Propagation of regularity in $L^p$-spaces for Kolmogorov-type hypoelliptic operators

ZHENG-QING CHEN AND XICHENG ZHANG

Abstract. Consider the following Kolmogorov-type hypoelliptic operator

$$\mathcal{L}_t := \sum_{j=2}^{n} x_j \cdot \nabla x_{j-1} + \text{tr}(a_t \cdot \nabla^2 x_n)$$

on $\mathbb{R}^{nd}$, where $n \geq 2$, $d \geq 1$, $x = (x_1, \ldots, x_n) \in (\mathbb{R}^d)^n = \mathbb{R}^{nd}$ and $a_t$ is a time-dependent constant symmetric $d \times d$-matrix that is uniformly elliptic and bounded. Let $\{T_{s,t}; t \geq s\}$ be the time-dependent semigroup associated with $\mathcal{L}_t$; that is, $\partial_s T_{s,t} f = -\mathcal{L}_s T_{s,t} f$. For any $p \in (1, \infty)$, we show that there is a constant $C = C(p, n, d) > 0$ such that for any $f(t, x) \in L^p(\mathbb{R} \times \mathbb{R}^{nd}) = L^p(\mathbb{R}^{1+nd})$ and every $\lambda \geq 0$,

$$\|\Delta_{x_j}^{1/(1+2(n-j))} \int_0^\infty e^{-\lambda t} T_{s,t} f(t+s, x) dt \|_p \leq C\|f\|_p, \quad j = 1, \ldots, n,$$

where $\|\cdot\|_p$ is the usual $L^p$-norm in $L^p(\mathbb{R} \times \mathbb{R}^{nd}; \text{d}s \times \text{d}x)$. To show this type of estimates, we first study the propagation of regularity in $L^2$-space from variable $x_n$ to $x_j$, $1 \leq j \leq n-1$, for the solution of the transport equation $\partial_t u + \sum_{j=2}^{n} x_j \cdot \nabla x_{j-1} u = f$.

1. Introduction

Let $n \geq 2$ and $d \in \mathbb{N}$. Denote by $\mathbb{M}_d^{\text{sym}}$ the set of all symmetric $d \times d$-matrices. In this paper, we consider the following Kolmogorov-type hypoelliptic operator on $\mathbb{R}^{nd}$:

$$\mathcal{L}_t := \sum_{i,j=1}^{d} a_{ij}^t \partial_{x_{ni}} \partial_{x_{nj}} + \sum_{j=2}^{n} x_j \cdot \nabla x_{j-1}, \quad (1.1)$$

where $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^{nd}$ with $x_j = (x_j, \ldots, x_{jd}) \in \mathbb{R}^d$ for each $j = 1, \ldots, n$, $\nabla x_j = (\partial_{x_{j1}}, \ldots, \partial_{x_{jd}})$, and $a_t = (a_{ij}^t) : \mathbb{R} \to \mathbb{M}_d^{\text{sym}}$ is a measurable map (independent of $x$) having

$$\kappa^{-1} \|d \times d \| \leq a_t \leq \kappa \|d \times d \| \quad (1.2)$$

Mathematics Subject Classification: Primary 42B20, 60H30; Secondary 35H10, 35Q20

Keywords: Kolmogorov’s hypoelliptic operators, Fefferman–Stein’s theorem, Propagation of $L^p$-regularity.

This work is partially supported by Simons Foundation Grant 520542 and by NNSFC Grant of China (No. 11731009).
for some $\kappa \geq 1$. Here $I_{d \times d}$ stands for the $d \times d$ identity matrix. Let $\nabla := (\nabla_{x_1}, \ldots, \nabla_{x_n})$, $\nabla^2_{x_n} := (\partial_{x_{ni}} \partial_{x_{nj}})_{i,j=1,\ldots,d}$ and

$$A = A_n = \begin{pmatrix}
I_{d \times d} & \mathbb{I}_{d \times d} & 0_{d \times d} & \cdots & \cdots \\
0_{d \times d} & I_{d \times d} & \mathbb{I}_{d \times d} & \cdots & \cdots \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
\vdots & \cdots & \ddots & \cdots & \ddots \\
0_{d \times d} & \cdots & \cdots & 0_{d \times d} & \mathbb{I}_{d \times d}
\end{pmatrix}_{nd \times nd}. \tag{1.3}
$$

We can rewrite $\mathcal{L}_t$ in the following compact form:

$$\mathcal{L}_t = \text{tr}(a_t \cdot \nabla^2_{x_n}) + Ax \cdot \nabla,$$

where “tr” denotes the trace of a matrix. The hypoelliptic operator $\mathcal{L}_t$ is a differential operator of second order in $x_n$ variable but is of order 1 in variables $x_1, \ldots, x_{n-1}$. These variables are connected through the drift terms $x_j \cdot \nabla_{x_{j-1}}$ for $2 \leq j \leq n$. In this paper, we study the regularity of the resolvent functions of $\mathcal{L}$. Roughly speaking, we show that Laplacian in $x_n$-variable of the resolvent of $\mathcal{L}$ is a bounded operator in $L^p$-space, and that this regularity in $x_n$ propagates to other variables in such a way that the fractional Laplacian in $x_j$-variable of power $1/(1 + 2(n-j))$ of the resolvent of $\mathcal{L}$ is a bounded operator in $L^p$-space for $1 \leq j \leq n-1$; see Theorem 1.1 for a precise statement.

The operator $\mathcal{L}_t$ corresponds to a degenerate diffusion process on $\mathbb{R}^{nd}$. Consider the following linear stochastic differential equations (SDEs) in $\mathbb{R}^{nd}$:

$$dX_t^{s,x} = AX_t^{s,x} dt + \sigma_t^a dW_t \quad \text{for } t > s \text{ with } X_s^{s,x} = x \in (\mathbb{R}^d)^n, \tag{1.4}
$$

where $(W_t)_{t \in \mathbb{R}}$ is a standard $nd$-dimensional Brownian motion and

$$\sigma_t^a := \begin{pmatrix}
0_{(n-1)d \times (n-1)d} & 0_{(n-1)d \times d} \\
0_{d \times (n-1)d} & (\sqrt{2}a_t)_{d \times d}
\end{pmatrix}_{nd \times nd}. \tag{1.5}
$$

It is easy to see that the solution $X_t^{s,x}$ of (1.4) is explicitly given by

$$X_t^{s,x} = e^{(t-s)A}x + \int_s^t e^{(t-r)A} \sigma_r^a dW_r,$$

where $e^{tA}$ is the exponential matrix of $A$ with the expression

$$e^{tA} = \begin{pmatrix}
\mathbb{I}_{d \times d} & t \mathbb{I}_{d \times d} & \frac{t^2}{2} \mathbb{I}_{d \times d} & \cdots & \frac{t^{n-1}}{(n-1)!} \mathbb{I}_{d \times d} \\
0_{d \times d} & \mathbb{I}_{d \times d} & t \mathbb{I}_{d \times d} & \cdots & \frac{t^{n-2}}{(n-2)!} \mathbb{I}_{d \times d} \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
\vdots & \cdots & 0_{d \times d} & \mathbb{I}_{d \times d} & t \mathbb{I}_{d \times d} \\
0_{d \times d} & \cdots & \cdots & 0_{d \times d} & \mathbb{I}_{d \times d}
\end{pmatrix}_{nd \times nd} \tag{1.6}.$$
Notice that if }_t = a does not depend on } (i.e., time homogeneous), then
\[ X_t^{x, x} \overset{(d)}{=} Z_t^{x, x}, \text{ where } Z_t^x := e^{tA}x + \int_0^t e^{(t-r)A} \sigma^d dW_r. \tag{1.7} \]

In this case, }_t is an }{(nd)}-dimensional Gaussian random variable with density
\[ y \mapsto p_t(x, y) = \frac{\sqrt{(2\pi)^{nd} t^{nd} \det(\Sigma)}}{\sqrt{2\pi}^n} e^{\frac{-1}{2} (y - e^{tA}x)^T \Sigma^{-1} (y - e^{tA}x)}, \]
where }_t : \mathbb{R}^{nd} \to \mathbb{R}^{nd} is the dilation operator defined by
\[ \Theta_r(x) = (r^{2n-1} x_1, r^{2n-3} x_2, \ldots, r x_n), \tag{1.8} \]
and } := \int_0^1 e^{rA} \sigma^d (\sigma^d)^* e^{rA^*} dr is the covariance matrix of } (see [8] or (2.2)).
Here and in the sequel, for either a vector or a matrix } , we use }^T to denote its transpose.
The above-defined dilation operator } enjoys the property that
\[ \Theta_r(e^{tA} \xi) = e^{r^2tA} \Theta_r(\xi) \quad \text{for every } r, t > 0. \tag{1.9} \]

Define the time-inhomogeneous transition semigroup \{T_{s,t}; s < t\} of } by
\[ T_{s,t} f(x) := \mathbb{E} f(X_t^{x, x}), \quad x \in \mathbb{R}^{nd}, \tag{1.10} \]
for } \in \mathcal{C}^2_b(\mathbb{R}^{nd}). It is easy to check that } is the infinitesimal generator of \{T_{s,t}; s < t\}, that is,
\[ \partial_s T_{s,t} f + \mathcal{L}_s T_{s,t} f = 0 \quad \text{for } f \in \mathcal{C}_b^2(\mathbb{R}^{nd}). \tag{1.11} \]

The goal of this paper is to establish the following }-maximal regularity estimate.

**Theorem 1.1.** Let } \in (1, \infty). Under the uniform ellipticity condition (1.2), there is a constant } = } } } that for any } \in \mathcal{L}^p(\mathbb{R} \times \mathbb{R}^{nd}) and } \geq 0,
\[ \left\| \Delta_j^{1/((1+2(n-j))} \int_0^\infty e^{-tA} T_{s,t+s} f(t+s, x) dt \right\| \leq C \| f \|_p, \quad j = 1, \ldots, n, \tag{1.12} \]
where }_j := (\Delta_j)^{-1/((1+2(n-j))} is the fractional Laplacian acting on the }th variable } \in \mathbb{R}^d.

Note that
\[ u(s, x) := \int_0^\infty e^{-tA} T_{s,t+s} f(t+s, x) dt = \int_0^\infty e^{-tA} T_{s,t} f(t, x) dt \]
satisfies
\[ \partial_s u(s, x) + (\mathcal{L}_s - \lambda) u(s, x) + f(s, x) = 0. \tag{1.13} \]

One of the motivation of studying the estimate (1.12) comes from the study of the following \(n+1\)-order stochastic differential equation:
\[ dX_t^{(n)} = b_t(X_t, X_t^{(1)}, \ldots, X_t^{(n)}) \, dr + \sigma_t(X_t, X_t^{(1)}, \ldots, X_t^{(n)}) \, d\tilde{W}_t, \tag{1.14} \]
where $X_t^{(n)}$ denotes the $n$th-order derivative of $X_t$ with respect to the time variable, $b : \mathbb{R}_+ \times \mathbb{R}^{(n+1)d} \to \mathbb{R}^d$ and $\sigma : \mathbb{R}_+ \times \mathbb{R}^{(n+1)d} \to \mathbb{R}^d \otimes \mathbb{R}^d$ are measurable functions, and $\tilde{W}_t$ is a $d$-dimensional Brownian motion. Notice that if we let

$$X_t := (X_t, X_t^{(1)}, \ldots, X_t^{(n)}),$$

then $X_t$ solves the following one-order stochastic differential equation

$$dX_t = (X_t^{(1)}, \ldots, X_t^{(n)}, b_t(X_t))dt + (0, \ldots, 0, \sigma_t(X_t)d\tilde{W}_t), \quad X_0 = x,$$

where $x = (x_i)_{i=0,\ldots,n} = ((x_{ij})_{j=1,\ldots,d})_{i=0,\ldots,n}$. In particular, the infinitesimal generator of Markov process $X_t(x)$ is given by

$$L_t f(x) = \sum_{i,j,k=1}^d (\sigma_{ik}^j \sigma_{jk}^i)(x) \partial_{x_{ni}} \partial_{x_{nj}} f(x) + \sum_{j=1}^n x_j \cdot \nabla_{x_{j-1}} f(x) + b_t(x) \cdot \nabla_{x_n} f(x),$$

which is of the form similar to (1.1). Thus, the estimate (1.12) could be used to study the well-posedness of SDE (1.14) with rough coefficients $b$ and $\sigma$. Indeed, when $n = 1$ and $\sigma$ is bounded and uniformly nondegenerate, the second named author [12] studied the strong well-posedness of SDE (1.14) with both $(I - \Delta_{x_1})^{1/3} b$ and $\nabla \sigma$ in $L^p_{loc}(\mathbb{R}_+ \times \mathbb{R}^{2d})$ for some $p > 4d + 2$. See also [6] for similar results when $\sigma_t = I_{d \times d}$.

We now recall some related results in the literature about the estimate (1.12). In [9], Lunardi showed the Hölder regularity estimate for degenerate equation (1.1). In [3], the authors adopted Coifman–Weiss’ theorem to show the estimate (1.12) for $j = n$. Variable coefficients case is considered in [4] by the same authors. When $n = 2$, in [5], we established a version of Fefferman–Stein’s theorem and then used it to show the estimate (1.12) for $j = 1, 2$ even for nonlocal operators. It should be noticed that the methods used in [3, 5] are quite different. In [3], the key point is to show some weak 1–1-type estimate. While in [5], the main point is to show that the operator in (1.12) is bounded from $L^\infty$ to some BMO spaces. In particular, to show the propagation of the regularity from the nondegenerate component to the degenerate component, in [5], we have used Bouchut’s result [2]. More precisely, Bouchut studied the following transport equation:

$$\partial_t u + x_2 \cdot \nabla_{x_1} u = f,$$

and showed that for any $\alpha > 0$,

$$\| \Delta_{x_1}^{\alpha/(2(1+\alpha))} u \|_2 \leq C \| \Delta_{x_2}^{\alpha/2} u \|_2^{1/(1+\alpha)} \| f \|_2^{\alpha/(1+\alpha)},$$

where $C = C(\alpha, d) > 0$. A simplified proof of this type estimate was provided in [1]. Thus, the first goal of this paper is to extend the above estimate to the following more general transport equation:

$$\partial_t u + \sum_{j=2}^n x_j \cdot \nabla_{x_{j-1}} u = f.$$
That is, we want to show that for any $j = 1, \ldots, n - 1$ and $\alpha > 0$, there is a constant $C = C(\alpha, d, j, n) > 0$ such that
\[
\|\Delta_{x_j}^{\alpha/(2(1+\alpha))} u\|_2 \leq C\|\Delta_{x_{j+1}}^{\alpha/2} u\|_2^{1/(1+\alpha)} \| f \|_2^{\alpha/(1+\alpha)}.
\]
Such an extension from $n = 2$ to $n \geq 3$ is non-trivial; see Sect. 3.

Although the above result is proven for Laplacian operator, we can extend it to more general nonlocal operator as in [5] without any difficulty. Indeed, consider the following nonlocal operator:
\[
\tilde{L}_\nu f(x) := \int_{\mathbb{R}^d} \left[ f(x + \sigma y) + f(x - y) - 2f(x) \right] \nu(dy),
\]
where $\sigma \in M^d$ is a $d \times d$ matrix and $\nu$ is a symmetric Lévy measure on $\mathbb{R}^d$ (that is, $\nu$ is a measure on $\mathbb{R}^d \setminus \{0\}$ with $\nu(A) = \nu(-A)$ and $\int_{\mathbb{R}^d \setminus \{0\}} (1 \wedge |z|^2) \nu(dz) < \infty$). Let $n \geq 2$ and
\[
\mathcal{L}_t f(x) := \tilde{L}_{\sigma_t, x_n} f(x) + \sum_{j=2}^n x_j \cdot \nabla_{x_{j-1}} f(x),
\]
where $\tilde{L}_{\sigma_t, x_n}$ means that the operator acts on the variable $x_n$. Suppose that
\[
\|\sigma\|_\infty + \|\sigma^{-1}\|_\infty < \infty
\]
and for some $\alpha \in (0, 2)$,
\[
\nu_1^{(\alpha)} \leq \nu_s \leq \nu_2^{(\alpha)},
\]
where $\nu_1^{(\alpha)}$ and $\nu_2^{(\alpha)}$ are two symmetric and nondegenerate $\alpha$-stable Lévy measures (see [5]). Under the above assumptions, we can show as in [5] that for any $j = 1, \ldots, n$,
\[
\left\| \Delta_{x_j}^{\alpha/(2(1+\alpha(n-j)))} \int_0^\infty e^{-\lambda t} T_{s,t+s}^{\nu, \sigma} f(t + s, x) dt \right\|_p \leq C \| f \|_p,
\]
(1.15)
where $T_{s,t}^{\nu, \sigma}$ is defined as in (1.10) by using the time-inhomogeneous Markov process $\{\{Z_{t}^{x,s}; t \geq 0\}; (s, x) \in \mathbb{R} \times \mathbb{R}^d\}$ determined by the family of Lévy measures $\{\nu_s, s \in \mathbb{R}\}$ in place of Brownian motion. We note that at almost the same time, Huang, Menozzi and Priola [7] have obtained (1.15) for time-independent $\sigma$ and $\nu$ by using Coifman–Weiss’ theorem. As mentioned above, our proof is based on a tailored version of Fefferman–Stein’s theorem.

This paper is organized as follows: In Sect. 2, we prepare some estimates about the probability density function of $X_t^{x,s}$ and establish a Fefferman–Stein-type theorem. In Sect. 3, we show the propagation of the regularity for transport equation. In Sect. 4, we prove our main result.

Throughout this paper, we use the following convention: The letters $C$ and $c$ with or without subscripts will denote a positive constant, whose value may change in different places. Moreover, we use $A \lesssim B$ to denote $A \leq CB$ for some constant $C > 0$. 
2. Preliminaries

2.1. Density estimate for $X_t^{s,x}$

In this subsection, we establish some estimates on the density of $X_t^{s,x}$, which will be used later.

**Lemma 2.1.** Under (1.2), $X_t^{s,0}$ of (1.7) has a smooth density function $p_{s,t}^{(0)}(y)$. For each $\beta = (\beta_1, \ldots, \beta_n) \in \mathbb{N}_0^n$, where $\mathbb{N}_0 = \{0\} \cup \mathbb{N}$, there are constants $C$, $c > 0$ only depending on $n$, $\beta$, $d$ and $\kappa$ such that for all $s < t$ and $y \in \mathbb{R}^{nd}$,

$$|\nabla_{y_1}^{\beta_1} \cdots \nabla_{y_n}^{\beta_n} p_{s,t}^{(0)}(y)| \leq C (t - s)^{-(n^2d + \sum_{i=1}^n (2n-i+1)\beta_i)/2} e^{-c|\Theta(t-s)|-1/2 y|^2},$$

(2.1)

where $\Theta$ is the dilation operator defined by (1.8).

**Proof.** Since $(c W, \sigma^2)_{(d)} = W$ for $c \neq 0$, by a change of variables, we have

$$X_t^{s,0} = \int_s^t e^{(t-r)A} \sigma^a_r dW_r \overset{(d)}{=} (t-s)^{1/2} \int_0^1 e^{(t-s)(1-r)A} \sigma^a_{s+(t-s)r} dW_r.$$

Hence, by definitions (1.5), (1.9) and (1.8),

$$\Theta^{(t-s)-1/2} X_t^{s,0} \overset{(d)}{=} \int_0^1 e^{(1-r)A} \sigma^a_{s+(t-s)r} dW_r \overset{(d)}{=} \int_0^1 e^{rA} \tilde{\sigma}_r^a dW_r =: Z,$$

where $\tilde{\sigma}_r^a := \sigma^a_{s+(t-s)(1-r)}$. Since $Z$ is a $nd$-dimensional Gaussian random variable with mean value zero and covariance matrix

$$\Sigma = \int_0^1 e^{rA} \tilde{\sigma}_r^a (\tilde{\sigma}_r^a)^* e^{rA^*} dr,$$

(2.2)

we have

$$p_{s,t}^{(0)}(y) = \frac{e^{-(\Theta(t-s)-1/2 y)^* \Sigma^{-1} (\Theta(t-s)-1/2 y)}}{(2\pi)^{nd} (t-s)^{n^2d} \det(\Sigma)^{1/2}}.$$  

(2.3)

On the other hand, by (1.5), (1.6) and (1.3), we have for all $y \in \mathbb{R}^{nd}$,

$$y^* \Sigma y = \int_0^1 |y^* e^{rA} \tilde{\sigma}_r^a|^2 dr \geq 2\kappa^{-1} \int_0^1 \left| \sum_{j=1}^n \frac{r^{n-j}}{(n-j)!} y_j \right|^2 dr$$

$$\geq 2\kappa^{-1} |y|^2 \inf_{|\omega| = 1} \int_0^1 \left| \sum_{j=1}^n \frac{r^{n-j}}{(n-j)!} \omega_j \right|^2 dr.$$

Since the unit sphere in $\mathbb{R}^{nd}$ is compact, and for each $\omega \in \mathbb{R}^{nd} \backslash \{0\}$,

$$\delta(\omega) := \int_0^1 \left| \sum_{j=1}^n \frac{r^{n-j}}{(n-j)!} \omega_j \right|^2 dr > 0,$$
we have \( c_0 := \inf_{|\omega|=1} \delta(\omega) > 0 \), and so
\[
y^* \Sigma y \geq 2\kappa^{-1}|y|^2 \inf_{|\omega|=1} \delta(\omega) \geq 2\kappa^{-1}c_0 |y|^2.
\] (2.4)

The desired estimate now follows by the chain rule, (2.3) and (2.4). \( \square \)

For \( \alpha \in (0, 2] \), the fractional Laplacian \( \Delta^{\alpha/2} \) in \( \mathbb{R}^d \) is defined by Fourier’s transform as
\[
\hat{\Delta^{\alpha/2}} f(\xi) = -|\xi|^\alpha \hat{f}(\xi), \quad f \in \mathcal{S}(\mathbb{R}^d),
\]
where \( \mathcal{S}(\mathbb{R}^d) \) is the space of Schwartz rapidly decreasing functions. For \( \alpha \in (0, 2) \), up to a multiplying constant, an alternative definition of \( \Delta^{\alpha/2} \) is given by the following integral form (cf. [10]):
\[
\Delta^{\alpha/2} f(x) = \int_{\mathbb{R}^d} \frac{\delta_z^{(2)} f(x)}{|z|^{d+\alpha}} \, dz,
\] (2.5)
where \( \delta_z^{(2)} f(x) := f(x+z) + f(x-z) - 2f(x) \). Observe that for \( \alpha \in (0, 1) \),
\[
\Delta^{\alpha/2} f(x) = 2 \int_{\mathbb{R}^d} \frac{\delta_z^{(1)} f(x)}{|z|^{d+\alpha}} \, dz \quad \text{with} \quad \delta_z^{(1)} f(x) := f(x+z) - f(x).
\] (2.6)

**Corollary 2.2.** For any \( j = 1, \ldots, n \), \( \alpha \in [0, 2] \) and \( \beta = (\beta_1, \ldots, \beta_n) \in \mathbb{N}_0^n \), there is a constant \( C > 0 \) such that for all \( f \in C_b^\infty (\mathbb{R}^{nd}) \) and \( s < t \),
\[
\| \Delta^{\alpha/2}_{x_j} \nabla_{x_1} \cdots \nabla_{x_n} T_{s,t} f \|_{\infty} \leq C |t-s|^{-\left(\sum_{i=1}^{n}(2(n-i)+1)\beta_i+(2(n-j)+1)\alpha/2\right)} \| f \|_{\infty}.
\] (2.7)
\[
\| \nabla_{x_1} \cdots \nabla_{x_n} T_{s,t} \Delta^{\alpha/2}_{x_j} f \|_{\infty} \leq C |t-s|^{-\left(\sum_{i=1}^{n}(2(n-i)+1)\beta_i+(2(n-j)+1)\alpha/2\right)} \| f \|_{\infty}.
\] (2.8)
where \( \Delta^{\alpha/2}_{x_j} \) means that the fractional Laplacian acts on the variable \( x_j \), and \( \Delta^0_{x_j} \equiv \mathbb{I} \).

**Proof.** Let \( p_{s,t}(x,y) \) be the probability density function of \( X_{t}^{s,x} = e^{(t-s)A}x + X_{t}^{s,0} \).

The transition semigroup \( T_{s,t} \) has a density kernel
\[
p_{s,t}(x,y) = p_{s,t}^{(0)}(y - e^{(t-s)A}x).
\] (2.9)

When \( \alpha = 0 \) or \( \alpha = 2 \), the desired estimates follow directly from (2.9) and Lemma 2.1.

So it suffices to show (2.7) and (2.8) for \( \alpha \in (0, 2) \). For notational convenience, we write
\[
h_y(x) := \nabla_{x_1} \cdots \nabla_{x_n} p_{s,t}(x,y), \quad \gamma := \sum_{i=1}^{n}(2(n-i) + 1)\beta_i/2
\]
and
\[
\delta_{z_j}^{(2)} h_y(x) = h_y(x + \tilde{z}_j) + h_y(x - \tilde{z}_j) - 2h_y(x), \quad \tilde{z}_j = (0, \ldots, z_j, \ldots, 0).
\]
By (2.1), (2.9) and the chain rule, we have
\[
|\delta_{z_j}^2 h_y(x)| \leq C(t-s)^{-n^2d/2-\gamma} \left( e^{-c(\Theta_{(t-s)}-1/2)(y-e^{(t-s)}A(x+z_j))} + e^{-c(\Theta_{(t-s)}-1/2)(y-e^{(t-s)}Ax)} \right)^2.
\]

On the other hand, using the mean value theorem twice, we have
\[
\delta_{z_j}^2 h_y(x) = z_j \int_0^1 (\partial_j h_y(x+s\tilde{z}_j) - \partial_j h_y(x-s\tilde{z}_j)) \, ds
= z_j (\partial_j h_y(x+s_0\tilde{z}_j) - \partial_j h_y(x-s_0\tilde{z}_j)) = 2s_0\hat{z}_j^2 \partial_j^2 h_y(\tilde{x})
\]
for some \(s_0 \in [0, 1]\) and then for some \(\tilde{x}\) in the line segment connecting \(x-s_0\tilde{z}_j\) and \(x+s_0\tilde{z}_j\). Hence, by (2.9) and (2.1),
\[
|\delta_{z_j}^2 h_y(x)| \leq 2|z_j|^2 |\partial_j^2 h_y(\tilde{x})| \leq C(t-s)^{-n^2d/2-\gamma-2(n-j)+1} \times e^{-c(\Theta_{(t-s)}-1/2)(y-e^{(t-s)}Ax)} |z_j|^2.
\]

By formula (2.5), we have
\[
\Delta_{x_j}^{\alpha/2} h_y(x) = \left( \int_{|z_j| > (t-s)^{(2(n-j)+1)/2}} + \int_{|z_j| \leq (t-s)^{(2(n-j)+1)/2}} \right) \frac{\delta_{z_j}^2 h_y(x)}{|z_j|^{d+\alpha}} \, dz_j
=: I_1(x, y) + I_2(x, y).
\]

For \(I_1(x, y)\), we have by (2.10)
\[
\int_{\mathbb{R}^d} |I_1(x, y)| \, dy \lesssim (t-s)^{-\gamma} \int_{|z_j| > (t-s)^{(2(n-j)+1)/2}} \frac{dz_j}{|z_j|^{d+\alpha}} \lesssim (t-s)^{-\gamma-(2(n-j)+1)\alpha/2},
\]
where we have used that
\[
(t-s)^{-n^2d/2} \int_{\mathbb{R}^d} e^{-c(\Theta_{(t-s)}-1/2)(y)^2} \, dy = \int_{\mathbb{R}^d} e^{-c|y|^2} \, dy.
\]

For \(I_2(x, y)\), we have by (2.11)
\[
\int_{\mathbb{R}^d} |I_2(x, y)| \, dy \lesssim (t-s)^{-\gamma-(2(n-j)+1)} \int_{|z_j| \leq (t-s)^{(2(n-j)+1)/2}} \frac{|z_j|^2}{|z_j|^{d+\alpha}} \, dz_j
\]
\[
\lesssim (t-s)^{-\gamma-(2(n-j)+1)\alpha/2}.
\]

Combining the above calculations, we obtain
\[
|\Delta_{x_j}^{\alpha/2} \nabla_{x_1} \partial_1 h_{x_1} \cdots \nabla_{x_n} \partial_n h_{x_n} T_s f(x)| = \left| \int_{\mathbb{R}^d} \Delta_{x_j}^{\alpha/2} h_y(x) f(y) \, dy \right|
\]
\[
\lesssim C(t-s)^{-\gamma-(2(n-j)+1)\alpha/2} \|f\|_{\infty}.
\]

Thus, we proved (2.7). Similarly, we can show (2.8). \qed
2.2. Fefferman–Stein’s theorem

First of all, we introduce a family of “balls” in $\mathbb{R}^{1+nd}$ that matches the geometry induced by the hypoelliptic operator (1.1). For any $r > 0$ and point $(t_0, x_0) \in \mathbb{R}^{1+nd}$, we define

$$Q_r(t_0, x_0) := \left\{(t, x) : \ell(t - t_0, x - e^{(t-t_0)A}x_0) \leq r\right\},$$

where

$$\ell(t, x) := \max\left\{|t|^{1/2}, |x_1|^{1/(2n-1)}, |x_2|^{1/(2n-3)}, \ldots, |x_{n-1}|^{1/3}, |x_n|\right\}.$$

We use $\mathbb{Q}$ to denote the set of all such balls.

**Lemma 2.3.**

(i) $\ell(r^2t, \Theta_r x) = r\ell(t, x)$ for any $r > 0$, where $\Theta_r$ is the dilation operator defined by (1.8).

(ii) $|Q_r(t_0, x_0)| = \omega_d^n r^{n+2}$, where $\cdot$ denotes the Euclidean volume and $\omega_d$ is the volume of the unit ball in $(\mathbb{R}^d, \cdot)$.

(iii) For all $(t, x), (s, y), (r, z) \in \mathbb{R}^{1+nd}$, we have

$$\ell(s - t, y - e^{(s-t)A}x) \leq 3\ell(t - s, x - e^{(t-s)A}y) \leq 12\left(\ell(t - r, x - e^{(t-r)A}z) + \ell(r - s, z - e^{(r-s)A}y)\right).$$  \hfill (2.12)

(iv) Suppose that $Q_r(t_0, x_0) \cap Q_r(t'_0, x'_0) \neq \emptyset$, then

$$Q_r(t_0, x_0) \subset Q_{20r}(t'_0, x'_0).$$ \hfill (2.13)

(v) For $(t, x), (s, y) \in \mathbb{R}^{1+nd}$, define

$$\rho((t, x), (s, y)) := \ell(t - s, x - e^{(t-s)A}y) + \ell(s - t, y - e^{(s-t)A}x),$$

and for $(t_0, x_0) \in \mathbb{R}^{1+nd}$ and $r > 0$,

$$\tilde{Q}_r(t_0, x_0) := \{(t, x) : \rho((t, x), (t_0, x_0)) \leq r\}.$$

Then, $Q_r(t_0, x_0) \subset \tilde{Q}_r(t_0, x_0) \subset Q_{4r}(t_0, x_0)$.

**Proof.** (i) and (ii) follow directly from the definition of $\ell(t, x)$.

(iii) We only prove the second inequality in (2.12). The first one is similar. Observing that for all $(t, x), (s, y) \in \mathbb{R}^{1+nd}$,

$$\ell(t + s, x + y) \leq \ell(t, x) + \ell(s, y),$$ \hfill (2.14)

we have

$$\ell(t - s, x - e^{(t-s)A}y) \leq \ell(t - r, x - e^{(t-r)A}z) + \ell(r - s, e^{(t-r)A}z - e^{(t-s)A}y).$$
For simplicity, we write
\[ a := \ell(t-r, x - e^{(t-r)A}z), \quad b := \ell(r-s, z - e^{(r-s)A}y). \]

By the definition of \( \ell \), we have
\[ |t-r| \leq a^2, \quad |(z - e^{(r-s)A}y)_j| \leq b^{1+2(n-j)}, \quad j = 1, \ldots, n. \]

Hence, for each \( i = 1, \ldots, n \),
\[
|e^{(t-r)A}z - e^{(t-s)A}y|_i = \left| \sum_{j=1}^{n} (e^{(t-r)A})_{ij} (z - e^{(r-s)A}y)_j \right|
\leq \sum_{j=i}^{n} \frac{|t-r|^{j-i}}{(j-i)!} b^{1+2(n-j)} \leq \sum_{j=i}^{n} \frac{a^{2(j-i)}}{(j-i)!} b^{1+2(n-j)}
\leq (a \vee b)^{1+2(n-i)} \sum_{j=i}^{n} \frac{1}{(j-i)!} \leq 3(a \vee b)^{1+2(n-i)},
\tag{2.15}
\]
and
\[ \ell(t-s, x - e^{(t-s)A}y) \leq a + 3(a \vee b) \leq 4(a + b). \]

(iv) and (v) are easy consequences of (iii).

For \( f \in L^1_{loc}(\mathbb{R}^{1+nd}) \), we define the Hardy–Littlewood maximal function by
\[ \mathcal{M}f(t, x) := \sup_{r>0} \frac{1}{Q_r(t, x)} \int_{Q_r(t, x)} |f(t', x')| \, dx' \, dt', \]
and the sharp function by
\[ \mathcal{M}^\# f(t, x) := \sup_{r>0} \frac{1}{Q_r(t, x)} \int_{Q_r(t, x)} |f(t', x') - f_{Q_r(t, x)}| \, dx' \, dt'. \]

Here, for a \( Q \in \mathcal{Q} \),
\[ f_Q := \int_Q f(t', x') \, dx' \, dt' := \frac{1}{|Q|} \int_Q f(t', x') \, dx' \, dt'. \]

One says that a function \( f \in BMO(\mathbb{R}^{1+nd}) \) if \( \mathcal{M}^\# f \in L^\infty(\mathbb{R}^{1+nd}) \). Clearly, \( f \in BMO(\mathbb{R}^{1+nd}) \) if and only if there exists a constant \( C > 0 \) such that for any \( Q \in \mathcal{Q} \), and for some \( c_Q \in \mathbb{R} \),
\[ \int_Q |f(t', x') - c_Q| \, dx' \, dt' \leq C. \]

Using Lemma 2.3, the following version of Fefferman–Stein-type theorem can be established in a similar way as that for [5, Theorem 2.12]. We omit the details here.
THEOREM 2.4. Suppose $q \in (1, \infty)$, and $\mathcal{P}$ is a bounded linear operator from $L^q(\mathbb{R}^{1+nd})$ to $L^q(\mathbb{R}^{1+nd})$ and also from $L^\infty(\mathbb{R}^{1+nd})$ to $\text{BMO}(\mathbb{R}^{1+nd})$. Then, for any $p \in [q, \infty)$, there is a constant $C > 0$ depending only on $p$, $q$ and the norms of $\|\mathcal{P}\|_{L^q \to L^q}$ and $\|\mathcal{P}\|_{L^\infty \to \text{BMO}}$ so that

$$\|\mathcal{P}f\|_p \leq C \|f\|_p \quad \text{for every } f \in L^p(\mathbb{R}^{1+nd}).$$

3. Propagation of regularity in $L^2$-space for transport equations

Fix $n \geq 2$ and $\lambda \geq 0$. Let $A$ be as in (1.3). In this section, corresponding to (1.13), we consider the following linear transport equation in $\mathbb{R}^{nd}$ for $u = u(s, x)$:

$$\partial_s u + Ax \cdot \nabla u - \lambda u + f = 0. \quad (3.1)$$

Taking Fourier’s transform in the spatial variable $x \in \mathbb{R}^{nd}$, we obtain for any $(s, \xi) \in \mathbb{R} \times \mathbb{R}^{nd}$,

$$\partial_s \hat{u}(s, \xi) - A^* \xi \cdot \nabla |\hat{u}|^2 - \lambda |\hat{u}|^2 + \hat{f}(s, \xi) = 0. \quad (3.2)$$

Recall that $A^*$ is the transpose of $A$. Multiplying both sides by the complex conjugate of $\hat{u}$, we get

$$\partial_s |\hat{u}|^2 - A^* \xi \cdot \nabla |\hat{u}|^2 - \lambda |\hat{u}|^2 + 2\text{Re}(\hat{f}, \overline{\hat{u}}) = 0.$$

Let $\phi(\xi)$ be a smooth function and define

$$g_\phi := 2\phi \text{Re}(\hat{f}, \overline{\hat{u}}) + A^*(\xi \cdot \nabla \phi)|\hat{u}|^2. \quad (3.3)$$

We have

$$\partial_s (|\hat{u}|^2 \phi) - A^* \xi \cdot \nabla (|\hat{u}|^2 \phi) - \lambda (|\hat{u}|^2 \phi) + g_\phi = 0, \quad (3.4)$$

and if $\hat{u}$ has compact support, then

$$|\hat{u}|^2 \phi)(s, \xi) = -\int_0^\infty \partial_t \left( e^{-\lambda t} (|\hat{u}|^2 \phi)(t + s, e^{-tA^*} \xi) \right) dt$$

$$= \int_0^\infty e^{-\lambda t} g_\phi(t + s, e^{-tA^*} \xi) dt, \quad (3.5)$$

where $e^{-tA^*} = (e^{-tA})^*$ is the transpose of exponential matrix in (1.6).

The following result on propagation of regularity in $L^2$-spaces is of independent interest and also the key step in the proof of Theorem 1.1.

THEOREM 3.1. Let $f, u \in L^2(\mathbb{R}^{1+nd})$ so that (3.1) holds in the weak sense that for any $\varphi \in C^\infty_c(\mathbb{R}^{1+nd})$,

$$\langle u, \partial_s \varphi \rangle + \langle u, Ax \cdot \nabla \varphi \rangle + \lambda \langle u, \varphi \rangle = \langle f, \varphi \rangle.$$
For any \( \alpha > 0 \) and \( j = 1, 2, \ldots, n-1 \), there is a constant \( C = C(\alpha, j, d) > 0 \) such that

\[
\| \Delta_{x_j}^{\alpha \over 2(1+\sigma)} u \|^2_2 \leq C \| \Delta_{x_{j+1}}^{\alpha \over 2} u \|_{2}^{1 \over 1+\sigma} \| f \|_2^{\alpha \over 1+\sigma}. \tag{3.6}
\]

In particular,

\[
\| \Delta_{x_j}^{\alpha \over 2(1+\sigma)} u \|^2_2 \leq C \| \Delta_{x_{j+1}}^{\alpha \over 2} u \|_{2}^{1 \over 1+\sigma} \| f \|_2^{\alpha \over 1+\sigma}. \tag{3.7}
\]

**Proof.** Let \( \rho : [0, \infty) \rightarrow [0, 1] \) be a smooth function with \( \rho(s) = 1 \) for \( s < 1 \) and \( \rho(s) = 0 \) for \( s > 2 \). For \( R > 0 \), let \( \eta_R(s, \xi) := \rho(|s|/R)\rho(|\xi|/R) \) and

\[
\hat{u}_R := \eta_R \hat{u}, \quad \hat{f}_R := (A^\ast \xi \cdot \nabla \eta_R - \partial_s \eta_R) \hat{u} + \eta_R \hat{f}.
\]

Since \( \hat{u}_R \) satisfies

\[
\partial_s \hat{u}_R - A^\ast \xi \cdot \nabla \hat{u}_R - \lambda \hat{u}_R + \hat{f}_R = 0,
\]

if we can show that for some \( C = C(\alpha, j, d) > 0 \),

\[
\| \Delta_{x_j}^{\alpha \over 2(1+\sigma)} u_R \|_2 \leq C \| \Delta_{x_{j+1}}^{\alpha \over 2} u_R \|_{2}^{1 \over 1+\sigma} \| f \|_2^{\alpha \over 1+\sigma},
\]

where \( u_R \) and \( f_R \) are the inverse Fourier transforms of \( \hat{u}_R \) and \( \hat{f}_R \), then letting \( R \rightarrow \infty \), we get (3.6). Hence, without loss of generality, in the following, we may and do assume that \( \hat{u} \) has compact support and

\[
0 < \| f \|_2 < \infty, \quad 0 < \| \Delta_{x_{j+1}}^{\alpha/2} u \|_2 < \infty.
\]

We use induction method. Let us first look at the case of \( j = 1 \). We follow the simple argument of Alexander [1]. For any \( \varepsilon > 0 \), by Plancherel’s identity, we have

\[
\| \Delta_{x_1}^{\alpha \over 2(1+\sigma)} u \|^2_2 = \int_{-\infty}^{\infty} \int_{\mathbb{R}^d} |\xi_1|^{2\alpha \over 1+\sigma} |\hat{u}(s, \xi)|^2 \, d\xi \, ds
\]

\[
= \int_{-\infty}^{\infty} \int_{\mathbb{R}^d} |\xi_1|^{2\alpha \over 1+\sigma} 1_{\{\xi_2^2 > |\xi_1|^2/(1+\alpha)\}} |\hat{u}(s, \xi)|^2 \, d\xi \, ds
\]

\[
+ \int_{-\infty}^{\infty} \int_{\mathbb{R}^d} |\xi_1|^{2\alpha \over 1+\sigma} 1_{\{\xi_2^2 \leq |\xi_1|^2/(1+\alpha)\}} |\hat{u}(s, \xi)|^2 \, d\xi \, ds
\]

\[
=: I_1(\varepsilon) + I_2(\varepsilon).
\]

For \( I_1(\varepsilon) \), we have

\[
I_1(\varepsilon) \leq \varepsilon^{2\alpha} \int_{-\infty}^{\infty} \int_{\mathbb{R}^d} |\xi_2|^{2\alpha} |\hat{u}(s, \xi)|^2 \, d\xi \, ds = \varepsilon^{2\alpha} \| \Delta_{x_2}^{\alpha \over 2} u \|_2^2.
\]
For $I_2(\varepsilon)$, letting $\Omega := \{\varepsilon | \xi_2 | \leq \xi_1 |^{1/(1+\alpha)} \}$ and by (3.5) with $\phi = 1$, we have

$$I_2(\varepsilon) = 2 \int_{-\infty}^{\infty} \int_{\mathbb{R}^d} |\xi_1|^{2\alpha/(1+\alpha)} 1_{\Omega}(\xi) \int_0^\infty e^{-\lambda t} \text{Re}(\tilde{f}, \tilde{u})(t + s, e^{-t A^*} \xi) dr d\xi ds$$

$$= 2 \int_{-\infty}^{\infty} \int_{\mathbb{R}^d} \left[ |(e^{t A^*} \xi)|_1^{2\alpha/(1+\alpha)} 1_{\Omega}(e^{t A^*} \xi) e^{-\lambda t} \text{Re}(\tilde{f}, \tilde{u})(t + s, \xi) \right] dr d\xi ds$$

$$= 2 \int_{-\infty}^{\infty} \int_{\mathbb{R}^d} \left( \int_0^\infty |(e^{t A^*} \xi)|_1^{2\alpha/(1+\alpha)} 1_{\Omega}(e^{t A^*} \xi) e^{-\lambda t} dt \right) \text{Re}(\tilde{f}, \tilde{u})(s, \xi) d\xi ds.$$

Observing that

$$e^{t A^*} \xi = \left( \xi_1, t \xi_1 + \xi_2, \frac{t^2}{2} \xi_1 + t \xi_2 + \xi_3, \ldots, \frac{t^{n-1}}{(n-1)!} \xi_1 + \cdots + t \xi_{n-1} + \xi_n \right),$$

since $\lambda \geq 0$, we have

$$\int_0^\infty |(e^{t A^*} \xi)|_1^{2\alpha/(1+\alpha)} 1_{\Omega}(e^{t A^*} \xi) e^{-\lambda t} dt = |\xi_1|^{2\alpha/(1+\alpha)} 1_{\Omega}(\xi)$$

$$\leq |\xi_1|^{2\alpha/(1+\alpha)} \int_{-\infty}^\infty |\xi_2|/|\xi_1|^{\alpha/(1+\alpha)} |\xi_1|^{2\alpha/(1+\alpha)} d\xi ds = 2 |\xi_1|^{\alpha/(1+\alpha)} \varepsilon.$$

Hence, by Young’s inequality we have

$$I_2(\varepsilon) \leq \frac{4}{\varepsilon} \int_{-\infty}^\infty \int_{\mathbb{R}^d} |\xi_1|^{\alpha/(1+\alpha)} \int_{\mathbb{R}^d} |\tilde{f}| \tilde{u}(s, \xi) d\xi ds$$

$$\leq \frac{1}{2} \int_{-\infty}^\infty \int_{\mathbb{R}^d} |\xi_1|^{2\alpha/(1+\alpha)} |\tilde{u}|^2(s, \xi) d\xi ds + \frac{8}{\varepsilon^2} \int_{-\infty}^\infty \int_{\mathbb{R}^d} |\tilde{f}|^2(s, \xi) d\xi ds$$

$$= \frac{1}{2} \|\Delta_{x_1}^{\alpha/(1+\alpha)} u\|_2^2 + \frac{8}{\varepsilon^2} \|f\|_2^2.$$
We first treat $K_1(\delta)$. By the induction hypothesis, one sees that for $j = 1, 2, \ldots, k$,

$\|\Delta_{x_j}^{2(1+(k-j+2)\alpha)} u\|_2 \leq C \|\Delta_{x_k+1}^{2(1+\alpha)} u\|_2^{1+\alpha} \| f\|_2^{(k-j+1)\alpha}$. (3.9)

Observing that

$1 - \phi_\delta(\xi) \leq \sum_{j=1}^k 1\{|\xi| \geq \delta^{k-j+1}|\xi_{k+1}|^{1+(k-j+2)\alpha)/(1+\alpha)\}$,

by (3.9) and Young’s inequality, we have

$K_1(\delta) \leq \sum_{j=1}^k \int_{-\infty}^{\infty} \int_{\mathbb{R}^d} |\xi_{k+1}|^{2\alpha/(1+\alpha)} 1\{|\xi| \geq \delta^{k-j+1}|\xi_{k+1}|^{1+(k-j+2)\alpha)/(1+\alpha)} |\hat{u}(s, \xi)|^2 d\xi ds$

$\leq \sum_{j=1}^k \delta^{-\alpha/(1+\alpha)} \int_{-\infty}^{\infty} \int_{\mathbb{R}^d} |\xi|^{2\alpha/(1+\alpha)} |\hat{u}(s, \xi)|^2 d\xi ds$

$\leq \sum_{j=1}^k \delta^{-\alpha/(1+\alpha)} \|\Delta_{x_j}^{2(1+\alpha)} u\|_2\| f\|_2^{(k-j+1)\alpha}$

$\leq C \sum_{j=1}^k \|\Delta_{x_k}^{2(1+\alpha)} u\|_2^{1+\alpha} \| f\|_2^{2(k-j+1)\alpha}$

$\leq \frac{1}{4} \|\Delta_{x_k+1}^{2(1+\alpha)} u\|_2^2 + C\delta^{-2} \| f\|_2^2$.

Next, we treat the trouble term $K_2(\delta)$. Let

$\xi(k) := (\xi_1, \xi_2, \ldots, \xi_k), \xi^{(k)}(k) := (\xi_{k+1}, \xi_{k+2}, \ldots, \xi_n)$

and define

$h_\delta(s, \xi^{(k)}) := \int_{\mathbb{R}^k} (|\hat{u}|^2 \phi_\delta)(s, \xi(k), \xi^{(k)}) d\xi(k)$.

Integrating both sides of (3.4) with respect to the first $k$-variables $\xi(k)$, we get

$\partial_t h_\delta = A_{n-k} \xi^{(k)} \cdot \nabla \xi^{(k)} h_\delta - \lambda h_\delta$

$= \int_{\mathbb{R}^k} \left( \sum_{j=1}^k \xi_j \cdot \nabla \xi_{j+1} (|\hat{u}|^2 \phi_\delta) - g_\phi \right) d\xi(k) =: g^{(k)}_{\phi^\delta},$

where $A_{n-k}$ is defined by (1.3) but with $n-k$ in place of $n$, and by (3.5),

$h_\delta(s, \xi^{(k)}) = \int_0^\infty e^{-\lambda t} g_{\phi^\delta}^{(k)} (t+s, e^{-tA_{n-k}} \xi^{(k)}) dt.$
Given $\varepsilon > 0$, let
\[ \psi_\varepsilon(\xi^{(k)}) := \rho(\varepsilon|\xi_{k+2}^{(2)}|/|\xi_{k+1}^{(1)}|^{1/(1+\alpha)}) = \rho(\varepsilon|\xi_{k+2}^{2}/|\xi_{k+1}^{1}|^{1/(1+\alpha)}) \]

As above, we make the following decompositions:
\[ K_2(\delta) = \int_{-\infty}^\infty \int_{\mathbb{R}^{(n-1)d}} |\xi_{k+1}^{(2)}|^{2a/(1+\alpha)} \left(1 - \psi_\varepsilon(\xi^{(k)})\right) h_\delta(s, \xi^{(k)}) d\xi^{(k)} ds \]
\[ + \int_{-\infty}^\infty \int_{\mathbb{R}^{(n-1)d}} |\xi_{k+1}^{(2)}|^{2a/(1+\alpha)} \psi_\varepsilon(\xi^{(k)}) h_\delta(s, \xi^{(k)}) d\xi^{(k)} ds \]
\[ =: K_{21}(\delta, \varepsilon) + K_{22}(\delta, \varepsilon). \]

For $K_{21}(\delta, \varepsilon)$, we have
\[ K_{21}(\delta, \varepsilon) \leq \int_{-\infty}^\infty \int_{\mathbb{R}^{(n-1)d}} |\xi_{k+1}^{(2)}|^{2a/(1+\alpha)} \left[|\varepsilon|\xi_{k+2}^{(2)}|/|\xi_{k+1}^{(1)}|^{1/(1+\alpha)}\right] h_\delta(s, \xi^{(k)}) d\xi^{(k)} ds \]
\[ \leq \varepsilon^{2a} \int_{-\infty}^\infty \int_{\mathbb{R}^{nd}} |\xi_{k+2}^{(2)}|^{2a} \hat{u}^2(s, \xi) d\xi ds = \varepsilon^{2a} \|\Delta_{k+2}^\alpha u\|_{L_2}. \]

For $K_{22}(\delta, \varepsilon)$, by a change of variable and Fubini’s theorem, we have
\[ K_{22}(\delta, \varepsilon) = \int_{-\infty}^\infty \int_{\mathbb{R}^{(n-1)d}} |\xi_{1}^{(k)}|^{2a/(1+\alpha)} \psi_\varepsilon(\xi^{(k)}) \int_0^\infty e^{-\lambda t} g_{\phi_\lambda}^{(k)}(t + s, e^{-tA_{n-k}^*} \xi^{(k)}) dr d\xi^{(k)} ds \]
\[ = \int_{-\infty}^\infty \int_{\mathbb{R}^{(n-1)d}} \left(\int_0^\infty |(e^{tA_{n-k}^*} \xi^{(k)})_{1}|^{2a/(1+\alpha)} \psi_\varepsilon(e^{tA_{n-k}^*} \xi^{(k)}) e^{-\lambda t} dr \right) \]
\[ \times s_{\phi_\lambda}^{(k)}(s, \xi^{(k)}) d\xi^{(k)} ds. \]

Letting
\[ \gamma_\lambda(\xi^{(k)}) := \int_0^\infty |(e^{tA_{n-k}^*} \xi^{(k)})_{1}|^{2a/(1+\alpha)} \psi_\varepsilon(e^{tA_{n-k}^*} \xi^{(k)}) e^{-\lambda t} dr \]
\[ = |\xi_{k+1}^{(2)}|^{2a/(1+\alpha)} \int_0^\infty \rho(\varepsilon|\xi_{k+2}^{2}/|\xi_{k+1}^{1}|^{1/(1+\alpha)}) e^{-\lambda t} dr, \]
and recalling
\[ g_{\phi_\lambda}^{(k)} = - \int_{\mathbb{R}^{kd}} g_{\phi_\lambda}(\xi^{(k)}, \cdot) d\xi^{(k)} + \sum_{j=1}^k \int_{\mathbb{R}^{kd}} \xi_j \cdot \nabla \xi_{j+1}(|\hat{u}|^2 \phi_\lambda) d\xi^{(k)}, \]
we may write
\[ K_{22}(\delta, \varepsilon) = - \int_{-\infty}^\infty \int_{\mathbb{R}^{(n-1)d}} \gamma_\lambda(\xi^{(k)}) \int_{\mathbb{R}^{kd}} g_{\phi_\lambda}(s, \xi^{(k)}, \xi^{(k)}) d\xi^{(k)} d\xi^{(k)} ds \]
\[ + \sum_{j=1}^k \int_{-\infty}^\infty \int_{\mathbb{R}^{(n-1)d}} \gamma_\lambda(\xi^{(k)}) \int_{\mathbb{R}^{kd}} \xi_j \cdot \nabla \xi_{j+1}(|\hat{u}|^2 \phi_\lambda) d\xi^{(k)} d\xi^{(k)} ds \]
\[ =: K_{221}(\delta, \varepsilon) + K_{222}(\delta, \varepsilon). \]
As in estimating (3.8), we have

$$
\gamma_e(\xi^{(k)}) \leq 4|\xi_{k+1}|^{\frac{2}{1+\alpha}} / \varepsilon.
$$

and also by the definition of $g_{\phi_\delta}$,

$$
|g_{\phi_\delta}| \leq 2|\hat{f}| |\hat{u}| + \sum_{j=1}^{n-1} |\xi_j| \cdot |\nabla j+1 \phi_\delta| \cdot |\hat{u}|^2 \leq 2|\hat{f}| |\hat{u}| + C\delta |\xi_{k+1}|^{\frac{2}{1+\alpha}} |\hat{u}|^2.
$$

Hence,

$$
K_{221}(\delta, \varepsilon) \leq \frac{8}{\varepsilon} \int_{-\infty}^{\infty} \int_{\mathbb{R}^d} |\xi_{k+1}|^{\frac{\alpha}{1+\alpha}} |\hat{f}| |\hat{u}| + \frac{C\delta}{\varepsilon} \int_{-\infty}^{\infty} \int_{\mathbb{R}^d} |\xi_{k+1}|^{\frac{2\alpha}{1+\alpha}} |\hat{u}|^2
\leq \left(\frac{1}{4} + \frac{C\delta}{\varepsilon}\right) \|\Delta \frac{\alpha}{21+\alpha} u\|_2^2 + C\varepsilon^{-2} \|f\|_2.
$$

On the other hand, by elementary calculations, we also have

$$
|\nabla \xi_{k+1} \gamma_e(\xi^{(k)})| \leq C|\xi_{k+1}|^{-\frac{1}{1+\alpha}} / \varepsilon.
$$

Thus, since $\hat{u}$ has compact support, by the integration by parts formula, we have

$$
|K_{222}(\delta, \varepsilon)| = \left| \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n-k}} \xi_k \cdot \nabla \xi_{k+1} \gamma_e(\xi^{(k)}) (|\hat{u}|^2 \phi_\delta) d\xi^{(k)} d\xi^{(k)} ds \right|
\leq \frac{C}{\varepsilon} \int_{-\infty}^{\infty} \int_{\mathbb{R}^d} |\xi_k| |\xi_{k+1}|^{-\frac{1}{1+\alpha}} (|\hat{u}|^2 \phi_\delta) d\xi ds
\leq \frac{C}{\varepsilon} \int_{-\infty}^{\infty} \int_{\mathbb{R}^d} |\xi_k| |\xi_{k+1}|^{-\frac{1}{1+\alpha}} |\hat{u}|^2 1_{[|\xi_k| \leq \delta |\xi_{k+1}|]^{1+2\alpha/(1+\alpha)}} d\xi ds
\leq \frac{C\delta}{\varepsilon} \int_{-\infty}^{\infty} \int_{\mathbb{R}^d} |\xi_{k+1}|^{\frac{2\alpha}{1+\alpha}} |\hat{u}|^2 ds = \frac{C\delta}{\varepsilon} \|\Delta \frac{\alpha}{21+\alpha} u\|_2^2.
$$

Combining the above calculations, we obtain

$$
\|\Delta \frac{\alpha}{21+\alpha} u\|_2^2 \leq \left(\frac{1}{4} + \frac{C' \delta}{\varepsilon}\right) \|\Delta \frac{\alpha}{21+\alpha} u\|_2^2 + \varepsilon^{2\alpha} \|\Delta \frac{\alpha}{21+\alpha} u\|_2^2 + C_2' (\varepsilon^{-2} + \delta^{-2}) \|f\|_2^2.
$$

Choosing $\delta = \varepsilon / (4C')$, we get

$$
\|\Delta \frac{\alpha}{21+\alpha} u\|_2^2 \leq 2\varepsilon^{2\alpha} \|\Delta \frac{\alpha}{21+\alpha} u\|_2^2 + 2C_2' (1 + 16(C')^2) \varepsilon^{-2} \|f\|_2^2,
$$

which then yields (3.6) for $j = k + 1$ by letting $\varepsilon = (\|f\|_2 / \|\Delta \frac{\alpha}{21+\alpha} u\|_2)^{1/(1+\alpha)}$. \quad \Box
4. Proof of Theorem 1.1

4.1. Case: $p = 2$

Without loss of generality, we may assume $f \in C_c^\infty(\mathbb{R}^{1+nd})$. It follows from Fourier's transform and Hölder’s inequality that

$$
\int_{-\infty}^{\infty} \left\| \Delta_{s_n} \int_{0}^{\infty} e^{-\lambda t} T_{s,t+s} f(t + s, \cdot) \, dt \right\|_2^2 
= \int_{-\infty}^{\infty} \int_{\mathbb{R}^{nd}} |\xi_n|^2 \left| \int_{0}^{\infty} e^{-\lambda t} T_{s,t+s} f(t + s, \xi) \, dt \right|^2 \, d\xi 
\leq \int_{-\infty}^{\infty} \int_{\mathbb{R}^{nd}} \left( \int_{0}^{\infty} |\xi_n|^2 e^{-\frac{1}{2} \int_{0}^{t} |(\sigma_s^a)^* e^{-\tau A_s^a} \xi|^2 \, d\tau} \right)^2 \, d\xi 
\times \left( \int_{0}^{\infty} |\xi_n|^2 \left| \int_{0}^{\infty} e^{\frac{1}{2} \int_{0}^{t} s f(t, \cdot) \, dt} f(t + s, \xi) \, dt \right|^2 \, d\xi \right) 
\, d\xi 
= \int_{-\infty}^{\infty} \int_{\mathbb{R}^{nd}} \left( \int_{0}^{\infty} |\xi_n|^2 e^{-\frac{1}{2} \int_{0}^{t} |(\sigma_s^a)^* e^{-\tau A_s^a} \xi|^2 \, d\tau} \right)^2 \, d\xi \, d\xi.
$$

Let $\Theta_t$ be the dilation operator defined by (1.8). By definitions (1.5), (1.8) and (1.9), it is easy to see that

$$
t^{1/2} (\sigma_s^a)^* e^{-rt A_s^a} \xi = (\sigma_s^a)^* e^{-r A_s^a} \Theta_{t^{1/2}}(\xi).
$$

Thus, by a change of variable and an argument similar to that for (2.4), we have for some $c > 0$,

$$
\int_{0}^{t} |(\sigma_s^a)^* e^{-r A_s^a} \xi|^2 \, dr = \int_{0}^{1} |(\sigma_s^a)^* e^{-r A_s^a} \Theta_{t^{1/2}}(\xi)|^2 \, dr \geq c |\Theta_{t^{1/2}}(\xi)|^2,
$$

and similarly,

$$
\int_{0}^{t} |(\sigma_s^a)^* e^{(t-r) A_s^a} \xi|^2 \, dr \geq c |\Theta_{t^{1/2}}(\xi)|^2.
$$

Hence,

$$
\int_{0}^{\infty} |\xi_n|^2 e^{-\frac{1}{2} \int_{0}^{t} |(\sigma_s^a)^* e^{-\tau A_s^a} \xi|^2 \, d\tau} \, dt 
\leq \int_{0}^{\infty} |\xi_n|^2 e^{-c |\Theta_{t^{1/2}}(\xi)|^2} \, dt 
\leq \int_{0}^{\infty} |\xi_n|^2 e^{-c t |\xi_n|^2} \, dt = c^{-1},
$$

and

$$
\int_{0}^{\infty} |(e^{t A_s^a} \xi)|^2 e^{-\frac{1}{2} \int_{0}^{t} |(\sigma_s^a)^* e^{(t-r) A_s^a} \xi|^2 \, d\tau} \, dt 
\leq \int_{0}^{\infty} |(e^{t A_s^a} \xi)|^2 e^{-c t |\Theta_{t^{1/2}}(\xi)|^2} \, dt 
\leq \sum_{j=1}^{n} \frac{1}{(n-j)!} \int_{0}^{\infty} \left| t^{n-j} \xi_j \right|^2 e^{-c t (2(n-j)+1) |\xi_j|^2} \, dt 
= \sum_{j=1}^{n} \frac{c^{-1}(2(n-j)+1)^{-1}}{(n-j)!} \leq 2c^{-1}.
$$
Thus, by the change of variables and Fubini’s theorem, we further have
\[
\int_{-\infty}^{\infty} \left\| \Delta_{x_n} \int_{0}^{\infty} e^{-\lambda t} T_{s,t+s} f(t+s, \cdot) dt \right\|_2^2 ds 
\leq c^{-1} \int_{-\infty}^{\infty} \int_{\mathbb{R}^n} \left( \int_{0}^{\infty} |\xi_n|^2 e^{-\frac{1}{2} \left( t \int_{0}^{s} |(\sigma^{u})^{*} e^{-tA^{*} \xi}|^2 dr \right) |\hat{f}(t+s, e^{-tA^{*} \xi})|^2 dt \right) d\xi ds 
= c^{-1} \int_{-\infty}^{\infty} \int_{\mathbb{R}^n} \left( \int_{0}^{\infty} |(e^{tA^{*} \xi})_n|^2 e^{-\frac{1}{2} \left( t \int_{0}^{s} |(\sigma^{u})^{*} e^{(t-s)A^{*} \xi}|^2 dr \right) |\hat{f}(t+s, \xi)|^2 dt \right) d\xi ds 
= c^{-1} \int_{-\infty}^{\infty} \int_{\mathbb{R}^n} \left( \int_{0}^{\infty} |(e^{tA^{*} \xi})_n|^2 e^{-\frac{1}{2} \left( t \int_{0}^{s} |(\sigma^{u})^{*} e^{(t-s)A^{*} \xi}|^2 dr \right) |\hat{f}(s, \xi)|^2 d\xi ds 
\leq 2c^{-2} \int_{-\infty}^{\infty} \int_{\mathbb{R}^n} |\hat{f}(s, \xi)|^2 d\xi ds = 2c^{-2} \| f \|_2^2.
\]
This together with (3.7) completes the proof of (1.12) for \( p = 2 \).

4.2. Case: \( p \in (2, \infty) \)

Let \( \varrho \in C^\infty_c(\mathbb{R}^{nd}) \) be nonnegative with \( \int \varrho = 1 \). We use it to define a family of mollifiers
\[
\varrho_\varepsilon(x) = \varepsilon^{-nd} \varrho(x/\varepsilon); \quad \varepsilon > 0.
\]
For a function \( f(t, x) \) defined on \( \mathbb{R} \times \mathbb{R}^{nd} \) and \( \varepsilon > 0 \), let
\[
f_\varepsilon(t, x) := f(t, \cdot) * \varrho_\varepsilon(x) := \int_{\mathbb{R}^d} f(t, y) \varrho_\varepsilon(x - y) dy.
\]
For \( j = 1, \ldots, n \) and \( \varepsilon \in (0, 1) \), define
\[
\mathcal{P}_j^\varepsilon f := \mathcal{P}_j^a f_\varepsilon(s, x) := \Delta_{x_j}^{1/\left(1+2(n-j)\right)} \int_{0}^{\infty} e^{-\lambda t} T_{s,t+s}^a f_\varepsilon(t+s, x) dt,
\]
where the superscript \( a \) denotes the dependence on the diffusion coefficient \( a \). To use Theorem 2.4, our main task is to show that \( \mathcal{P}_j^\varepsilon \) is a bounded linear operator from \( L^\infty(\mathbb{R}^{1+nd}) \) to \( BMO \). More precisely, we want to prove that for any \( f \in L^\infty(\mathbb{R}^{1+nd}) \) with \( \| f \|_\infty \leq 1 \), and any \( Q = Q_r(t_0, x_0) \in \mathcal{Q} \),
\[
\int_{\mathcal{Q}} \| \mathcal{P}_j^a f_\varepsilon(s, x) - c^Q_\varepsilon f_\varepsilon(s, x) \|_2^2 \leq C, \quad (4.1)
\]
where \( c^Q_\varepsilon \) is a constant depending on \( Q \) and \( f_\varepsilon \), and \( C \) only depends on \( n, \kappa, p, d, \) and not on \( \varepsilon \).

**Lemma 4.1.** (Scaling Property) For any \( Q = Q_r(t_0, x_0) \in \mathcal{Q} \), we have
\[
\int_{Q_r(t_0, x_0)} \| \mathcal{P}_j^a f_\varepsilon(s, x) - c \|_2^2 = \int_{Q_1(0)} \| \mathcal{P}_j^a \tilde{f}_\varepsilon(s, x) - c \|_2^2, \quad (4.2)
\]
where \( c \in \mathbb{R} \), \( \tilde{a}_s := a_r^{2s+t_0} \) and \( \tilde{f}_\varepsilon(t, x) := f_\varepsilon\left(r^2t + t_0, \Theta_r x + e^{tA}x_0\right) \). Here, \( \Theta_r \) is the dilation operator defined in (1.8).
Proof. Let us write

\[ u_\varepsilon(s, x) := \int_0^\infty e^{-\lambda t} T_{s,t+s}^a f_\varepsilon(t + s, x)dt. \]

By the change of variables, we have

\[ \tilde{u}_\varepsilon(s, x) := r^{-2}u_\varepsilon\left(r^2s + t_0, \Theta_r x + e^{sA}x_0\right) = \int_0^\infty e^{-\lambda t} T_{s,t+s}^\varepsilon \tilde{f}_\varepsilon(t + s, x)dt. \]

Noticing that

\[ (\Delta_{x_j}^{1/(1+2(n-j))} u_\varepsilon)(s, x) = (\Delta_{x_j}^{1/(1+2(n-j))} u_\varepsilon)\left(r^2s + t_0, \Theta_r x + e^{sA}x_0\right), \]

by the change of variables again, we have

\[ \int_{Q_r(0, x_0)} |\Delta_{x_j}^{1/(1+2(n-j))} u_\varepsilon(s, x) - c|^2 = \int_{Q_1(0)} |\Delta_{x_j}^{1/(1+2(n-j))} \tilde{u}_\varepsilon(s, x) - c|^2. \]

The proof is finished. \( \square \)

Noticing that

\[ \int_0^\infty e^{-\lambda t} T_{s,t+s}^a f_\varepsilon(t + s, x)dt = \int_s^\infty e^{\lambda(s-t)} T_{s,t}^a f_\varepsilon(t, x)dt, \]

for \( s \in [-1, 1] \), we split \( \mathcal{P}_j^\varepsilon f(s, x) = \mathcal{P}_{j1}^\varepsilon f(s, x) + \mathcal{P}_{j2}^\varepsilon f(s, x) \) with (see [3, 11])

\[ \mathcal{P}_{j1}^\varepsilon f(s, x) := \Delta_{x_j}^{1/(1+2(n-j))} \int_s^2 e^{\lambda(s-t)} T_{s,t} f_\varepsilon(t, x)dt, \]

\[ \mathcal{P}_{j2}^\varepsilon f(s, x) := \Delta_{x_j}^{1/(1+2(n-j))} \int_2^\infty e^{\lambda(s-t)} T_{s,t} f_\varepsilon(t, x)dt. \]

In the rest of this paper, unless otherwise specified, all the constants contained in “\( \lesssim \)” will depend only on \( n, \kappa, p, d \).

Lemma 4.2. Under (1.2), there is a constant \( C = C(n, \kappa, p, d) > 0 \) such that for all \( f \in L^\infty(\mathbb{R}^{1+nd}) \) with \( \|f\|_\infty \leq 1 \),

\[ \sup_{\varepsilon \in (0,1)} \int_{Q_1(0)} |\mathcal{P}_{j1}^\varepsilon f(s, x)|^2 \leq C. \] (4.3)

Proof. For \( s \in [-1, 1] \), let

\[ u_\varepsilon(s, x) := \int_s^2 e^{\lambda(s-t)} T_{s,t} f_\varepsilon(t, x)dt = \int_s^\infty e^{\lambda(s-t)} T_{s,t} ((1[-1,1]f_\varepsilon)(t))(x)dt. \]

Since \( \|f\|_\infty \leq 1 \), we have

\[ \|u_\varepsilon(s)\|_\infty \leq 3, \quad s \in [-1, 1]. \] (4.4)
By (2.7), we have for any $s \in [-1, 1]$,
\[
\|\nabla_{x_n} u_\varepsilon(s)\|_\infty \leq \int_s^2 \|\nabla_{x_n} T_{s,t} f_\varepsilon(t)\|_\infty \, dt \lesssim \int_s^2 (t - s)^{-1/2} \, dt \lesssim 1. \tag{4.5}
\]
Let $\varphi$ be a nonnegative smooth cutoff function in $\mathbb{R}^d$ with $\varphi(x) = 1$ for $|x| \leq n$ and $\varphi(x) = 0$ for $|x| > 2n$. Noticing that
\[
\partial_s u_\varepsilon + \mathcal{L}_s u_\varepsilon + f_\varepsilon 1_{[-1,2]}(s) = 0,
\]
we have
\[
\partial_s (u_\varepsilon \varphi) + \mathcal{L}_s (u_\varepsilon \varphi) = \varphi \partial_s u_\varepsilon + \mathcal{L}_s (u_\varepsilon \varphi) = \mathcal{L}_s (u_\varepsilon \varphi) - \varphi \mathcal{L}_s u_\varepsilon - f_\varepsilon \varphi 1_{[-1,2]}(s)
\]
\[
= u_\varepsilon \mathcal{L}_s \varphi + 2 \sum_{i,j=1}^d a_{ij} \partial_{x_{n_i}} u_\varepsilon \partial_{x_{n_j}} \varphi - f_\varepsilon \varphi 1_{[-1,2]}(s) =: g_\varepsilon^{\varphi}.
\]
Since $\varphi$ is compactly supported, it follows from (4.4)–(4.5) that $\|g_\varepsilon^{\varphi}\|_2 \lesssim 1$. On the other hand, we have by (1.13) with $\lambda = 0$ that
\[
(u_\varepsilon \varphi)(s, x) = \int_s^\infty e^{\lambda(s-t)} T_{s,t} g_\varepsilon^{\varphi}(t, x) \, dt.
\]
By the definition of $\mathcal{D}_j^{\varepsilon,1}$, we have
\[
\int_{Q_1(0)} |\mathcal{D}_j^{\varepsilon,1} f\|^2 = \int_{Q_1(0)} |\Delta_x^{1/(1+2(n-j))} u_\varepsilon|^2 \leq \int_{\mathbb{R}^{1+nd}} |\Delta_x^{1/(1+2(n-j))} u_\varepsilon \varphi|^2
\]
\[
\leq 2 \|\Delta_x^{1/(1+2(n-j))} (u_\varepsilon \varphi)\|^2_2 + 2 \|\Delta_x^{1/(1+2(n-j))} (u_\varepsilon \varphi) - \varphi \Delta_x^{1/(1+2(n-j))} u_\varepsilon\|^2_2
\]
\[
=: I_1 + I_2.
\]
For $I_1$, by (1.12) for $p = 2$,
\[
I_1 = \left\| \Delta_x^{1/(1+2(n-j))} \int_s^\infty e^{\lambda(s-t)} T_{s,t} g_\varepsilon^{\varphi}(t, x) \, dt \right\|_2 \lesssim \|g_\varepsilon^{\varphi}\|_2 \lesssim 1.
\]
For $I_2$, if $j = n$, then by (4.4)–(4.5) and that $\varphi$ is compactly supported,
\[
I_2 = 2 \|u_\varepsilon \Delta_{x_n} \varphi + 2 \nabla_{x_n} u_\varepsilon \cdot \nabla_{x_n} \varphi\|^2_2 \lesssim \|u_\varepsilon\|^2_\infty \|\Delta_{x_n} \varphi\|^2_2 + \|\nabla_{x_n} u_\varepsilon\|_\infty \|\nabla_{x_n} \varphi\|^2_2 \lesssim 1;
\]
if $j = 1, \ldots, n - 1$, then by definition (2.6) and (4.4),
\[
I_2 = 2 \left\| \int_{\mathbb{R}^d} \frac{2 \delta_z^{(1)} u_\varepsilon \delta_z^{(1)} \varphi}{|z_j|^{d+2/(1+2(n-j))}} \, dz_j + u_\varepsilon \Delta_x^{1/(1+2(n-j))} \varphi \right\|_2^2
\]
\[
\leq 4 \left\| \int_{\mathbb{R}^d} \frac{2 \|u_\varepsilon\|_\infty \|\delta_z^{(1)} \varphi\|^2_2}{|z_j|^{d+2/(1+2(n-j))}} \, dz_j + 2 \|u_\varepsilon\|_\infty \|\Delta_x^{1/(1+2(n-j))} \varphi\|^2_2 \right\|
\]
\[
\lesssim \int_{\mathbb{R}^d} \frac{|z_j| \wedge 1}{|z_j|^{d+2/(1+2(n-j))}} \, dz_j + 1 \lesssim 1.
\]
The proof is complete. \qed
The next lemma is crucial for treating $\mathcal{P}^f_{\partial \Omega} f$.

**Lemma 4.3.** Under (1.3), there is a constant $C = C(n, \kappa, p, d) > 0$ such that for all $f \in L^\infty(\mathbb{R}^{1+n})$ with $\|f\|_\infty \leq 1$ and all $s \in [-1, 1]$,

$$
\sup_{s \in (0, 1)} \int_2^s |I^f_{\partial \Omega}(s, t)| dt = \int_2^s \left| \int_0^s \left( (\Delta^1_{\tau \Omega} f_{\partial \Omega} + 2(\nabla_{\tau \Omega} f_{\partial \Omega}) ) f_{\partial \Omega} \right)(t, 0) dr \right| dt \\
\leq \kappa \int_2^s \int_0^s \left| (\Delta^1_{\tau \Omega} f_{\partial \Omega} + 2(\nabla_{\tau \Omega} f_{\partial \Omega}) ) f_{\partial \Omega} \right|(t, 0) dr dt \\
\leq \int_2^s (t-r)^{-3/2} dr dt \lesssim 1.
$$

For $j = n$, we have by (4.8) that for all $s \in [-1, 1]$,

$$
\int_2^s |I^f_{\partial \Omega}(s, t)| dt = \int_2^s \left| \int_0^s \left( (\Delta_{\tau \Omega} tr(a_{\tau \Omega} \cdot \nabla_{\tau \Omega}) ) f_{\partial \Omega} \right)(t, 0) dr \right| dt \\
\leq \int_2^s \int_0^s \left| (\Delta_{\tau \Omega} tr(a_{\tau \Omega} \cdot \nabla_{\tau \Omega}) ) f_{\partial \Omega} \right|(t, 0) dr dt \\
\leq \int_2^s (t-r)^{-2} dr dt \lesssim 1.
$$

For $j = 2, \ldots, n-1$, since $\Delta^1_{\tau \Omega} f_{\partial \Omega}$ is a nonlocal operator, we have to carefully treat the trouble term $\Delta^1_{\tau \Omega} f_{\partial \Omega}(x_j \cdot \nabla_{x_{j-1}} T_{\tau \Omega} f_{\partial \Omega})(t, 0)$. Fix $\gamma \in \left( \frac{2(n-j)+1}{2}, \frac{(2(n-j)+1)^2}{4(n-j)} \right)$.
By (2.5) and (4.7), we may write
\[
I^e_j(s, t) = \int_{|z_j| > t^\gamma} \left( \delta_{z_j} T_{0,t} f_e(t, 0) - \delta_{z_j} T_{t,t} f_e(t, 0) \right) \frac{dz_j}{|z_j|^{d+2/(1+2(n-j))}} \\
+ \int_{|z_j| \leq t^\gamma} \left( \int_0^s \delta_{z_j} \text{tr}(a_r \cdot \nabla^2 s_n) T_{r,t} f_e(t, 0) dr \right) \frac{dz_j}{|z_j|^{d+2/(1+2(n-j))}} \\
+ \int_{|z_j| \leq t^\gamma} \left( \int_0^s \delta_{z_j} (Ax \cdot \nabla T_{r,t} f_e)(t, 0) dr \right) \frac{dz_j}{|z_j|^{d+2/(1+2(n-j))}}
\]
\[= I^e_{j1}(s, t) + I^e_{j2}(s, t) + I^e_{j3}(s, t). \]

For \(I^e_{j1}(s, t)\), thanks to \(\gamma > \frac{2(n-j)+1}{2}\), we have for all \(s \in [-1, 1]\),
\[
\int_2^\infty |I^e_{j1}(s, t)| dt \lesssim \int_2^\infty \int_{|z_j| > t^\gamma} \frac{dz_j}{|z_j|^{d+2/(1+2(n-j))}} dt \lesssim \int_2^\infty |t|^{-2\gamma/(1+2(n-j))} dt \lesssim 1.
\]

For \(I^e_{j2}(s, t)\), by (2.7) and \(\gamma < \frac{(2(n-j)+1)^2}{2(2(n-j)-1)}\), we have for all \(s \in [-1, 1]\),
\[
\int_2^\infty |I^e_{j2}(s, t)| dt \lesssim \int_2^\infty \int_0^s \left( \int_{|z_j| \leq t^\gamma} \| \nabla x_j \|_{\infty} \| \nabla^2 s_n T_{r,t} f_e \|_{\infty} \frac{dz_j}{|z_j|^{d+2/(1+2(n-j))}} \right) dr dt \\
\lesssim \int_2^\infty \int_0^s (t-r)^{-2(n-j)+3/2} \left( \int_{|z_j| \leq t^\gamma} \frac{|z_j| dz_j}{|z_j|^{d+2/(1+2(n-j))}} \right) dr dt \\
\lesssim \int_2^\infty (t-1)^{-2(n-j)+3/2} t^{-\gamma(1-2/(1+2(n-j)))} dt \lesssim 1.
\]

For \(I^e_{j3}(s, t)\), letting \(\tilde{z}_j = (0, \ldots, 0, z_j, 0, \ldots, 0)\) and observing that
\[
\int_{|z_j| \leq t^\gamma} \frac{z_j dz_j}{|z_j|^{d+2/(1+2(n-j))}} = 0,
\]
by (2.7) and \(\gamma < \frac{(2(n-j)+1)^2}{4(n-j)}\), we have for all \(s \in [-1, 1]\),
\[
\int_2^\infty |I^e_{j3}(s, t)| dt \\
= \int_2^\infty \left| \int_0^s \int_{|z_j| \leq t^\gamma} z_j \cdot \nabla x_{j-1} T_{r,t} f_e(t, \tilde{z}_j) \frac{dz_j}{|z_j|^{d+2/(1+2(n-j))}} dr \right| dt \\
= \int_2^\infty \left| \int_0^s \int_{|z_j| \leq t^\gamma} z_j \cdot \nabla x_{j-1} (T_{r,t} f_e(t, \tilde{z}_j) - T_{t,t} f_e(t, 0)) \frac{dz_j}{|z_j|^{d+2/(1+2(n-j))}} dr \right| dt \\
\leq \int_2^\infty \left| \int_0^s \int_{|z_j| \leq t^\gamma} |z_j|^2 \| \nabla x_j \|_{\infty} \| \nabla x_{j-1} T_{r,t} f_e \|_{\infty} \frac{dz_j}{|z_j|^{d+2/(1+2(n-j))}} dr \right| dt.
\]
By definition, we have

$$\sup_{x \in (0, 1)} \int_{Q_1(0)} |\mathcal{P}_{j_2}^s f(s, x) - \mathcal{P}_{j_2}^s f(0, 0)|^2 \leq C. \quad (4.9)$$

**Proof.** By definition, we have

$$|\mathcal{P}_{j_2}^s f(s, x) - \mathcal{P}_{j_2}^s f(0, 0)| \leq \int_2^\infty |e^{\lambda(s-t)} - e^{-\lambda t}| \| \Delta_{x_j}^{1/2(n-j)} \mathcal{T}_{s,t} f_\varepsilon(t) \|_\infty dt$$

$$+ \int_2^\infty e^{-\lambda t} |\Delta_{x_j}^{1/2(n-j)} \mathcal{T}_{s,t} f_\varepsilon(t, x) - \Delta_{x_j}^{1/2(n-j)} \mathcal{T}_{s,t} f_\varepsilon(t, 0)| dt$$

$$+ \int_2^\infty e^{-\lambda t} |\nabla_{x_k} \Delta_{x_j}^{1/2(n-j)} \mathcal{T}_{s,t} f_\varepsilon(t) - \Delta_{x_j}^{1/2(n-j)} \mathcal{T}_{0,t} f_\varepsilon(t, 0)| dt$$

$$=: I_1(s) + I_2(s, x) + I_3(s).$$

Noticing that by (2.7),

$$\| \Delta_{x_j}^{1/2(n-j)} \mathcal{T}_{s,t} f_\varepsilon(t) \|_\infty \lesssim (t-s)^{-1},$$

$$\| \nabla_{x_k} \Delta_{x_j}^{1/2(n-j)} \mathcal{T}_{s,t} f_\varepsilon(t) \|_\infty \lesssim (t-s)^{-(2(n-k)+3)/2},$$

we have for all $s \in [-1, 1],$

$$I_1 \lesssim \int_2^\infty |e^{\lambda(s-t)} - e^{-\lambda t}|(t-s)^{-1} dt$$

$$\lesssim |e^{\lambda s} - 1| \int_2^\infty e^{-\lambda t} dt = |e^{\lambda s} - 1|e^{-2\lambda}/\lambda \lesssim 1,$$

and for all $(s, x) \in Q_1(0),$

$$I_2(s, x) \lesssim \int_2^\infty \sum_{k=1}^n (t-s)^{-(2(n-k)+3)/2} dt \lesssim 1.$$ 

Moreover, by (4.6), we have for all $s \in [-1, 1],$

$$I_3(s) \lesssim 1.$$

Combining the above calculations, we obtain (4.9). \hfill \Box

Now, we can give

**Proof of (1.12) for $p \in (2, \infty).** By Lemmas 4.1, 4.2 and 4.4, it is easy to see that

$$\mathcal{P}_{j_2}^s : L^\infty(\mathbb{R}^{1+nd}) \to BMO$$

is a bounded linear operator with bound independent of $\varepsilon.$ Estimate (1.12) for $p \in (2, \infty)$ follows by Theorem 2.4 and the well-proved estimate for $p = 2.$ \hfill \Box
4.3. Case: \( p \in (1, 2) \)

We shall use duality argument to show (1.12) for \( p \in (1, 2) \). Let \( T_{s,t}^* \) be the adjoint operator of \( T_{s,t} \), that is,

\[
\int g T_{s,t} f = \int f T_{s,t}^* g.
\]

By the definition of \( T_{s,t} \), we have

\[
T_{s,t}^* f(x) := T_{s,t}^{*,a} f(x) := \mathbb{E} f \left( e^{(s-t)A} x - \int_t^s e^{(s-r)A} \sigma_r^a dW_r \right).
\]

For \( j = 1, \ldots, n \), we introduce a new operator

\[
Q^\varepsilon_j f := Q_j f^\varepsilon(s, x) := \int_{-\infty}^t e^{\lambda(s-t)} T_{s,t}^{*,a} \Delta_{x_j}^{1/(1+2(n-j))} f^\varepsilon(s, x) ds,
\]

where \( f^\varepsilon(t, x) = f(t, \cdot) * \varrho^\varepsilon(x) \) so that \( Q^\varepsilon_j f \) is well defined for \( f \in L^\infty(\mathbb{R}^{1+nd}) \). Notice that \( Q^\varepsilon_j \) can be considered as the adjoint operator of \( P^\varepsilon_j \) in the sense that

\[
\int \mathcal{P}_j^a f g = \int Q_j^a g f.
\]

As in the previous subsection, we want to show that

\( Q^\varepsilon_j \) is a bounded linear operator from \( L^\infty(\mathbb{R}^{1+nd}) \) to \( BMO \).

First of all, as in Lemma 4.1 we have

\[
\int_{Q_{\varepsilon}(t_0, x_0)} \left| Q_j^a f^\varepsilon(s, x) - c \right|^2 = \int_{Q_{\varepsilon}(0)} \left| \hat{Q}_j^a \tilde{f}^\varepsilon(s, x) - c \right|^2,
\]

where \( \hat{a} \) and \( \tilde{f} \) are defined as in Lemma 4.1. We aim to prove that there is a constant \( C = C(n, \kappa, p, d) > 0 \) independent of \( \varepsilon \in (0, 1) \) such that for all \( f \in L^\infty(\mathbb{R}^{1+nd}) \) with \( \| f \|_\infty \leq 1 \),

\[
\int_{Q_{\varepsilon}(0)} \left| \hat{Q}_j^a \tilde{f}^\varepsilon(s, x) - c \right|^2 \leq C.
\]

Below, we drop \( \hat{a} \) and the tilde, and make the following decomposition as above:

\[
Q^\varepsilon_j f(t, x) = \left( \int_{-2}^t + \int_{-\infty}^{t-2} \right) e^{\lambda(s-t)} T_{s,t}^{*,a} \Delta_{x_j}^{1/(1+2(n-j))} f^\varepsilon(s, x) ds
\]

\[
=: Q^\varepsilon_{j1} f(t, x) + Q^\varepsilon_{j2} f(t, x).
\]

The following lemma is crucial for treating \( Q^\varepsilon_{j1} \).
LEMMA 4.5. Let \( \varphi \in C^\infty_c(\mathbb{R}^d) \). For any \( p \in [1, 2] \), there are constants \( C, \beta > 0 \) such that for all \( h \in L^2(\mathbb{R}^d) \) and \( 0 < t - s \leq 3 \),

\[
\begin{align*}
\| \Delta_x^{1/((1+2(n-j))} (T_{s,t} \varphi^2 h) - \varphi_{s,t} T_{s,t} (\varphi h)) \|_p & \leq C \| (t - s)^{\beta-1} \|_2, \\
\| \Delta_x^{1/((1+2(n-j))} (\varphi_{s,t} T_{s,t} (\varphi h)) - \varphi_{s,t} \Delta_x^{1/((1+2(n-j))} T_{s,t} (\varphi h)) \|_p & \leq C \| (t - s)^{\beta-1} \|_2,
\end{align*}
\]

(4.10) (4.11)

where \( \varphi_{s,t}(x) := \varphi(e^{(t-s)A} x) \).

Proof. (1) Let \( p_{s,t}^{(0)}(y) \) be the distribution density of \( X_t^s \). By the definition and the chain rule, we have

\[
\begin{align*}
\nabla_x T_{s,t} f(x) & = \int_{\mathbb{R}^d} f(y) \nabla_x p_{s,t}^{(0)}(y - e^{(t-s)A} x)(x) dy \\
& = \int_{\mathbb{R}^d} f(y) \sum_{i=1}^j (\nabla_{y_i} p_{s,t}^{(0)}(y - e^{(t-s)A} x)) \frac{(t-s)^{j-i}}{(j-i)!} dy \\
& = \int_{\mathbb{R}^d} f(y + e^{(t-s)A} x) \sum_{i=1}^j (\nabla_{y_i} p_{s,t}^{(0)}(y)) \frac{(t-s)^{j-i}}{(j-i)!} dy.
\end{align*}
\]

By this formula, we have

\[
\begin{align*}
\| \nabla_x T_{s,t} (\varphi^2 h) - \varphi_{s,t} \nabla_x T_{s,t} (\varphi h) \|_p & \leq \| \nabla \varphi \|_\infty \| \varphi h \|_p \sum_{i=1}^j \int_{\mathbb{R}^d} |y| \| \nabla_{y_i} p_{s,t}^{(0)}(y) \| \frac{(t-s)^{j-i}}{(j-i)!} dy.
\end{align*}
\]

(4.12)

By (2.1), we have

\[
\begin{align*}
\int_{\mathbb{R}^d} |y| \| \nabla_{y_i} p_{s,t}^{(0)}(y) \| dy & \lesssim (t-s)^{-\frac{n^2d+2(n-i)+1}{2}} \sum_{k=1}^n \int_{\mathbb{R}^d} |y_k| e^{-c|\Theta_{(t-s)-1/2} y|^2} dy \\
& \lesssim \sum_{k=1}^n (t-s)^{-\frac{n^2d+2(k-i)}{2}} \int_{\mathbb{R}^d} e^{-c|\Theta_{(t-s)-1/2} y|^2} dy \lesssim \sum_{k=1}^n (t-s)^{j-k}.
\end{align*}
\]

Substituting this into (4.12), we obtain

\[
\| \nabla_x T_{s,t} (\varphi^2 h) - \varphi_{s,t} \nabla_x T_{s,t} (\varphi h) \|_p \lesssim \| \varphi h \|_p (t-s)^{j-n} \lesssim \| h \|_2 (t-s)^{j-n},
\]

and further,

\[
\| \nabla_x (T_{s,t} (\varphi^2 h) - \varphi_{s,t} T_{s,t} (\varphi h)) \|_p \lesssim \| h \|_2 (t-s)^{j-n}.
\]
Hence, for \( j = 1, 2, \ldots, n - 1 \),
\[
\| \Delta_x^{1/(1+2(n-j))} (T_{s,t} (\varphi^2 h) - \varphi_s T_{s,t} (\varphi h)) \|_p \\
\leq \int_{\mathbb{R}^d} \| \delta_{z_j} (T_{s,t} (\varphi^2 h) - \varphi_s T_{s,t} (\varphi h)) \|_p \frac{dz_j}{|z_j|^{d+2/(1+2(n-j))}} \\
\leq 2 \int_{|z_j| > (t-s)^{n-j}} \left( \| T_{s,t} (\varphi^2 h) \|_p + \| \varphi_s T_{s,t} (\varphi h) \|_p \right) \frac{dz_j}{|z_j|^{d+2/(1+2(n-j))}} \\
+ \int_{|z_j| \leq (t-s)^{n-j}} \left( \| \nabla_x \delta_{z_j} (T_{s,t} (\varphi^2 h) - \varphi_s T_{s,t} (\varphi h)) \|_p \right) \frac{|z_j| dz_j}{|z_j|^{d+2/(1+2(n-j))}} \\
\lesssim (\| \varphi^2 h \|_p + \| \varphi h \|_p) \int_{|z_j| > (t-s)^{n-j}} \frac{dz_j}{|z_j|^{d+2/(1+2(n-j))}} \\
+ \| h \|_2 (t-s)^{j-n} \int_{|z_j| \leq (t-s)^{n-j}} \frac{|z_j| dz_j}{|z_j|^{d+2/(1+2(n-j))}} \lesssim \| h \|_2 (t-s)^{-2(n-j)/(1+2(n-j))}.
\]

Thus, we get (4.10). For \( j = n \), (4.10) is direct by the chain rule.

(2) For \( j = 1, \ldots, n - 1 \), by (2.8) we have
\[
\| \Delta_x^{1/(1+2(n-j))} (\varphi_{s,s} T_{s,t} (\varphi h)) - \varphi_s T_{s,t} (\Delta_x^{1/(1+2(n-j))} T_{s,t} (\varphi h)) \|_p \\
\leq 2 \int_{\mathbb{R}^d} \| \delta_{z_j} (\varphi_{s,s} T_{s,t} (\varphi h)) - \varphi_s T_{s,t} (\Delta_x^{1/(1+2(n-j))} T_{s,t} (\varphi h)) \|_p \frac{dz_j}{|z_j|^{d+2/(1+2(n-j))}} \\
\lesssim \| \varphi h \|_p \int_{\mathbb{R}^d} \frac{dz_j}{|z_j|^{d+2/(1+2(n-j))}} \lesssim \| h \|_2 (t-s)^{-2(n-j)/(1+2(n-j))},
\]

which gives (4.11). For \( j = n \), (4.11) is direct by the chain rule.

\[ \square \]

**Lemma 4.6.** Under (1.3), there is a constant \( C = C(n, \kappa, p, d) > 0 \) such that for all \( f \in L^\infty (\mathbb{R}^{1+nd}) \) with \( \| f \|_\infty \leq 1 \),
\[
\sup_{s \in (0,1)} \int_{Q_1(0)} |g^s f (s, x)|^2 \leq C.
\] (4.13)

**Proof.** For \( t \in [-2, 1] \), define
\[
u_s(t, x) := \mathcal{Q}_{1}^s f (t, x) = \int_{-\infty}^t e^{\lambda(s-t)} \mathcal{T}_{s,t} \Delta_x^{1/(1+2(n-j))} (1_{[-2,1])} f_s(s)) \varphi^2 h ds.
\]

Let \( \varphi \) be a nonnegative smooth cutoff function in \( \mathbb{R}^{nd} \) with \( \varphi(x) = 1 \) for \( |x| \leq 1 \) and \( \varphi(x) = 0 \) for \( |x| > 2 \). We have
\[
\| u \|_{L^2(Q_1(0))} \leq \| 1_{[-2,1]} u \varphi^2 \|_2 = \sup_{\| h \|_2 \leq 1} \int_{\mathbb{R}^{1+nd}} 1_{[-2,1]} u \varphi^2 h \\
= \sup_{\| h \|_2 \leq 1} \int_{\mathbb{R}^{1+nd}} 1_{[-2,1]} f_s \Delta_x^{1/(1+2(n-j))} \int_0^1 e^{\lambda(t-s)} \mathcal{T}_{s,t} (\varphi^2 h(t)) dt.
\]
\[ \|1_{[-2,1]}\Delta^{1/(1+2(n-j))}_{x_j} \int_1^1 e^{\lambda(s-t)} T_{s,t}(\varphi^2 h(t)) dt \|_1 \leq \sup_{\|h\|_2 \leq 1} \|1_{[-2,1]}\Delta^{1/(1+2(n-j))}_{x_j} \int_{-\infty}^1 e^{\lambda(s-t)} T_{s,t}(\varphi^2 h(t)) dt \|_1. \]

Since the support of \( \varphi_{s,t}(x) = \varphi(e^{(t-s)A} x) \) is contained in \( \{ x : |x| \leq n \} \) for \( |t-s| \leq 3 \), by (4.10) and (4.11), we have for any \( h \in L^2(\mathbb{R}^{1+nd}) \) with \( \|h\|_2 \leq 1 \),

\begin{align*}
\|1_{[-2,1]}\Delta^{1/(1+2(n-j))}_{x_j} \int_{-\infty}^1 e^{\lambda(s-t)} \varphi_{s,t} \Delta^{1/(1+2(n-j))}_{x_j} T_{s,t}(\varphi h(t)) dt \|_1 & \lesssim 1, \\
\|1_{[-2,1]}\Delta^{1/(1+2(n-j))}_{x_j} \int_{-\infty}^1 e^{\lambda(s-t)} \varphi_{s,t} \Delta^{1/(1+2(n-j))}_{x_j} T_{s,t}(\varphi h(t)) dt \|_2 & \lesssim 1, \\
\|\Delta^{1/(1+2(n-j))}_{x_j} \int_{-\infty}^1 e^{\lambda(s-t)} T_{s,t}(1_{[-2,1]}\varphi^2 h(t)) dt \|_2 & \lesssim 1, \\
\|\varphi^2 h\|_2 & \leq 1, \\
\|\varphi^2 h\|_2 & \leq 1.
\end{align*}

where the last step is due to (1.12) for \( p = 2 \). The proof is complete. \( \square \)

The following lemma is the same as in Lemma 4.4.

**Lemma 4.7.** Under (1.3), there is a constant \( C = C(n, \kappa, p, d) > 0 \) such that for all \( f \in L^\infty(\mathbb{R}^{1+nd}) \) with \( \|f\|_\infty \leq 1 \),

\[ \sup_{\varepsilon \in (0,1)} \int_{Q_1(0)} |\partial_{j_2} f(t, x) - \partial_{j_2} f(0, 0)|^2 \leq C. \]

**Proof.** By (2.8), we have for all \( t \in [-1, 1] \),

\begin{align*}
\int_{-\infty}^{-2} |\mathcal{T}_{s,t}^* \Delta^{1/(1+2(n-j))}_{x_j} f_\varepsilon(s, 0) - \mathcal{T}_{s,0}^* \Delta^{1/(1+2(n-j))}_{x_j} f_\varepsilon(s, 0)| ds & \leq \int_{-\infty}^{-2} \int_0^t |\partial_r \mathcal{T}_{s,r}^* \Delta^{1/(1+2(n-j))}_{x_j} f_\varepsilon(s, 0)| dr ds \\
& \leq \int_{-\infty}^{-2} \int_0^t |\mathcal{L}_r^* \mathcal{T}_{s,r}^* \Delta^{1/(1+2(n-j))}_{x_j} f_\varepsilon(s, 0)| dr ds \\
& = \int_{-\infty}^{-2} \int_0^t |\mathcal{L}_r^* \mathcal{T}_{s,r}^* \Delta^{1/(1+2(n-j))}_{x_j} f_\varepsilon(s, 0)| dr ds \\
& \leq C \int_{-\infty}^{-2} \int_0^t (r-s)^{-2} dr ds \leq C.
\end{align*}

Using this estimate and (2.8), as in the proof of Lemma 4.4, we obtain the desired estimate. \( \square \)
Proof of (1.12) for $p \in (1, 2)$. By Lemmas 4.6 and 4.7, the operator
\[ \mathcal{Q}_\varepsilon : L^\infty(\mathbb{R}^{1+2d}) \to BMO \] is bounded with norm independent of $\varepsilon$.

Moreover, by duality, we have
\[ \mathcal{Q}_\varepsilon : L^2(\mathbb{R}^{1+2d}) \to L^2(\mathbb{R}^{1+2d}) \] is bounded with norm independent of $\varepsilon$.

Hence, for $q = p/(p - 1) \in (2, \infty)$, by Theorem 2.4, we have for some $C > 0$ independent of $\varepsilon$,
\[ \| \mathcal{Q}_\varepsilon f \|_q = \left\| \int_{-\infty}^t e^{\lambda(s-t)} T_{s,t}^* \Delta_j^{1/(1+2(n-j))} f_s ds \right\|_q \leq C \| f \|_q. \]

Now, for $p \in (1, 2)$, by Fatou’s lemma, we get
\[
\| \mathcal{P}_\varepsilon f \|_p \leq \| f \|_p \sup_{\| h \|_q \leq 1} \left\| \int_{-\infty}^t e^{\lambda(s-t)} T_{s,t}^* \Delta_j^{1/(1+2(n-j))} h_s ds \right\|_q
\leq \| f \|_p \lim_{\varepsilon \to 0} \sup_{\| h \|_q \leq 1} \left\| \int_{-\infty}^t e^{\lambda(s-t)} T_{s,t}^* \Delta_j^{1/(1+2(n-j))} h_{\varepsilon} ds \right\|_q \leq C \| f \|_p,
\]
which gives (1.12) for $p \in (1, 2)$.

Acknowledgements

We thank the referees for helpful comments.

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

REFERENCES
[1] R. Alexander: Fractional order kinetic equations and hypoellipticity. Anal. Appl. (Singap.) 10, no.3 (2012), 237-247.
[2] F. Bouchut: Hypoelliptic regularity in kinetic equations. J. Math. Pures Appl. 81 (2002), 1135-1159.
[3] M. Bramanti, G. Cupini, E. Lanconelli and E. Priola: Global $L^p$-estimate for degenerate Ornstein-Uhlenbeck operators. Math Z. 266 (2010), 789-816.
[4] M. Bramanti, G. Cupini, E. Lanconelli and E. Priola: Global $L^p$-estimates for degenerate Ornstein-Uhlenbeck operators with variable coefficients. Math. Nachr. 286 (2013), no. 11-12, 1087-1101.
[5] Z.-Q. Chen and X. Zhang: $L^p$-maximal hypoelliptic regularity of nonlocal kinetic Fokker-Planck operators. J. Math. Pures et Appliquées 116 (2018), 52-87.
[6] E. Fedrizzi, F. Flandoli, E. Priola and J. Vovelle: Regularity of stochastic kinetic equations. Electron. J. Probab. Volume 22 (2017), paper no. 48, 42 pp.
[7] L. Huang, S. Menozzi and E. Priola: $L^p$-estimates for degenerate non-local Kolmogorov operators. J. Math. Pures Appl. 121 (2019), 162-215.
[8] R. A. N. Kolmogorov: Zufällige Bewegungen. Ann. Math. 35 (1934), 116-117.
[9] E. Lunardi: Schauder’s estimates for a class of degenerate elliptic and parabolic operators with unbounded coefficients in $\mathbb{R}^n$. Annali della Scuola Normale Superiore di Pisa, Vol. 24 (1997), 133-164.
[10] E. M. Stein: *Harmonic Analysis: Real-Variable Methods, Orthogonality and Oscillatory Integrals*. Princeton University Press, Princeton 1993.

[11] X. Zhang: $L^p$-maximal regularity of nonlocal parabolic equations and applications. *Ann. I. H. Poincare-AN* **30** (2013), 573-614.

[12] X. Zhang: Stochastic Hamiltonian flows with singular coefficients. *Sci. China. Math.* **61** (2018), 1353-1384.

Zhen-Qing Chen
Department of Mathematics
University of Washington
Seattle, WA 98195
USA
E-mail: zqchen@uw.edu

Xicheng Zhang
School of Mathematics and Statistics
Wuhan University
Wuhan 430072, Hubei
People’s Republic of China
E-mail: XichengZhang@gmail.com