_0^t p^\nu_{s,t} \text{tr}(\tilde{\nu}_{x_0} \cdot \nabla^2 \nu_{x_0})(s, x)ds + \int_0^t P^\nu_{s,t}f(s, x)ds =: I_1(t, x) + I_2(t, x).
$$

Below, without loss of generality, we assume $x_0 = 0$ and drop the subscript and superscript $x_0$. First of all, for $I_1(t, x)$, by $\mathbf{(H^\beta)}$ and Lemma 3.1, we have

$$
|\mathcal{R}_t I_1(t, 0)| \leq \int_0^t \|\mathcal{R}_t p^\nu_{s,t} \text{tr}(\tilde{\nu} \cdot \nabla^2 u)(s, 0)\|ds \leq \int_0^t \left( \int_{\mathbb{R}^d} |x|^\beta \mathcal{R}_t p^\nu_{s,t}(x)dx \right)ds \|\nabla^2 u\|_{L^\infty_T} \leq 2^{-2j-\beta} \|\nabla^2 u\|_{L^\infty_T} \leq \varepsilon 2^{-2j-\beta} ||u||_{L^\infty_T(B^\beta_{\infty\infty})} + 2^{-2j-\beta} ||u||_{L^\infty_T},
$$

where $\varepsilon > 0$ and the last inequality is due to the interpolation and Young’s inequalities. For $I_2(t, x)$, by (2.5) and Lemma 3.1 again, we have

$$
|\mathcal{R}_t I_2(t, 0)| \leq \int_0^t \|\mathcal{R}_t p^\nu_{s,t} f(s, 0)\|ds = \int_0^t \left( \int_{\mathbb{R}^d} |x|^\beta \mathcal{R}_t p^\nu_{s,t}(y) f(s, y)dy \right)ds \leq 2^{-2j-\beta} ||f||_{L^\infty_T(B^\beta_{\infty\infty})}.
$$

Combining the above estimates, we obtain that for any $\varepsilon \in (0, 1)$ and $j \in \mathbb{N}$,

$$2^{2j+\beta} |\mathcal{R}_t \mu(t, x_0)| = 2^{2j+\beta} |\mathcal{R}_t \mu_{x_0}(t, 0)| \leq \varepsilon ||u||_{L^\infty_T(B^\beta_{\infty\infty})} + ||u||_{L^\infty_T} + ||f||_{L^\infty_T(B^\beta_{\infty\infty})}, \quad \text{(3.6)}
$$

Moreover, for $j = 0$, it is easy to see that

$$|\mathcal{R}_t \mu(t, x_0)| \leq ||u||_{L^\infty_T}.
$$

Thus by the definition of Besov space, we arrive at

$$
||u||_{L^\infty_T(B^\beta_{\infty\infty})} = \sup_{t \in [0, T]} \sup_{j \in \mathbb{N}_0} 2^{j+\beta} |\mathcal{R}_t \mu(t, \cdot)|_{\infty} \leq \varepsilon ||u||_{L^\infty_T(B^\beta_{\infty\infty})} + C |u||_{L^\infty_T} + C ||f||_{L^\infty_T(B^\beta_{\infty\infty})},
$$

which gives the desired estimate by choosing $\varepsilon = 1/2$ and Theorem 2.2. □
4. Commutator estimates

In the sequel, we shall only consider the following case of anisotropic Besov spaces:

\[ n = 2, \ m_1 = m_2 = d, \ a = (1 + \alpha, 1), \text{ where } \alpha \in (0, 2). \]

For \( h \in \mathbb{R}^d \) and \( f(x, v) : \mathbb{R}^{2d} \to \mathbb{R} \), we introduce

\[
\delta_{h,1} f(x, v) := \delta_h f(\cdot, v)(x), \ \delta_{h,2} f(x, v) := \delta_h f(x, \cdot)(v),
\]

\[
\mathcal{R}^f_j(x, v) := \mathcal{R}_j f(\cdot, v)(x), \ \mathcal{R}^f_j(x, v) := \mathcal{R}_j f(x, \cdot)(v),
\]

and for \( \beta > 0 \),

\[
\|f\|_{C^\beta_x} := \sup_{v \in \mathbb{R}^d} \|f(\cdot, v)\|_{C^\beta_v}, \quad \|f\|_{C^\beta_v} := \sup_{x \in \mathbb{R}^d} \|f(x, \cdot)\|_{C^\beta_v},
\]

\[
\|f\|_{B_2^{\alpha,\beta}} := \sup_{v \in \mathbb{R}^d} \|f(\cdot, v)\|_{B_2^\alpha}, \quad \|f\|_{B_2^{\beta}} := \sup_{x \in \mathbb{R}^d} \|f(x, \cdot)\|_{B_2^\beta}.
\]

Moreover, for \( \beta \in (0, 1) \), we introduce the following semi-norm for later use:

\[
\|f\|_{C^{1+\beta}_v} := \|f\|_{C^{1+\beta}_v} = \|\nabla_v f\|_{C^\beta_v}.
\] (4.1)

For \( \gamma, \beta > 0 \), we define the mixed norm

\[
C^\gamma_x C^\beta_v := \left\{ f(x, v) : \|f\|_{C^\gamma_x} := \|f\|_{C^\gamma_x} + \sup_{\alpha_h, \beta} \|\delta_{h,1} f(\alpha_h, v)\|_{C^\gamma_v} < \infty \right\},
\]

and for \( \gamma \in \mathbb{R} \) and \( \beta > 0 \),

\[
B^{\gamma,\beta}_x C^\beta_v := \left\{ f(x, v) : \|f\|_{B^{\gamma,\beta}_x} := \sup_{j \in \mathbb{N}_0} 2^j \|\mathcal{R}^f_j f\|_{C^\beta_v} < \infty \right\}.
\]

In particular, by Theorem 2.2, we have for \( \gamma \in \mathbb{R} \) and \( \beta > 0 \),

\[
\sup_{j \in \mathbb{N}_0} 2^{j(1+\alpha)} \|\mathcal{R}^f_j f\|_{C^\beta_v} \leq \sup_j 2^{j(1+\alpha)} \|\mathcal{R}^f_j f\|_{C^{1+\beta}_v} = \|f\|_{B^{\gamma,\beta}_x} + \|f\|_{B^{\gamma,\beta}_x}.
\] (4.2)

and for \( \gamma > 0 \) ans \( \beta > 0 \),

\[
B^{\gamma,\beta}_x C^\beta_v \simeq C^\gamma_x C^\beta_v.
\]

We list some easy properties for later use.

**Lemma 4.1.** (i) For any \( \theta \in [0, 1] \) and \( \beta, \gamma > 0 \), it holds that for some \( C = C(\theta, \gamma, \beta) > 0 \),

\[
\|f\|_{C^{1+\theta}_v} \leq C \|f\|_{C^\beta_v} \|f\|_{C^\gamma_v}^{1-\theta}.
\] (4.3)

(ii) For all \( j \in \mathbb{N}_0 \), it holds that for some \( C = C(\alpha) > 0 \),

\[
\|\nabla_v \mathcal{R}^f_j f\|_{C^\beta_v} \leq C 2^{(1+\alpha)j} \|\mathcal{R}^f_j f\|_{C^\beta_v}, \quad \|\nabla_v \mathcal{R}^f_j f\|_{C^\beta_v} \leq C 2^{j} \|\mathcal{R}^f_j f\|_{C^\beta_v}.
\] (4.4)

(iii) For any \( \beta \in (0, 2) \), it holds that for some \( C = C(\alpha, \beta) > 0 \),

\[
\|\mathcal{R}^f_j f\|_{C^\beta_v} \leq C 2^{-\beta j} \|f\|_{C^\beta_v}, \quad j \geq 1.
\] (4.5)

(iv) For any \( \beta \in (0, 1 \wedge \alpha) \), it holds that for some \( C = C(\alpha, \beta) > 0 \),

\[
\|\nabla_v f\|_{C^\beta_v} \leq C \|f\|_{C^{1+\beta}_v},
\] (4.6)

where \( \|f\|_{C^{1+\beta}_v} \) is defined by (4.1).
Proof. (i) Notice that
\[ \|\delta|^{[y]+1}\delta|^{[y]+1}_{\mathcal{H}_2^2} f\|_{\infty} \leq \|f\|_{C^0_d}|h|^\gamma, \quad \|\delta|^{[y]+1}\delta|^{[y]+1}_{\mathcal{H}_2^2} f\|_{\infty} \leq \|f\|_{C^0_d}|h|^\beta. \]
Hence,
\[ \|\delta|^{[y]+1}\delta|^{[y]+1}_{\mathcal{H}_2^2} f\|_{\infty} \leq \|f\|_{C^0_d}\|f\|_{C^0_d}^{1-\theta}\|h\|^\theta[|h|^{(1-\theta)\beta}. \]
From this we obtain the desired estimate (4.3).

(ii) It is a direct consequence of (2.4).

(iii) Noticing that for \( j \geq 1 \),
\[ \mathcal{R}_j^f(x) = \int_{\mathbb{R}^d} \hat{\phi}_j^f(h)f(x+h)dh, \quad \int_{\mathbb{R}^d} \hat{\phi}_j^f(h)dh = \phi_j^f(0) = 0, \]
by (2.2) and \( \hat{\phi}_j^f(-h) = \hat{\phi}_j^f(h) \), we have
\[ \mathcal{R}_j^f(x) = \frac{1}{2} \int_{\mathbb{R}^d} \hat{\phi}_j^f(h)\delta_j^\beta_\delta f(x)dh. \]
Hence,
\[ \|\mathcal{R}_j^f\|_{\infty} \leq \frac{1}{2} \sup_h \|\delta_j^\beta_\delta f\|_{\mathcal{H}_2^2}/|h|^{\beta_\delta} \int_{\mathbb{R}^d} \hat{\phi}_j^f(h)|h|^\beta_\delta dh \leq 2^{-\beta_\delta}[f]_{\mathcal{C}_3}. \]

(iv) By Theorem 2.2 and definition, we have
\[ \|\nabla_v f\|_{\mathcal{E}_u} \leq \|\nabla_v f\|_{\infty} + \sup_{j \in \mathbb{N}} 2^{\beta_\delta j}\|\mathcal{R}_j^\delta f\|_{\infty} \leq \|\nabla_v f\|_{\infty} + \sup_{j \in \mathbb{N}} 2^{1+\beta_\delta j}\|\mathcal{R}_j^\delta f\|_{\infty} \]
\[ \leq \|\nabla_v f\|_{\infty} + [f]_{\mathcal{C}_1} \leq \|\nabla_v f\|_{\infty} + [f]_{\mathcal{C}_1} + [f]_{\mathcal{C}_1} \leq \|f\|_{\mathcal{C}_1}. \]
The proof is complete. \( \square \)

We now show several commutator estimates, which are extensions of [13, Lemma 2.3], and will play a key role in showing the Schauder estimate below.

Lemma 4.2. (i) For any \( \beta \in (0, \alpha) \) and \( \gamma \in (-1-\beta, 0] \), there is a constant \( C = C(d, \beta, \gamma) > 0 \) such that for all \( x, v \in \mathbb{R}^d \) and \( j \geq 5 \),
\[ \|\mathcal{R}_j^f, \tilde{f}\|g(x, v) \leq C 2^{-j(\beta+1)}(2^{-j\beta} + |x|^\beta + |v|^\beta)\|f\|_{\mathcal{C}_1}\|g\|_{\mathcal{C}_3}, \]
where \( \|f\|_{\mathcal{C}_1} \) is defined by (4.1), and
\[ \tilde{f}(x, v) := f(x, v) - f(0, 0) - v \cdot \nabla_v f(0, 0). \]

(ii) For any \( \beta \in (0, 1) \) and \( \gamma \in (-\beta, 0] \), there is a constant \( C = C(d, \beta, \gamma) > 0 \) such that
\[ \|\mathcal{R}_j^f, f\|g(x, v) \leq C 2^{-j(\beta+\gamma)}[f]_{\mathcal{C}_1}\|g\|_{\mathcal{C}_1}, \quad j \geq 5. \]

Proof. We only prove (i) since (ii) is similar and easier. First of all, we have
\[ |\tilde{f}(x, v) - \tilde{f}(x_v)| = |f(x, v) - f(x, v) + (v - 0) \cdot \nabla_v f(0, 0)| \]
\[ \leq |x - \tilde{x}|^{\beta_\delta} [f]_{\mathcal{C}_1} + |v - \tilde{v}|(|x|^\beta + |v| + |v|^\beta)\|\nabla_v f\|_{\mathcal{E}_u} \]
\[ \leq (|x - \tilde{x}|^{\beta_\delta} + |v - \tilde{v}|(|x|^\beta + |v| + |v|^\beta))\|f\|_{\mathcal{C}_1}, \]

11
where the last step is due to (4.6). Since for \( \gamma = 0 \), (4.7) is easily derived from the above estimate, we assume \( \gamma < 0 \) below. Noting that by (2.6),

\[
S_{k-1} \tilde{f}(x, v) = 2^{(2+\alpha)kd} \int_{\mathbb{R}^{2d}} \tilde{\phi}_0'(2^{(1+\alpha)k}x', 2^k v') \tilde{f}(x - x', v - v') dx' dv',
\]

we have

\[
|S_{k-1} \tilde{f}(x, v) - S_{k-1} \tilde{f}(x, v)| \leq (|x - \bar{x}|^{4+\alpha} + |v - \bar{v}|(2^{-k\beta} + |x|^{1+\alpha} + |v|^\beta)) \|f\|_{C^{1+\beta}}.
\]

(4.9)

On the other hand, noting that

\[
[R^a_j, S_{k-1} \tilde{f}]g(x, v) = \int_{\mathbb{R}^{2d}} \tilde{\phi}_0'(x - \bar{x}, v - \bar{v})(S_{k-1} \tilde{f}(\bar{x}, \bar{v}) - S_{k-1} \tilde{f}(x, v))g(\bar{x}, \bar{v}) dx dv,
\]

we have by (4.9) and (2.4),

\[
||[R^a_j, S_{k-1} \tilde{f}]g||_{C^{1+\beta}} \leq \left(2^{-j(1+\beta)} + 2^{-j(2^{-k\beta} + |x|^{1+\beta} + |v|^\beta))} \|f\|_{C^{1+\beta}}\|g\|_{\infty}.
\]

(4.10)

Now by using Bony’s decomposition (2.7), we can write

\[
[R^a_j, \tilde{f}]g = [R^a_j, T_j]g + R^a_j(T_{-j} \tilde{f}) - T_{-j}R^a_j \tilde{f} + R^a_j R(\tilde{f}, g) - R(\tilde{f}, R^a_j g).
\]

For the first term, by (2.8) we have

\[
||[R^a_j, T_j]g||_{C^{1+\beta}} \leq \sum_{|k - j| \leq 4} \|R^a_j(S_{k-1} \tilde{f} R^a_k g) - S_{k-1} \tilde{f} R^a_k R^a_j g||_{C^{1+\beta}} \leq \sum_{k < j \leq 4} \|R^a_j(S_{k-1} \tilde{f} R^a_k g)||_{\infty}
\]

(4.10)

\[
\leq \sum_{k < j \leq 4} \left(2^{-j(1+\beta)} + 2^{-j(2^{-k\beta} + |x|^{1+\beta} + |v|^\beta))} \|f\|_{C^{1+\beta}}\|g\|_{\infty}
\]

\[
\leq 2^{-j(1+\beta)} + 2^{-j(2^{-k\beta} + |x|^{1+\beta} + |v|^\beta))} \|f\|_{C^{1+\beta}}||g||_{B^{1+\beta}_\infty}.
\]

On the other hand, we also have

\[
||R^a_j(T_{-j} \tilde{f})||_{\infty} \leq \sum_{|k - j| \leq 4} \|R^a_j(S_{k-1} g R^a_k \tilde{f})||_{\infty} \leq \sum_{k < j \leq 4} \|R^a_j(S_{k-1} g R^a_k \tilde{f})||_{\infty}
\]

\[
\leq \sum_{k < j \leq 4} \|S_{k-1} g R^a_k \tilde{f}||_{\infty} \leq \sum_{k < j \leq 4} \sum_{m < k - 2} \|R^a_m g||_{\infty} \|R^a_k \tilde{f}||_{\infty}.
\]

Since \( j \geq 5 \) and \( |k - j| \leq 4 \), by (4.5), we further have

\[
||R^a_j(T_{-j} \tilde{f})||_{\infty} \leq ||g||_{B^{1+\beta}_\infty} ||f||_{C^{1+\beta}} \sum_{|k - j| \leq 4} \sum_{m < k - 2} 2^{-m(1+\beta)} \leq ||g||_{B^{1+\beta}_\infty} ||f||_{C^{1+\beta}} 2^{-j(1+\beta)},
\]

where the last step is due to \( \gamma < 0 \) and \( ||f||_{C^{1+\beta}} \leq ||f||_{C^{1+\beta}} \). Similarly,

\[
||T_{-j} R^a_j \tilde{f}||_{\infty} \leq \sum_{k > j - 2} \|S_{k-1} g R^a_k \tilde{f}||_{\infty} \leq \sum_{k > j - 2} \|S_{k-1} g R^a_k \tilde{f}||_{\infty}
\]

\[
\leq \sum_{k > j - 2} 2^{-k(1+\beta)} ||f||_{C^{1+\beta}} \leq 2^{-j(1+\beta)} ||f||_{C^{1+\beta}} ||g||_{B^{1+\beta}_\infty}.
\]

Finally, since \( 1 + \beta + \gamma > 0 \), we have

\[
||R^a_j R(\tilde{f}, g)||_{\infty} \leq \sum_{|i| \leq 1, k \geq j - 4} \|R^a_j(R^a_i \tilde{f} R^a_k g)||_{\infty} \leq \sum_{|i| \leq 1, k \geq j - 4} \|R^a_i \tilde{f}||_{\infty} \|R^a_k g||_{\infty}
\]

\[
\leq \sum_{k \geq j - 4} 2^{-k(1+\beta)} ||f||_{C^{1+\beta}} \leq 2^{-j(1+\beta)} ||f||_{C^{1+\beta}} ||g||_{B^{1+\beta}_\infty},
\]

\[12\]
\[ \|R(\tilde{f}, R^d g)\|_\infty = \left\| \sum_{|\ell| < 1, |\beta| < 1} R_{\ell}^d \tilde{f} R_{\ell}^d R_{\ell}^d g \right\|_\infty \leq [f]_{C_1^{\beta, \gamma}} [g]_{B_1^{2\gamma}} 2^{-\beta(1+\beta+\gamma)}. \]

Combining the above calculations, we complete the proof. □

**Lemma 4.3.** For any \(0 < \beta \leq \gamma < 1\) and \(\eta \in (-\gamma, 0)\), there is a \(C = C(\beta, \gamma, \eta) > 0\) such that
\[ \| [R_{\ell}^d R^f g] \|_c \leq C2^{(\beta-\gamma-\eta)}[f]_{C_1'} [g]_{C_1''}, \quad j \geq 5. \] (4.11)

**Proof.** If \(\ell \leq j + 1\), then by (4.8),
\[ \| R_{\ell}^d [R_{\ell}^d, f] g \|_c \leq \| [R_{\ell}^d, f] g \|_c \leq 2^{-\beta[\eta]} [f]_{C_1'} [g]_{C_1''} \leq 2^{-\beta[\eta]} 2^{(\beta-\gamma-\eta)}[f]_{C_1'} [g]_{C_1''}. \]
If \(\ell > j + 1\), since \(\hat{R}_{\ell}^d \hat{R}_{\ell}^d \hat{f} = \phi_{\ell} \hat{\phi}_{\ell} \hat{f} \equiv 0\), we have
\[ R_{\ell}^d [R_{\ell}^d, f] g = R_{\ell}^d R_{\ell}^d (f R_{\ell}^d g) - R_{\ell}^d (f R_{\ell}^d g) - f R_{\ell}^d R_{\ell}^d g = -[R_{\ell}^d, f] R_{\ell}^d g. \]
Thus by (4.8) again, we have
\[ \| R_{\ell}^d [R_{\ell}^d, f] g \|_c \leq \| R_{\ell}^d [R_{\ell}^d, f] g \|_c \leq 2^{-\gamma}[f]_{C_1'} [R_{\ell}^d] g \|_c \leq 2^{-\gamma}[f]_{C_1'} [g]_{C_1''}. \]
Hence,
\[ \| [R_{\ell}^d R^f g] \|_c \leq \sup_{\ell \in N_0} 2^{\beta[\eta]} \| R_{\ell}^d [R_{\ell}^d, f] g \|_c \leq 2^{(\beta-\gamma-\eta)[f]_{C_1'} [g]_{C_1''}.} \]
The proof is complete. □

**Lemma 4.4.** Let \(\beta, \gamma_2, \theta \in (0, 1)\) and \(\gamma, \gamma_1 \in (0, 1 + \alpha)\). Under the conditions
\[ \gamma \lor \gamma_2 < \gamma_1, \quad \theta \gamma_2 < \gamma + \beta \leq (1-\theta) \gamma_1 + \theta \gamma_2, \quad \beta \leq \theta \gamma_2, \] (4.12)
there is a constant \(C > 0\) such that for all \(j \geq 5\),
\[ \| [R_{\ell}^d, f] g \|_{c_1} \leq 2^{-\beta[\eta]} \left( [f]_{C_1^{\gamma_1/(1+\alpha)}} + [f]_{C_1''} \right) \left( \|g\|_{C_1^{\gamma_1/(1+\alpha)}} + \|g\|_{C_1''} \right). \] (4.13)

**Proof.** First of all, by applying (4.11) with \((\beta, \gamma, \gamma, \gamma, \gamma_2, \theta \gamma_2)\) in place of \((\beta, \gamma, \gamma, \beta, \gamma, \gamma_2)\), we have
\[ \| [R_{\ell}^d, f] g \|_{c_1} \leq 2^{-\beta[\eta]} \left( [f]_{C_1^{\gamma_1/(1+\alpha)}} \right) \|g\|_{c_1^{\gamma_1/(1+\alpha)}}. \]
Thus, by definition it suffices to prove
\[ \| [R_{\ell}^d, f] g \|_{c_1} \leq \sup_{\ell \in N_0} 2^{-\beta[\eta]} \| R_{\ell}^d [R_{\ell}^d, f] g \|_c \leq \text{RHS of (4.13)}. \] (4.14)

(Case: \(\ell \leq \frac{j}{1+\alpha}\)). Since \(\gamma_2 < \gamma_1\) and \(\gamma + \beta < (1-\theta) \gamma_1 + \theta \gamma_2 \leq \gamma_1\), by (4.8), we have
\[ \| R_{\ell}^d [R_{\ell}^d, f] g \|_c \leq \| R_{\ell}^d [R_{\ell}^d, f] g \|_c \leq 2^{-\beta[\eta]} [f]_{C_1^{\gamma_1/(1+\alpha)}} \|g\|_{c_1^{\gamma_1/(1+\alpha)}} \leq 2^{-\beta[\eta]} 2^{-\beta[\gamma]} [f]_{C_1^{\gamma_1/(1+\alpha)}} \|g\|_{c_1^{\gamma_1/(1+\alpha)}}. \]

(Case: \(\ell > \frac{j}{1+\alpha}\)). Notice that
\[ R_{\ell}^d [R_{\ell}^d, f] g = [R_{\ell}^d, f] R_{\ell}^d g + [R_{\ell}^d, [R_{\ell}^d, f] g] =: I_1 + I_2. \]
For \(I_1\), since \(\gamma < \gamma_1\), by (4.8), we have
\[ |I_1| \leq 2^{-\beta[\eta]} [f]_{C_1^{\gamma_1/(1+\alpha)}} \|g\|_{c_1^{\gamma_1/(1+\alpha)}} \leq 2^{-\beta[\eta]} 2^{-\beta[\gamma]} [f]_{C_1^{\gamma_1/(1+\alpha)}} \|g\|_{c_1^{\gamma_1/(1+\alpha)}}. \]
For \(I_2\), by definition, (4.12) and (4.8), we have
\[ |I_2| = \left| \int_{\mathbb{R}^d} \hat{\phi}_{\ell}(\tilde{v})[R_{\ell}^d, \delta_2 f(\cdot, v)] g(\cdot, v - \tilde{v}) d\tilde{v} \right| \]
\[
\leq 2^{-\frac{y \beta - y \varphi_2}{1+\alpha}} \int_{\mathbb{R}^d} \delta_t(\tilde{\varphi})[\delta_{v,2} f]_{c_{x}^{(1-\theta)\gamma_1/(1+\alpha)}} \|g\| \|c_{x}^{(y \beta - (1-\theta)\gamma_2)/(1+\alpha)} d\tilde{\varphi}
\]
\[
\leq 2^{-\frac{y \beta - y \varphi_2}{1+\alpha}} \left( \int_{\mathbb{R}^d} \delta_t(\tilde{\varphi})|\tilde{\varphi}|^{y \varphi_2} d\tilde{\varphi} \right) [f]_{c_{x}^{(1-\theta)\gamma_1/(1+\alpha)}} [\delta]_{c_{x}^{(1-\theta)\gamma_2/(1+\alpha)}},
\]
\[
\leq 2^{-\frac{y \beta - y \varphi_2}{1+\alpha}} 2^{-y \varphi_2 \ell} [f]_{c_{x}^{(1-\theta)\gamma_1/(1+\alpha)}} [\delta]_{c_{x}^{(1-\theta)\gamma_2/(1+\alpha)}},
\]
\[
\leq 2^{-\frac{y \beta - y \varphi_2}{1+\alpha}} \ell 2^{-y \varphi_2} (|f|_{c_{x}^{(1-\theta)\gamma_1/(1+\alpha)}} + |f|_{c_{x}^{\gamma}}) [\delta]_{c_{x}^{(1-\theta)\gamma_2/(1+\alpha)}},
\]

Hence, for all \( \ell \in \mathbb{N}_0 \) and \( j \geq 5 \),
\[
\|R_x^{(\ell)}[R_j, f]g\|_{\infty} \leq \|R_x^{(\ell)}[R_j, f]g\|_{\infty} \leq 2^{-\frac{y \beta - y \varphi_2}{1+\alpha}} 2^{-y \varphi_2} (|f|_{c_{x}^{(1-\theta)\gamma_1/(1+\alpha)}} + |f|_{c_{x}^{\gamma}}) [\delta]_{c_{x}^{(1-\theta)\gamma_2/(1+\alpha)}},
\]
which gives (4.14). The proof is complete. \( \Box \)

**Corollary 4.5.** Let \( \theta \in (0, \alpha - 1) \) and \( 0 < \beta < \gamma < 1 + \theta \). For any \( \varepsilon \in (0, 1) \), there are \( \theta > 0 \) close to zero and constants \( C_{\varepsilon}, C > 0 \) such that for all \( j \geq 5 \),
\[
\|R^\varepsilon_j, b \cdot \nabla_x \|_{C^0_{\theta}} \leq 2^{-\frac{(1-\theta)\varphi}{1+\alpha}} \|b\|_{C_{\theta}^{1+\theta}} \|\nabla_x \|_{C_{\theta}^{(1-\theta)\gamma_1 + \theta \varphi_2}} + C_{\varepsilon} \|u\|_{C^{\gamma + \varphi}},
\]

where \( \mathcal{L}^{(\alpha)}_{K_N} \) is defined by (1.7).

**Proof.** Let \( \theta \in (0, 1) \) be fixed, which will be determined below.

(i) By applying Lemma 4.4 with \( (\theta, \gamma, \gamma_1, \gamma_2, \theta) \) in place of \( (\beta, \gamma, \gamma_1, \gamma_2, \theta) \),
\[
\|R^\varepsilon_j, b \cdot \nabla_x \|_{C^0_{\theta}} \leq \|b\|_{C_{\theta}^{1+\theta}} \|\nabla_x \|_{C_{\theta}^{(1-\theta)\gamma_1 + \theta \varphi_2}} + C_{\varepsilon} \|u\|_{C^{\gamma + \varphi}},
\]

Choosing \( \theta > 0 \) small enough so that
\[
\frac{\theta (\alpha + \beta)}{(1-\theta)\gamma_1 + \theta \varphi_2} > \theta \beta,
\]
by (4.3) and Young’s inequality, for any \( \varepsilon \in (0, 1) \), there is a constant \( C_{\varepsilon} > 0 \) such that
\[
\|u\|_{C_{x}^{(1-\theta)\gamma_1 + \theta \varphi_2}} \leq \varepsilon \|u\|_{C_{x}^{(1-\theta)\gamma_1 + \theta \varphi_2}} + C_{\varepsilon} \|u\|_{C^{\gamma + \varphi}},
\]

and also,
\[
\|u\|_{C_{x}^{(1-\theta)\gamma_1 + \theta \varphi_2}} \leq \varepsilon \|u\|_{C_{x}^{(1-\theta)\gamma_1 + \theta \varphi_2}} + \|u\|_{\infty}.
\]

Substituting these two estimates into (4.18), we obtain (4.15).

(ii) By Lemma 4.4 with \( (\theta, \gamma, \gamma_1, \gamma_2, \theta) \) in place of \( (\beta, \gamma, \gamma_1, \gamma_2, \theta) \), we obtain (4.16).

(iii) Recalling (1.7) and (2.2), and noticing that
\[
[R_j, \mathcal{L}^{(\alpha)}_{K_N}] u(x, v) = \int_{\mathbb{R}^d} \left( [R_j, \kappa(\cdot, \cdot, w)] \delta_{w_2}^\varepsilon \delta_{w_2}^\varepsilon \right) u(x, v) \frac{dw}{|w|^{d+\alpha}},
\]
by Lemma 4.4 with \( (\theta, \gamma, \gamma_1, \gamma_2, \theta) \) in place of \( (\beta, \gamma, \gamma_1, \gamma_2, \theta) \), we obtain
\[
\|R^\varepsilon_j, \mathcal{L}^{(\alpha)}_{K_N} \|_{C^0_{\theta}} \leq \int_{\mathbb{R}^d} \|R^\varepsilon_j, \kappa(\cdot, \cdot, w)] \delta_{w_2}^\varepsilon \delta_{w_2}^\varepsilon \|_{C^0_{\theta}} \frac{dw}{|w|^{d+\alpha}}.
\]
\[\begin{align*}
\int_{\mathbb{R}^d} |x|^\beta |v|^{\gamma} |\nabla_x^m \nabla_v^n p_{s,t}(x, v)| \, dx \, dv &\leq C(1-s)^{-((1+\alpha)(\beta+\gamma+(q+1)\nu)/\alpha)}
\end{align*}\]
Hence, 
\[ p^{\kappa, \ell}_{s,t}(2^{-(1+\alpha)}\tilde{x} + \Pi_{s,t}2^{-j}\tilde{v}, 2^{-j}\tilde{v}) = (t-s)^{-\frac{d}{\alpha} - d} p^{\kappa, \ell}_{0,1}(\tilde{h}^{\alpha+1} \tilde{x} + \tilde{h}\theta_{s,t}\tilde{v}, \tilde{h}\tilde{v}), \]
where 
\[ \tilde{h} := (t-s)^{-\frac{1}{\alpha}} 2^{-j}, \quad \theta_{s,t} := \Pi_{s,t}/(t-s). \]
Since the support of \( \phi_i^\alpha \) is contained in the annulus, by Fourier’s transform, 
\[ (\Delta^\alpha_{x,v} \phi_i^\alpha)(\xi, \eta) := (|\xi|^2 + |\eta|^2)^{-\alpha} \phi_i^\alpha(\xi, \eta) \in \mathcal{S}(\mathbb{R}^2d), \]
so that \( \Delta^\alpha_{x,v} \phi_i^\alpha \) is a well-defined Schwartz function. Thus we have 
\[
\mathcal{U} := \int_{\mathbb{R}^{2d}} |x|^\beta |v|^\gamma \left| \int_{\mathbb{R}^{2d}} \Delta^\alpha_{x,v} \phi_i^\alpha(x - \tilde{x}, v - \tilde{v}) p^{\kappa, \ell}_{0,1}(\tilde{h}^{\alpha+1} \tilde{x} + \tilde{h}\theta_{s,t}\tilde{v}, \tilde{h}\tilde{v}) d\tilde{x}d\tilde{v} \right| dx dv 
= \int_{\mathbb{R}^{2d}} |x|^\beta |v|^\gamma \left| \int_{\mathbb{R}^{2d}} \Delta^\alpha_{x,v} \phi_i^\alpha(x - \tilde{x}, v - \tilde{v}) \Delta^\alpha_{x,v} p^{\kappa, \ell}_{0,1}(\tilde{h}^{\alpha+1} \tilde{x} + \tilde{h}\theta_{s,t}\tilde{v}, \tilde{h}\tilde{v}) d\tilde{x}d\tilde{v} \right| dx dv 
\leq \int_{\mathbb{R}^{2d}} |x|^\beta |v|^\gamma |\Delta^\alpha_{x,v} \phi_i^\alpha(x, v)| dx dv \int_{\mathbb{R}^{2d}} |\Delta^\alpha_{x,v} p^{\kappa, \ell}_{0,1}(\tilde{h}^{\alpha+1} \tilde{x} + \tilde{h}\theta_{s,t}\tilde{v}, \tilde{h}\tilde{v})| d\tilde{x} d\tilde{v} 
+ \int_{\mathbb{R}^{2d}} |\Delta^\alpha_{x,v} \phi_i^\alpha(x, v)| dx dv \int_{\mathbb{R}^{2d}} |x|^\beta |v|^\gamma |\Delta^\alpha_{x,v} p^{\kappa, \ell}_{0,1}(\tilde{h}^{\alpha+1} \tilde{x} + \tilde{h}\theta_{s,t}\tilde{v}, \tilde{h}\tilde{v})| d\tilde{x} d\tilde{v}. \]
By the chain rule, (5.5) and cumbersome calculations, we have 
\[ \int_{\mathbb{R}^{2d}} |\Delta^\alpha_{x,v} p^{\kappa, \ell}_{0,1}(\tilde{h}^{\alpha+1} \tilde{x} + \tilde{h}\theta_{s,t}\tilde{v}, \tilde{h}\tilde{v})| dx dv \leq \tilde{h}^{(\alpha+1)(n-d) - d} + \tilde{h}^{n-(\alpha+2)d}, \]
and 
\[ \int_{\mathbb{R}^{2d}} |x|^\beta |v|^\gamma |\Delta^\alpha_{x,v} p^{\kappa, \ell}_{0,1}(\tilde{h}^{\alpha+1} \tilde{x} + \tilde{h}\theta_{s,t}\tilde{v}, \tilde{h}\tilde{v})| dx dv \leq \tilde{h}^{(\alpha+1)(n-d) - d} + \tilde{h}^{n-(\alpha+2)d} \tilde{h}^{-(\alpha+1)\beta - \gamma}. \]
Therefore, 
\[ \mathcal{U} \leq \left( \tilde{h}^{(\alpha+1)(n-d) - d} + \tilde{h}^{n-(\alpha+2)d} \right) \left( 1 + \tilde{h}^{-(\alpha+1)\beta - \gamma} \right), \]
and 
\[ \mathcal{J}_{s,t} \leq 2^{-(1+\alpha)\beta + \gamma) j} \tilde{h}^{(\alpha+1)\beta + \gamma) j} \left( \tilde{h}^{(\alpha+1)(n-d) - d} + \tilde{h}^{n-(\alpha+2)d} \right) \left( 1 + \tilde{h}^{-(\alpha+1)\beta - \gamma} \right) \]
\[ = \left( \tilde{h}^{(\alpha+1)n} + \tilde{h}^{n} \right) \left( 2^{-(1+\alpha)\beta + \gamma) j} + j \frac{(\alpha+1)\beta + \gamma}{\alpha} \right). \]
(5.9)
Without loss of generality, assume \( t > 2^{-\alpha j} \). We denote the left hand side of (5.6) by \( \mathcal{J} \), and make the following decomposition:
\[ \mathcal{J} = \left( \int_{t-2^{-\alpha j}}^t + \int_0^{t-2^{-\alpha j}} \right) (t-s)^{q} \mathcal{J}_{s,t} ds =: \mathcal{J}_1 + \mathcal{J}_2. \]
For \( \mathcal{J}_1 \), using (5.9) with \( n = 0 \), and by the change of variables, we have 
\[ \mathcal{J}_1 \leq \int_{t-2^{-\alpha j}}^t (t-s)^q \left( 2^{-((1+\alpha)\beta + \gamma) j} + j \frac{(\alpha+1)\beta + \gamma}{\alpha} \right) ds \]
\[ \leq \int_0^{2^{-\alpha j}} s^q \left( 2^{-((1+\alpha)\beta + \gamma) j} + s \frac{(\alpha+1)\beta + \gamma}{\alpha} \right) ds \leq 2^{-(1+\alpha)\beta + \gamma (q+1) \alpha} j. \]
For \( \mathcal{J}_2 \), choosing \( n \) large enough in (5.9) so that 
\[ 1 + q - \frac{n}{\alpha} + \frac{(\alpha+1)\beta + \gamma}{\alpha} + \frac{\gamma}{\alpha} < 0, \]
by similar calculations as above, we also have
\[ J_2 \lesssim \int_{2^{-\alpha}a_j}^a s^d \left((s^{-\frac{1}{2}}2^{-\gamma})^{(\alpha+1)\eta} + (s^{-\frac{1}{2}}2^{-\gamma})^\eta \right)(2^{-(1+\alpha)\beta+\gamma}j^j + s^\eta \gamma a_j) ds \lesssim 2^{-(1+\alpha)\beta+\gamma+(q+1)\alpha}. \]
Combining the above calculations, we obtain the desired estimate. □

6. Schauder’s estimates for non-local degenerate equations

In this section we consider the following nonlocal degenerate equation in \(\mathbb{R}^d\):
\[ \partial_t u = \mathcal{L}_{\kappa, \theta}^{(\alpha)} u + b \cdot \nabla u - \lambda u + f, \quad \lambda \geq 0, \tag{6.1} \]
where \(\mathcal{L}_{\kappa, \theta}^{(\alpha)}\) is defined by (1.7) and \(b\) is a measurable function with the form
\[ b(t, x, v) = (b^{(1)}(t, x, v), b^{(2)}(t, x, v)). \]

Throughout this section we assume \((H_{\beta, \gamma}^{(\alpha, \vartheta)})\) For some \(c_0 \geq 1\) and \(\vartheta \in (0, \alpha - 1), \beta \in (0, 1)\), it holds that for all \(t \geq 0\) and \(x, v, w \in \mathbb{R}^d\),
\[ c_0^{-1} \leq k(t, x, v, w) \leq c_0, \quad \left[k(t, \cdot, w)\right]_{C^\gamma} + \left[b^{(1)}(t, \cdot)\right]_{C^\gamma} + \left[b^{(2)}(t, \cdot)\right]_{C^\gamma} \lesssim c_0, \]
where \(\| \cdot \|_{C^{\gamma, \vartheta}}\) is defined by (4.1), and for some \(\gamma \in [\beta, 1 + \alpha]\),
\[ \left[k(t, \cdot, v)\right]_{C^{\gamma/(1+\vartheta)}} + \left[b^{(1)}(t, \cdot)\right]_{C^{\gamma/(1+\vartheta)}} + \left[b^{(2)}(t, \cdot)\right]_{C^{\gamma/(1+\vartheta)}} \lesssim c_0, \]
and for some closed and convex subset \(\mathcal{E} \subset GL_d(\mathbb{R})\), where \(GL_d(\mathbb{R})\) is the set of all invertible \(d \times d\)-matrices,
\[ \nabla_s^\gamma b^{(1)}(t, x, v) \in \mathcal{E}. \tag{6.2} \]

**Definition 6.1 (Classical solutions).** Let \(\lambda \geq 0\). We call a bounded continuous function \(u\) defined on \(\mathbb{R}_+ \times \mathbb{R}^d\) a classical solution of PDE (6.1) if for some \(\varepsilon \in (0, 1)\),
\[ u \in C([0, \infty); C^{(\alpha+1)\varepsilon}_\kappa \cap C^{1+\varepsilon}_x), \]
and for all \(t \geq 0\) and \(x, v \in \mathbb{R}^d\),
\[ u(t, x, v) = \int_0^t \left(\mathcal{L}_{\kappa, \theta}^{(\alpha)} u + b \cdot \nabla u - \lambda u + f\right)(s, x, v) ds. \]

We have the following maximum principle for classical solutions.

**Theorem 6.2 (Maximum principle).** Let \(\lambda, T > 0\). For any classical solution \(u\) of PDE (6.1) in the sense of Definition 6.1, it holds that
\[ \|u\|_{L^\infty_T} \leq (1 - e^{-\lambda T}) \|f\|_{L^\infty_T} / \lambda. \tag{6.3} \]

**Proof.** Let
\[ \tilde{u}(t, x, v) := -u(t, x, v)e^{\lambda t} + \int_0^t \|f(s, \cdot, \cdot)\|_{L^\infty} e^{\lambda s} ds. \tag{6.4} \]
By (6.1), it is easy to see that for Lebesgue almost all \(t > 0\),
\[ \partial_t \tilde{u} - \mathcal{L}_{\kappa, \theta}^{(\alpha)} \tilde{u} - b \cdot \nabla \tilde{u} \geq 0. \]
Since \(\lim_{\varepsilon \to 0} \tilde{u}(t, x, v) = 0\), by [9, Theorem 6.1], we have
\[ \tilde{u}(t, x, v) \geq 0. \]
Thus, by (6.4), we get
\[ u(t, x, v) \leq e^{-\lambda t} \int_0^t \|f(s, \cdot, \cdot)\|_\infty e^{\lambda s} ds \leq (1 - e^{-\lambda t}) \|f\|_{L^p(T)} / \lambda. \]

By symmetry, we obtain (6.3). \( \square \)

The goal of this section is to prove the following Schauder’s apriori estimate.

**Theorem 6.3.** Let \( \alpha \in (1, 2) \) and \( \beta \in (0, 1), \theta \in (0, \alpha - 1), \gamma \in [\beta, 1 + \alpha). \) Under (H\( \beta, \gamma \)), for any \( T > 0 \), there is a constant \( C > 0 \) only depending on \( T, c_0, \beta, \epsilon, d, \alpha, \theta \) such that for any \( \lambda \geq 0 \) and any classical solution \( u \) of (6.1),
\[ \|u\|_{L^\infty(T; C^{(\beta + \alpha)(1+\alpha), \beta + \gamma}([0, T]) \cap C([0, T]))} \leq C \|f\|_{L^p(T)} \gamma^6. \] (6.5)

**Remark 6.4.** Although our result is stated for \( \alpha \in (1, 2) \), it in fact also works for \( \alpha = 2 \). In this case, under (H\( \beta, \gamma \)), Chaudru, Honoré and Menozzi [6, Theorem 1] has proven (6.5) for \( \gamma = \beta \), When \( \gamma = \beta \), our assumption on \( b^{(1)} \) is weaker since we only assume (H\( \beta, \gamma \)) for some \( \theta \in (0, 1) \).

To prove this theorem we use the perturbation argument by freezing the coefficients along the characterization curve as usual. We need the following well-known fact from ODE.

**Lemma 6.5.** Let \( b : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^d \) be a time-dependent measurable vector field. Suppose that for each \( t > 0 \), \( x \mapsto b(t, x) \) is continuous and for some \( C > 0 \) and all \( (t, x) \in \mathbb{R}^+ \times \mathbb{R}^d \),
\[ |b(t, x)| \leq C(1 + |x|). \]

Then for each \( x \in \mathbb{R}^d \), there is a global solution \( \theta_t \) to the following ODE:
\[ \dot{\theta}_t = b(t, \theta_t), \quad \theta_0 = x. \]

Moreover, if we denote by \( S_x : = \{ \theta : \theta_0 = x \} \) the set of all solutions with starting point \( x \), then for each \( T > 0 \),
\[ \cup_{x \in \mathbb{R}^d} \cup_{\theta \in S_x} \{ \theta_T \} = \mathbb{R}^d. \] (6.6)

**Proof.** We only show (6.6). Fix \( y \in \mathbb{R}^d \) and \( T > 0 \). Let \( (\tilde{\theta}_t)_{t \in [0, T]} \) be the solution of ODE:
\[ \dot{\tilde{\theta}}_t = -b(T - t, \tilde{\theta}_t), \quad \tilde{\theta}_0 = y, \]

and \( (\tilde{\theta}_t)_{t \geq 0} \) solve the ODE
\[ \dot{\tilde{\theta}}_t = b(T + t, \tilde{\theta}_t), \quad \tilde{\theta}_0 = y. \]

Define
\[ \theta_t : = \tilde{\theta}_{T-t} 1_{t \leq T} + \tilde{\theta}_{T-t} 1_{t > T}. \]

It is easy to see that \( \theta_T = y \) and \( \theta \in S_x \) with \( x = \tilde{\theta}_T \). \( \square \)

Fix \( (x_0, v_0) \in \mathbb{R}^{2d} \). Let \( \theta_t \) solve the following ODE in \( \mathbb{R}^{2d} \):
\[ \dot{\theta}_t = b(t, \theta_t), \quad \theta_0 = (x_0, v_0). \]

Define
\[ \tilde{u}(t, x, v) := u(t, x + \theta^{(1)}_t, v + \theta^{(2)}_t), \quad \tilde{f}(t, x, v) := f(t, x + \theta^{(1)}_t, v + \theta^{(2)}_t), \]
\[ \kappa_0(t, w) := \kappa(t, \theta_t, w), \quad \tilde{\kappa}(t, x, v, w) := \kappa(t, x + \theta^{(1)}_t, v + \theta^{(2)}_t, w) - \kappa(t, \theta^{(1)}_t, \theta^{(2)}_t, w), \]

and
\[ U_t := \nabla_x b^{(1)}(t, \theta_t), \quad \tilde{b}(t, x, v) := b(t, x + \theta^{(1)}_t, v + \theta^{(2)}_t) - b(t, \theta_t) - (U_t, 0). \]
By (6.2), there is a constant $c_1 \geq 1$ only depending on $\mathcal{E}$ such that for all $0 \leq s < t$,

$$|U_s| (t-s) |\Pi(s,t)| \leq c_1,$$

where $\Pi(s,t) := \int_s^t U_r \, dr$. (6.7)

It is easy to see that $\tilde{u}$ satisfies the following freezing equation:

$$\partial_t \tilde{u} = \mathcal{L}^{(\alpha)}_{\kappa_0 : V} \tilde{u} + U V \cdot \nabla \tilde{u} - \lambda \tilde{u} + \mathcal{L}^{(\alpha)}_V \tilde{u} + \tilde{b} \cdot \nabla \tilde{u} + \tilde{f},$$

where

$$\mathcal{L}^{(\alpha)}_{\kappa_0 : V} (x, v) := \int \mathbb{E}^{(2)}_w u(x, v) \kappa_0(t, w) \frac{dw}{|w|^d + \alpha},$$

$$\delta^{(2)}_w \ := \delta^{(2)}_{w:2}.$$  

Below, without loss of generality, we drop the tilde over $u, f, \kappa, b$ and assume $x_0 = v_0 = 0$ and

$$|\kappa(t, x, v, w)| \leq |\kappa(t, \cdot, w)|_{C^{(1+\alpha)}(10+x)} + |\kappa(t, \cdot, w)|_{C^\beta(10+x)}$$

and

$$|\tilde{b}^{(1)}(t, x, v)| \leq \tilde{b}^{(1)}(t, \cdot)|C^{1+\alpha} + |v|^\beta,$$

$$|\tilde{b}^{(2)}(t, x, v)| \leq \tilde{b}^{(2)}(t, \cdot)|C^{1+\alpha} + |v|^\beta.$$  

Let $p_{s,t}^{(\alpha)}(x, v)$ be the heat kernel of $\mathcal{L}^{(\alpha)}_{\kappa_0 + U : V}$. Define for $\lambda \geq 0$,

$$P_{s,t}^{\lambda} f(x, v) := (\Gamma_{s,t} P_{s,t}^{\lambda} \ast \Gamma_{s,t} f)(x, v),$$

where $\Gamma_{s,t}$ is defined by (5.4). By Duhamel’s formula we have (see (5.3))

$$u(t, x, v) = \int_0^t P_{s,t}^{\lambda} \mathcal{L}^{(\alpha)}_{\kappa_0 : V} u(s, x, v) ds + \int_0^t P_{s,t}^{\lambda} (b \cdot \nabla u)(s, x, v) ds + \int_0^t P_{s,t}^{\lambda} f(s, x, v) ds.$$  

We prepare the following lemmas.

**Lemma 6.6.** Under $(H_{\rho, \gamma})$, for any $\varepsilon \in (0, 1)$, there is a constant $C > 0$ such that for all $j \in \mathbb{N}$ and $\lambda \geq 0$, $t \in [0, T]$,

$$\int_0^t \| R_j \mathcal{L}^{(\alpha)}_{\kappa_0 : V} u(s, 0, 0) ds \leq C 2^{-(\alpha + \beta)} \| u \|_{L^\infty T(C^{\alpha + \varepsilon}_V)},$$

$$\int_0^t \| R_j \mathcal{L}^{(\alpha)}_{\kappa_0 : V} u(s, 0, 0) ds \leq C 2^{-\gamma \alpha_j} \| u \|_{L^\infty T(C^{(1+\alpha)\varepsilon}_V)},$$

Proof. (i) First of all, by (6.8), we have for $u \in C^{\alpha + \varepsilon}_V$,

$$|\mathcal{L}^{(\alpha)}_{\kappa_0 : V} u(x, v)| = \left| \int_\mathbb{R}^{d} \delta^{(2)}_w u(x, v) \cdot \kappa(t, x, v, w) \frac{dw}{|w|^d + \alpha} \right| \leq \left| x \right|^\alpha \left| v \right|^\beta \int_\mathbb{R}^{d} \delta^{(2)}_w u(x, v) \frac{dw}{|w|^d + \alpha} \leq \left( \left| x \right|^\alpha + \left| v \right|^\beta \right) \| u \|_{C^{\alpha + \varepsilon}_V} \int_\mathbb{R}^{d} (1 + \left| w \right|^d) \frac{dw}{|w|^d + \alpha} \leq \left( \left| x \right|^\alpha + \left| v \right|^\beta \right) \| u \|_{C^{\alpha + \varepsilon}_V}.$$  

Thus by definition (6.11), we have

$$\int_0^t \| R_j \mathcal{L}^{(\alpha)}_{\kappa_0 : V} u(s, 0, 0) ds = \int_0^t \left| \int_\mathbb{R}^{2d} \mathcal{R}_j \mathcal{L}^{(\alpha)}_{\kappa_0 : V} u(s, x, v) \cdot (\Gamma_{s,t} \mathcal{L}^{(\alpha)}_{\kappa_0 : V} u)(s, x, v) dx d\nu \right| ds$$

$$\leq \| u \|_{L^\infty T(C^{\alpha + \varepsilon}_V)} \int_0^t \left| \int_\mathbb{R}^{d} \mathcal{R}_j \mathcal{L}^{(\alpha)}_{\kappa_0 : V} u(s, x, v) \cdot \left( \lambda + \Pi s + \left| \Pi s \right| \right) dx d\nu \right| ds,$$

which in turn gives (6.13) by direct application of (5.6) and $\gamma \geq \beta$.

(ii) Notice that by (2.5) and $\mathcal{R}_j \Gamma_{s,t} = \Gamma_{s,t} \mathcal{R}_j$,  

$$\int_0^t \| R_j \mathcal{L}^{(\alpha)}_{\kappa_0 : V} u(s, 0, 0) ds = \int_0^t \left| \int_\mathbb{R}^{d} \mathcal{R}_j \mathcal{L}^{(\alpha)}_{\kappa_0 : V} u(s, x, v) \cdot (\Gamma_{s,t} \mathcal{L}^{(\alpha)}_{\kappa_0 : V} u)(s, x, v) dx d\nu \right| ds.$$  

\[
= \int_0^\alpha \left| \int_{\mathbb{R}^{2d}} \overline{R}_j^{\gamma} \Gamma_{s,j} P_{s,j}^4(x, v) \cdot \left( \Gamma_{s,j} \mathcal{R}^x_{\epsilon \nu} u \right)(s, x, v) \, dx \, dv \right| \, ds =: \mathcal{I}_1 + \mathcal{I}_2,
\]

where
\[
\mathcal{I}_1 := \int_0^\alpha \left| \int_{\mathbb{R}^{2d}} \overline{R}_j^{\gamma} \Gamma_{s,j} P_{s,j}^4(x, v) \cdot \left( \Gamma_{s,j} \mathcal{R}^x_{\epsilon \nu} \mathcal{R}_j^x u \right)(s, x, v) \, dx \, dv \right| \, ds,
\]
\[
\mathcal{I}_2 := \int_0^\alpha \left| \int_{\mathbb{R}^{2d}} \overline{R}_j^{\gamma} \Gamma_{s,j} P_{s,j}^4(x, v) \cdot \left( \Gamma_{s,j} [\mathcal{R}_j^x, \mathcal{R}^x_{\epsilon \nu}] u \right)(s, x, v) \, dx \, dv \right| \, ds.
\]

For \(\mathcal{I}_1\), by the assumptions, we have
\[
\left| \mathcal{L}^{x}_{\epsilon \nu} \mathcal{R}_j^x u(x, v) \right| = \left| \int_{\mathbb{R}^{d}} \delta^{(2)}_{\nu} \mathcal{R}_j^x u(x, v) \cdot \kappa(t, x, v, w) \frac{dw}{|w|^{d+\alpha}} \right| \\
\lesssim (|x|^{1+\delta} + |v|^\delta) \int_{\mathbb{R}^{d}} \left| \delta^{(2)}_{\nu} \mathcal{R}_j^x u(x, v) \right| \frac{dw}{|w|^{d+\alpha}} \\
\lesssim (|x|^{1+\delta} + |v|^\delta) ||\mathcal{R}_j^x u||_{C^{\alpha+\epsilon}},
\]

and thus, by (5.7) and \(\gamma \geq \beta\),
\[
\mathcal{I}_1 \lesssim ||\mathcal{R}_j^x u||_{L_2^\gamma(C^{\alpha+\epsilon})} \int_0^\alpha \int_{\mathbb{R}^{2d}} \left| \overline{R}_j^{\gamma} \Gamma_{s,j} P_{s,j}^4(x, v) \left( |x + \Pi_{s,j} v|^{1+\delta} + |v|^\delta \right) \right| \, dx \, dv \, ds \\
\lesssim ||\mathcal{R}_j^x u||_{L_2^\gamma(C^{\alpha+\epsilon})} \left( 2^{-\frac{\beta t}{2\alpha+1}} j + 2^\frac{\alpha t}{2\alpha+1} \right) \lesssim 2^{-\frac{\beta t}{2\alpha+1}} ||u||_{L_2^\gamma(C^{\beta(1+\alpha)})} \alpha^{2\alpha+1} \epsilon.
\]

For \(\mathcal{I}_2\), by definition, we have
\[
||\mathcal{R}_j^x, \mathcal{L}^{x}_{\epsilon \nu} u||_{\infty} \leq \int_{\mathbb{R}^{2d}} \left| \delta^{(2)}_{\nu} u \right|_{\infty} \frac{dw}{|w|^{d+\alpha}} \int_{\mathbb{R}^{d}} |x|^{1+\delta} \phi_j(x) \, dx \leq ||u||_{C^{\alpha+\epsilon}} 2^{-\frac{T}{1+\delta} j}.
\]

Hence, by (5.7) again,
\[
\mathcal{I}_2 \lesssim ||u||_{L_2^\gamma(C^{\epsilon+\alpha})} 2^{-\frac{\beta t}{2\alpha+1}} j \int_0^\alpha \int_{\mathbb{R}^{2d}} \left| \overline{R}_j^{\gamma} \Gamma_{s,j} P_{s,j}^4(x, v) \right| \, dx \, dv \, ds \lesssim ||u||_{L_2^\gamma(C^{\alpha+\epsilon})} 2^{-\frac{\beta t}{2\alpha+1}} j.
\]

Combining the above calculations, we obtain (6.14). \(\Box\)

To treat the other terms in (6.12), we need the following lemma.

**Lemma 6.7.** Let \(c_1 \geq 1\) be the same as in (6.7). For \(t \geq 0\) and \(j \in \mathbb{N}\), define
\[
\Theta_j^t := \left\{ \ell \in \mathbb{N}_0 : 2^\ell \leq 2^4 c_1 (2^j + t 2^{(1+\alpha) j}), 2^j \leq 2^4 c_1 (2^\ell + t 2^{(1+\alpha) \ell}) \right\}.
\]

(i) Let \(0 \leq s < t\) and \(j \in \mathbb{N}\). For any \(\ell \not\in \Theta_j^{t-s}\), it holds that
\[
\langle \mathcal{R}_j^x f, \Gamma_{s,j} \mathcal{R}_j^x g \rangle = \int_{\mathbb{R}^{2d}} \mathcal{R}_j^x f(x, v) \cdot \Gamma_{s,j} \mathcal{R}_j^x g(x, v) \, dx \, dv = 0. \tag{6.15}
\]

(ii) For any \(\beta > 0\), there is a constant \(C = C(c_1, \beta) > 0\) such that for all \(j \in \mathbb{N}\) and \(t \geq 0\),
\[
\sum_{\ell \in \Theta_j^t} 2^{-\beta \ell} \leq C \left( 2^{-j} + t 2^{(1+\alpha) j} \right)^\beta, \quad \sum_{\ell \in \Theta_j^t} 2^{\beta \ell} \leq C \left( 2^j + t 2^{(1+\alpha) j} \right)^\beta. \tag{6.16}
\]

(iii) For any \(T > 0\), there exists a \(j_0 = j_0(c_1, \alpha, T) \in \mathbb{N}\) such that for all \(j \geq j_0\) and \(t \in [0, T]\),
\[
\inf \{ \ell : \ell \in \Theta_j^t \} \geq 5.
\]
Proof. (i) To prove (6.15), by Fourier’s transform we have
\[ \langle \mathcal{R}_j^a f, \Gamma_{s_j} \mathcal{R}_a^j g \rangle = \int_{\mathbb{R}^d} \phi_j^a(\xi, \eta) \hat{f}(\xi, \eta) \phi_j^a(\xi, \eta - \Pi_{s_j} \xi) \hat{g}(\xi, \eta - \Pi_{s_j} \xi) d\xi d\eta. \]

Notice that
\[ \text{supp} \phi_j^a \subset \{ (\xi, \eta) : 2^{-j+1} \leq |\xi|^{1/(1+\alpha)} + |\eta| \leq 2^{j+1} \} =: I_j. \]
Assuming \( \langle \mathcal{R}_j^a \Gamma_{s_j} f, \Gamma_{s_j} \mathcal{R}_a^j g \rangle \neq 0 \) for \( j, \ell \in \mathbb{N} \), we must have
\[ (\xi, \eta) \in I_j \text{ and } (\xi, \eta - \Pi_{s_j} \xi) \in I_{\ell}, \]
which implies that
\[ |\xi| \leq 2^{1/(1+\alpha)(j+1)}, \quad |\eta| \leq 2^{j+1}, \]
and
\[ 2^{j-1} \leq |\xi|^{1/(1+\alpha)} + |\eta - \Pi_{s_j} \xi| \leq 2 \cdot 2^{j+1} + c_1 (t - s) 2^{1/(1+\alpha)(j+1)} \leq 2^3 c_1 (2^j + (t - s) 2^{1/(1+\alpha)j}). \]
By symmetry we also have
\[ 2^{-j} \leq c_1 (2^j + (t - s) 2^{1/(1+\alpha)j}). \]
If \( \langle \mathcal{R}_j^a \Gamma_{s_j} f, \Gamma_{s_j} \mathcal{R}_a^j g \rangle \neq 0 \) for \( j \in \mathbb{N} \), we still have
\[ 2^{j-1} \leq |\xi|^{1/(1+\alpha)} + |\eta - \Pi_{s_j} \xi| + |\Pi_{s_j} \xi| \leq 2 + c_1 (t - s) 2^{1/(1+\alpha)}. \]
Combining the above calculations, one sees that for \( \ell \notin \Theta_j^{-s} \), (6.15) holds.

(ii) We only prove the first estimate in (6.16). If \( \ell > j \), then \( 2^{-\ell} \leq 2^{-j} \). If \( \ell \leq j \), then by the definition of \( \Theta_j^\ell \),
\[ 2^{-\ell} \leq 2^4 c_1 2^{-j} (1 + (t - s) 2^{\alpha\ell}) \leq 2^4 c_1 (2^{-j} + (t - s) 2^{(\alpha-1)j}) =: D, \]
which implies \( \ell \geq -\ln D/\ln 2 \). Thus, we have
\[ \sum_{\ell \in \Theta_j^\ell} 2^{-\beta \ell} \leq \sum_{\ell \geq -\ln D/\ln 2} 2^{-\beta \ell} \leq (2D)^{\beta}/(1 - 2^{-\beta}) \leq (2^{-j} + (t - s) 2^{(\alpha-1)j})^\beta. \]

(iii) By definition of \( \Theta_j^\ell \), it suffices to take \( j_0 > \ln(2^4 c_1 (2^5 + T 2^{1/(1+\alpha)5})) / \ln 2. \)

Lemma 6.8. Let \( T > 0 \) and \( j_0 \) be as in (iii) of Lemma 6.7. Under \( (H_{\beta, \gamma}^0) \), there is a constant \( C > 0 \) such that for all \( j \geq j_0, \lambda \geq 0 \) and \( t \in [0, T] \),
\[ \int_0^T |\mathcal{R}_j^a P_{s_j}^1 (b \cdot \nabla u)(s, 0, 0) ds \leq C 2^{-(\alpha+\beta)j} \|u\|_{L^\infty_t(L^2_x(C^{\alpha+\beta}_{\beta, \gamma} + 1))}, \]
\[ \int_0^T |\mathcal{R}_j^a P_{s_j}^1 (b^{(1)} \cdot \nabla u)(s, 0, 0) ds \leq C 2^{-\frac{j}{1+\alpha}} \|u\|_{L^\infty_t(L^2_x(C^{\alpha+\beta}_{\beta, \gamma} + 1))}, \]
\[ \int_0^T |\mathcal{R}_j^a P_{s_j}^1 (b^{(2)} \cdot \nabla u)(s, 0, 0) ds \leq C 2^{-\frac{j}{1+\alpha}} \|\nabla u\|_{L^\infty_t(L^2_x(C^{\alpha+\beta}_{\beta, \gamma} + 1))}. \]

Proof. (i) Let \( \Theta = \Theta_j^{-s} \) as in Lemma 6.7. By definition (6.11) and (6.15), we have
\[ \int_0^\infty \int_{\mathbb{R}^d} |\mathcal{R}_j^a P_{s_j}^1 (b \cdot \nabla u)(s, 0, 0) ds = \int_0^\infty \int_{\mathbb{R}^d} \mathcal{R}_j^a \Gamma_{s_j} P_{s_j}^1(x, v) \cdot \mathcal{R}_j^a \Gamma_{s_j} (b \cdot \nabla u)(s, x, v) dxdv ds \]
\[ \quad = \int_0^\infty \left| \sum_{\ell \in \Theta_j^\ell} \int_{\mathbb{R}^d} \mathcal{R}_j^a \Gamma_{s_j} P_{s_j}^1(x, v) \cdot \mathcal{R}_j^a (b \cdot \nabla u)(s, x, v) dxdv \right| ds. \]
Below we drop the time variable and write
\[ R^\alpha_t(b \cdot \nabla u)(x, v) = (b \cdot \nabla R^\alpha_t u)(x, v) + [R^\alpha_t, b \cdot \nabla]u(x, v) =: \mathcal{J}_1 + \mathcal{J}_2. \]

For \( \mathcal{J}_1 \), by (6.9) and (4.4), we have
\[ |(b^{(1)} \cdot \nabla)R^\alpha_t u(x, v)| \lesssim \left(|x|^{\alpha_0} + |v|^{1+\theta}\right)2^{(1+\alpha j)}\|R^\alpha_t u\|_\infty \lesssim \left(|x|^{\alpha_0} + |v|^{1+\theta}\right)2^{(1-\beta+\theta)\ell}\|u\|_{B^{a+\beta}_{2,\infty}}, \]
and by (6.10) and (4.4),
\[ |(b^{(2)} \cdot \nabla)R^\alpha_t u(x, v)| \lesssim \left(|x|^{\alpha_0} + |v|^{1+\theta}\right)2^{\ell}\|R^\alpha_t u\|_\infty \lesssim \left(|x|^{\alpha_0} + |v|^{1+\theta}\right)2^{(1-\alpha-\beta+\theta)\ell}\|u\|_{B^{a+\beta}_{2,\infty}}. \]
Hence, by (2.9),
\[ |\mathcal{J}_1| \lesssim \|u\|_{C^{\alpha_0+\beta-\delta}_{2,\infty}}\left(|x|^{\alpha_0} + |v|^{1+\theta}\right)2^{(1-\beta+\theta)\ell} + \left(|x|^{\alpha_0} + |v|^{1+\theta}\right)2^{(1-\alpha-\beta+\theta)\ell}. \]  
(6.21)

For \( \mathcal{J}_2 \), due to \( j \geq j_0 \) and by (iii) of Lemma 6.7, we have \( \ell \geq 5 \) for \( \ell \in \Theta_j^{-5} \). Thus we can use (4.7) with \((\theta, \beta, 1 - \theta)\) in place of \((\beta, \gamma)\) to derive that
\[
\left|\left[R^\alpha_t, b^{(1)} \cdot \nabla\right]u(x, v)\right| \leq 2^{-(\beta-\theta)}\left(2^{-\ell} + |x|^{\alpha_0} + |v|^\theta\right)\left|\nabla x u\right|_{C^{1-\theta}_{\alpha_0}} \lesssim \left(2^{-\ell} + 2^{-(\beta-\theta)}(|x|^{\alpha_0} + |v|^\theta)\right)\|u\|_{C^{\alpha_0+\beta-\delta}_{2,\infty}}.
\]

Moreover, by \( \gamma \geq \beta, \alpha - \theta > 1 \) and the definition, we also have
\[
\left|\left[R^\alpha_t, b^{(2)} \cdot \nabla\right]u(x, v)\right| = \left|\int_{\mathbb{R}^{2d}} \tilde{\phi}^\alpha(x - \tilde{x}, v - \tilde{v})\left(b^{(2)}(\tilde{x}, \tilde{v}) - b^{(2)}(x, v)\right)\nabla v u(\tilde{x}, \tilde{v}) d\tilde{x} d\tilde{v}\right| \lesssim \left|\nabla v u\right|_{\infty} \int_{\mathbb{R}^{2d}} |\tilde{\phi}^\alpha(\tilde{x}, \tilde{v})| \cdot \left(|\tilde{x}|^{\alpha_0} + |\tilde{v}|^{1+\theta}\right) d\tilde{x} d\tilde{v} \lesssim 2^{-(\beta)}\|u\|_{C^{\alpha_0+\beta-\delta}_{2,\infty}}.
\]
Therefore,
\[ |\mathcal{J}_2| \lesssim \|u\|_{C^{\alpha_0+\beta-\delta}_{2,\infty}}\left(2^{-\beta} + 2^{-(\beta-\theta)}(|x|^{\alpha_0} + |v|^\theta)\right). \]  
(6.22)

Combining (6.20)-(6.22), and by (6.16) we get
\[
\int_0^\infty \left|\mathcal{R}^\alpha_t P^A_{s,t}(b \cdot \nabla u)(s, 0, 0)\right| ds \lesssim \|u\|_{C^{\alpha_0+\beta-\delta}_{\infty}} \int_0^\infty \left|\mathcal{R}^\alpha_t \Gamma_{s,t} P^A_{s,t}(x, v)\right| ds \times \left(|x + \Pi_{s,t} v|^\alpha_0 + |v|^{1+\theta}\right)2^{(1+\alpha j)} \left|2^j + (t-s)2^{(1+\alpha j)}\right|^{-\beta+\theta} \left(|x + \Pi_{s,t} v|^{\alpha_0} + |v|^\theta\right)2^{-j} + (t-s)2^{(\alpha-1)j} s^{\alpha-\beta-\theta} \left|2^{-j} + (t-s)2^{(\alpha-1)j}\right|^{\beta-\theta} |\theta|\left(|x + \Pi_{s,t} v|^{\alpha_0} + |v|^\theta\right) + (2^{-j} + (t-s)2^{(\alpha-1)j})^{\beta-\theta} dxdvd\sigma,
\]
which in turn yields (6.17) by using (6.7) and (5.6) item by item, as well as \( \gamma \geq \beta. \)

(ii) As above by definition (6.11) and (2.5), we have
\[
\int_0^\infty \left|\mathcal{R}^\alpha_t P^A_{s,t}(b^{(1)} \cdot \nabla u)(s, 0, 0)\right| ds = \int_0^\infty \left|\int_{\mathbb{R}^{2d}} \mathcal{R}^\alpha_t \Gamma_{s,t} P^A_{s,t}(x, v) \cdot \Gamma_{s,t} (b^{(1)} \cdot \nabla u)(s, x, v) dxdv\right| ds \times \left|\int_{\mathbb{R}^{2d}} \mathcal{R}^\alpha_t \Gamma_{s,t} P^A_{s,t}(x, v) \cdot \Gamma_{s,t} (b^{(1)} \cdot \nabla u)(s, x, v) dxdv\right| ds = \mathcal{J}_1 + \mathcal{J}_2,
\]
where we have used \( \mathcal{R}^\alpha_t \mathcal{R}^\beta_t = \Gamma_{s,t} \mathcal{R}^\gamma_t \) and
\[
\mathcal{J}_1 := \int_0^\infty \left|\int_{\mathbb{R}^{2d}} \mathcal{R}^\alpha_t \Gamma_{s,t} P^A_{s,t}(x, v) \cdot \Gamma_{s,t} (b^{(1)} \cdot \nabla \mathcal{R}^\alpha_t u)(s, x, v) dxdv\right| ds,
\]
and
\[ J_2 := \left| \int_0^T \left| \int_{\mathbb{R}^d} \overline{R}^2 r_j, p_s, (x, v) \cdot \Gamma_s, [R_j^x, b^{(1)}] \cdot \nabla_s u(s, x, v) \right| \, ds \right|. \]

For \( J_1 \), noticing that by (6.9) and the definition of \( R_j^x \),

\[
|b^{(1)} \cdot \nabla_s R_j^x u|(x, v) \lesssim (|x|^\frac{4\theta}{1+\theta} + |v|^{1+\theta}) 2^j ||R_j^x u||_\infty,
\]

we have

\[
J_1 \leq 2^j ||R_j^x u||_\infty \int_0^T \int_{\mathbb{R}^d} |\overline{R}^2 r_j, p_s, (x, v) | \left( |x + \Pi_s v|^\frac{4\theta}{1+\theta} + |v|^{1+\theta} \right) \, dx \, dv \, ds
\]

\[
\leq \frac{1}{2} \left[ 2^{-\frac{4\theta}{1+\theta}} ||R_j^x u||_\infty \right] \int_0^T \int_{\mathbb{R}^d} \frac{(1+\theta)^\gamma}{\beta^{1+\gamma}} \, dx \, dv \, ds
\]

For \( J_2 \), by (4.8) with \( \frac{(1+\theta)^\gamma}{\beta^{1+\gamma}} = (1+\theta)^\gamma \frac{1}{\beta^{1+\gamma}} \) in place of \( (\beta, \gamma) \) and (5.7), we have

\[
J_2 \leq \int_0^T \int_{\mathbb{R}^d} |\overline{R}^2 r_j, p_s, (x, v) |||R_j^x, b^{(1)} \cdot \nabla_s u||_\infty \, dx \, dv \, ds
\]

\[
\leq 2^{-\frac{4\theta}{1+\theta}} ||\nabla_s u||_\infty ||R_j^x u||_\infty \left( \frac{(1+\theta)^\gamma}{\beta^{1+\gamma}} \right)
\]

Combining the above two estimates and Theorem 2.2, we obtain (6.18).

(iii) As above by definition we have

\[
\int_0^T \left| \int_{\mathbb{R}^d} \overline{R}^2 r_j, p_s, (b^{(2)} \cdot \nabla_s u)(s, x, v) \right| \, dx \, dv \, ds = \int_0^T \left| \int_{\mathbb{R}^d} \overline{R}^2 r_j, p_s, (x, v) \cdot \Gamma_s, (b^{(2)} \cdot \nabla_s R_j^x u)(s, x, v) \right| \, dx \, dv \, ds
\]

\[
= \int_0^T \left| \int_{\mathbb{R}^d} \overline{R}^2 r_j, p_s, (x, v) \cdot \Gamma_s, (b^{(2)} \cdot \nabla_s R_j^x u)(s, x, v) \right| \, dx \, dv \, ds
\]

\[
J_2 := \int_0^T \left| \int_{\mathbb{R}^d} \overline{R}^2 r_j, p_s, (x, v) \cdot \Gamma_s, (b^{(2)} \cdot \nabla_s R_j^x u)(s, x, v) \right| \, dx \, dv \, ds.
\]

For \( J_1 \), noticing that by (6.10),

\[
|b^{(2)} \cdot \nabla_s R_j^x u|(x, v) \lesssim (|x|^\frac{4\theta}{1+\theta} + |v|^{1+\theta}) ||\nabla_s R_j^x u||_\infty,
\]

we have

\[
J_1 \leq \frac{1}{2} \left[ 2^{-\frac{4\theta}{1+\theta}} ||\nabla_s R_j^x u||_\infty \right] \int_0^T \int_{\mathbb{R}^d} \overline{R}^2 r_j, p_s, (x, v) | \left( |x + \Pi_s v|^\frac{4\theta}{1+\theta} + |v|^{1+\theta} \right) \, dx \, dv \, ds
\]

\[
\leq \frac{1}{2} \left[ 2^{-\frac{4\theta}{1+\theta}} ||\nabla_s R_j^x u||_\infty \right] \int_0^T \int_{\mathbb{R}^d} \frac{(1+\theta)^\gamma}{\beta^{1+\gamma}} \, dx \, dv \, ds
\]

For \( J_2 \), by the commutator estimate (4.8), we have

\[
J_2 \leq \int_0^T \int_{\mathbb{R}^d} \left| \overline{R}^2 r_j, p_s, (x, v) [R_j^x, b^{(2)} \cdot \nabla_s u] \right|||R_j^x u||_\infty \, dx \, dv \, ds \leq 2^{-\frac{4\theta}{1+\theta}} ||\nabla_s u||_\infty.
\]

Combining the above calculations, we obtain (6.19). □

**Lemma 6.9.** For any \( \beta \in (0, 1) \), there is a constant \( C > 0 \) such that for all \( j \geq 5 \) and \( \lambda \geq 0 \), \( t \in [0, T] \),

\[
\int_0^T \left| [R_j^x p_s, f](s, x, v) \right| \, dx \, dv \, ds \leq C2^{-(\alpha+\beta)T} ||f||_{L^q_T} \, \langle c \rangle,
\]

(6.23)
\[
\int_0^t |R^i_j f(s,0,0) ds \leq C 2^{-\frac{\varepsilon t}{\varepsilon'}} \|f\|_{L^{\infty}_T(K_i^{(1+\alpha)})}.
\] (6.24)

**Proof.** We only prove the first one. The second one is similar and easier by (5.7). Let \( \Theta = \Theta^{t-s}_j \) be as in Lemma 6.7. By definition (6.11) and Lemma 6.7, we have

\[
\int_0^t |R^i_j f(s,0,0) ds = \int_0^t \int_{\mathbb{R}^d} |R^i_j f(x,v)| dx dv ds
\]

which gives (6.23) by application of (5.6).

Now we are in a position to give

**Proof of Theorem 6.3.** (i) Fix \( \varepsilon \in (0,1) \). We first show the following estimates:

\[
\|u\|_{L^{\infty}_T(K_i^{(1+\alpha)})} \leq C \|f\|_{L^{\infty}_T(c^\alpha_i)},
\] (6.25)

and

\[
\|u\|_{L^{\infty}_T(K_i^{(1+\alpha)})} \leq C \left( \|f\|_{L^{\infty}_T(c^\gamma_i)} + \|u\|_{L^{\infty}_T(c^{(1+\alpha)}_i)} \right).
\] (6.26)

By Lemmas 6.6, 6.8 and 6.9, we have

\[
|R^i_j u(t,\theta)| \leq 2^{-(\alpha+\beta)j} \|u\|_{L^{\infty}_T(c^\alpha_i)} + 2^{-(\alpha+\beta)j} \|f\|_{L^{\infty}_T(c^\alpha_j)}, \quad j \geq j_0,
\]

and

\[
|R^i_j u(t,\theta)| \leq 2^{-\varepsilon\frac{2^j}{2^j}} \left( \|u\|_{L^{\infty}_T(c^{(1+\alpha)}_i)} + \|u\|_{L^{\infty}_T(c^{(1+\alpha)}_i)} + \|f\|_{L^{\infty}_T(c^{(1+\alpha)}_i)} \right), \quad j \geq j_0.
\]

Moreover, it is clear that

\[
|R^i_j u(t,\theta)| \leq \|u\|_{L^{\infty}_T}, \quad j \in \mathbb{N}_0.
\]

By (6.6), Theorem 2.2 and (2.10), for any \( \varepsilon' > 0 \), the above estimates lead to

\[
\|u(t)\|_{c^{(1+\alpha)}_i} \leq \|u(t)\|_{B^{(1+\alpha)}_c} = \sup_{j \in \mathbb{N}_0} 2^{(\alpha+\beta)j} \|R^i_j u(t)\|_{L^{\infty}} \leq \varepsilon' \|u\|_{L^{\infty}_T(c^\alpha_i)} + \|u\|_{L^{\infty}_T} + \|f\|_{L^{\infty}_T(c^\alpha_i)},
\]

and

\[
\|u(t)\|_{c^{(1+\alpha)}_i} \leq \|u(t)\|_{B^{(1+\alpha)}_c} = \sup_{j \in \mathbb{N}_0} 2^{(\alpha+\beta)j} \|R^i_j u(t)\|_{L^{\infty}} \leq \varepsilon' \|u\|_{L^{\infty}_T(c^\alpha_i)} + \|u\|_{L^{\infty}_T(c^{(1+\alpha)}_i)} + \|f\|_{L^{\infty}_T(c^{(1+\alpha)}_i)}.
\]

which in turn yield (6.25) and (6.26) by taking \( \varepsilon' = 1/2 \) and (6.3).

(ii) For \( j \geq 1 \), we have

\[
\partial_t R^j u = \mathcal{L}^{(\alpha)}_{K_v} R^j u + b \cdot \nabla R^j u + R^j f + [R^j, \mathcal{L}^{(\alpha)}_{K_v}] u + [R^j, b \cdot \nabla] u.
\]

24
For \( \theta \in (0, 1] \) being small enough, by Schauder’s estimate (6.25) and Corollary 4.5, we have
\[
||R^x_j u||_{L^p_\infty(C^{1+|\beta|}_0)} \leq ||R^x_j f||_{L^p_\infty(C^{1+|\beta|}_0)} + ||R^x_j, \mathcal{L}^{(\alpha)}[u]||_{L^p_\infty(C^{1+|\beta|}_0)} + ||[R^x_j, b \cdot \nabla] u||_{L^p_\infty(C^{1+|\beta|}_0)}
\]
\[
\leq ||R^x_j f||_{L^p_\infty(C^{1+|\beta|}_0)} + 2^{1+|\alpha|} \left( \varepsilon ||u||_{L^p_\infty(C^{1+|\beta|+2|\eta|+|\gamma|+|\alpha|+1)_0)} + ||u||_{L^p_\infty(C^{1+|\beta|}_0)} \right),
\]
where \( j \geq j_0 \) and \( \varepsilon \in (0, 1) \). On the other hand, for \( j = 0, \ldots, j_0 \),
\[
||R^x_j u||_{L^p_\infty(C^{1+|\beta|}_0)} \leq ||R^x_j f||_{L^p_\infty(C^{1+|\beta|}_0)} + ||R^x_j u||_{L^p_\infty(C^{1+|\beta|+|\gamma|+|\alpha|+1}_0)} \leq ||u||_{L^p_\infty(C^{1+|\beta|}_0)},
\]
and also,
\[
\sup_{j \in \mathbb{N}_0} 2^{1+|\alpha|} ||R^x_j f||_{L^p_\infty(C^{1+|\beta|}_0)} (4.2) \leq ||f||_{L^p_\infty(C^{1+|\beta|+2|\eta|+|\gamma|+|\alpha|+1}_0)} + ||f||_{L^p_\infty(C^{1+|\beta|}_0)} (4.3)
\]
\[
\leq ||f||_{L^p_\infty(C^{1+|\beta|}_0)} + ||f||_{L^p_\infty(C^{1+|\beta|}_0)} + \varepsilon ||u||_{L^p_\infty(C^{1+|\beta|+2|\eta|+|\gamma|+|\alpha|+1}_0)},
\]
which implies by taking \( \varepsilon \) small enough,
\[
||u||_{L^p_\infty(C^{1+|\beta|+2|\eta|+|\gamma|+|\alpha|+1}_0)} \leq ||f||_{L^p_\infty(C^{1+|\beta|}_0)} + ||f||_{L^p_\infty(C^{1+|\beta|}_0)}.
\]
Substituting this into (6.26) with \( \theta \) being small enough, we obtain the desired estimate (6.5). \( \square \)

**Remark 6.10.** The restriction of \( \alpha \in (1, 2) \) is only used in Lemma 6.8, which is caused by the moment problem due to \( 1 + \vartheta < \alpha \). In particular, if \( b^{(1)}(t, x, v) = v + b^{(1)}(t, x) \), then under the following restrictions:
\[
\frac{1+\vartheta}{1+\alpha} < \alpha, \quad \frac{\vartheta}{1+\vartheta} < \alpha, \quad \alpha + \beta > 1,
\]
which implies \( \alpha > \frac{\sqrt{\vartheta} - 1}{2} \), Theorem 6.3 still holds for \( \alpha \in (\frac{\sqrt{\vartheta} - 1}{2}, 1] \). Here we conjecture that the moment restriction is superfluous. At this moment we do not know how to drop it. Such a problem also appears in [8]. Moreover, if \( b(t, x, v) = (v, 0) \), which corresponds to the kinetic equation (1.9), then Theorem 6.3 holds for all \( \alpha \in (0, 2) \).

We have the following existence of classical solutions.

**Theorem 6.11.** Let \( \alpha \in (1, 2) \) and \( \beta \in (0, 1) \), \( \vartheta \in (0, \alpha - 1) \), \( \gamma \in (1, 1 + \alpha) \). Under \( \mathbf{H}_{\mu, \gamma}^{\alpha, \vartheta}(t) \), for any \( f \in L^{\infty}_\text{loc}(C^{1+|\beta|}_x \cap C^{0}_v) \), there is a unique classical solution \( u \) in the sense of Definition 6.1 such that for any \( T > 0 \) and some \( C > 0 \) being independent of \( \lambda > 0 \),
\[
||u||_{L^p_\infty(C^{1+|\beta|+2|\eta|+|\gamma|+|\alpha|+1}_0)} \leq C ||f||_{L^p_\infty(C^{1+|\beta|}_0)} + ||u||_{L^p_\infty} \leq \lambda^{-1} ||f||_{L^p_\infty}.
\]

**Proof.** Let \( \rho_n \in \mathbb{N} \) and \( \rho'_n \in \mathbb{N} \) be the usual mollifiers in \( \mathbb{R}^{3d} \) and \( \mathbb{R}^{2d} \) respectively. Define
\[
\kappa_n(t, x, v, w) := \kappa(t, \cdot) \ast \rho_n(x, v, w), \quad n \in \mathbb{N},
\]
and
\[
b_n(t, x, v) := b(t, \cdot) \ast \rho'_n(x, v), \quad f_n(t, x, v) := f(t, \cdot) \ast \rho'_n(x, v), \quad n \in \mathbb{N}.
\]
Fix \( \lambda, T > 0 \). By Theorem 8.2 in appendix, there is a unique smooth \( u_n \in C([0, T]; C^{2}(\mathbb{R}^{2d})) \) solving the following PDE:
\[
\partial_t u_n = \mathcal{L}^{(\alpha)}_{\kappa_n, \gamma} u_n + b_n \cdot \nabla u_n - \lambda u_n + f_n, \quad u_n(0) = 0.
\]
Under \((H_{\beta,\gamma}^{\alpha,\beta})\), by Theorem 6.3, there is a constant \(C > 0\) such that for all \(n \in \mathbb{N}\),
\[
\|u_n\|_{L_T^\infty(C_x^{(\gamma + \alpha)/(1 + \alpha)}(\mathbb{C}_v^{\alpha + \beta}))} \leq C\|f_n\|_{L_T^\infty(C_x^{(\gamma + \alpha)/(1 + \alpha)}(\mathbb{C}_v^{\alpha + \beta}))} \leq C\|f\|_{L_T^\infty(C_x^{(\gamma + \alpha)/(1 + \alpha)}(\mathbb{C}_v^{\alpha + \beta}))},
\]
(6.32)
Moreover, since \(\alpha \in (1, 2)\) and \(\gamma \in (1, 1 + \vartheta)\), we also have for some \(\varepsilon > 0\),
\[
\|\nabla u_n\|_{L_T^\infty(C_v^\gamma)} \leq C\|u_n\|_{L_T^\infty(C_x^{(\gamma + \alpha)/(1 + \alpha)}(\mathbb{C}_v^{\alpha + \beta}))} \leq C\|f\|_{L_T^\infty(C_x^{(\gamma + \alpha)/(1 + \alpha)}(\mathbb{C}_v^{\alpha + \beta}))}.
\]
Hence, from approximation equation (6.31) and the above uniform estimates, one sees that
\[
\sup_n \|\partial_t u_n \cdot 1_{|x| + |v| \leq m}\|_{L_T^\infty} \leq C_m, \ m \in \mathbb{N}.
\]
Thus by Ascoli-Arzela’s theorem and a standard diagonalization argument, there are subsequence \(n_k\) and continuous function \(u : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}\) such that for each \(m \in \mathbb{N}\),
\[
\lim_{k \rightarrow \infty} \sup_{t \in [0,T], |x| + |v| \leq m} |u_{n_k}(t, x, v) - u(t, x, v)| = 0.
\]
Moreover, we also have
\[
u \in L_T^{\infty}(C_x^{(\gamma + \alpha)/(1 + \alpha)} \cap C_v^{\alpha + \beta}).
\]
In fact, by (2.9) and Fatou’s lemma, we have
\[
\|u(t)\|_{C_v^{\alpha + \beta}} \leq \sup_{j \geq 0} 2^{(\alpha + \beta)j} \|R_j^v u(t)\|_{\infty} \leq \sup_{j \geq 0} 2^{(\alpha + \beta)j} \lim_{n \rightarrow \infty} \|R_j^v u_n(t)\|_{\infty} \leq \lim_{n \rightarrow \infty} \|u_n(t)\|_{C_v^{\alpha + \beta}} \leq \|f\|_{L_T^\infty(C_x^{(\gamma + \alpha)/(1 + \alpha)}(\mathbb{C}_v^{\alpha + \beta}))}
\]
(6.32)
Noticing the following interpolation inequality (see [30, Theorem 3.2.1]),
\[
\|\nabla f\|_{\infty} \leq C\|f\|_{C_x^{1/(1 + \varepsilon)}(\mathbb{C}_v^{\alpha + \beta})} \|f\|_{L_x^{\alpha + \beta}(1/\varepsilon)},
\]
we further have
\[
\lim_{k \rightarrow \infty} \sup_{t \in [0,T], |x| + |v| \leq m} |\nabla u_{n_k}(t, x, v) - \nabla u(t, x, v)| = 0.
\]
By taking limits for equation (6.31), we obtain that \(u\) satisfies (6.1) in the sense of Definition 6.1. By (6.3), we complete the proof.

\[\square\]

**Remark 6.12.** If we do not assume \(\gamma > 1\) in Theorem 6.11, then under \((H_{\beta,\gamma}^{\alpha,\vartheta})\), for any \(f \in L_T^\infty(C_\alpha^\beta)\), we can show the existence of \(u \in L_T^\infty(C_\alpha^{\alpha + \beta})\) solving PDE (6.1) in the distributional sense since \(b \cdot \nabla u\) is a distribution under the above regularity.

### 7. Degenerate SDEs with H"older Drifts

#### 7.1. Pathwise uniqueness of SDEs with multiplicative Lévy noises

Let \(L_t^{(\alpha)}\) be a symmetric and rationally invariant \(\alpha\)-stable process with \(\alpha \in (1, 2)\) on some probability space \((\Omega, \mathcal{F}, \mathbb{P})\), so that whose generator is given by the fractional Laplacian \(\Delta^{\alpha/2}\). In this section we consider the following degenerate SDE with jumps in \(\mathbb{R}^d\):
\[
dZ_{s,t} = b(t, Z_{s,t}) dt + (0, \sigma(t, Z_{s,t}) dL_t^{(\alpha)}), \quad Z_{s,s} = z \in \mathbb{R}^d, \ t \geq s \geq 0,
\]
(7.1)
where \(\sigma : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d \otimes \mathbb{R}^d\) and \(b : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d\) are measurable functions satisfying
(\( \widehat{H}_{p,y}^{\alpha,\beta} \)) \( \sigma \) is Lipschitz continuous in \( x \) uniformly in \( t \), and for some \( c_0 > 1 \) and all \( t \geq 0 \),

\[
c_0^{-1} |\xi| \leq |\sigma(t, z)\xi| \leq c_0 |\xi|, \quad \xi \in \mathbb{R}^d, \ z \in \mathbb{R}^{2d},
\]

and for some \( \vartheta \in (0, \alpha - 1) \), \( \gamma \in (1, 1 + \alpha) \) and \( \beta \in (0, 1) \),

\[
|b(t, 0)| + |b(t, \cdot)|_{C_{\gamma/|1+\alpha|}(\mathbb{R}^d)} + \|\nabla_v b^{(1)}(t, \cdot)\|_{C_{\beta}^0} + |b^{(2)}(t, \cdot)|_{C_{\beta}^0} \leq c_0.
\]

Moreover, (6.2) holds.

Let \( N(dt, dw) \) be the Poisson random measure associated with \( L^{(\alpha)} \), i.e.,

\[
N((0, t] \times \Gamma) := \sum_{0 < s \leq t} 1_{\Gamma}(L^{(\alpha)}_s - L^{(\alpha)}_s), \quad t > 0, \ \Gamma \in \mathcal{B}(\mathbb{R}^d \setminus \{0\}).
\]

Let \( \tilde{N}(dt, dw) := N(dt, dw) - 
\]

\[
dw|z|^{d+\alpha} \text{ be the compensated Poisson random measure. By the Lévy-Itô decomposition, we can write for each } t > 0,
\]

\[
L^{(\alpha)}_t = \int_0^t \int_{\mathbb{R}^d} w \tilde{N}(ds, dw).
\]

Thus, if we let \( Z_{s,t} = (X_{s,t}, V_{s,t}) \), then SDE (7.1) can be written as

\[
\left\{ \begin{array}{l}
\, dX_{s,t} = b^{(1)}(t, Z_{s,t})dt, \quad (X_{s,t}, V_{s,t}) = (x, v), \\
\, dV_{s,t} = b^{(2)}(t, Z_{s,t})dt + \int_{\mathbb{R}^d} \sigma(t, Z_{s,t})w \tilde{N}(dt, dw).
\end{array} \right.
\]

In particular, the generator of this SDE is given by \( \mathcal{L}^{(\alpha)}_{\sigma, v} + b \cdot \nabla \) with

\[
\mathcal{L}^{(\alpha)}_{\sigma, v} f(x, v) = \text{p.v.} \int_{\mathbb{R}^d} \left( f(x, v + \sigma(t, z)w) - f(x, v) \right) \frac{dw}{|w|^{d+\alpha}}
\]

\[
= \text{p.v.} \int_{\mathbb{R}^d} \left( f(x, v + w) - f(x, v) \right) \kappa(t, z, w) \frac{dw}{|w|^{d+\alpha}},
\]

where \( z = (x, v) \) and

\[\kappa(t, z, w) := \det(\sigma^{-1}(t, z))|w|^{d+\alpha}/|\sigma^{-1}(t, z)w|^{d+\alpha}.\]

Under \( (\widehat{H}_{p,y}^{\alpha,\beta}) \), it is easy to see that for some \( c_0 > 1 \),

\[
c_0^{-1} \leq \kappa(t, z, w) \leq c_0, \quad |\kappa(t, z, w) - \kappa(t, z', w)| \leq c_0 |z - z'|.
\]

We have the following result.

**Theorem 7.1.** Let \( \alpha \in (1, 2) \), \( \beta \in (0, \alpha - 1) \), \( \gamma \in (1, 1 + \alpha) \) and \( \beta \in (1, 1 + \alpha) \). Under \( (\widehat{H}_{p,y}^{\alpha,\beta}) \), for each \( s \geq 0 \) and \( z \in \mathbb{R}^{2d} \), there exists a unique strong solution \( (Z_{s,t})_{t \geq s} \) to SDE (7.1).

**Proof.** Since the coefficients are continuous and linear growth, the existence of a solution is well-known (cf. [29]). By Yamada-Watanabe’s theorem (cf.[27]), it suffices to show the pathwise uniqueness. Without loss of generality, we assume \( s = 0 \) and simply write

\[
Z_t := Z_{0,t}.
\]

Since \( b \) is unbounded, to construct Zvonkin’s transformation (cf. [51]), we need to cutoff \( b \). For \( m \in \mathbb{N} \), let \( \chi_m \) be a smooth cutoff function in \( \mathbb{R}^{2d} \) with

\[
\chi_m(z) = 1, \ |z| \leq m, \ \chi_m(z) = 0, \ |z| > m + 1.
\]

Fix \( T > 0 \) and \( m \in \mathbb{N} \). Consider the following backward equation:

\[
\partial_t u^m_\lambda + \mathcal{L}^{(\alpha)}_{\sigma, v} u^m_\lambda - \lambda u^m_\lambda + b \cdot \nabla u^m + b^m(\cdot) = 0, \quad u^m_\lambda(T, \cdot) = 0. \quad (7.2)
\]
Under \( \mathbf{H}_{\beta,\gamma}^{\alpha} \), by Theorem 6.11, there is a unique classical solution \( u^m_t \) to the above equation with regularity: for some \( C_m \geq 1 \) and all \( \lambda \geq 1 \),

\[
\|u^m_t\|_{L^\infty_t(C^{1+\alpha}_x)} \leq C_m, \quad \|u^m_t\|_{L^\infty_t} \leq C_m^{-1}. 
\]

Since \( \gamma \in (1 + \frac{\alpha}{2}, 1 + \alpha) \) and \( \beta \in (1 - \frac{\alpha}{2}, 1) \), by interpolation inequalities (2.10) and (4.3), there is an \( \varepsilon_0 > 0 \) small enough such that

\[
\|u^m_t\|_{L^\infty_t(C^{1+\varepsilon_0}_x)} + \|u^m_t\|_{L^\infty_t(C^{1+\alpha+\varepsilon_0}_x)} \leq C_\lambda,
\]

where \( C_\lambda \rightarrow 0 \) as \( \lambda \rightarrow \infty \). In particular, for \( \lambda \geq 1 \) large enough,

\[
\|\nabla u^m_t\|_{L^\infty_t(C^{1+\varepsilon}_x)} \leq \frac{1}{2}, \quad (7.3)
\]

Define

\[
\Phi^m_t(z) := z + u^m_t(t, z).
\]

By (7.3), one sees that

\[
z \mapsto \Phi^m_t(z)
\]

is a diffeomorphism on \( \mathbb{R}^{2d} \) and

\[
\|\nabla \Phi^m_t\|_{\infty} + \|\nabla(\Phi^m_t)^{-1}\|_{\infty} \leq 2.
\]

Moreover, by (7.2) we have

\[
\partial_s \Phi^m + \mathcal{L}^{(\alpha)}_{\sigma,v} \Phi^m + b \cdot \nabla \Phi^m = b(1 - \chi^m) + \lambda u^m_t. \quad (7.4)
\]

Let \( Z_t \) and \( Z'_t \) be two solutions of SDE (7.1) defined on the same probability space with the same starting point \( z \). Define a stopping time

\[
\tau_m := \inf \left\{ t > 0 : |Z_s| \wedge |Z'_s| \geq m + 1 \text{ or } |\Delta L_t| > m \right\},
\]

and let

\[
g^m_s(z, w) := \Phi^m_s(z + (0, \sigma(s, z)w)) - \Phi^m_s(z).
\]

By Itô’s formula and (7.4), one sees that

\[
\Phi^m_{t \wedge \tau_m}(Z_{t \wedge \tau_m}) - \Phi^m_0(z) = \int_0^{t \wedge \tau_m} \int_{\mathbb{R}^d} \left( \Phi^m_s(Z_{s-} + (0, \sigma(s, Z_{s-})w)) - \Phi^m_s(Z_{s-}) \right) \tilde{N}(ds, dw)
\]

\[
+ \int_0^{t \wedge \tau_m} (\partial_s \Phi^m + b \cdot \nabla \Phi^m + \mathcal{L}^{(\alpha)}_{\sigma,v}(\Phi^m))(s, Z_s)ds
\]

\[
= \int_0^{t \wedge \tau_m} \int_{B_m} g^m_s(Z_{s-}, w) \tilde{N}(ds, dw) + \lambda \int_0^{t \wedge \tau_m} u^m_t(s, Z_s)ds,
\]

where we have used that \( (b(1 - \chi^m))(s, Z_s) = 0 \) for \( s < \tau_m \) and

\[
\int_0^{t \wedge \tau_m} \int_{B_m} g^m_s(Z_{s-}, w) \tilde{N}(ds, dw) = \sum_{0 < s \leq t \wedge \tau_m} g^m_{s-}(Z_{s-}, \Delta L_s) \cdot 1_{|\Delta L_s| > m} = 0.
\]

Noticing that

\[
\nabla g^m_s(z, w) = (\nabla \Phi^m_s)(z + (0, \sigma(s, z)w))(\tilde{1} + (0, \nabla \sigma(s, z)w)) - \nabla \Phi^m_s(z),
\]

by (7.3) we have

\[
\|\nabla g^m_s(z, w)\|_{\infty} \leq 2(\|\sigma\|_{\infty}|w|)^{\alpha/2 + \varepsilon_0} + 2\|\nabla \sigma\|_{\infty}|w|.
\]

Hence, by the isometry formula of stochastic integrals,

\[
\mathbb{E}[Z_{t \wedge \tau_m} - Z'_{t \wedge \tau_m}]^2 \leq 4\mathbb{E}[\Phi^m_{t \wedge \tau_m}(Z_{t \wedge \tau_m}) - \Phi^m_{t \wedge \tau_m}(Z'_{t \wedge \tau_m})]^2
\]

\[28\]
\[
\begin{align*}
\mathbb{E} \int_{0}^{\tau_{m}^{C}} \int_{B_{m}} |g_{s}^{m}(Z_{s-}, w) - g_{s}^{m}(Z'_{s-}, w)|^{2} \frac{dw}{|w|^{d+\alpha}} ds
+ \lambda \mathbb{E} \int_{0}^{\tau_{m}^{C}} |u_{1}^{m}(s, Z_{s}) - u_{1}^{m}(s, Z'_{s})|^{2} ds
\leq \mathbb{E} \int_{0}^{\tau_{m}^{C}} |Z_{s} - Z'_{s}|^{2} ds \left( \int_{B_{m}} \frac{(|w|^{\alpha+2\sigma} + |w|^{2}) dw}{|w|^{d+\alpha}} + \lambda \right),
\end{align*}
\]
which yields by Gronwall’s inequality
\[
Z_{\tau_{m}^{C}} = Z'_{\tau_{m}^{C}}, \quad t \geq 0.
\]
Finally, letting \( m \to \infty \), we obtain the pathwise uniqueness.

**Remark 7.2.** By suitable localization technique, we can directly construct the solution by Zvonkin’s transformation without using Yamada-Watanabe’s theorem.

### 7.2. \( C^{1} \)-stochastic diffeomorphism flows for SDEs with additive Lévy noise

In this subsection we consider the \( C^{1} \)-stochastic diffeomorphism flows property for SDE (7.1) with additive Lévy noises. We introduce the following spaces: For a Fréchet space \( \mathbb{F} \) and time interval \( I \), define
\[
C(I; \mathbb{F}) := \{ f : I \to \mathbb{F} \text{ is continuous} \}, \quad D(I; \mathbb{F}) := \{ f : I \to \mathbb{F} \text{ is càdlàg} \}.
\]
For \( k \in \mathbb{N}_{0} \), let \( \mathbb{C}^{k} \) be the Fréchet space of all \( k \)-order continuous differentiable functions with Fréchet metric:
\[
d(f, g) := \sum_{j=0}^{k} \sum_{m \in \mathbb{N}} 2^{-m} \left( 1 \wedge \sup_{|x| < n} |\nabla^{j}f(x) - \nabla^{j}g(x)| \right).
\]
We have

**Theorem 7.3.** Let \( \alpha \in (1, 2), \gamma \in (1 + \frac{\alpha}{2}, 1 + \alpha) \) and \( \beta \in (1 - \frac{\alpha}{2}, 1) \). Assume \( \sigma \equiv 1 \) and
\[
b(t, x, v) = (v + b^{(1)}(t, x), b^{(2)}(t, x, v)), \quad (7.5)
\]
with
\[
b^{(1)} \in L_{x, \text{loc}}^{\gamma/\alpha}(\mathbb{C}^{\gamma/(1+\alpha)}), \quad b^{(2)} \in L_{x, \text{loc}}^{\gamma/\alpha}(\mathbb{C}^{\gamma/(1+\alpha)} \cap \mathbb{C}^{\beta}_{v}).
\]
Then the unique strong solution \([Z_{s,t}(z, \omega), t > s \geq 0, z \in \mathbb{R}^{2d}]\) of SDE (7.1) forms a \( C^{1} \)-stochastic diffeomorphism flow. More precisely, there is a null set \( N \) such that for all \( \omega \notin N \),
\[
(i) \text{ For all } 0 \leq s < r < t, \text{ it holds that } Z_{s,t}(z, \omega) = Z_{r,t}(Z_{s,r}(z, \omega), \omega), \quad \forall z \in \mathbb{R}^{2d},
\]
and
\[
z \mapsto Z_{s,t}(z, \omega) \text{ is a } C^{1}-\text{diffeomorphism on } \mathbb{R}^{2d}.
\]
(ii) \( t \mapsto Z_{s,t}(\cdot, \omega) \in D([s, \infty); \mathbb{C}^{1}) \) and \( s \mapsto Z_{s,t}(\cdot, \omega) \in D([0, t]; \mathbb{C}^{1}) \).

**Proof.** Fix \( T, \lambda > 0 \). Consider the following backward equation:
\[
\partial_{s} u_{1} + \Delta_{v}^{\alpha/2} u_{1} - \lambda u_{1} + b \cdot \nabla u_{1} + b = 0, \quad u_{1}(T, \cdot) = 0,
\]
where \( \Delta_{v}^{\alpha/2} \) is the fractional Laplacian acting on the variable \( v \). Under the assumptions of the theorem, by Theorem 6.11, there is a unique classical solution \( u_{1}^{m} \) to the above equation with regularity: for some \( C \geq 1 \) and all \( \lambda \geq 1 \),
\[
\|u_{1}\|_{L_{T}^{\gamma/\alpha}((1+\alpha)/\gamma^{\alpha}),\mathbb{C}^{\gamma/\alpha} \cap \mathbb{C}^{\beta}_{v})} \leq C, \quad \|u_{1}\|_{L_{T}^{\gamma/\alpha}} \leq C\lambda^{-1}.
\]
As in Theorem 7.1, by (2.10) and (4.3), there are \( \lambda \geq 1 \) large enough and \( \varepsilon_0 > 0 \) such that
\[
\| \nabla u_{3}(t, \cdot) \|_{C_{v}^{a/2+\varepsilon_0}} \leq \frac{1}{2}, \quad \| u_{3} \|_{L_{T}^{2}(C_{v}^{a/2+\varepsilon_0})} + \| u_{3} \|_{L_{T}^{2}(C_{v}^{a/2+\varepsilon_0})} < \infty. \tag{7.6}
\]
Define
\[
\Phi_{t}(z) := z + u_{3}(t, z).
\]
By (7.6), one sees that
\[
z \mapsto \Phi_{t}(z) \text{ is a diffeomorphism on } \mathbb{R}^{2d}
\]
and
\[
\| \nabla \Phi_{t} \|_{\infty} + \| \nabla \Phi_{t}^{-1} \|_{\infty} \leq 2.
\]
Let
\[
\tilde{Z}_{t} := \Phi_{t}(Z_{t}).
\]
By Itô’s formula, one sees that
\[
\tilde{Z}_{t} = \Phi_{t}(Z_{t}) = \Phi_{0}(z) + \int_{0}^{t} \left( \partial_s \Phi + b \cdot \nabla \Phi + \Delta_{v}^{a/2} \Phi \right)(s, Z_{s}) ds \\
+ \int_{0}^{t} \int_{\mathbb{R}^{d}} \left( \Phi_{s}(Z_{s-} + (0, w)) - \Phi_{s}(Z_{s-}) \right) \tilde{N}(ds, dw)
\]
\[
= \tilde{z} + \int_{0}^{t} b(s, \tilde{Z}_{s}) ds + \int_{0}^{t} \int_{\mathbb{R}^{d}} \tilde{g}_{s}(\tilde{Z}_{s-}, w) \tilde{N}(ds, dw), \tag{7.7}
\]
where \( \tilde{z} := \Phi_{0}(z) \) and
\[
\tilde{b}(t, z) := \lambda u_{3}(t, \Phi_{t}^{-1}(z)), \quad \tilde{g}_{s}(z, w) := \Phi_{t}(\Phi_{t}^{-1}(z)) + (0, w) - z.
\]
In particular, \( \{Z_{s,t}, t \geq 0\} \) solves SDE (7.1) if and only if \( \{\tilde{Z}_{s,t}, t \geq 0\} \) solves SDE (7.7) (see [10, Lemma 3.4]).

**Claim:** There are \( \varepsilon_1, \varepsilon_2 \in (0, \varepsilon_0) \) and constant \( C > 0 \) such that for all \( t \in [0, T] \) and \( w \in \mathbb{R}^{d} \),
\[
\| \nabla \tilde{g}_{t}(\cdot, w) \|_{C_{v}^{a/2+\varepsilon_1}} \leq C(|w|^{a/2+\varepsilon_2} \wedge 1), \quad \| \nabla \tilde{b} \|_{L_{T}^{2}(C_{v}^{a/2})} < \infty.
\]
Indeed, noticing that
\[
\nabla \tilde{g}_{t}(z, w) = (\nabla \Phi_{t})(\Phi_{t}^{-1}(z)) + (0, w) \cdot \nabla \Phi_{t}^{-1}(z) - I
\]
\[
= \left( (\nabla \Phi_{t})(\Phi_{t}^{-1}(z)) + (0, w) \right) - \nabla \Phi_{t} \circ \Phi_{t}^{-1}(z) \cdot \nabla \Phi_{t}^{-1}(z),
\]
by (7.6) we have
\[
\| \nabla \tilde{g}_{t}(\cdot, w) \|_{\infty} \leq \| \nabla \Phi_{t} \|_{C_{v}^{a/2+\varepsilon_0}} \frac{1}{2} \| \nabla \Phi_{t}^{-1} \|_{\infty} \leq 1 \wedge |w|^{a/2+\varepsilon_0},
\]
and by definition,
\[
\| \nabla \tilde{g}_{t}(\cdot, w) \|_{C_{v}^{a/2+\varepsilon_0}} \leq \| \nabla \Phi_{t}(\Phi_{t}^{-1}(\cdot)) + (0, w) \|_{C_{v}^{a/2+\varepsilon_0}} \| \nabla \Phi_{t}^{-1} \|_{\infty} + \| \nabla \Phi_{t} \|_{\infty} \| \nabla \Phi_{t}^{-1} \|_{C_{v}^{a/2+\varepsilon_0}} \leq 1.
\]
Thus we obtain the first claim by standard interpolation technique. The second one is easy by definition and (7.6).

By the above claim and [23, Theorem 4.1], the unique solutions of SDE (7.7) and so SDE (7.1) define a \( C^{1} \)-stochastic diffeomorphism flow and (i), (ii) hold. See also [10, 37] for more details.
Remark 7.4. Although we here only consider the symmetric and rotationally invariant \(\alpha\)-stable additive noise, by the same argument as used in [10], it is possible to consider more general additive Lévy noises, even cylindrical Lévy noises. On the other hand, by Remark 6.10, the \(\alpha\) in Theorem 7.3 in fact can be less than 1, but with a lower bound \((\sqrt{17} - 1)/4 \approx 0.78078\). Indeed, by restriction (6.29) and \(\gamma \in (1 + \frac{\alpha}{2}, 1 + \alpha)\), we need to require
\[
\alpha^2 + \alpha > 1 + \frac{\alpha}{2} \Rightarrow \alpha > (\sqrt{17} - 1)/4.
\]

7.3. Application to random transport equations with Hölder coefficients. In this subsection we apply Theorem 7.3 to a random transport equation with Hölder coefficient. First of all, the following corollary is an easy consequence of Theorem 7.3.

Corollary 7.5. Let \(\alpha \in (1, 2)\) and \(\gamma \in (\frac{2 + \alpha}{2(1 + \alpha)}, 1)\). Assume that
\[
b^{(1)} : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \in L_{\text{loc}}^\infty(\mathcal{C}^\gamma).
\]
Consider the following random ODE
\[
dY_{s,t}(x, \omega)/dt = b^{(1)}(t, Y_{s,t}(x, \omega)) + L_t^{(\alpha)}(\omega), \ t > s, \ Y_{s,s} = x.
\]  
(7.8)

For \(\mathbb{P}\)-almost all \(\omega\), there exists a family of solutions \(\{Y_{s,t}(x, \omega), x \in \mathbb{R}^d, 0 \leq s < t < \infty\}\) to the above random ODE so that
(i) For each \(s < r < t\), it holds that
\[
Y_{s,r}(x, \omega) = Y_{r,t}(Y_{s,r}(x, \omega), \omega), \ x \in \mathbb{R}^d,
\]
and
\[
x \mapsto Y_{s,t}(x, \omega) \text{ is a } C^1 \text{-diffeomorphism on } \mathbb{R}^d.
\]
(7.10)
(ii) For each \(s \geq 0, t \mapsto Y_{s,t}(\cdot, \omega) \in C([s, \infty), \mathcal{C}^0) \cap D([s, \infty), \mathcal{C}^1)\).

Proof. For \(z = (x, v)\), by Theorem 7.3 let
\[
Z_{s,t}(z) := (X_{s,t}(x, v), L_t - L_s + v)
\]
solve the degenerate SDE (7.1) with
\[
\sigma \equiv 1, \ b(t, x, v) := (v + b^{(1)}(t, x), 0).
\]

Now we define
\[
Y_{s,t}(x) := X_{s,t}(x, L_s)
\]
Clearly, it solves ODE (7.8). Now we check (i) and (ii) hold. By (i) of Theorem 7.3, we have
\[
(Y_{s,r}(x) , L_t) = (X_{s,r}(x, L_s), L_s) = Z_{s,r}(x, L_s) = Z_{s,t} \circ Z_{r,t}(x, L_s)
\]
\[
= Z_{t,t}(X_{s,t}(x, L_s), L_t) = (X_{t,t}(X_{s,t}(x, L_s), L_t), L_t) = (Y_{t,t} \circ Y_{s,t}(x), L_t),
\]
which implies (7.9) and (7.10) by (i) of Theorem 7.3. Finally, (ii) follows by (ii) of Theorem 7.3 and equation (7.8). \(\square\)

Remark 7.6. Here an open question is to show Davie’s uniqueness [15] for the above random ODE, that is, for almost all \(\omega\), ODE (7.8) has a unique solution. See [38] for the study of random ODE \(dY_t/dt = b(t, Y_t + L_t^{(\alpha)})\). We will study this problem in a future work.

We need the following real analysis result, which can be proven by the completely same method as in [39, p149, Theorem 7.21]. We omit the details.
Lemma 7.7. Let \( f : [a, b] \to \mathbb{R} \) be a continuous function. Assume that for each point \( t \in [a, b) \), the right derivative \( f'_r^+ := \lim_{\varepsilon \to 0^+} (f(t + \varepsilon) - f(t))/\varepsilon \) exists and \( f'_r^+ \in L^1([a, b]) \). Then \( f \) is absolutely continuous on \([a, b]\).

Now we can state the following result.

Theorem 7.8. Let \( \alpha \in (1, 2) \). Assume \( b \in C(\mathbb{R}_+ \times \mathbb{R}^d) \cap L^\infty_{loc}(\mathbb{R}^d) \) with \( \gamma \in (\frac{2+\alpha}{2(1+\alpha)}, 1) \). For any \( \varphi \in \mathscr{C}_1 \) and almost all \( \omega \), there is a unique function \((t, x) \mapsto u(t, x, \omega) \in C(\mathbb{R}_+; \mathbb{C}^0) \cap D(\mathbb{R}_+; \mathbb{C}^1)\) so that for each \( x \in \mathbb{R}^d \), \( t \mapsto u(t, x, \omega) \) is absolutely continuous and
\[
\partial_t u(t, x, \omega) + (b(t, x) + L_x(\omega)) \cdot \nabla_x u(t, x, \omega) = 0, \quad u(0, x) = \varphi(x). \tag{7.11}
\]

Proof. We only show the existence since the uniqueness is standard by the maximum principle (see [9, Theorem 6.1]). By Corollary 7.5, for each \( x \in \mathbb{R}^d \), let \( X_{s,t}(x, \omega) \) solve the following random ODE:
\[
X_{s,t}(x, \omega) = x + \int_s^t (b(r, X_{s,r}(x, \omega)) + L_x^{(a)}(\omega)) dr, \quad 0 \leq s \leq t. \tag{7.12}
\]

Define
\[
u(t, x, \omega) := \varphi(X_{t,0}^{-1}(x, \omega)).
\]
Clearly, by (ii) of Corollary 7.5, we have for almost all \( \omega \),
\[
u(t, \cdot, \omega) \in C(\mathbb{R}_+; \mathbb{C}^0) \cap D(\mathbb{R}_+; \mathbb{C}^1), \tag{7.13}
\]
and by (i) of Corollary 7.5, for \( \varepsilon > 0 \),
\[
u(t + \varepsilon, x) = \varphi(X_{t,0}^{-1}(x) \circ X_{t,t+\varepsilon}^{-1}(x)) = \nu(t, X_{t,t+\varepsilon}^{-1}(x)).
\]

Here and below we drop the \( \omega \). Hence,
\[
\frac{\nu(t + \varepsilon, x) - \nu(t, x)}{\varepsilon} = \frac{X_{t,t+\varepsilon}^{-1}(x) - x}{\varepsilon} \int_0^1 (\nabla_x u)(t, \theta X_{t,t+\varepsilon}^{-1}(x) + (1 - \theta)x) d\theta. \tag{7.14}
\]

Since \( \varepsilon \mapsto X_{t,t+\varepsilon}^{-1}(\cdot) \in D(\mathbb{R}_+; \mathbb{C}^1) \) and \( \nabla X_{t,t+\varepsilon}^{-1}(x) = (\nabla X_{t,t+\varepsilon})^{-1} \circ X_{t,t+\varepsilon}^{-1}(x) \), we have
\[
\limsup_{\varepsilon \to 0^+} \| \nabla X_{t,t+\varepsilon}^{-1}(x) - \| = 0, \quad \forall R > 0. \tag{7.15}
\]

Noticing that
\[
X_{t,t+\varepsilon}^{-1}(x) - x = (X_{t,t+\varepsilon}(x)) \cdot \int_0^1 \nabla X_{t,t+\varepsilon}^{-1}(\theta x + (1 - \theta)X_{t,t+\varepsilon}(x)) d\theta,
\]
since \((t, x) \mapsto b(t, x)\) is continuous and \( t \mapsto L_t^{(a)} \) is right continuous, by (7.12) and (7.15), for each \((t, x) \in \mathbb{R}_+ \times \mathbb{R}^d\), we have
\[
\lim_{\varepsilon \to 0^+} (X_{t,t+\varepsilon}^{-1}(x) - x)/\varepsilon = -b(t, x) - L_t^{(a)}.
\]

Therefore, by (7.14) and the continuity of \( x \mapsto \nabla_x u(t, x) \),
\[
\partial_t^* u(t, x) := \lim_{\varepsilon \to 0^+} \frac{\nu(t + \varepsilon, x) - \nu(t, x)}{\varepsilon} = -(b(t, x) + L_t^{(a)}) \cdot \nabla_x u(t, x), \quad \forall (t, x) \in \mathbb{R}_+ \times \mathbb{R}^d,
\]
where \( \partial_t^* u \) stands for the right derivative. Since \( t \mapsto (b(t, x) + L_t^{(a)}) \cdot \nabla_x u(t, x) \) is bounded, by Lemma 7.7, \( t \mapsto u(t, x) \) is absolutely continuous. The proof is complete. \( \square \)

Remark 7.9. If \( \varphi \in L^\infty(\mathbb{R}^d) \) and \( \text{div} b \in L^1_{loc}(\mathbb{R}_+ \times \mathbb{R}^d) \), then as in [22, Theorem 20], we can show that \( u(t, x) := \varphi(X_{t,0}^{-1}(x)) \) is the unique bounded weak solution of transport equation (7.11).
8. Appendix

In this section we use a probabilistic method to show the existence of a smooth solution to PDE (6.1) when the coefficients are smooth. We first recall the following result proved in [25].

**Lemma 8.1.** Given \( d_0 \in \mathbb{N} \) and \( c_0 > 1 \), let \( \kappa(x, z) : \mathbb{R}^{d_0} \times B_1 \to [c_0^{-1}, c_0] \) be a smooth function with bounded derivatives of all orders. For any \( \alpha \in (0, 2) \), there is a measurable map \( \Phi(x, z) : \mathbb{R}^{d_0} \times B_1 \to B_1 \) such that for any nonnegative measurable function \( f \),

\[
\int_{B_1} f \circ \Phi(x, z) \frac{dz}{|z|^{d+\alpha}} = \int_{B_1} f(z) \kappa(x, z) \frac{dz}{|z|^{d+\alpha}}.
\]

Moreover, \( \Phi \) enjoys the following properties:

(i) \( \Phi(x, 0) = 0 \) and if \( \kappa(x, -z) = \kappa(x, z) \), then \( \Phi(x, -z) = -\Phi(x, z) \).

(ii) For all \( i, j \in \mathbb{N}_0 \), there is a \( C_{ij} > 0 \) such that for all \( x \in \mathbb{R}^{d_0} \) and \( z \in B_1 \),

\[
|\nabla_i \nabla_j \Phi(x, z)| \leq C_{ij}|z|^{-j},
\]

where \( C_{ij} \) is a polynomial of \( \|\nabla^m \nabla^s \kappa\|_{\infty}, m = 1, \cdots, i, n = 0, \cdots, j \).

**Theorem 8.2.** Suppose that \( \kappa \) and \( b \) satisfy that for any \( m \in \mathbb{N} \) and \( t > 0 \),

\[
c_0^{-1} \leq \kappa(t, x, v, w) \leq c_0, \quad \|\nabla^m \kappa(t, \cdot)\|_{\infty} + \|\nabla^m b(t, \cdot)\|_{\infty} \leq c_m.
\]

Then there is a classical solution \( u \in \cap_{m \in \mathbb{N}} C(\mathbb{R}^+; \mathcal{C}^m) \) to PDE (6.1).

**Proof.** We decompose the operator \( \mathcal{L}^{(a)}_{k, \nu} \) as two parts: small jumps \( \mathcal{L}^{(a)}_{k, \nu} \) and large jumps \( \mathcal{L}^{(a)}_{\nu} \).

\[
\mathcal{L}^{(a)}_{k, \nu} f = \left( \int_{B_1} + \int_{B_1} \right) \delta^{(2)}_w f(x, v) \kappa(t, x, v, w) \frac{dw}{|w|^{d+\alpha}} =: \mathcal{L}^{(a)}_{k, \nu} f + \mathcal{L}^{(a)}_{\nu} f,
\]

where

\[
\delta^{(2)}_w f(x, v) := f(x, v + w) + f(x, v - w) - 2f(x, v).
\]

By Lemma 8.1, there is a measurable map \( g(t, x, v, w) : \mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}^d \times B_1 \to B_1 \) with

\[
|\nabla_i \nabla_j g(t, x, v, w)| \leq C_{ij}|w|^{-j},
\]

and so that

\[
\mathcal{L}^{(a)}_{\nu} f(x, v) = \int_{B_1} \delta^{(2)}_w f(x, v) \frac{dw}{|w|^{d+\alpha}}.
\]

Now we consider the following SDE:

\[
dZ_{s,t} = \int_{B_1} g(t, Z_{s,t-}, w) \bar{N}(dt, dw) + b(t, Z_{s,t}) dt, \quad Z_{s,s} = z.
\]

Since the coefficients are smooth and have bounded derivatives of all orders greater than 1, it is well-known that there is a unique solution \( Z_{s,t}(z) \), which forms a \( C^\infty \)-stochastic flows (cf. [23, Theorem 4.1]). More precisely, it holds that for any \( s < r < t \) and \( z \in \mathbb{R}^d \),

\[
Z_{s,r}(z) = Z_{r,t} \circ Z_{s,t}(z), \quad a.s.,
\]

and for any \( j \in \mathbb{N} \) and \( p \geq 1 \),

\[
\sup_{z \in \mathbb{R}^d} \sup_{0 \leq s < r \leq T} \mathbb{E}[|\nabla^j Z_{s,r}(z)|^p] < \infty. \tag{8.1}
\]
Moreover, let \( f : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R} \) be a measurable function with \( \| \nabla_j f \|_\infty < \infty \) for each \( j \in \mathbb{N} \). It is also well-known that

\[
(x, v) \mapsto u(s, x, v) := \int_s^T e^{\lambda(s-t)} \mathbb{E} f(t, Z_{s,t}(x, v)) dt \in \mathcal{E}_m(\mathbb{R}^d)
\]
solves the following PDE:

\[
\partial_s u + \mathcal{L}^{(a)}_{k,v} u + b \cdot \nabla u - \lambda u = f, \quad u(T, \cdot) = 0.
\]

From the representation, by the chain rule and (8.1), it is easy to see that for any \( m \in \mathbb{N} \), there is a constant \( C > 0 \) such that for all \( s \in [0, T] \) and \( \lambda \geq 0 \),

\[
\| u(s) \|_{\mathcal{E}^m} \leq C \int_s^T e^{\lambda(s-t)}\| f(t, \cdot) \|_{\mathcal{E}^m} dt \leq C \int_s^T \| f(t, \cdot) \|_{\mathcal{E}^m} dt,
\]

and by the definition of \( \mathcal{L}^{(a)}_{k,v} u \),

\[
\| \mathcal{L}^{(a)}_{k,v} u \|_{\mathcal{E}^m} \leq C \| u \|_{\mathcal{E}^m}.
\]

Next we consider the following Picard’s iteration: Fix \( m \geq 2 \). Let \( u_0 \equiv 0 \). For \( n \in \mathbb{N} \), let \( u_n \) be the unique \( \mathcal{E}^m \)-valued solution of the following PDE:

\[
\partial_s u_n + \mathcal{L}^{(a)}_{k,v} u_n + b \cdot \nabla u_n - \lambda u_n = f - \mathcal{L}^{(a)}_{k,v} u_{n-1}, \quad u_n(T, \cdot) = 0. \tag{8.3}
\]

By (8.2) we have

\[
\| u_n(s) \|_{\mathcal{E}^m} \leq C \int_s^T \| f(t) - \mathcal{L}^{(a)}_{k,v} u_{n-1}(t) \|_{\mathcal{E}^m} dt \leq C \int_s^T \left( \| f(t) \|_{\mathcal{E}^m} + \| u_{n-1}(t) \|_{\mathcal{E}^m} \right) dt,
\]

which yields by Gronwall’s inequality that for some \( C > 0 \),

\[
\sup_{n} \sup_{s \in [0, T]} \| u_n(s) \|_{\mathcal{E}^m} \leq C \sup_{s \in [0, T]} \| f(s) \|_{\mathcal{E}^m}. \tag{8.4}
\]

Similarly, we also have

\[
\| u_n(s) - u_k(s) \|_{\mathcal{E}^m} \leq C \int_s^T \left( \| u_{n-1}(t) - u_{k-1}(t) \|_{\mathcal{E}^m} \right) dt, \quad n, k \in \mathbb{N}.
\]

By (8.4) and Fatou’s lemma, we have

\[
\lim sup_{n,k \to \infty} \sup_{t \in [s,T]} \| u_n(t) - u_k(t) \|_{\mathcal{E}^m} \leq C \int_s^T \lim sup_{n,k \to \infty} \left( \| u_{n-1}(t) - u_{k-1}(t) \|_{\mathcal{E}^m} \right) dt,
\]

which yields by Gronwall’s inequality again

\[
\lim sup_{n,k \to \infty} \sup_{t \in [0,T]} \| u_n(t) - u_k(t) \|_{\mathcal{E}^m} = 0.
\]

Therefore, there exists a \( u \in C([0, T]; \mathcal{E}^m) \) so that

\[
\lim_{n \to \infty} \sup_{t \in [0, T]} \| u_n(t) - u(t) \|_{\mathcal{E}^m} = 0.
\]

By taking limits for equation (8.3), we obtain the existence of \( \mathcal{E}^m \)-valued solution to PDE (6.1). Since \( m \) is arbitrary, we complete the proof. \( \square \)
REFERENCES

[1] L. Ambrosio, Transport equation and Cauchy problem for BV vector fields. *Invent. Math.*, **158** (2004), 227–260.

[2] J. Bae and M. Kassmann, Schauder estimates in generalized Hölder spaces. Available at arXiv:1505.05498.

[3] H. Bahouri, J.-Y. Chemin, and R. Danchin, *Fourier analysis and nonlinear partial differential equations*. Grundlehren der Mathematischen Wissenschaften, Vol.343. Springer, Heidelberg, 2011.

[4] R. F. Bass, Regularity results for stable-like operators. *J. Funct. Anal.*, **257** (2009), 2693–2722.

[5] P. E. Chaudru de Raynal, Strong existence and uniqueness for stochastic differential equation with Hölder drift and degenerate noise. *Annales de l’Institut Henri Poincaré*, **53** (1) (2017), 259–286.

[6] P. E. Chaudru de Raynal, I. Honoré, and S. Menozzi, Sharp Schauder estimates for some degenerate Kolmogorov equations. Available at arXiv:1810.12227.

[7] P. E. Chaudru de Raynal, I. Honoré, and S. Menozzi, Strong regularization by Brownian noise propagating through a weak Hörmander structure. Available at arXiv:1810.12225.

[8] P. E. Chaudru de Raynal, S. Menozzi, and E. Priola, Schauder estimates for drifted fractional operators in the supercritical case. Available at arXiv:1902.02616.

[9] Z.-Q. Chen, E. Hu, L. Xie, and X. Zhang, Heat kernels for non-symmetric diffusion operators with jumps. *J. Differential Equations*, **263** (2017), 6576–6634.

[10] Z.-Q. Chen, R. Song, and X. Zhang, Stochastic flows for Lévy processes with Hölder drifts. *Rev. Mat. Iberoam.*, **34** (2018), 1755–1788.

[11] Z.-Q. Chen and X. Zhang, $L^p$-maximal hypoelliptic regularity of nonlocal kinetic Fokker-Planck operators. *J. Math. Pures Appl.* (9), **116** (2018), 52–87.

[12] Z.-Q. Chen and X. Zhang, Propagation of regularity in $L^p$-spaces for Kolmogorov type hypoelliptic operators. Available at arXiv:1706.02181.

[13] Z.-Q. Chen, X. Zhang, and G. Zhao, Well-posedness of supercritical SDE driven by Lévy processes with irregular drifts. Available at arXiv:1709.04632.

[14] S. Dachkovski, Anisotropic function spaces and related semi-linear hypoelliptic equations. *Math. Nachr.*, **248/249** (2003), 40–61.

[15] A.M. Davie, Uniqueness of solutions of stochastic differential equations. *Int. Math. Res. Notices*, **124** (2007), no.24, 26pages.

[16] R. J. Diperna and P. L. Lions, Ordinary differential equations, transport theory and Sobolev spaces. *Invent. Math.*, **98** (1989), 511–547.

[17] H. Dong and D. Kim, Schauder estimates for a class of non-local elliptic equations. *Discrete Contin. Dyn. Syst.*, **33** (2013), 2319–2347.

[18] S. Fang, D. Luo, and A. Thalmaier, Stochastic differential equations with coefficients in Sobolev spaces. *J. Func. Anal.*, **259** (2011), 1129-1168.

[19] A. Figalli, Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients. *J. Func. Anal.*, **254** (2008), 109-153.

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

[21] E. Fedrizzi, F. Flandoli, E. Priola, and J. Voivelle, Regularity of stochastic kinetic equations. *Electron. J. Probab.*, **22** (2017), Paper No. 48, 42.

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

[23] T. Fujiwara and H. Kunita, Canonical SDE’s based on semimartingales with spatial parameters. I. Stochastic flows of diffeomorphisms. *Kyushu J. Math.*, **53** (1999), 265–300.

[24] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*. Classics in Mathematics. Springer-Verlag, Berlin, 2001.

[25] Z. Hao, X. Peng, and X. Zhang, Hörmander’s hypoelliptic theorem for nonlocal operators. Available at arXiv:1901.06621.

[26] L. Huang, S. Menozzi, and E. Priola, $L^p$ Estimates for degenerate non-local Kolmogorov operators. *J. Math. Pures Appl.*, **121** (2019), 162–215.
[27] N. Ikeda and S. Watanabe, *Stochastic differential equations and diffusion processes*. North-Holland Mathematical Library, Vol.24. North-Holland Publishing Co., Amsterdam; Kodansha, Ltd., Tokyo, 1989.

[28] C. Imbert, T. Jin, and R. Shvydkoy, Schauder estimates for an integro-differential equation with applications to a nonlocal Burgers equation. *Ann. Fac. Sci. Toulouse Math.* (6), **27** (2018), 667–677.

[29] J. Jacod and A. N. Shiryaev, *Limit theorems for stochastic processes*. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Vol.288. Springer-Verlag, Berlin, 2003.

[30] N. V. Krylov, *Lectures on elliptic and parabolic equations in Hölder spaces*. Graduate Studies in Mathematics, Vol.12. American Mathematical Society, Providence, RI, 1996.

[31] N. V. Krylov and E. Priola, Poisson stochastic process and basic Schauder and Sobolev estimates in the theory of parabolic equations. *Arch. Ration. Mech. Anal.*, **225** (2017), 1089–1126.

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

[33] H. Kunita, *Stochastic flows and stochastic differential equations*. Cambridge Studies in Advanced Mathematics, Vol.24. Cambridge University Press, Cambridge, 1990.

[34] L. Lorenzi, Schauder estimates for degenerate elliptic and parabolic problems with unbounded coefficients in $\mathbb{R}^N$. *Differential Integral Equations*, **18** (2005), 531–566.

[35] A. Lunardi, Schauder estimates for a class of degenerate elliptic and parabolic operators with unbounded coefficients in $\mathbb{R}^n$. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4), **24** (1997), 133–164.

[36] E. Priola, Global Schauder estimates for a class of degenerate Kolmogorov equations. *Studia Math.*, **194** (2009), 117–153.

[37] E. Priola, Pathwise uniqueness for singular SDEs driven by stable processes. *Osaka J. Math.*, **49** (2012), 421–447.

[38] E. Priola, Davie’s type uniqueness for a class of SDEs with jumps. *Ann. Inst. Henri Poincaré Probab. Stat.*, **54** (2018), 694–725.

[39] W. Rudin, *Real and complex analysis*. McGraw-Hill Book Co., New York, 1987.

[40] H. Triebel, *Theory of function spaces*. Monographs in Mathematics, Vol.78. Birkhäuser Verlag, Basel, 1983.

[41] H. Triebel, *Theory of function spaces. II*. Monographs in Mathematics, Vol.84. Birkhäuser Verlag, Basel, 1992.

[42] C. Villani, *A review of mathematical topics in collisional kinetic theory*. Handbook of Fluid Mechanics. Ed. S. Friedlander, D.Serre, 2002.

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

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

[45] X. Zhang, Strong solutions of SDEs with singular drift and Sobolev diffusion coefficients. *Stoch. Proc. and Appl.*, **115** (2005), no.11, 1805-1818.

[46] X. Zhang, Stochastic flows of SDEs with irregular coefficients and stochastic transport equations. *Bull. Sci. Math. France*, **134** (2010), 340–378.

[47] X. Zhang, Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients. *Electron. J. Probab.*, **16** (2011), no. 38, 1096–1116.

[48] X. Zhang, Well-posedness and large deviation for degenerate SDEs with Sobolev coefficients. *Rev. Mat. Iberoam.*, **29** (2013), no. 1, 25-52.

[49] X. Zhang and G. Zhao, Heat kernel and ergodicity of SDEs with distributional drifts. Available at arXiv:1710.10537.

[50] X. Zhang and G. Zhao, Dirichlet problem for supercritical nonlocal operators. Available at arXiv:1809.05712.

[51] A. K. Zvonkin, A transformation of the phase space of a diffusion process that removes the drift. *Mat. Sb.* (N.S.), **93**(135) (1974), 129–149, 152.

School of Mathematics and Statistics, Wuhan University, Wuhan, Hubei 430072, P.R.China, Emails: zimo-hao@whu.edu.cn, mywu@whu.edu.cn, XichengZhang@gmail.com