Calderón-Zygmund theory of nonlocal parabolic equations with discontinuous coefficients

SUN-SIG BYUN, KYEONGBAE KIM AND DEEPAK KUMAR

Abstract. We prove Calderón-Zygmund type estimates of weak solutions to non-homogeneous nonlocal parabolic equations under a minimal regularity requirement on kernel coefficients. In particular, the right-hand side is presented by a sum of fractional Laplacian type data and a non-divergence type data. Interestingly, even though the kernel coefficients are discontinuous, we obtain a significant increment of fractional differentiability for the solutions, which is not observed in the corresponding local parabolic equations.

1. Introduction

1.1. **Overview.** In this paper, we study higher regularity properties for weak solutions to the following non-homogeneous nonlocal parabolic equation:

$$u_t + \mathcal{L}_A^{\Phi} u = (-\Delta)^{\frac{s}{2}} f + g \quad \text{ in } \Omega_T \equiv \Omega \times (0,T), \tag{1.1}$$

where $s \in (0,1)$, $T > 0$ and Ω is an open and bounded set in $\mathbb{R}^n$ with $n \geq 2$. The nonlocal operators appearing in problem (1.1) are defined by

$$\mathcal{L}_A^{\Phi} u(x,t) = \text{p.v.} \int_{\mathbb{R}^n} \Phi\left(\frac{u(x,t)-u(y,t)}{|x-y|^s}\right) \frac{A(x,y,t)}{|x-y|^{n+s}}\, dy$$

and

$$(-\Delta)^{\frac{s}{2}} f(x,t) = \text{p.v.} \int_{\mathbb{R}^n} (f(x,t)-f(y,t)) \frac{1}{|x-y|^{n+s}}\, dy.$$

Here, $f : \mathbb{R}^n \times (0,T) \to \mathbb{R}$ and $g : \Omega_T \to \mathbb{R}$ are given measurable functions, $A : \mathbb{R}^n \times \mathbb{R}^n \times (0,T) \to \mathbb{R}$ is a given kernel coefficient satisfying

$$L^{-1} \leq A(x,y,t) \leq L \quad \text{and} \quad A(x,y,t) = A(y,x,t) \quad \text{for } (x,y,t) \in \mathbb{R}^n \times \mathbb{R}^n \times (0,T) \tag{1.2}$$

and for some constant $L \geq 1$, and $u : \mathbb{R}^n \times (0,T) \to \mathbb{R}$ is the unknown. In addition, $\Phi : \mathbb{R} \to \mathbb{R}$ is a measurable function satisfying $\Phi(0) = 0$ and

$$\begin{cases} (\Phi(\xi) - \Phi(\xi'))(\xi - \xi') \geq L^{-1}|\xi - \xi'|^2, \\ |\Phi(\xi) - \Phi(\xi')| \leq L|\xi - \xi'| \quad \text{for any } \xi, \xi' \in \mathbb{R}. \end{cases} \tag{1.3}$$

Nonlocal parabolic problems appear naturally in the physical world, e.g., as in anomalous diffusion processes from the areas of physics, finance, biology, ecology, geophysics, and many others. Particularly, the nonlocal nonlinear operators of the above types find their application in image processing [27] and phase transition models [25].

It is known that when the leading operator appearing in (1.1) is linear and the right-hand side is regular enough, then solutions enjoy higher Hölder regularity, which in turn, yields improved Sobolev regularity. These assertions can be justified by the functional analysis tools along with a precise integral representation of the solution through suitable heat kernel type estimates. Unfortunately, when the operator is nonlinear, the aforementioned techniques fail to apply. To this end, our objective is to obtain a fine fractional Sobolev regularity for weak solutions to (1.1) by using purely analytic and geometric techniques. In particular, we introduce a unified approach of covering arguments to obtain some uniform measure density estimates for some level sets involving the solution, which we will explain in the sequel. More precisely, our aim is to establish Calderón-Zygmund type estimates for weak solutions to (1.1)

2010 *Mathematics Subject Classification.* 35A01, 35B65, 35D30, 35R05, 35R09.

Key words and phrases. Nonlocal parabolic equations, Sobolev Regularity, Calderón-Zygmund type estimates, Nonlinear equations.

S. Byun was supported by NRF-2022R1A2C1009312, K. Kim and D. Kumar were supported by NRF-2021R1A4A1027378.

under a minimal regularity requirement on the kernel coefficient $A = A(x,y,t)$. More precisely, we want to find an extra condition on A besides (1.2) under which the following implication holds

$$\begin{aligned} &f(x,t)-f(y,t)\in L^q_{\mathrm{loc}}\left(\Omega\times\Omega_T\,;\,\frac{dx\,dy\,dt}{|x-y|^{n+\sigma q}}\right) \text{ with } f\in L^q_{\mathrm{loc}}(0,T;L^1_s(\mathbb{R}^n)) \quad\text{and}\quad g\in L_{\mathrm{loc}}^{\frac{q\left(n+2s+\frac{2\sigma q}{q-2}\right)}{n+4s}}(\Omega_T) \\ &\Longrightarrow \frac{u(x,t)-u(y,t)}{|x-y|^s}\in L^q_{\mathrm{loc}}\left(\Omega\times\Omega_T\,;\,\frac{dx\,dy\,dt}{|x-y|^{n+\sigma q}}\right) \end{aligned} \tag{1.4}$$

for any $q\in(2,\infty)$ and $\sigma\in\left(0,\left(1-\frac{2}{q}\right)\min\left\{s-\frac{2s}{q},1-s\right\}\right)$ with the desired Calderón-Zygmund type estimates like (1.9). In particular, using the notion of fractional gradients introduced in [13,39]; that is,

$$d_\beta u(x,y,t) := \frac{u(x,t)-u(y,t)}{|x-y|^\beta} \quad \text{for any } \beta\in[0,1),$$

the implication (1.4) can be rewritten as

$$\begin{aligned} &d_0 f\in L^q_{\mathrm{loc}}\left(\Omega\times\Omega_T\,;\,\frac{dx\,dy\,dt}{|x-y|^{n+\sigma q}}\right) \text{ with } f\in L^q_{\mathrm{loc}}(0,T;L^1_s(\mathbb{R}^n)) \quad\text{and}\quad g\in L_{\mathrm{loc}}^{\frac{q\left(n+2s+\frac{2\sigma q}{q-2}\right)}{n+4s}}(\Omega_T) \\ &\Longrightarrow d_s u\in L^q_{\mathrm{loc}}\left(\Omega\times\Omega_T\,;\,\frac{dx\,dy\,dt}{|x-y|^{n+\sigma q}}\right). \end{aligned}$$

1.2. **Some known results.** For the elliptic problems, a self-improving property of a weak solution to the problem:

$$\mathcal{L}^\Phi_A u = (-\Delta)^{\frac{s}{2}} f + g \tag{1.5}$$

is obtained by Kuusi, Mingione and Sire [35] by introducing the notion of dual pairs. When $\Phi(\xi)=\xi$ and $A = A(x,y)$ is Hölder continuous, Calderón-Zygmund type estimate for (1.5) is established by Mengesha, Schikorra and Yeepo [40] via commutator estimates. In addition, the aforementioned articles deal with more general equations as source terms involve $\tilde{s}$-fractional Laplacian with $\tilde{s}\neq\frac{s}{2}$. For operators with possibly discontinuous coefficients, such as VMO coefficients, Nowak [42,43] obtain Calderón-Zygmund type estimates when $f=0$ by using the maximal function and the notion of dual pairs. We refer to [1] for the global Calderón-Zygmund type estimate of (1.5) with $A=1$, $\Phi(\xi)=\xi$, $f=0$ and the zero Dirichlet condition on the exterior of the domain. The main tool employed in this work is the Green function representation of the solution.

We now mention some related results for the case of nonlinear nonlocal operators. When $\Phi(\xi) = |\xi|^{p-2}\xi$ with $p>2$ and $f=0$, Nowak and Diening [19] obtain sharp regularity results containing borderline cases by establishing precise pointwise bounds in terms of fractional sharp maximal functions. On the other hand, Calderón-Zygmund type estimates of solutions to the problem:

$$\mathcal{L}^\Phi_A u = (-\Delta_p)^{\frac{s}{p}} f,$$

are established in [13] via a maximal function free technique which was first introduced in [2]. We mention the work [22] as well for L^p-theory of a strong solution to nonlocal elliptic equations. For additional regularity results related to nonlocal elliptic equations, we refer to [5,8,9,12,14,15,17,18,23,24,32,34,36,37,41,44–46] and references therein.

For the parabolic problems, Auscher, Bortz, Egert and Saari [4] prove a self-improving property of solutions to (1.1) with $f=0$ by using functional analysis techniques. When $\Phi(\xi)=\xi$, $A(x,y,t)\equiv 1$ and $f=0$, Biccari, Warma and Zuazua [6] provide optimal regularity results of a weak solution by using a cut off argument. For the L^p-theory of strong solutions to nonlocal parabolic equations, we refer to [21,51]. We further mention [3,10,16,26,29–31,38,48–50] and references therein for various regularity results of nonlocal parabolic equations.

1.3. **Main results.** To explain the desired Carderón-Zygmund type estimate (1.4), we first introduce the notion of dual pairs. For a measurable function $F:\mathbb{R}^n\times\mathbb{R}^n\times(0,T)$ and $\tau\in(0,1)$, we define

$$D^\tau F(x,y,t) := \frac{F(x,y,t)}{|x-y|^\tau} \quad\text{and}\quad \mu_\tau(\mathcal{A}) := \int_{\mathcal{A}} \frac{dx\,dy}{|x-y|^{n-2\tau}}, \quad \text{for any measurable set } \mathcal{A}\subset\mathbb{R}^n\times\mathbb{R}^n.$$

Furthermore, we write $d\mu_{\tau,t} = d\mu_\tau\,dt$. Then we observe that

$$d_s u\in L^q_{\mathrm{loc}}\left(\Omega\times\Omega_T\,;\,\frac{dx\,dy\,dt}{|x-y|^{n+\sigma q}}\right) \Longleftrightarrow D^\tau d_s u\in L^q_{\mathrm{loc}}\left(\Omega\times\Omega_T\,;\,d\mu_{\tau,t}\right), \text{ where } \tau=\tfrac{q\sigma}{q-2}. \tag{1.6}$$

From this observation, we deduce that the solution u improves its integrability order as well as the differentiability order by achieving the same integrability as that of the associated non-homogeneous term in the Sobolev scale and a substantial gain in the differentiability order.

We now introduce a nonlocal tail space. We say that $u \in L^p(0,T;L^1_{2s}(\mathbb{R}^n))$ if

$$\left\| \int_{\mathbb{R}^n} \frac{|u(x,\cdot)|}{(1+|x|)^{n+2s}}\,dx \right\|_{L^p(0,T)} < \infty$$

for any $p \in [1,\infty]$. In particular, we write

$$\mathrm{Tail}_{p,2s}\left(u; B_r(x_0)\times I\right) = \left(\fint_I \left(r^{2s} \int_{\mathbb{R}^n\setminus B_r(x_0)} \frac{|u(y,t)|}{|y-x_0|^{n+2s}}\,dy \right)^p dt \right)^{\frac{1}{p}} \quad \text{if } p\in[1,\infty)$$

and

$$\mathrm{Tail}_{\infty,2s}\left(u; B_r(x_0)\times I\right) = \sup_{t\in I} \left(r^{2s} \int_{\mathbb{R}^n\setminus B_r(x_0)} \frac{|u(y,t)|}{|y-x_0|^{n+2s}}\,dy \right),$$

where $I \subset \mathbb{R}$ is a bounded time interval. We note from Hölder's inequality that for any $1\le p_1 \le p_2 \le \infty$,

$$\mathrm{Tail}_{p_1,2s}\left(u; B_r(x_0)\times I\right) \le \mathrm{Tail}_{p_2,2s}\left(u; B_r(x_0)\times I\right). \tag{1.7}$$

As usual, a solution to (1.1) is defined in the weak sense as below.

Definition 1.1. Let $f \in L^2_{\mathrm{loc}}\left(0,T;L^1_s(\mathbb{R}^n)\right)$ with $d_0 f \in L^2_{\mathrm{loc}}\left(\Omega\times\Omega_T\,;\,\frac{dx\,dy\,dt}{|x-y|^n}\right)$ and $g \in L^{\frac{2(n+2s)}{n+4s}}_{\mathrm{loc}}(\Omega_T)$. We say that

$$u \in L^2_{\mathrm{loc}}\left(0,T;W^{s,2}_{\mathrm{loc}}(\Omega)\right) \cap C_{\mathrm{loc}}\left(0,T;L^2_{\mathrm{loc}}(\Omega)\right) \cap L^\infty_{\mathrm{loc}}\left(0,T;L^1_{2s}(\mathbb{R}^n)\right)$$

is a weak solution to (1.1) if

$$\begin{aligned}
&\int_{t_1}^{t_2}\int_\Omega -u\phi_t\,dx\,dt + \int_{t_1}^{t_2}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \Phi\left(\frac{u(x,t)-u(y,t)}{|x-y|^s}\right)(\phi(x,t)-\phi(y,t))\frac{A(x,y,t)}{|x-y|^{n+s}}\,dx\,dy\,dt \\
&= -\int_\Omega (u\phi)(x,t)\,dx\bigg|_{t=t_1}^{t=t_2} + \int_{t_1}^{t_2}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} (f(x,t)-f(y,t))(\phi(x,t)-\phi(y,t))\frac{1}{|x-y|^{n+s}}\,dx\,dy\,dt \\
&\quad + \int_{t_1}^{t_2}\int_\Omega g\phi\,dx\,dt
\end{aligned}$$

holds for any $\phi \in L^2(t_1,t_2;W^{s,2}(\Omega)) \cap W^{1,2}\left(t_1,t_2;L^2(\Omega)\right)$ with compact spatial support contained in Ω and $(t_1,t_2) \Subset (0,T)$.

We next introduce the notion of (δ,R)-vanishing condition on A, for some δ and $R>0$. We say that A is (δ,R)-vanishing in Ω_T, if

$$\sup_{0<r\le R,\, Q_r(z_0)\subset\Omega_T} \fint_{\Lambda_r(t_0)}\fint_{B_r(x_0)}\fint_{B_r(x_0)} |A(x,y,t)-(A)_{r,x_0}(t)|\,dx\,dy\,dt \le \delta,$$

where $z_0=(x_0,t_0)$ and

$$(A)_{r,x_0}(t) \coloneqq \fint_{B_r(x_0)}\fint_{B_r(x_0)} A(x,y,t)\,dx\,dy. \tag{1.8}$$

We now observe the following scaling invariance property for the problem (1.1).

Lemma 1.2. *Let $Q_r(z_0) \Subset \Omega_T$. Suppose that A is (δ,R)-vanishing in Ω_T. Then*

$$\tilde{u}(x,t) = \frac{u(rx+x_0, r^{2s}t+t_0)}{r^s}$$

is a weak solution to

$$\tilde{u}_t + \mathcal{L}^\Phi_{\tilde{A}}\tilde{u} = (-\Delta)^{\frac{s}{2}}\tilde{f} + \tilde{g} \quad \text{in } Q_1,$$

where

$$\tilde{f}(x,t) = f(rx+x_0, r^{2s}t+t_0), \quad \tilde{g}(x,t) = r^s g(rx+x_0, r^{2s}t+t_0)$$

and $\tilde{A}(x,y,t) = A(rx+x_0, ry+x_0, r^{2s}t+t_0)$ is $\left(\delta,\frac{R}{r}\right)$-vanishing in Q_1.

We now introduce our main results.

Theorem 1.1. *Let u be a weak solution to problem (1.1). Let $R > 0$ and $q \in (2, \infty)$ be given, and fix $\sigma \in \left(0, \left(1-\frac{2}{q}\right)\min\left\{s-\frac{2s}{q}, 1-s\right\}\right)$. Then there is a constant $\delta = \delta(n,s,L,q,\sigma) \in (0,1)$ such that if A is (δ,R)-vanishing in Ω_T, $f \in L^q_{\mathrm{loc}}\left(0,T;L^1_s(\mathbb{R}^n)\right)$ with $d_0 f \in L^q_{\mathrm{loc}}\left(\frac{dx\,dy\,dt}{|x-y|^{n+\sigma q}}\,;\Omega\times\Omega_T\right)$ and $g \in L_{\mathrm{loc}}^{\frac{q\left(n+2s+\frac{2\sigma q}{q-2}\right)}{n+4s}}(\Omega_T)$, then $d_s u \in L^q_{\mathrm{loc}}\left(\frac{dx\,dy\,dt}{|x-y|^{n+\sigma q}}\,;\Omega\times\Omega_T\right)$. Moreover, there is a constant $c = c(n,s,L,q,\sigma)$ such that*

$$
\begin{aligned}
&\left(\fint_{\Lambda_{\frac{r}{2}}(t_0)}\fint_{B_{\frac{r}{2}}(x_0)}\int_{B_{\frac{r}{2}}(x_0)} |d_{s+\sigma}u|^q \frac{dx\,dy\,dt}{|x-y|^n}\right)^{\frac{1}{q}}\\
&\le c\left(\fint_{\Lambda_r(t_0)}\fint_{B_r(x_0)}\int_{B_r(x_0)}\left|\frac{d_s u}{r^\sigma}\right|^2\frac{dx\,dy\,dt}{|x-y|^n}\right)^{\frac{1}{2}} + c\,\mathrm{Tail}_{\infty,2s}\left(\frac{u-(u)_{B_r(x_0)}(t)}{r^{s+\sigma}}\,;Q_r(z_0)\right)\\
&\quad + c\sup_{t\in\Lambda_r(t_0)}\left(\fint_{B_r(x_0)}\frac{|u-(u)_{Q_r(z_0)}|^2}{r^{2s+2\sigma}}\,dx\right)^{\frac{1}{2}} + c\left(\fint_{Q_r(z_0)}\left(r^{s-\sigma}|g|\right)^{\frac{q\left(n+2s+\frac{2\sigma q}{q-2}\right)}{n+4s}}dx\,dt\right)^{\frac{n+4s}{q\left(n+2s+\frac{2\sigma q}{q-2}\right)}}\\
&\quad + c\left(\fint_{\Lambda_r(t_0)}\fint_{B_r(x_0)}\int_{B_r(x_0)}|d_\sigma f|^q\frac{dx\,dy\,dt}{|x-y|^n}\right)^{\frac{1}{q}} + c\,\mathrm{Tail}_{q,s}\left(\frac{f-(f)_{B_r(x_0)}(t)}{r^\sigma}\,;Q_r(z_0)\right),
\end{aligned}
\tag{1.9}
$$

whenever $Q_r(z_0) \Subset \Omega_T$ and $r \in (0,R]$.

Remark 1. A few comments are in order for the restricted range of σ in Theorem 1.1. In the elliptic case, we observe that a similar type of result holds for all $\sigma \in (0, \min\{s, 1-s\})$ (see [13, Theorem 1.2]). However, in our case, to handle the nonlocal parabolic tail induced by the non-homogeneous term f, we have to impose the condition $\tau \in \left(0, s-\frac{2s}{q}\right)$ (see (5.27) below). Therefore, from the observation (1.6), we deduce that the Calderón-Zygmund type estimate (1.9) holds under $\sigma \in \left(0, \left(1-\frac{2}{q}\right)\min\left\{s-\frac{2s}{q}, 1-s\right\}\right)$. In this regard, if $f = 0$, then this restriction is removed and the results hold for a broader range of σ (see Theorem 1.2 below).

Remark 2. As we pointed out earlier, in the elliptic case, from regularity results for (1.5) with $g = 0$, we deduce a higher regularity of a weak solution u to (1.5). However, in the parabolic case, it does not hold. More specifically, if $g \in L^{\tilde q}_{\mathrm{loc}}(\Omega_T)$ for some $\tilde q > 2$, then we find a solution $f \in L^{\tilde q}_{\mathrm{loc}}\left(0,T;H^{s,\tilde q}_{\mathrm{loc}}(\Omega)\right) \cap L^{\tilde q}_{\mathrm{loc}}(0,T;L^1_s(\mathbb{R}^n))$ to

$$(-\Delta)^{\frac{s}{2}} f(\cdot,t) = g(\cdot,t) \quad \text{in } \Omega$$

for a.e. $t \in (0,T)$ (see [13, Subsection 1.2]). This implies that u is a weak solution to

$$u_t + \mathcal{L}^\Phi_A u = (-\Delta)^{\frac{s}{2}} f.$$

By Theorem 1.1, we deduce that

$$u \in L^{\tilde q}_{\mathrm{loc}}\left(0,T;W^{s+\sigma,\tilde q}_{\mathrm{loc}}(\Omega)\right) \text{ for any } \sigma \in \left(0, \left(1-\frac{2}{\tilde q}\right)\min\left\{s-\frac{2s}{\tilde q}, 1-s\right\}\right).$$

We now select $\tilde q = \frac{q\left(n+2s+\frac{2\sigma q}{q-2}\right)}{n+4s}$. Since $\tilde q < q$, we do not obtain the desired result given in Theorem 1.2. Therefore, we consider a more general non-homogeneous term which consists of $(-\Delta)^{\frac{s}{2}} f$ and g.

On account of Remark 1, we obtain an improved Sobolev regularity when we consider only non-divergence data g.

Theorem 1.2. *Let u be a weak solution to problem (1.1) with $f = 0$. Let $R > 0$ and $q \in (2,\infty)$ be given, and fix $\sigma \in \left(0, \left(1-\frac{2}{q}\right)\min\{s, 1-s\}\right)$. Then there is a constant $\delta = \delta(n,s,L,q,\sigma) \in (0,1)$ such that if A is (δ,R)-vanishing in Ω_T and $g \in L_{\mathrm{loc}}^{\frac{q\left(n+2s+\frac{2\sigma q}{q-2}\right)}{n+4s}}(\Omega_T)$, then $u \in L^q_{\mathrm{loc}}\left(0,T;W^{s+\sigma,q}_{\mathrm{loc}}(\Omega)\right)$ with the estimate (1.9).*

Remark 3. Let us compare the local Calderón-Zygmund theory and the nonlocal ones. It is known that if v is a weak solution to

$$v_t - \mathrm{div}\,(BDv) = g,$$

then we obtain the following implication

$$g \in L_{\mathrm{loc}}^{\frac{q(n+2)}{n+4}}(\Omega_T) \Longrightarrow v \in L^q_{\mathrm{loc}}(0,T;W^{1,q}_{\mathrm{loc}}(\Omega)),$$

whenever the coefficient B is δ-vanishing for sufficiently small δ depending only on the given data (see, for instance, [7]). We observe that in the limiting case when $\sigma \to 0$ and $s \to 1$, the result in Theorem 1.2 is the same as the local one. However, in the nonlocal case, although the kernel coefficient is discontinuous, Theorem 1.2 implies that the solution u obtains not only higher integrability but also higher differentiability along the Sobolev-scale. This is in some sense a purely nonlocal phenomenon, since in order to observe such results in the local case, the coefficient B is assumed to have some fractional regularity in the literature (see [33]). A similar phenomenon for the nonlocal equations is observed in [4, 35, 42, 43] and references therein.

1.4. **Working methods and novelties.** We now briefly explain our approach to obtain the desired estimates (A.15) and (1.9). As usual, keeping the relation (1.6) in mind, to prove that the fractional gradient term $d_s u$ is in $L^q_{\mathrm{loc}}\left(\Omega\times\Omega_T\,;\,\frac{dx\,dy\,dt}{|x-y|^{n+\sigma q}}\right)$, it suffices to show that

$$\int_0^\infty \lambda^{q-1}\mu_{\tau,t}\left(\{(x,y,t)\in\mathcal{Q}\,:\,|D^\tau d_s u|(x,y,t)>\lambda\}\right)d\lambda<\infty$$

holds for any $\mathcal{Q}=B\times B\times\Lambda$, where $B\Subset\Omega$ is a ball and $\Lambda\Subset(0,T)$ is a time interval. To do this, we construct coverings for upper level sets of $|D^\tau d_s u|$ inspired by the maximal function free technique as introduced in [2]. Using an exit-time argument, we are able to construct coverings for the diagonal part of the upper level sets. For the off-diagonal part, we use a reverse Hölder-type inequality (see (5.48), below) which is obtained regardless of the information that u solves (1.1). As in [35, Lemma 5.3], this inequality contains additional correction terms involving diagonal cylinders which induce some serious difficulties, as such cylinders do not come from any exit-time argument. We would like to mention that in the elliptic case, Calderón-Zygmund cube decomposition and an involved combinatorial argument are used to overcome these difficulties (see [35]). However, in the parabolic case, there are additional difficulties, since the correction terms contain L^2-oscillation integrals by sup norm term (see the second term of the right-hand side in (5.32) and Lemma 5.2 below). To this end, we employ Vitali's covering lemma along with an exit-time argument instead of Calderón-Zygmund cube decomposition in order to construct coverings for upper-level sets of $|D^\tau d_s u|$. We would like to mention that this argument is new even in the elliptic case, and we believe that this argument can be applied to degenerate or singular nonlocal parabolic equations. We also point out that due to the appearance of the additional L^2-oscillation integrals by sup norm term, functionals used to apply an exit-time argument also contain a term of a similar kind which is usually not observed in the local parabolic problems (see (5.11)). We will elaborate on how to take care of this term while obtaining a good bound on the measure of exit-time cylinders in Remarks 6 and 7 below. Moreover, we use a non-trivial exit time radius in the covering arguments in light of the rigorous tail estimates as in (5.25) and (5.26), since the additional L^2-oscillation terms by sup norm are estimated by the sum of L^2-integral of $d_s u$ and $d_0 f$, $L^{\frac{2(n+2s)}{n+4s}}$-integral of g and tail terms of u and f (see Lemma 3.3). Consequently, by constructing suitable coverings, we are able to make use of comparison estimates, which further require some higher Hölder continuity estimates along with a self-improving property for limiting equations, and a boot strap argument to finally obtain the desired result (see Section 6).

We would like to remark that a similar covering argument along with Gehring's Lemma (in the spirit of [35]) can be used to obtain a self-improving property of weak solutions to (1.1) without imposing any regularity assumption on the kernel coefficient A. Indeed, for the sake of completion, we prove a self-improving property of weak solutions to (1.1) with $f=g=0$ (see Appendix A). In addition, this result generalizes the ones given in [4] by allowing nonlinear structure assumptions on the nonlocal operator.

1.5. **Plan of the paper.** This paper is organized as follows. In Section 2, we introduce some notations, embedding inequalities, properties of the measure $\mu_{\tau,t}$, tail estimates. In Section 3, we derive some energy estimates. Section 4 is devoted to establishing some comparison estimates. In Section 5, we construct coverings of upper level sets of fractional gradients for weak solutions. Section 6 contains the proof of the main theorem. We end the paper with two appendices. In the first appendix, we give the proof of a self-improving property for weak solutions to (1.1) with $f=g=0$, whereas the second appendix deals with the existence of a weak solution to the corresponding boundary value problem of (1.1).

2. Preliminaries and Notations

As usual, we write c to mean a general constant equal to or bigger than 1 and it possibly changes from line to line. In addition, we employ parentheses to denote the relevant dependencies on parameters such as $c = c(n, s)$, and we denote

$$\mathsf{data} = \mathsf{data}(n, s, L, q, \sigma).$$

For $a, b \in \mathbb{R}^+$, by the notation $a \approx_{n,s} b$, we mean that there is a constant $c = c(n, s)$ such that $\frac{b}{c} \le a \le cb$. A generic point $z \in \mathbb{R}^{n+1}$ will be denoted by

$$z = (x, t) \in \mathbb{R}^n \times \mathbb{R}.$$

We write parameters i, j, l, k and m to mean nonnegative integers. We denote a time interval as

$$\Lambda_r(t_0) := (t_0 - r^{2s}, t_0 + r^{2s}).$$

The parabolic cylinder is defined by

$$Q_r(z_0) = B_r(x_0) \times \Lambda_r(t_0),$$

where $B_r(x_0)$ denotes the ball in $\mathbb{R}^n$ centered at x_0 with radius r. We write

$$\mathcal{B}_r(x_0, y_0) = B_r(x_0) \times B_r(y_0), \quad \mathcal{B}_r(x_0) = B_r(x_0) \times B_r(x_0)$$

and

$$\mathcal{Q}_r(x_0, y_0, t_0) = \mathcal{B}_r(x_0, y_0) \times \Lambda_r(t_0), \quad \mathcal{Q}_r(x_0, t_0) = \mathcal{B}_r(x_0) \times \Lambda_r(t_0)$$

for any $x_0, y_0 \in \mathbb{R}^n$, $t_0 \in \mathbb{R}$ and $r > 0$. For a given measurable function $h : \Omega_T \to \mathbb{R}$, we write for any $Q_r(z_0) \subset \Omega_T$,

$$(h)_{B_r(x_0)}(t) = \fint_{B_r(x_0)} h(x, t)\, dx \quad \text{and} \quad (h)_{Q_r(x_0)} = \fint_{Q_r(z_0)} h(z)\, dz.$$

We denote the parabolic Sobolev conjugate of $p \in [1, \infty)$ by

$$p_{\#} = p\left(1 + \frac{2s}{n}\right). \tag{2.1}$$

We are going to mention some lemmas starting with the following embedding result.

Lemma 2.1. *(see [20, Lemma 2.3]) Let $p \in [1, \infty)$ and $h \in L^p(0, T; W^{s,p}(B_r)) \cap L^\infty(0, T; L^2(B_r))$. Then there is a constant $c = c(n, s, p)$ such that*

$$\int_0^T \fint_{B_r} |h|^{p_{\#}}\, dz \le c\left(r^{sp} \int_0^T \int_{B_r} \fint_{B_r} \frac{|h(x, t) - h(y, t)|^p}{|x - y|^{n+sp}}\, dx\, dy\, dt + \int_0^T \fint_{B_r} |h|^p\, dz\right) \times \left(\sup_{t \in (0,T)} \fint_{B_r} |h|^2\, dx\right)^{\frac{sp}{n}}.$$

In particular, we have that

$$\begin{aligned} \int_0^T \fint_{B_r} |h - (h)_{B_r}(t)|^{p_{\#}}\, dz \le c &\left(r^{sp} \int_0^T \int_{B_r} \fint_{B_r} \frac{|h(x, t) - h(y, t)|^p}{|x - y|^{n+sp}}\, dx\, dy\, dt\right) \\ &\times \left(\sup_{t \in (0,T)} \fint_{B_r} |h - (h)_{B_r}(t)|^2\, dx\right)^{\frac{sp}{n}}. \end{aligned}$$

Next, we list some properties of the measure $\mu_{\tau,t}$.

Lemma 2.2. *There exists a constant C_n depending only on n such that*

(1) *For any $x_0 \in \mathbb{R}^n$, $t_0 \in \mathbb{R}$ and $R > 0$, the following holds*

$$\mu_{\tau,t}\left(\mathcal{Q}_R(x_0, t_0)\right) = C_n \frac{R^{n+2s+2\tau}}{\tau}. \tag{2.2}$$

(2) *Let ρ and R be any positive numbers, and let $x_0 \in \mathbb{R}^n$. Then*

$$\frac{\mu_{\tau,t}\left(\mathcal{Q}_R(x_0, t_0)\right)}{\mu_{\tau,t}\left(\mathcal{Q}_\rho(x_0, t_0)\right)} = \left(\frac{R}{\rho}\right)^{n+2s+2\tau}. \tag{2.3}$$

(3) *Let $\mathcal{K}_r(x_0, y_0)$ be any cube in $\mathcal{B}_R$ for $r, R > 0$ and $x_0, y_0 \in \mathbb{R}^n$. Then*

$$\frac{\mu_{\tau,t}\left(\mathcal{Q}_R\right)}{\mu_{\tau,t}\left(\mathcal{K}_r(x_0, y_0) \times \Lambda_r\right)} \le 2^n \frac{C_n}{\tau}\left(\frac{R}{r}\right)^{2n+2s}. \tag{2.4}$$

(4) *Let $a \ge 1$. Then we have*

$$\mu_{\tau,t}\left(\mathcal{Q}_{aR}(x_0, y_0, t_0)\right) \le c a^{2n+2s} \tau^{-1} \mu_{\tau,t}\left(\mathcal{Q}_R(x_0, y_0, t_0)\right) \tag{2.5}$$

for some constant $c = c(n)$.

Proof. Since the proofs of (1)-(3) follow from [13, Lemma 2.1], we only give the proof of (4). If $x_0 = y_0$, (4) follows from (2). We assume $x_0 \neq y_0$. Let

$$D = \operatorname{diam}(B_{aR}(x_0), B_{aR}(y_0)).$$

If $D \leq 4a$, then $\mathcal{Q}_{aR}(x_0, y_0, t_0) \subset \mathcal{Q}_{10aR}(x_0, t_0)$. Thus by (3), we get

$$\mu_{\tau,t}\left(\mathcal{Q}_{aR}(x_0, y_0, t_0)\right) \leq \mu_{\tau,t}\left(\mathcal{Q}_{10aR}(x_0, t_0)\right) \leq \frac{ca^{2n+2s}}{\tau}\mu_{\tau,t}\left(\mathcal{Q}_{R}(x_0, y_0, t_0)\right)$$

for some constant $c = c(n)$. If $D > 4a$, we observe

$$D \leq |x-y| \leq 3D \quad \text{for any } x \in B_{aR}(x_0) \text{ and } y \in B_{aR}(y_0).$$

Thus we have

$$\mu_{\tau,t}\left(\mathcal{Q}_{aR}(x_0, y_0, t_0)\right) \leq \frac{(aR)^{2n+2s}}{D^{n-2\tau}} \leq ca^{2n+2s}\mu_{\tau,t}\left(\mathcal{Q}_{R}(x_0, y_0, t_0)\right)$$

for some constant $c = c(n)$. This completes the proof. □

We now give some useful estimates to control the parabolic tail.

Lemma 2.3. *Let $h \in L^p\left(\Lambda_2; L^1_{2\beta}(\mathbb{R}^n)\right)$ and $D^\tau d_{\tilde{s}} h \in L^p\left(\mathcal{Q}_2; d\mu_{\tau,t}\right)$ where $\beta \in (0,1)$, $\tilde{s} \in [0,\beta]$ and $p \in [1,\infty)$. Let $Q_\rho(z_0) \Subset Q_2$, where $z_0 \in Q_{r_1}$ with $0 < r_1 < 2$. Suppose that there is a natural number $l \geq 1$ such that*

$$Q_{2^l\rho}(z_0) \Subset Q_2.$$

Then for any integer $k \in [0,l]$, there are constants $c_p = c(n,p)$ and $\tilde{c} = \tilde{c}(n)$ independent of k and l such that

$$\begin{aligned}
&\operatorname{Tail}_{p,2\beta}\left(\frac{h-(h)_{B_\rho(x_0)}(t)}{\rho^{\tilde{s}+\tau}}; Q_\rho(z_0)\right) \\
&\leq c_p \mathbb{A}_k \sum_{i=1}^{k} 2^{i\left(-2\beta+\tilde{s}+\tau+\frac{2s}{p}\right)}\left(\frac{1}{\tau}\fint_{\mathcal{Q}_{2^i\rho}(z_0)} |D^\tau d_{\tilde{s}} h|^p \, d\mu_{\tau,t}\right)^{\frac{1}{p}} + c_p \mathbb{A}_{k+1-l}\frac{2^{-2\beta l+\tilde{s}+\tau+\frac{2s}{p}}}{\rho^{\tilde{s}+\tau+\frac{2s}{p}}}\left(\frac{1}{\tau}\fint_{\mathcal{Q}_2} |D^\tau d_{\tilde{s}} h|^p \, d\mu_{\tau,t}\right)^{\frac{1}{p}} \\
&\quad + \tilde{c}\mathbb{A}_{l-k}\sum_{j=k+1}^{l} 2^{j(-2\beta+\tilde{s}+\tau)} \sup_{t\in\Lambda_{2^j\rho}(t_0)} \fint_{B_{2^j\rho}(x_0)} \frac{|h-(h)_{B_{2^j\rho}(x_0)}(t)|}{(2^j\rho)^{\tilde{s}+\tau}}\,dx \\
&\quad + \tilde{c}\mathbb{A}_{l-k}\frac{2^{-2\beta l+\tilde{s}+\tau}}{\rho^{\tilde{s}+\tau}} \sup_{t\in\Lambda_2} \fint_{B_2} \frac{|h-(h)_{B_2}(t)|}{2^{s+\tau}}\,dx + \frac{\tilde{c}}{(2-r_1)^{n+2s}}\left(\frac{2}{\rho}\right)^{-2\beta+\tilde{s}+\tau-\frac{2s}{p}} \operatorname{Tail}_{p,2\beta}\left(\frac{h-(h)_{B_2}(t)}{2^{\tilde{s}+\tau}}; Q_2\right),
\end{aligned} \tag{2.6}$$

where

$$\mathbb{A}_m = \begin{cases} 1 & \text{if } m = 1,2,\dots, \\ 0 & \text{if } m = 0,-1,\dots. \end{cases}$$

Proof. Using Minkowski's inequality, we get that

$$\begin{aligned}
&\operatorname{Tail}_{p,2\beta}\left(\frac{h-(h)_{B_\rho(x_0)}(t)}{\rho^{\tilde{s}+\tau}}; Q_\rho(z_0)\right) \\
&\leq \sum_{i=1}^{l}\left(\fint_{\Lambda_\rho(t_0)}\left(\int_{B_{2^i\rho}(x_0)\setminus B_{2^{i-1}\rho}(x_0)} \rho^{2\beta-\tilde{s}-\tau}\frac{|h-(h)_{B_\rho(x_0)}(t)|}{|y-x_0|^{n+2\beta}}\,dy\right)^p dt\right)^{\frac{1}{p}} \\
&\quad + \left(\fint_{\Lambda_\rho(t_0)}\left(\int_{B_2\setminus B_{2^l\rho}(x_0)} \rho^{2\beta-\tilde{s}-\tau}\frac{|h-(h)_{B_\rho(x_0)}(t)|}{|y-x_0|^{n+2\beta}}\,dy\right)^p dt\right)^{\frac{1}{p}} \\
&\quad + \left(\fint_{\Lambda_\rho(t_0)}\left(\int_{\mathbb{R}^n\setminus B_2} \rho^{2\beta-\tilde{s}-\tau}\frac{|h-(h)_{B_\rho(x_0)}(t)|}{|y-x_0|^{n+2\beta}}\,dy\right)^p dt\right)^{\frac{1}{p}} =: \sum_{i=1}^{l} I_i + \mathcal{T}_1 + \mathcal{T}_2.
\end{aligned}$$

From the estimate of $T_k^{\frac{1}{p-1}}$ in [13, Lemma 2.6], we have

$$I_i \leq \frac{2^{-2\beta i}}{\rho^{\tilde{s}+\tau}}\sum_{j=1}^{i}\left(\fint_{\Lambda_\rho(t_0)}\left(\fint_{B_{2^j\rho}(x_0)} |h-(h)_{B_{2^j\rho}(x_0)}(t)|\,dy\right)^p dt\right)^{\frac{1}{p}} =: 2^{-2\beta i}\sum_{j=1}^{i} I_{i,j}.$$

Using Hölder's inequality and then following the same line as for (2.12) in [13, Lemma 2.6], we obtain

$$\sum_{j=1}^{i} I_{i,j} \le c_p \sum_{j=1}^{k} \mathbb{A}_k 2^{j\left(\tilde{s}+\tau+\frac{2s}{p}\right)} \left(\frac{1}{\tau} \fint_{\mathcal{Q}_{2^j\rho}(z_0)} |D^\tau d_{\tilde{s}} h|^p \, d\mu_{\tau,t}\right)^{\frac{1}{p}} + \sum_{j=k+1}^{i} \mathbb{A}_{i-k} 2^{j(\tilde{s}+\tau)} \sup_{t\in\Lambda_\rho(t_0)} \fint_{B_{2^j\rho}(x_0)} \frac{|h-(h)_{B_{2^j\rho}(x_0)}(t)|}{(2^j\rho)^{\tilde{s}+\tau}} \, dx,$$

for some constant $c_p = 2^{n+2p}$, where we have taken supremum for the time variable if $j \ge k+1$. Similarly, we estimate $\mathcal{T}_1$ as

$$\mathcal{T}_1 \le 2^{-2\beta l} \sum_{i=1}^{l} I_{l,i} + c_p \mathbb{A}_{k+1-l} 2^{-2\beta l} \left(\frac{2}{\rho}\right)^{\tilde{s}+\tau+\frac{2s}{p}} \left(\frac{1}{\tau} \fint_{\mathcal{Q}_2} |D^\tau d_{\tilde{s}} h|^p \, d\mu_{\tau,t}\right)^{\frac{1}{p}} + c\mathbb{A}_{l-k} 2^{-2\beta l} \left(\frac{2}{\rho}\right)^{\tilde{s}+\tau} \sup_{t\in\Lambda_2} \fint_{B_2} \frac{|h-(h)_{B_2}(t)|}{2^{\tilde{s}+\tau}} \, dx =: \mathcal{T}_{1,1},$$

where $c = c(n)$. Lastly, we estimate $\mathcal{T}_2$ as

$$\mathcal{T}_2 \le c\mathcal{T}_{1,1} + c\left(\fint_{\Lambda_\rho} \left(\int_{\mathbb{R}^n\setminus B_2} \left(\frac{2}{\rho}\right)^{-2\beta+\tilde{s}+\tau} \left(\frac{2}{2-r_1}\right)^{n+2\beta} 2^{-2\beta+\tilde{s}+\tau} \frac{|h-(h)_{B_2}(t)|}{|y|^{n+2\beta}} \, dy\right)^p dt\right)^{\frac{1}{p}}, \tag{2.7}$$

where for the last term we have used the fact that

$$|y-x_0| \ge |y| - |x_0| \ge \frac{|y|(2-r_1)}{2} \quad \text{for any } y \in \mathbb{R}^n \setminus B_2.$$

We combine all the estimates I_i, $\mathcal{T}_1$ and $\mathcal{T}_2$ and use Fubini's theorem as in [13, Lemma 2.6] to get the desired result (2.6). □

Remark 4. By tracking the choice of the constant c_p appearing in Lemma 2.3, we find that $c_p \le c_q$ if $p \le q$.

We end this section with the following iteration lemma.

Lemma 2.4. *(See [28, Lemma 6.1]) Let $\phi : [1,2] \to \mathbb{R}$ be a nonnegative bounded function. For $1 \le r_1 < r_2 \le 2$, we assume that*

$$\phi(r_1) \le \frac{1}{2}\phi(r_2) + \frac{\lambda_0}{(r_2-r_1)^{\frac{5n}{s}}},$$

where $\lambda_0 > 0$. Then, there is a constant $c = c(n,s)$ such that

$$\phi(1) \le c\lambda_0.$$

3. Energy estimates and the Sobolev-Poincaré inequalities

In this section, we give energy estimates and derive Sobolev-Poincaré type inequalities from the energy estimates. We first give an energy inequality of a weak solution u to (1.1).

Lemma 3.1. *Let u be a local weak solution to (1.1). Let $0 < \rho < r \le 2\rho$ with $Q_{2\rho}(z_0) \Subset \Omega_T$. Then, there is a constant $c = c(n,s,L)$ such that*

$$\begin{aligned}
&\left[\fint_{\Lambda_\rho(t_0)} \fint_{B_\rho(x_0)} \int_{B_\rho(x_0)} \frac{|u(x,t)-u(y,t)|^2}{|x-y|^{n+2s}} \, dx\, dy\, dt + \sup_{t\in\Lambda_\rho(t_0)} \fint_{B_\rho(x_0)} \frac{|u(x,t)-k|^2}{\rho^{2s}} \, dx\right] \\
&\le c\frac{r^{n+2-2s}}{(r-\rho)^{n+2}} \iint_{Q_r(z_0)} |u-k|^2 \, dz + c\left(\frac{r}{r-\rho}\right)^{2(n+2s)} \fint_{\Lambda_r(t_0)} \fint_{B_r(x_0)} \int_{B_r(x_0)} \frac{|f(x,t)-f(y,t)|^2}{|x-y|^n} \, dx\, dy\, dt \\
&\quad + c\left(\fint_{Q_r(z_0)} (r^s|g|)^\gamma \, dz\right)^{\frac{2}{\gamma}} + c\left(\frac{r}{r-\rho}\right)^{2(n+2s)} \mathrm{Tail}_{\gamma,2s}\left(\frac{u-(u)_{B_r(x_0)}(t)}{r^s}; Q_r(z_0)\right)^2 \\
&\quad + c\left(\frac{r}{r-\rho}\right)^{2(n+2s)} \mathrm{Tail}_{2,s}\left(f-(f)_{B_r(x_0)}(t); Q_r(z_0)\right)^2,
\end{aligned} \tag{3.1}$$

where $k \in \mathbb{R}$ and

$$\gamma = \frac{2(n+2s)}{n+4s}. \tag{3.2}$$

Proof. Since $u-k$ is also a weak solution to (1.1), we may assume that $k=0$. Let us take a cutoff function $\psi \in C_c^\infty\left(B_{\frac{\rho+r}{2}}\right)$ such that

$$\psi \equiv 1 \text{ on } B_\rho(x_0) \quad \text{and} \quad |D\psi| \le \frac{16}{r-\rho},$$

and a smooth function $\eta \in C^\infty(\mathbb{R})$ such that

$$\eta \equiv 1 \text{ on } \left(t_0 - \left(\frac{r^{2s}+\rho^{2s}}{2}\right), \infty\right), \quad \eta \equiv 0 \text{ on } \left(-\infty, t_0 - \left(\frac{3r^{2s}+\rho^{2s}}{4}\right)\right] \quad \text{and} \quad |\eta'| \le \frac{16}{r^{2s}-\rho^{2s}}.$$

With the aid of [11, Lemma 3.2], we deduce that

$$\begin{aligned}
I :&= \fint_{\Lambda_r(t_0)} \fint_{B_r(x_0)} \int_{B_r(x_0)} \frac{|(u\psi\eta)(x,t) - (u\psi\eta)(y,t)|^2}{|x-y|^{n+2s}} \, dx\, dy\, dt + \sup_{t\in\Lambda_r} \fint_{B_r(x_0)} \frac{|(u\psi\eta)(x,t)|^2}{r^{2s}} \, dx \\
&\le c\frac{r^{2(1-s)}}{(r-\rho)^2} \fint\!\!\fint_{Q_r(z_0)} |u|^2 \, dz + c\frac{1}{r^{2s}-\rho^{2s}} \fint\!\!\fint_{Q_r(z_0)} |u|^2 \, dz + c\fint\!\!\fint_{Q_r(z_0)} |g||u\psi\eta| \, dz \\
&\quad + c\left(\frac{r}{r-\rho}\right)^{n+2s} \fint_{\Lambda_r(t_0)} \int_{\mathbb{R}^n\setminus B_r(x_0)} \frac{|u(y,t)|}{|y-x_0|^{n+2s}} \, dy \fint_{B_r(x_0)} |(u\psi\eta)(x,t)| \, dx\, dt \\
&\quad + c\fint_{\Lambda_r(t_0)} \fint_{B_r(x_0)} \int_{B_r(x_0)} \frac{(f(x,t)-f(y,t))\left((u\psi\eta)(x,t) - (u\psi\eta)(y,t)\right)}{|x-y|^{n+s}} \, dx\, dy\, dt \\
&\quad + c\left(\frac{r}{r-\rho}\right)^{n+s} \fint_{\Lambda_r(t_0)} \int_{\mathbb{R}^n\setminus B_r(x_0)} \frac{|f(x,t)-f(y,t)|}{|y-x_0|^{n+s}} \, dy \fint_{B_r(x_0)} |(u\psi\eta)(x,t)| \, dx\, dt =: \sum_{i=1}^6 I_i.
\end{aligned} \tag{3.3}$$

After a few simple algebraic computations along with the fact that $\rho < r \le 2\rho$, we observe that $I_2 \le cI_1$. Using Hölder's inequality, Lemma 2.1 and Young's inequality, we have

$$I_3 \le \left(\fint\!\!\fint_{Q_r(z_0)} (r^s|g|)^\gamma\right)^{\frac{1}{\gamma}} \left(\fint\!\!\fint_{Q_r(z_0)} \left|\frac{u\psi\eta}{r^s}\right|^{2_\#}\right)^{\frac{1}{2_\#}} \le c\left(\fint\!\!\fint_{Q_r(z_0)} (r^s|g|)^\gamma\right)^{\frac{2}{\gamma}} + \frac{I}{8} + r^{-2s}\fint\!\!\fint_{Q_r(z_0)} |u|^2 \, dx\, dt$$

and

$$\begin{aligned}
I_4 &\le c\left(\frac{r}{r-\rho}\right)^{(n+2s)} \mathrm{Tail}_{\gamma,2s}\left(\frac{u-(u)_{B_r(x_0)}(t)}{r^s}; Q_r(z_0)\right) \left(\fint\!\!\fint_{Q_r(z_0)} \left|\frac{u\psi\eta}{r^s}\right|^{2_\#}\right)^{\frac{1}{2_\#}} + c\left(\frac{r}{r-\rho}\right)^n I_1 \\
&\le c\left(\frac{r}{r-\rho}\right)^{2(n+2s)} \mathrm{Tail}_{\gamma,2s}\left(\frac{u-(u)_{B_r(x_0)}(t)}{r^s}; Q_r(z_0)\right)^2 + \frac{I}{8} + r^{-2s}\fint\!\!\fint_{Q_r(z_0)} |u|^2 \, dx\, dt + c\left(\frac{r}{r-\rho}\right)^n I_1,
\end{aligned}$$

where the constant $2_\#$ is defined in (2.1). On the other hand, we obtain

$$I_5 \le c\fint_{\Lambda_r(t_0)} \fint_{B_r(x_0)} \int_{B_r(x_0)} \frac{|f(x,t)-f(y,t)|^2}{|x-y|^n} \, dx\, dy\, dt + \frac{I}{8}$$

by Hölder's inequality and Young's inequality. For the last two terms, we first observe that

$$\begin{aligned}
I_5 + I_6 &\le c\left(\frac{r}{r-\rho}\right)^{n+s} \fint_{\Lambda_r(t_0)} \int_{\mathbb{R}^n\setminus B_r(x_0)} \fint_{B_r(x_0)} \frac{|f(x,t)-(f)_{B_r(x_0)}(t)|}{|y-x_0|^{n+s}} |(u\psi\eta)(x,t)| \, dx\, dy\, dt \\
&\quad + c\left(\frac{r}{r-\rho}\right)^{n+s} \fint_{\Lambda_r(t_0)} \int_{\mathbb{R}^n\setminus B_r(x_0)} \frac{|f(y,t)-(f)_{B_r(x_0)}(t)|}{|y-x_0|^{n+s}} \, dy \fint_{B_r(x_0)} |(u\psi\eta)(x,t)| \, dx\, dt.
\end{aligned}$$

By following the same line as in the estimate of $I_{2,2}$ in [13, Lemma 3.3], we estimate I_5 as

$$\begin{aligned}
I_5 + I_6 &\le \frac{I}{8} + c\left(\frac{r}{r-\rho}\right)^{2(n+2s)} \fint_{\Lambda_r(t_0)} \fint_{B_r(x_0)} \int_{B_r(x_0)} \frac{|f(x,t)-f(y,t)|^2}{|x-y|^n} \, dx\, dy\, dt \\
&\quad + c\,\mathrm{Tail}_{2,s}\left(f-(f)_{B_r(x_0)}(t); Q_r(z_0)\right)^2.
\end{aligned}$$

Using the definitions of ψ and η, we estimate I as

$$I \ge \left[\fint_{\Lambda_\rho(t_0)} \fint_{B_\rho(x_0)} \int_{B_\rho(x_0)} \frac{|u(x,t)-u(y,t)|^2}{|x-y|^{n+2s}} \, dx\, dy\, dt + \sup_{t\in\Lambda_\rho(t_0)} \fint_{B_\rho(x_0)} \frac{|u(x,t)|^2}{\rho^{2s}} \, dx\right].$$

We combine the estimates I and I_i for each $i = 1, 2, \ldots, 6$ along with (3.3) to obtain (3.1). □

Next we give a gluing lemma.

Lemma 3.2. *(See [11, Lemma 4.5].) Let $Q_\rho(z_0) \Subset \Omega_T$. Take $\psi \in C_c^\infty\left(B_{\frac{3\rho}{4}}(x_0)\right)$ with $\psi \equiv 1$ in $B_{\frac{\rho}{2}}(x_0)$. Then we have that*

$$\begin{aligned}
&\sup_{t_1,t_2\in\Lambda_\rho(t_0)} \left|(u)^\psi_{B_\rho}(t_1) - (u)^\psi_{B_\rho}(t_2)\right| \\
&\leq c\rho^{2s-1}\fint_{Q_\rho(z_0)}\int_{B_\rho(x_0)} \frac{|u(x,t)-u(y,t)|}{|x-y|^{n+2s-1}}\,dy\,dz + c\rho^{2s}\fint_{Q_\rho(z_0)}\int_{\mathbb{R}^n\setminus B_\rho(x_0)} \frac{|u(x,t)-u(y,t)|}{|y-x_0|^{n+2s}}\,dy\,dz \\
&\quad + c\rho^{2s-1}\fint_{Q_\rho(z_0)}\int_{B_\rho(x_0)} \frac{|f(x,t)-f(y,t)|}{|x-y|^{n+s-1}}\,dy\,dz + c\rho^{2s}\fint_{Q_\rho(z_0)}\int_{\mathbb{R}^n\setminus B_\rho(x_0)} \frac{|f(x,t)-f(y,t)|}{|y-x_0|^{n+s}}\,dy\,dz \\
&\quad + c\rho^{2s}\fint_{Q_\rho(z_0)} |g|\,dz
\end{aligned}$$

for some constant $c = c(n,s)$, where $(u)^\psi_{B_\rho}(t) = \frac{1}{\|\psi\|_{L^1}}\int_{B_\rho} u(x,t)\psi(x)\,dx$.

We now show that L^2-oscillation integral by the sup-norm is estimated by the sum of L^2-integral of $d_s u$ and $d_0 f$, L^γ-integral of g, and tail terms of u and f. Before giving the estimate, we define a function $G : \Omega \times \Omega_T \to \mathbb{R}$ by

$$G(x,y,t) = g(x,t). \tag{3.4}$$

We directly deduce that

$$\int_{Q_r} |g|^p\,dz \approx_{n,s,p} \frac{1}{r^{2\tau}}\int_{\mathcal{Q}_r} |G|^p\,d\mu_{\tau,t}. \tag{3.5}$$

Lemma 3.3. *Let u be a weak solution to (1.1) and let $Q_{2\rho}(z_0) \Subset \Omega_T$. Then we have*

$$\begin{aligned}
\sup_{t\in\Lambda_\rho(t_0)}\fint_{B_\rho(x_0)} \frac{|u-(u)_{Q_\rho(z_0)}|^2}{\rho^{2s+2\tau}}\,dx \leq{}& \frac{c_0}{\tau}\fint_{\mathcal{Q}_{2\rho}(z_0)} |D^\tau d_s u|^2\,d\mu_{\tau,t} + c_0\,\mathrm{Tail}_{2,2s}\left(\frac{u-(u)_{B_{2\rho}(x_0)}(t)}{(2\rho)^{s+\tau}}; Q_{2\rho}(z_0)\right)^2 \\
&+ \frac{c_0}{\tau}\fint_{\mathcal{Q}_{2\rho}(z_0)} |D^\tau d_0 f|^2\,d\mu_{\tau,t} + c_0\,\mathrm{Tail}_{2,s}\left(\frac{f-(f)_{B_{2\rho}(x_0)}(t)}{(2\rho)^{\tau}}; Q_{2\rho}(z_0)\right)^2 \\
&+ c_0\left(\frac{1}{\tau}\fint_{\mathcal{Q}_{2\rho}(z_0)} \left((2\rho)^{s-\tau}|G|\right)^\gamma\,d\mu_{\tau,t}\right)^{\frac{2}{\gamma}},
\end{aligned} \tag{3.6}$$

where $c_0 = c_0(n,s,L)$ is a constant.

Proof. We may assume that $z_0 = 0$. Using (3.1) with $r = 2\rho$ and $k = (u)_{Q_\rho}$, we have

$$\begin{aligned}
\sup_{t\in\Lambda_\rho}\fint_{B_\rho} \frac{|u(x,t)-(u)_{Q_\rho}|^2}{\rho^{2s+2\tau}}\,dx \leq{}& c\fint\!\!\fint_{Q_{2\rho}} \frac{|u-(u)_{Q_{2\rho}}|^2}{(2\rho)^{2s+2\tau}}\,dz + \frac{c}{\tau}\fint_{\mathcal{Q}_{2\rho}} |D^\tau d_0 f|^2\,d\mu_{\tau,t} + c\left(\fint_{Q_{2\rho}} \left((2\rho)^{s-\tau}|g|\right)^\gamma\,dz\right)^{\frac{2}{\gamma}} \\
&+ c\left[\mathrm{Tail}_{\gamma,2s}\left(\frac{u-(u)_{B_{2\rho}}(t)}{(2\rho)^{s+\tau}}; Q_{2\rho}\right)^2 + \mathrm{Tail}_{2,s}\left(\frac{f-(f)_{B_{2\rho}}(t)}{(2\rho)^{\tau}}; Q_{2\rho}\right)^2\right].
\end{aligned}$$

Applying (A.3) and (2.2) to the first term and the third term in the right-hand side of the above inequality, respectively, we obtain the desired estimate (3.6). □

4. Comparison estimates

This section is devoted to establishing comparison estimates. We now assume

$$\tau \in (0, \min\{s, 1-s\}). \tag{4.1}$$

Before proving comparison estimates, we first give two lemmas. The first one is a self-improving property for weak solutions to the corresponding homogeneous problem of (1.1).

Lemma 4.1. *Let $w \in L^2\left(\Lambda_3; W^{s,2}(B_3)\right) \cap L^\infty\left(\Lambda_3; L^1_{2s}(\mathbb{R}^n)\right)$ be a weak solution to*

$$w_t + \mathcal{L}^\Phi_A w = 0 \quad \text{in } Q_3.$$

Then there are constants $\epsilon_0 = \epsilon_0(n,s,L) \in (0,1)$ and $c = c(n,s,L)$ such that

$$\begin{aligned}
\left(\frac{1}{\tau}\fint_{\mathcal{Q}_2} |D^\tau d_s w|^{2(1+\epsilon_0)}\,d\mu_{\tau,t}\right)^{\frac{1}{2(1+\epsilon_0)}} \leq{}& c\left(\frac{1}{\tau}\fint_{\mathcal{Q}_3} |D^\tau d_s w|^2\,d\mu_{\tau,t}\right)^{\frac{1}{2}} + \mathrm{Tail}_{\infty,2s}\left(\frac{w-(w)_{B_3}(t)}{3^{s+\tau}}; Q_3\right) \\
&+ c\left(\sup_{t\in\Lambda_3}\fint_{B_3} \frac{|w-(w)_{B_3}(t)|^2}{3^{2(s+\tau)}}\,dx\right)^{\frac{1}{2}}.
\end{aligned}$$

Proof. By Theorem A.1 below, we have

$$\left(\fint_{Q_r}\int_{B_r}|r^{\epsilon_1}d_{s+\epsilon_1}w|^{2+\epsilon_1}\frac{dx\,dz}{|x-y|^n}\right)^{\frac{1}{2+\epsilon_1}}\leq c\left(\fint_{Q_{2r}}\int_{B_{2r}}|d_sw|^2\frac{dx\,dz}{|x-y|^n}\right)+c\,\mathrm{Tail}_{\infty,2s}\left(\frac{w-(w)_{B_{2r}}(t)}{(2r)^s};Q_{2r}\right)$$
$$+c\left(\sup_{t\in\Lambda_{2r}}\fint_{B_{2r}}\frac{|w-(w)_{B_{2r}}(t)|^2}{(2r)^{2s}}\,dx\right)^{\frac{1}{2}}$$

for some constant $\epsilon_1=\epsilon_1(n,s,L)\in(0,1)$, where $Q_{2r}\subset Q_3$. By taking $\epsilon_0=\frac{\epsilon_1}{2}$ and using (2.2), we get

$$\left(\frac{1}{\tau}\fint_{\mathcal{Q}_r}|D^\tau d_sw|^{2(1+\epsilon_0)}\,d\mu_{\tau,t}\right)^{\frac{1}{2(1+\epsilon_0)}}\leq c\left(\frac{1}{\tau}\fint_{\mathcal{Q}_{2r}}|D^\tau d_sw|^2\,d\mu_{\tau,t}\right)^{\frac{1}{2}}+\mathrm{Tail}_{\infty,2s}\left(\frac{w-(w)_{B_{2r}}(t)}{(2r)^{s+\tau}};Q_{2r}\right)$$
$$+c\left(\sup_{t\in\Lambda_{2r}}\fint_{B_{2r}}\frac{|w-(w)_{B_{2r}}(t)|^2}{(2r)^{2(s+\tau)}}\,dx\right)^{\frac{1}{2}}.$$

The standard covering argument along with Lemma 5.2 gives the desired result (see [13, Lemma 3.1] for more details). □

The second one is a higher Hölder regularity of weak solutions to fractional parabolic equations with locally constant coefficients with respect to the spatial variables.

Lemma 4.2. *(See [11, Theorem 1.2]) Let* $v\in L^2\left(\Lambda_2;W^{s,2}(B_2)\right)\cap L^\infty\left(\Lambda_2;L^1_{2s}\left(\mathbb{R}^n\right)\right)$ *be a weak solution to*

$$v_t+\mathcal{L}^\Phi_{A_2(t)}v=0\quad\text{in } Q_2,$$

where we denote $A_2(t)=A_{2,0}(t)$ *which is defined in* (1.8)*. Then for any* $\alpha\in(0,\min\{2s,1\})$*, there is a constant* $c=c(n,s,L,\alpha)$ *such that*

$$[v]_{C^{\alpha,\frac{\alpha}{2s}}(Q_1)}\leq c\left(\frac{1}{\tau}\fint_{\mathcal{Q}_2}|D^\tau d_sv|^2\,d\mu_{\tau,t}\right)^{\frac{1}{2}}+c\,\mathrm{Tail}_{\infty,2s}\left(v-(v)_{B_2}(t);Q_2\right).$$

We are now in position to prove the following comparison lemma.

Lemma 4.3. *Let* $Q_4\Subset\Omega_T$*. For any* $\epsilon>0$*, there is a constant* $\delta=\delta(n,s,L,\epsilon)$ *such that for any weak solution* u *to* (1.1) *satisfying*

$$\frac{1}{\tau}\fint_{\mathcal{Q}_4}|D^\tau d_su|^2\,d\mu_{\tau,t}+\mathrm{Tail}_{\infty,2s}\left(\frac{u-(u)_{B_4}(t)}{4^{s+\tau}};Q_4\right)^2\leq 1 \tag{4.2}$$

and

$$\left(\frac{1}{\tau}\fint_{\mathcal{Q}_4}\left(4^{s-\tau}|G|\right)^\gamma\,d\mu_{\tau,t}\right)^{\frac{2}{\gamma}}+\frac{1}{\tau}\fint_{\mathcal{Q}_4}|D^\tau d_0f|^2\,d\mu_{\tau,t}+\mathrm{Tail}_{2,s}\left(\frac{f-(f)_{B_4}(t)}{4^\tau};Q_4\right)^2$$
$$+\left(\fint_{\mathcal{Q}_2}|A-(A)_2(t)|\,dx\,dy\,dt\right)^2\leq\delta^2, \tag{4.3}$$

there is a solution v *to*

$$v_t+\mathcal{L}^\Phi_{A_2(t)}v=0\quad\text{in } Q_2$$

such that

$$\frac{1}{\tau}\fint_{\mathcal{Q}_1}|D^\tau d_s(u-v)|^2\,d\mu_{\tau,t}\leq\epsilon^2\quad\text{and}\quad\|D^\tau d_sv\|_{L^\infty(\mathcal{Q}_1)}\leq c, \tag{4.4}$$

where $c=c(n,s,L,\tau)$*.*

Proof. The proof is divided into several steps for the ease of readability.
Step 1: *First comparison estimates.* For a fixed weak solution u to problem (1.1), we consider the following problem:

$$\begin{cases} w_t+\mathcal{L}^\Phi_Aw &=0\quad\text{in } Q_3,\\ \qquad\quad w&=u\quad\text{in } (\mathbb{R}^n\setminus B_3)\times\Lambda_3\cup B_3\times\{-3^{2s}\}.\end{cases} \tag{4.5}$$

We intend to apply Lemma B.1 for the existence and uniqueness of w. To this end, it remains to show that $u_t\in L^2(\Lambda_3;W^{s,2}(B_4))^*$. Indeed, using the fact that u is a solution to problem (1.1), we find that for all $\phi\in L^2(\Lambda_3;W^{s,2}(B_4))\cap C^1_0(\Lambda_3;L^2(B_4))$, there holds

$$\left|\int_{\Lambda_3}\langle u,\phi_t\rangle dt\right|\leq c\int_{\Lambda_3}\|\phi(\cdot,t)\|_{W^{s,2}(B_4)}dt,$$

for some c depending only on n, s, L, u, f and g. Thus, we have the existence of $w \in L^2\left(\Lambda_3; W^{s,2}(B_4)\right) \cap C\left(\Lambda_3; L^2(B_3)\right) \cap L^\infty\left(\Lambda_3; L^1_{2s}(\mathbb{R}^n)\right)$ satisfying (4.5). Then $\varphi := u - w$ solves

$$\varphi_t + \mathcal{L}_A^\Phi u - \mathcal{L}_A^\Phi w = (-\Delta)^{\frac{s}{2}} f + g \quad \text{in } Q_3.$$

With the help of an approximation argument, we take φ as a test function to the above equation to see that for every $\tilde{T} \in (-3^{2s}, 3^{2s}]$, setting $\tilde{\Lambda}_3 := (-3^{2s}, \tilde{T}]$, there holds

$$\begin{aligned}
&\frac{1}{2}\int_{B_3} (\varphi(x,\tilde{T}))^2\,dx + \frac{1}{L}\int_{\tilde{\Lambda}_3}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |\varphi(x,t)-\varphi(y,t)|^2 \frac{A(x,y,t)}{|x-y|^{n+2s}}\,dx\,dy\,dt \\
&\quad \le \int_{\tilde{\Lambda}_3}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |(f(x,t)-f(y,t))(\varphi(x,t)-\varphi(y,t))| \frac{dx\,dy\,dt}{|x-y|^{n+s}} + \int_{\tilde{\Lambda}_3}\int_{B_3} |g\varphi|\,dz,
\end{aligned}$$

where we have used the fact that $\varphi(\cdot, -3^{2s}) = 0$ in B_3 and the first condition in (1.3). Noting the bounds on A, the above expression yields

$$\begin{aligned}
&\frac{1}{2}\int_{B_3} (\varphi(x,\tilde{T}))^2\,dx + \int_{\tilde{\Lambda}_3}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|\varphi(x,t)-\varphi(y,t)|^2}{|x-y|^{n+2s}}\,dx\,dy\,dt \\
&\quad \le L^2 \underbrace{\int_{\tilde{\Lambda}_3}\int_{B_4}\int_{B_4} |f(x,t)-f(y,t)|\,|\varphi(x,t)-\varphi(y,t)| \frac{dx\,dy\,dt}{|x-y|^{n+s}}}_{=:J_1} \\
&\qquad + 2L^2 \underbrace{\int_{\tilde{\Lambda}_3}\int_{B_3}\int_{\mathbb{R}^n\setminus B_4} |f(x,t)-f(y,t)|\,|\varphi(x,t)| \frac{dx\,dy\,dt}{|x-y|^{n+s}}}_{=:J_2} + L^2 \underbrace{\int_{\tilde{\Lambda}_3}\int_{B_3} |g\varphi|\,dz}_{=:J_3}. \qquad (4.6)
\end{aligned}$$

For J_1, applying Hölder's inequality and Young's inequality, we observe that

$$\begin{aligned}
J_1 &\le \left(\int_{\Lambda_3}\int_{B_4}\int_{B_4} |f(x,t)-f(y,t)|^2 \frac{dx\,dy\,dt}{|x-y|^n}\right)^{\frac{1}{2}} \left(\int_{\tilde{\Lambda}_3}\int_{B_4}\int_{B_4} |\varphi(x,t)-\varphi(y,t)|^2 \frac{dx\,dy\,dt}{|x-y|^{n+2s}}\right)^{\frac{1}{2}} \\
&\le \frac{c}{\tau}\fint_{\mathcal{Q}_4} |D^\tau d_0 f|^2 d\mu_{\tau,t} + \frac{1}{4L^2}\int_{\tilde{\Lambda}_3}\int_{B_4}\int_{B_4} |\varphi(x,t)-\varphi(y,t)|^2 \frac{dx\,dy\,dt}{|x-y|^{n+2s}},
\end{aligned}$$

where $c = c(n,s,L)$. On a similar account, we deduce that

$$\begin{aligned}
J_2 &\le \int_{\tilde{\Lambda}_3}\int_{B_3}\int_{\mathbb{R}^n\setminus B_4} \left(|f(x,t)-(f)_{B_4}(t)| + |f(y,t)-(f)_{B_4}(t)|\right)|\varphi(x,t)| \frac{dx\,dy\,dt}{|x-y|^{n+s}} \\
&\le c\left(\int_{\Lambda_4}\int_{B_4} |f(x,t)-(f)_{B_4}(t)|^2 dx\,dt\right)^{\frac{1}{2}} \left(\int_{\tilde{\Lambda}_3}\int_{B_3} \varphi(x,t)^2 dx\,dt\right)^{\frac{1}{2}} \\
&\quad + c\left(\int_{\tilde{\Lambda}_3}\int_{B_3} |\varphi|^2\,dz\right)^{\frac{1}{2}} \left(\int_{\Lambda_3}\left(\int_{\mathbb{R}^n\setminus B_4} \frac{|f(y,t)-(f)_{B_4}(t)|}{|y|^{n+s}} dy\right)^2 dt\right)^{\frac{1}{2}}.
\end{aligned}$$

Noticing the fact that $\varphi(\cdot,t) = 0$ in $\mathbb{R}^n \setminus B_3$, using the Sobolev-Poincaré inequality, we get

$$J_2 \le \frac{1}{8L^2}\int_{\tilde{\Lambda}_3}\int_{B_4}\int_{B_4} |\varphi(x,t)-\varphi(y,t)|^2 \frac{dx\,dy\,dt}{|x-y|^{n+2s}} + \frac{c}{\tau}\fint_{\mathcal{Q}_4} |D^\tau d_0 f|^2 d\mu_{\tau,t} + c\,\mathrm{Tail}_{2,s}\left(\frac{f-(f)_{B_4}(t)}{4^\tau}; Q_4\right)^2,$$

where $c = c(n,s,L)$. Now for J_3, using Hölder's inequality, Lemma 2.1 and Young's inequality, we get

$$J_3 \le c\left(\int_{Q_3} |g(x,t)|^\gamma dx\,dt\right)^{\frac{2}{\gamma}} + \frac{1}{4L^2}\int_{\tilde{\Lambda}_3}\int_{B_3}\int_{B_3} \frac{|\varphi(x,t)-\varphi(y,t)|^2}{|x-y|^{n+2s}}\,dx\,dy\,dt + \frac{1}{4L^2}\sup_{t\in\Lambda_3}\int_{B_3} |\varphi(x,t)|^2 dx.$$

Therefore, using the estimates of J_1, J_2 and J_3 in (4.6), taking supremum over $\tilde{T} \in (-3^{2s}, 3^{2s}]$ and recalling the definition of G from (3.4) along with (3.5), we obtain

$$\begin{aligned}
&\sup_{t\in\Lambda_3}\fint_{B_3} (u-w)^2(x,t)\,dx + \frac{1}{\tau}\fint_{\mathcal{Q}_3} |D^\tau d_s(u-w)|^2\,d\mu_{\tau,t} \\
&\quad \le \left(\frac{1}{\tau}\fint_{\mathcal{Q}_4} \left(4^{s-\tau}|G|\right)^\gamma\,d\mu_{\tau,t}\right)^{\frac{2}{\gamma}} + \frac{c}{\tau}\fint_{\mathcal{Q}_4} |D^\tau d_0 f|^2 d\mu_{\tau,t} + c\,\mathrm{Tail}_{2,s}\left(\frac{f-(f)_{B_4}(t)}{4^\tau}; Q_4\right)^2 \le c\,\delta^2, \qquad (4.7)
\end{aligned}$$

where we have also used (2.2) and (4.3).
Step 2: *Uniform self-improving inequality for w.* We first observe from Lemma 4.1 that

$$\left(\frac{1}{\tau}\fint_{\mathcal{Q}_2}|D^\tau d_s w|^{2(1+\epsilon_0)}\,d\mu_{\tau,t}\right)^{\frac{1}{2(1+\epsilon_0)}} \le c\left(\frac{1}{\tau}\fint_{\mathcal{Q}_3}|D^\tau d_s w|^2\,d\mu_{\tau,t}\right)^{\frac{1}{2}} + c\,\mathrm{Tail}_{\infty,2s}\left(\frac{w-(w)_{B_3}(t)}{3^{s+\tau}};Q_3\right)$$
$$+\,c\left(\sup_{t\in\Lambda_3}\fint_{B_3}\frac{|w-(w)_{B_3}(t)|^2}{3^{2(s+\tau)}}\,dx\right)^{\frac{1}{2}},$$

where $\epsilon_0=\epsilon_0(n,s,L)\in(0,1)$ and $c=c(n,s,L)$. Furthermore, by using (4.7), we have

$$\left(\frac{1}{\tau}\fint_{\mathcal{Q}_3}|D^\tau d_s w|^2\,d\mu_{\tau,t}\right)^{\frac{1}{2}} \le \left(\frac{1}{\tau}\fint_{\mathcal{Q}_3}|D^\tau d_s (u-w)|^2\,d\mu_{\tau,t}\right)^{\frac{1}{2}} + \left(\frac{1}{\tau}\fint_{\mathcal{Q}_3}|D^\tau d_s u|^2\,d\mu_{\tau,t}\right)^{\frac{1}{2}} \le c.$$

Using (3.6) with a slight modification, (4.7) and (4.2), we next have

$$\left(\sup_{t\in\Lambda_3}\fint_{B_3}\frac{|w-(w)_{B_3}(t)|^2}{3^{2(s+\tau)}}\,dx\right)^{\frac{1}{2}} \le c\left(\sup_{t\in\Lambda_3}\fint_{B_3}\frac{|u-(u)_{B_3}(t)|^2}{3^{2(s+\tau)}}\,dx\right)^{\frac{1}{2}} + c\left(\sup_{t\in\Lambda_3}\fint_{B_3}\frac{|u-w|^2}{3^{2(s+\tau)}}\,dx\right)^{\frac{1}{2}} \le c.$$

For the tail term, a simple computation together with (4.7) and (4.2) yields

$$\mathrm{Tail}_{\infty,2s}\left(\frac{w-(w)_{B_3}(t)}{3^{s+\tau}};Q_3\right) \le \sup_{t\in\Lambda_3}\fint_{B_3}\frac{|w-u|^2}{3^{2(\tau+s)}}\,dx + \mathrm{Tail}_{\infty,2s}\left(\frac{u-(u)_{B_3}(t)}{3^{s+\tau}};Q_3\right) \le c.$$

Consequently,

$$\left(\frac{1}{\tau}\fint_{\mathcal{Q}_2}|D^\tau d_s w|^{2(1+\epsilon_0)}\,d\mu_{\tau,t}\right)^{\frac{1}{2(1+\epsilon_0)}} \le c. \tag{4.8}$$

Step 3: *Second comparison estimates.* For w as in Step 1, we consider the following problem:

$$\begin{cases} v_t+\mathcal{L}^{\Phi}_{A_2(t)}v &= 0 \quad \text{in } Q_2,\\ \qquad\qquad v &= w \quad \text{in } (\mathbb{R}^n\setminus B_2)\times\Lambda_2\cup B_2\times\{-2^{2s}\}. \end{cases} \tag{4.9}$$

Similar to Step 1, we have the existence of a unique solution $v\in L^2\left(\Lambda_2;W^{s,2}(B_3)\right)\cap C\left(\Lambda_2;L^2(B_2)\right)\cap L^\infty\left(\Lambda_2;L^1_{2s}(\mathbb{R}^n)\right)$ to the problem (4.9). Taking $\tilde\varphi:=v-w$ as a test function (upon approximation) to

$$\tilde\varphi_t+\mathcal{L}^{\Phi}_{A_2(t)}v-\mathcal{L}^{\Phi}_{A_2(t)}w=\mathcal{L}^{\Phi}_{A}w-\mathcal{L}^{\Phi}_{A_2(t)}w \quad \text{in } Q_2$$

and then using (1.3) and Hölder's inequality, we obtain

$$\begin{aligned}
&\frac{1}{\tau}\int_{\mathcal{Q}_2}|D^\tau d_s\tilde\varphi|^2\,d\mu_{\tau,t}+\sup_{t\in\Lambda_2}\int_{B_2}|\tilde\varphi(x,t)|^2dx\\
&\le c\int_{\Lambda_2}\int_{B_2}\int_{B_2}|A(x,t)-(A)_2(t)|\frac{|w(x,t)-w(y,t)||\tilde\varphi(x,t)-\tilde\varphi(y,t)|}{|x-y|^{n+2s}}dx\,dy\,dt\\
&\le c\left(\frac{1}{\tau}\int_{\mathcal{Q}_2}|D^\tau d_s\tilde\varphi|^2\,d\mu_{\tau,t}\right)^{\frac{1}{2}}\left(\frac{1}{\tau}\int_{\mathcal{Q}_2}|D^\tau d_s w|^{2(1+\epsilon_0)}\,d\mu_{\tau,t}\right)^{\frac{1}{2(1+\epsilon_0)}}\\
&\quad\times\left(\int_{\Lambda_2}\int_{B_2}\int_{B_2}|A-(A)_2(t)|^{2\frac{1+\epsilon_0}{\epsilon_0}}\,dx\,dy\,dt\right)^{\frac{\epsilon_0}{2(1+\epsilon_0)}}.
\end{aligned}$$

Finally, up on using the vanishing condition on A and (4.8), the above expression yields

$$\frac{1}{\tau}\int_{\mathcal{Q}_2}|D^\tau d_s(v-w)|^2\,d\mu_{\tau,t}\le c\,\delta^{\frac{\epsilon_0}{2(1+\epsilon_0)}}, \tag{4.10}$$

for some $c=c(n,s,L)$. Coupling (4.7) with (4.10) and using triangle inequality, we get the first part of (4.4) by taking δ sufficiently small depending on n,s,L and ϵ.
Step 4: *Uniform bound on $|D^\tau d_s v|$.* From Lemma 4.2 along with (4.1), we observe that

$$\|D^\tau d_s v\|_{L^\infty(\mathcal{Q}_1)}=\sup_{(x,y,t)\in\mathcal{Q}_1}\frac{|v(x,t)-v(y,t)|}{|x-y|^{s+\tau}}\le c\left(\frac{1}{\tau}\fint_{\mathcal{Q}_2}|D^\tau d_s v|^2\,d\mu_{\tau,t}\right)^{\frac{1}{2}}+c\,\mathrm{Tail}_{\infty,2s}\left(v-(v)_{B_2}(t);Q_2\right),$$

where $c=c(n,s,L,\tau)$. Proceeding as in Step 2 and [13, Lemma 3.3], it can be shown that the right-hand side quantity of the above expression is bounded by a uniform constant $c=c(n,s,L,\tau)$. This completes the proof of the lemma. □

We finish this section by giving a non-scaled version of the above lemma and this directly follows from Lemma 4.3 along with a scaling argument (see Lemma 1.2).

Lemma 4.4. *Let $Q_{20\rho_i}(z_i) \Subset \Omega_T$. For any $\epsilon > 0$, there is a constant $\delta = \delta(n, s, L, \epsilon)$ such that for any weak solution u to (1.1) satisfying*

$$\frac{1}{\tau}\fint_{\mathcal{Q}_{20\rho_i}(z_i)} |D^\tau d_s u|^2 \, d\mu_{\tau,t} + \mathrm{Tail}_{\infty,2s}\left(\frac{u-(u)_{B_{20\rho_i(x_i)}}(t)}{(20\rho_i)^{s+\tau}}\,;\, Q_{20\rho_i}(z_i)\right)^2 \leq 1$$

and

$$\left(\frac{1}{\tau}\fint_{\mathcal{Q}_{20\rho_i}(z_i)} \left((20\rho_i)^{s-\tau}|G|\right)^\gamma \, d\mu_{\tau,t}\right)^{\frac{2}{\gamma}} + \frac{1}{\tau}\fint_{\mathcal{Q}_{20\rho_i}(z_i)} |D^\tau d_0 f|^2 \, d\mu_{\tau,t} + \mathrm{Tail}_{2,s}\left(\frac{f-(f)_{B_{20\rho_i}(x_i)}(t)}{(20\rho_i)^{\tau}}\,;\, Q_{20\rho_i}(z_i)\right)^2$$
$$+ \left(\fint_{\mathcal{Q}_{10\rho_i}(z_i)} |A - (A)_{10\rho_i, x_i}(t)| \, dx\, dy\, dt\right)^2 \leq \delta^2,$$

then there exists a solution v to

$$v_t + \mathcal{L}^{\Phi}_{A_{10\rho_i,x_i}(t)} v = 0 \quad \textit{in } Q_{10\rho_i}(z_i)$$

such that

$$\frac{1}{\tau}\fint_{\mathcal{Q}_{5\rho_i}(z_i)} |D^\tau d_s(u-v)|^2 \, d\mu_{\tau,t} \leq \epsilon^2 \quad \textit{and} \quad \|D^\tau d_s v\|_{L^\infty\left(\mathcal{Q}_{5\rho_i}(z_i)\right)} \leq c, \tag{4.11}$$

where $c = c(n, s, L, \tau)$.

5. Coverings of upper level sets

In this section, we construct parabolic cylinders covering the upper level set of $d_s u$, where

$$u \in L^2\left(\Lambda_2\,;\, W^{s,2}(B_2)\right) \cap C\left(\Lambda_2; L^2(B_2)\right) \cap L^\infty\left(\Lambda_2\,;\, L^1_{2s}(\mathbb{R}^n)\right)$$

is a weak solution to the localized problem:

$$u_t + \mathcal{L}^{\Phi}_A u = (-\Delta)^{\frac{s}{2}} f + g \quad \text{in } Q_2. \tag{5.1}$$

In addition, we assume that $f \in L^q\left(\Lambda_2; L^1_s\left(\mathbb{R}^n\right)\right)$ and

$$\int_{\mathcal{Q}_2} |D^\tau d_s u|^p + |D^\tau d_0 f|^q + |G|^\gamma \, d\mu_{\tau,t} < \infty,$$

where $p \in [2, q]$,

$$\tau \in \left(0, s - \frac{2s}{q}\right) \tag{5.2}$$

and the constant γ is defined in (3.2). Let us denote

$$\tilde{q} = \frac{1}{2}\left(q + \frac{2s}{s-\tau}\right) \tag{5.3}$$

to see that

$$s > \tau + \frac{2s}{\tilde{q}} \tag{5.4}$$

and

$$\tilde{q} < q, \tag{5.5}$$

which follow from the choice of τ given in (5.2). We point out that (5.4) and (5.5) are needed to handle the tail induced by the right-hand side f and to employ Fubini's theorem, respectively (see (5.27) and (6.10)). We now present the main proposition of this section.

Proposition 5.1. *Let $1 \leq r_1 < r_2 \leq 2$, $\delta > 0$ and u be a weak solution to (5.1). Then, there are two families of countable disjoint cylinders $\{\mathcal{Q}_{\rho_i}(z_i)\}_{i\geq 0}$ and $\left\{\mathcal{Q}_{\tilde{r}_j}\left(x_{1,j}, x_{2,j}, t_{0,j}\right)\right\}_{j\geq 0}$ such that*

$$U_\lambda := \{(x,y,t) \in \mathcal{Q}_{r_1} \,:\, |D^\tau d_s u(x,y,t)| \geq \lambda\} \subset \left(\bigcup_{i\geq 0} \mathcal{Q}_{5^{\frac{2}{s}}\rho_i}(z_i)\right) \bigcup \left(\bigcup_{j\geq 0} \mathcal{Q}_{5^{\frac{1}{s}}\tilde{r}_j}\left(x_{1,j}, x_{2,j}, t_{0,j}\right)\right) \tag{5.6}$$

whenever $\lambda \geq \lambda_0$, *where*

$$\lambda_0 := \frac{c\tau^{\frac{1}{q}-\frac{1}{\gamma}}}{(r_2-r_1)^{\frac{5n}{s}}}\left(\left(\frac{1}{\tau}\fint_{\mathcal{Q}_2}|D^\tau d_s u|^p\,d\mu_{\tau,t}\right)^{\frac{1}{p}} + \mathrm{Tail}_{\infty,2s}\left(\frac{u-(u)_{B_2}(t)}{2^{s+\tau}};Q_2\right) + \left(\sup_{t\in\Lambda_2}\fint_{B_2}\frac{|u-(u)_{Q_2}|^2}{2^{2s+2\tau}}\,dx\right)^{\frac{1}{2}}\right)$$
$$+\frac{c\tau^{\frac{1}{q}-\frac{1}{\gamma}}}{(r_2-r_1)^{\frac{5n}{s}}}\frac{1}{\delta}\left(\left(\frac{1}{\tau}\fint_{\mathcal{Q}_2}|D^\tau d_0 f|^q\,d\mu_{\tau,t}\right)^{\frac{1}{q}} + \mathrm{Tail}_{q,s}\left(\frac{f-(f)_{B_2}(t)}{2^{\tau}};Q_2\right) + \left(\frac{1}{\tau}\fint_{\mathcal{Q}_2}(2^{s-\tau}|G|)^\gamma\,d\mu_{\tau,t}\right)^{\frac{1}{\gamma}}\right) \tag{5.7}$$

for some constant $c = c(n,s,L,q,\tau)$. *In particular, there exist constants* $a_u = a_u(n,s,L,q,\tau) \in (0,1]$, $a_f = a_f(n,s,L,q,\tau) \in (0,1]$ *and* $a_g = a_g(n,s,L,q,\tau) \in (0,1]$ *such that*

$$\begin{aligned}
&\sum_{i\geq 0}\mu_{\tau,t}\left(\mathcal{Q}_{\rho_i}(z_i)\right) + \sum_{j\geq 0}\mu_{\tau,t}\left(\mathcal{Q}_{\tilde{r}_j}(x_{1,j},x_{2,j},t_{0,j})\right)\\
&\leq \frac{c}{\lambda^p}\int_{\mathcal{Q}_{r_2}\cap\{|D^\tau d_s u|>a_u\lambda\}}|D^\tau d_s u|^p\,d\mu_{\tau,t} + \frac{c}{(\delta\lambda)^{\tilde{q}}}\int_{\mathcal{Q}_{r_2}\cap\{|D^\tau d_0 f|>a_f\delta\lambda\}}|D^\tau d_0 f|^{\tilde{q}}\,d\mu_{\tau,t}\\
&\quad+\frac{c}{(\delta\lambda)^{b_\tau}}\int_{\mathcal{Q}_{r_2}\cap\{|G|^\gamma>(a_g\delta\lambda)^{b_\tau}G_0^{-1}\}}|G|^\gamma G_0\,d\mu_{\tau,t},
\end{aligned} \tag{5.8}$$

where we denote

$$b_\tau = \frac{2n+4s}{n+2s+2\tau} \quad \text{and} \quad G_0 = \left(\int_{Q_2}|g|^\gamma\,dz\right)^{\frac{b_\tau-\gamma}{\gamma}}. \tag{5.9}$$

In addition, we get that

$$\left(\fint_{\mathcal{Q}_{5^{\frac{1}{s}}\tilde{r}_j}(x_{1,j},x_{2,j},t_{0,j})}|D^\tau d_s u|^{p_\#}\,d\mu_{\tau,t}\right)^{\frac{1}{p_\#}} \leq c_{od}\lambda \quad \text{for any } j \tag{5.10}$$

and for some constant $c_{od} = c_{od}(n,s,p)$, *where the constant* $p_\#$ *is defined in* (2.1),

Remark 5. As we pointed out earlier, (5.2) is only employed to control the tail term of f. Thus if $f=0$, we can remove the condition (5.2).

Proof. We first define the functional

$$\begin{aligned}
\Theta_D(z_0,r) &= \left(\fint_{\mathcal{Q}_r(z_0)}|D^\tau d_s u|^p\,d\mu_{\tau,t}\right)^{\frac{1}{p}} + \tau^{\frac{1}{\gamma}}\left(\sup_{t\in\Lambda_r(t_0)}\fint_{B_r(x_0)}\frac{|u-(u)_{Q_r(z_0)}|^2}{r^{2s+2\tau}}\,dx\right)^{\frac{1}{2}}\\
&\quad+\frac{1}{\delta}\left(\fint_{\mathcal{Q}_r(z_0)}|D^\tau d_0 f|^{\tilde{q}}\,d\mu_{\tau,t}\right)^{\frac{1}{\tilde{q}}} + \frac{1}{\delta}\left(\fint_{\mathcal{Q}_r(z_0)}(r^{s-\tau}|G|)^\gamma\,d\mu_{\tau,t}\right)^{\frac{1}{\gamma}}
\end{aligned} \tag{5.11}$$

for any $z_0 \in Q_{r_1}$ and $r>0$ with $Q_r(z_0) \subset Q_2$. The rest of the proof is divided into 8 steps.

Step 1. Coverings for the diagonal part. Let us take

$$\lambda_0 = \frac{M\kappa^{-1}\tau^{\frac{1}{q}}}{(r_2-r_1)^{\frac{5n}{s}}}\left(\left(\frac{1}{\tau}\fint_{\mathcal{Q}_2}|D^\tau d_s u|^p\,d\mu_{\tau,t}\right)^{\frac{1}{p}} + \mathrm{Tail}_{\infty,2s}\left(\frac{u-(u)_{B_2}(t)}{2^{s+\tau}};Q_2\right) + \left(\sup_{t\in\Lambda_2}\fint_{B_2}\frac{|u-(u)_{Q_2}|^2}{2^{2s+2\tau}}\,dx\right)^{\frac{1}{2}}\right)$$
$$+\frac{M\kappa^{-1}\tau^{\frac{1}{q}}}{(r_2-r_1)^{\frac{5n}{s}}}\frac{1}{\delta}\left(\left(\frac{1}{\tau}\fint_{\mathcal{Q}_2}|D^\tau d_0 f|^q\,d\mu_{\tau,t}\right)^{\frac{1}{q}} + \mathrm{Tail}_{q,s}\left(\frac{f-(f)_{B_2}(t)}{2^{\tau}};Q_2\right) + \left(\frac{1}{\tau}\fint_{\mathcal{Q}_2}(2^{s-\tau}|G|)^\gamma\,d\mu_{\tau,t}\right)^{\frac{1}{\gamma}}\right), \tag{5.12}$$

where $M\geq 1$ and $\kappa\in(0,1]$ are free parameters which we will determine later (see (5.18) and (5.35)). More precisely, the parameter M will be used to handle the diagonal part and the parameter κ will be used to handle the non-diagonal part. We next take a positive integer $j_0 \geq 5$ such that

$$\frac{16(c_0+\tilde{c}+2c_q)^2}{1-2^{-s+\tau+\frac{2s}{\tilde{q}}}} \leq 2^{j_0\left(s-\tau-\frac{2s}{\tilde{q}}\right)}, \tag{5.13}$$

where c_0 is the constant determined in Lemma 3.3, and $\tilde{c}$ and c_q are the constants determined in (2.6). Using (5.3), we observe that the number j_0 depends only on n,s,L,q and τ. We then note for any $z_0 \in Q_{r_1}$,

$$Q_{5^{\frac{2}{s}}\times 2^{j_0+3}\mathcal{R}_{1,2}}(z_0) \subset Q_{r_2},$$

where we denote

$$\mathcal{R}_{1,2} = 2^{-j_0-3} \times 5^{-\frac{2}{s}} \times (s(r_2 - r_1))^{\frac{1}{s}}. \tag{5.14}$$

Let us now define for $\lambda \geq \lambda_0$,

$$D_{\kappa\lambda} = \left\{ z_0 \in Q_{r_1} : \sup_{0<\rho\leq \mathcal{R}_{1,2}} \Theta_D(z_0, \rho) > \kappa\lambda \right\}.$$

Since $\tau < 1$, we observe that for any $z_0 \in Q_{r_1}$ and $r \in \left[\mathcal{R}_{1,2}, 5^{\frac{2}{s}} \times 2^{j_0+3}\mathcal{R}_{1,2}\right]$,

$$\left(\fint_{\mathcal{Q}_r(z_0)} |D^\tau d_s u|^p \, d\mu_{\tau,t} \right)^{\frac{1}{p}} + \frac{1}{\delta} \left(\fint_{\mathcal{Q}_r(z_0)} |D^\tau d_0 f|^{\tilde{q}} \, d\mu_{\tau,t} \right)^{\frac{1}{\tilde{q}}} + \frac{1}{\delta} \left(\fint_{\mathcal{Q}_r(z_0)} (r^{s-\tau}|G|)^\gamma \, d\mu_{\tau,t} \right)^{\frac{1}{\gamma}} \leq \frac{\kappa\lambda}{4}$$

holds by assuming

$$M \geq 2^{10n(j_0+4+5s^{-1})} s^{-\frac{5n}{s}}. \tag{5.15}$$

In addition, using Lemma 3.3 and the fact that $\tau < 1$ with $\gamma < 2$, we have

$$\begin{aligned} &\tau^{\frac{1}{\gamma}} \left(\sup_{t\in\Lambda_r(t_0)} \fint_{B_r(x_0)} \frac{|u-(u)_{Q_r(z_0)}|^2}{r^{2s+2\tau}} \, dx \right)^{\frac{1}{2}} \\ &\leq c_0 \left(\fint_{\mathcal{Q}_{2r}(z_0)} |D^\tau d_s u|^2 \, d\mu_{\tau,t} \right)^{\frac{1}{2}} + c_0 \left(\fint_{\mathcal{Q}_{2r}(z_0)} |D^\tau d_0 f|^2 \, d\mu_{\tau,t} \right)^{\frac{1}{2}} + c_0 \left(\fint_{\mathcal{Q}_{2r}(z_0)} \left((2r)^{s-\tau}|G|\right)^\gamma \, d\mu_{\tau,t} \right)^{\frac{1}{\gamma}} \\ &\quad + \underbrace{c_0 \tau^{\frac{1}{2}} \operatorname{Tail}_{2,2s}\left(\frac{u-(u)_{B_{2r}(x_0)}(t)}{(2r)^{s+\tau}}; Q_{2r}(z_0) \right)}_{T_1} + \underbrace{c_0 \tau^{\frac{1}{2}} \operatorname{Tail}_{2,s}\left(\frac{f-(f)_{B_{2r}(x_0)}(t)}{(2r)^{\tau}}; Q_{2r}(z_0) \right)}_{T_2}, \end{aligned} \tag{5.16}$$

where the constant c_0 is determined in Lemma 3.3. By Hölder's inequality and (2.6), we further estimate T_1 and T_2 as

$$\begin{aligned} T_1 + T_2 &\leq c_0 \tau^{\frac{1}{2}} \operatorname{Tail}_{p,2s}\left(\frac{|u-(u)_{B_{2r}(x_0)}(t)|}{(2r)^{s+\tau}}; Q_{2r}(z_0) \right) + c_0 \tau^{\frac{1}{2}} \operatorname{Tail}_{\tilde{q},s}\left(\frac{|f-(f)_{B_{2r}(x_0)}(t)|}{(2r)^{\tau}}; Q_{2r}(z_0) \right) \\ &\leq c_0 c_p \left(\frac{2}{\mathcal{R}_{1,2}} \right)^{s+\tau+\frac{2s}{p}} \left(\fint_{\mathcal{Q}_2} |D^\tau d_s u|^p \right)^{\frac{1}{p}} + c_0 c_q \left(\frac{2}{\mathcal{R}_{1,2}} \right)^{\tau+\frac{2s}{q}} \left(\fint_{\mathcal{Q}_2} |D^\tau d_0 f|^q \right)^{\frac{1}{q}} \\ &\quad + c_0 \tilde{c} \left(\frac{2}{\mathcal{R}_{1,2}} \right)^{5n} \operatorname{Tail}_{p,2s}\left(\frac{u-(u)_{B_2}(t)}{2^{s+\tau}}; Q_4 \right) + c_0 \tilde{c} \left(\frac{2}{\mathcal{R}_{1,2}} \right)^{5n} \operatorname{Tail}_{q,s}\left(\frac{f-(f)_{B_2}(t)}{2^{\tau}}; Q_2 \right), \end{aligned} \tag{5.17}$$

where the constants c_p, c_q and $\tilde{c}$ are determined in Lemma 2.3. Using Remark 4 and Hölder's inequality to the third term on the right-hand side of (5.17), we deduce from (5.16) that

$$\tau^{\frac{1}{\gamma}} \left(\sup_{t\in\Lambda_r(t_0)} \fint_{B_r(x_0)} \frac{|u-(u)_{Q_r(z_0)}|^2}{r^{2s+2\tau}} \, dx \right)^{\frac{1}{2}} \leq \frac{\kappa\lambda}{4}$$

holds by taking

$$M = (c_0 c_q + c_0 \tilde{c}) 2^{10n(j_0+4+5s^{-1})} s^{-\frac{5n}{s}} \tag{5.18}$$

which clearly satisfies (5.15). As a result, we observe that for any $z_0 \in Q_{r_1}$ and $r \in \left[\mathcal{R}_{1,2}, 5^{\frac{2}{s}} \times 2^{j_0+3}\mathcal{R}_{1,2}\right]$,

$$\Theta_D(z_0, r) \leq \kappa\lambda. \tag{5.19}$$

Therefore, for each $z \in D_{\kappa\lambda}$, there is an exit radius $\rho_z \leq \mathcal{R}_{1,2}$ such that

$$\Theta_D(z, \rho_z) \geq \kappa\lambda \quad \text{and} \quad \Theta_D(z, \rho) \leq \kappa\lambda \quad \text{if } \rho_z \leq \rho \leq 5^{\frac{2}{s}} \times 2^{j_0+3}\mathcal{R}_{1,2}. \tag{5.20}$$

We first observe that if $Q_\rho(z_1) \cap Q_r(z_2) \neq \emptyset$ with $\frac{\rho}{2} \leq r \leq 2\rho$, then we have $Q_r(z_2) \subset Q_{5^{\frac{1}{s}}\rho}(z_1)$. Thus we apply Vitali's covering lemma to the collection $\{Q_{2^{j_0}\rho_z}(z)\}_{z\in D_{\kappa\lambda}}$, in order to find a family of mutually disjoint countable cylinders

$$\left\{ Q_{2^{j_0}\rho_{z_i}}(z_i) \right\}_{i\geq 0} \text{ such that } D_{\kappa\lambda} \subset \bigcup_{z_0\in D_{\kappa\lambda}} Q_{\rho_{z_0}}(z_0) \subset \bigcup_{i=0}^{\infty} Q_{5^{\frac{1}{s}}\times 2^{j_0}\rho_{z_i}}(z_i). \tag{5.21}$$

In addition, by the proof of Vitali's covering lemma, we get that for any $z \in D_{\kappa\lambda}$, there is i such that

$$\frac{2^{j_0}\rho_{z_i}}{2} \leq 2^{j_0}\rho_z \leq 2^{j_0+1}\rho_{z_i} \quad \text{and} \quad Q_{2^{j_0}\rho_z}(z) \subset Q_{5^{\frac{1}{s}}\times\rho_i}(z_i), \tag{5.22}$$

where we denote

$$\rho_i = 2^{j_0}\rho_{z_i} \quad \text{for each } i. \tag{5.23}$$

From (5.20), we have

$$\begin{aligned}\kappa\lambda \le & \left(\fint_{\mathcal{Q}_{\rho_{z_i}}(z_i)} |D^\tau d_s u|^p \, d\mu_{\tau,t}\right)^{\frac{1}{p}} + \tau^{\frac{1}{\gamma}} \left(\sup_{t\in\Lambda_{\rho_{z_i}}(t_i)} \fint_{B_{\rho_{z_i}}(x_i)} \frac{|u-(u)_{Q_{\rho_{z_i}}(z_i)}|^2}{\rho_{z_i}^{2s+2\tau}}\, dx\right)^{\frac{1}{2}} \\ & + \frac{1}{\delta}\left(\fint_{\mathcal{Q}_{\rho_{z_i}}(z_i)} |D^\tau d_0 f|^{\tilde{q}} \, d\mu_{\tau,t}\right)^{\frac{1}{\tilde{q}}} + \frac{1}{\delta}\left(\fint_{\mathcal{Q}_{\rho_{z_i}}(z_i)} \left((\rho_{z_i})^{s-\tau}|G|\right)^\gamma d\mu_{\tau,t}\right)^{\frac{1}{\gamma}}.\end{aligned} \tag{5.24}$$

We first note from (2.6) with $\beta = s$, $\tilde{s} = s$ and $k = j_0$ that

$$\begin{aligned}\tau^{\frac{1}{\gamma}}\,\mathrm{Tail}_{p,2s}\left(\frac{u-(u)_{B_{2\rho}(x_i)}(t)}{(2\rho)^{s+\tau}}; Q_{2\rho}(z_i)\right) \le\ & c_p \sum_{j=2}^{j_0} 2^{i\left(-s+\tau+\frac{2s}{p}\right)} \left(\fint_{\mathcal{Q}_{2^j\rho}(z_i)} |D^\tau d_s u|^p\, d\mu_{\tau,t}\right)^{\frac{1}{p}} \\ & + \tilde{c}\tau^{\frac{1}{\gamma}} \sum_{j=j_0+1}^{l} 2^{j(-s+\tau)} \left(\sup_{t\in\Lambda_{2^j\rho}(t_i)} \fint_{B_{2^j\rho}(x_i)} \frac{|u-(u)_{B_{2^j\rho}(x_i)}(t)|^2}{(2^j\rho)^{2s+2\tau}}\, dx\right)^{\frac{1}{2}} \\ & + \tilde{c}\tau^{\frac{1}{\gamma}} 2^{-2sl}\left(\frac{2}{\rho}\right)^{s+\tau} \left(\sup_{t\in\Lambda_2}\fint_{B_2} \frac{|u-(u)_{B_2}(t)|^2}{2^{2s+2\tau}}\, dx\right)^{\frac{1}{2}} \\ & + \frac{\tilde{c}}{(2-r_1)^{n+2s}}\left(\frac{2}{\rho}\right)^{-s+\tau} \mathrm{Tail}_{\infty,2s}\left(\frac{u-(u)_{B_2}(t)}{2^{\tilde{s}+\tau}}; Q_2\right),\end{aligned} \tag{5.25}$$

where $\rho = \rho_{z_i}$ and l is the positive integer such that $2^{j_0+1}\mathcal{R}_{1,2} \le 2^l\rho_{z_i} < 2^{j_0+2}\mathcal{R}_{1,2}$. We point out that for the last tail term in (5.25), we have taken supremum in the time variable for the expression in (2.7). Similarly, we observe

$$\begin{aligned}& \tau^{\frac{1}{\gamma}}\,\mathrm{Tail}_{\tilde{q},s}\left(\frac{f-(f)_{B_{2\rho}(x_i)}(t)}{(2\rho)^{\tau}}; Q_{2\rho}(z_i)\right) \\ & \le c_{\tilde{q}} \sum_{j=2}^{j_0} 2^{i\left(-s+\tau+\frac{2s}{\tilde{q}}\right)} \left(\fint_{\mathcal{Q}_{2^j\rho}(z_i)} |D^\tau d_0 f|^{\tilde{q}}\, d\mu_{\tau,t}\right)^{\frac{1}{\tilde{q}}} + c_{\tilde{q}} \sum_{j=j_0+1}^{l} 2^{i\left(-s+\tau+\frac{2s}{\tilde{q}}\right)} \left(\fint_{\mathcal{Q}_{2^j\rho}(z_i)} |D^\tau d_0 f|^{\tilde{q}}\, d\mu_{\tau,t}\right)^{\frac{1}{\tilde{q}}} \\ & + c_{\tilde{q}} 2^{-sl}\left(\frac{2}{\rho}\right)^{\tau+\frac{2s}{\tilde{q}}} \left(\fint_{\mathcal{Q}_2} |D^\tau d_0 f|^q\, d\mu_{\tau,t}\right)^{\frac{1}{q}} + \frac{\tilde{c}}{(2-r_1)^{n+2s}}\left(\frac{2}{\rho}\right)^{-s+\tau+\frac{2s}{q}} \mathrm{Tail}_{q,s}\left(\frac{f-(f)_{B_2}(t)}{2^\tau}; Q_2\right),\end{aligned} \tag{5.26}$$

where we have used Hölder's inequality for the third and fourth terms in the right-hand side of (5.26). We combine (5.16) with $r = \rho_{z_i}$, (5.25) and (5.26) together with (5.4), (5.23) and Remark 4 to get

$$\begin{aligned}& \tau^{\frac{1}{\gamma}} \sup_{t\in\Lambda_{\rho_{z_i}}(t_i)} \left(\fint_{B_{\rho_{z_i}}(x_i)} \frac{|u-(u)_{Q_{\rho_{z_i}}(z_i)}|^2}{\rho_{z_i}^{2s+2\tau}}\, dx\right)^{\frac{1}{2}} \\ & \le c\left(\fint_{\mathcal{Q}_{\rho_i}(z_i)} |D^\tau d_s u|^p\, d\mu_{\tau,t}\right)^{\frac{1}{p}} + c\left(\fint_{\mathcal{Q}_{\rho_i}(z_i)} |D^\tau d_0 f|^{\tilde{q}}\, d\mu_{\tau,t}\right)^{\frac{1}{\tilde{q}}} + c\left(\fint_{\mathcal{Q}_{\rho_i}(z_i)} \left((\rho_i)^{s-\tau}|G|\right)^\gamma d\mu_{\tau,t}\right)^{\frac{1}{\gamma}} \\ & + c_0\tilde{c}\left[\sum_{j=j_0+1}^{l} 2^{i(-s+\tau)}\Theta_D\left(z_j, 2^j\rho_{z_i}\right) + \tau^{\frac{1}{\gamma}}\left(\frac{2}{2^{j_0}\mathcal{R}_{1,2}}\right)^{s+\tau}\left(\sup_{t\in\Lambda_2}\fint_{B_2}\frac{|u-(u)_{B_2}(t)|^2}{2^{2s+2\tau}}\, dx\right)^{\frac{1}{2}}\right] \\ & + c_0 c_q\left[\sum_{j=j_0+1}^{l} 2^{i\left(-s+\tau+\frac{2s}{\tilde{q}}\right)}\Theta_D\left(z_j, 2^j\rho_{z_i}\right) + \left(\frac{2}{2^{j_0}\mathcal{R}_{1,2}}\right)^{\tau+\frac{2s}{\tilde{q}}}\left(\fint_{\mathcal{Q}_2} |D^\tau d_0 f|^q\, d\mu_{\tau,t}\right)^{\frac{1}{q}}\right] \\ & + \frac{c_0\tilde{c}}{(r_2-r_1)^{5n}}\left(\mathrm{Tail}_{\infty,2s}\left(\frac{u-(u)_{B_2}(t)}{2^{s+\tau}}; Q_2\right) + \mathrm{Tail}_{q,s}\left(\frac{f-(f)_{B_2}(t)}{2^\tau}; Q_2\right)\right),\end{aligned} \tag{5.27}$$

where $c = c(n, s, L, q, \tau)$. We here highlight that (5.4) is necessary to handle the sixth term in the right-hand side of (5.27), as $\sum\limits_{j=j_0+1}^{\infty} 2^{i\left(-s+\tau+\frac{2s}{\tilde{q}}\right)} < \frac{2^{j_0\left(-s+\tau+\frac{2s}{\tilde{q}}\right)}}{1-2^{-s+\tau+\frac{2s}{\tilde{q}}}}$. We further estimate the right-hand side of (5.27) using (5.4), (5.13), (5.14), (5.18) and (5.20) as

$$\tau^{\frac{1}{\gamma}} \sup_{t\in\Lambda_{\rho_{z_i}}(t_i)} \left(\fint_{B_{\rho_{z_i}}(x_i)} \frac{|u-(u)_{Q_{\rho_{z_i}}(z_i)}|^2}{\rho_{z_i}^{2s+2\tau}}\,dx \right)^{\frac{1}{2}}$$
$$\leq c\left(\fint_{\mathcal{Q}_{\rho_i}(z_i)} |D^\tau d_s u|^p \,d\mu_{\tau,t} \right)^{\frac{1}{p}} + c\left(\fint_{\mathcal{Q}_{\rho_i}(z_i)} |D^\tau d_0 f|^{\tilde{q}} \,d\mu_{\tau,t} \right)^{\frac{1}{\tilde{q}}} + c\left(\fint_{\mathcal{Q}_{\rho_i}(z_i)} \left((\rho_i)^{s-\tau}|G|\right)^\gamma \,d\mu_{\tau,t} \right)^{\frac{1}{\gamma}} + \frac{\kappa\lambda}{4}.$$

Plugging the above estimate into (5.24) along with (5.23), we find that

$$\kappa\lambda \leq c\left(\fint_{\mathcal{Q}_{\rho_i}(z_i)} |D^\tau d_s u|^p \,d\mu_{\tau,t} \right)^{\frac{1}{p}} + \frac{c}{\delta}\left(\fint_{\mathcal{Q}_{\rho_i}(z_i)} |D^\tau d_0 f|^{\tilde{q}} \,d\mu_{\tau,t} \right)^{\frac{1}{\tilde{q}}} + \frac{c}{\delta}\left(\fint_{\mathcal{Q}_{\rho_i}(z_i)} \left((\rho_i)^{s-\tau}|G|\right)^\gamma \,d\mu_{\tau,t} \right)^{\frac{1}{\gamma}}$$

for some constant $c = c(n,s,L,q,\tau)$. Therefore we deduce that one of the following must hold:

$$\begin{aligned} &\frac{\kappa\lambda}{3} \leq c\left(\fint_{\mathcal{Q}_{\rho_i}(z_i)} |D^\tau d_s u|^p \,d\mu_{\tau,t} \right)^{\frac{1}{p}}, \quad \frac{\kappa\lambda}{3} \leq \frac{c}{\delta}\left(\fint_{\mathcal{Q}_{\rho_i}(z_i)} |D^\tau d_0 f|^{\tilde{q}} \,d\mu_{\tau,t} \right)^{\frac{1}{\tilde{q}}}, \\ &\frac{\kappa\lambda}{3} \leq \frac{c}{\delta}\left(\fint_{\mathcal{Q}_{\rho_i}(z_i)} \left((\rho_i)^{s-\tau}|G|\right)^\gamma \,d\mu_{\tau,t} \right)^{\frac{1}{\gamma}}. \end{aligned} \tag{5.28}$$

If the first inequality or the second inequality in (5.28) holds, then we get

$$\mu_{\tau,t}\left(\mathcal{Q}_{\rho_i}(z_i)\right) \leq \frac{c}{(\kappa\lambda)^p} \int_{\mathcal{Q}_{\rho_i}(z_i)} |D^\tau d_s u|^p \,d\mu_{\tau,t} \quad \text{or} \quad \mu_{\tau,t}\left(\mathcal{Q}_{\rho_i}(z_i)\right) \leq \frac{c}{(\kappa\delta\lambda)^{\tilde{q}}} \int_{\mathcal{Q}_{\rho_i}(z_i)} |D^\tau d_0 f|^{\tilde{q}} \,d\mu_{\tau,t}. \tag{5.29}$$

On the other hand, if the third inequality in (5.28) holds, we observe that

$$\mu_{\tau,t}\left(\mathcal{Q}_{\rho_i}(z_i)\right) \leq \frac{c}{(\kappa\delta\lambda)^\gamma} \int_{\mathcal{Q}_{\rho_i}(z_i)} \left((\rho_i)^{s-\tau}|G|\right)^\gamma \,d\mu_{\tau,t} = \frac{c}{(\kappa\delta\lambda)^\gamma} \int_{\mathcal{Q}_{\rho_i}(z_i)} \left((\rho_i)^{s-\tau}|G|\right)^\gamma G_0 G_0^{-1} \,d\mu_{\tau,t}.$$

We note from (5.9) that

$$\begin{aligned} \rho_i^{\gamma(s-\tau)} G_0^{-1} \leq \rho_i^{\gamma(s-\tau)} \left(\int_{Q_{\rho_i}(z_i)} |g|^\gamma \right)^{\frac{\gamma-b_\tau}{\gamma}} &\leq c\rho_i^{\gamma(s-\tau)} \left(\fint_{\mathcal{Q}_{\rho_i}(z_i)} \rho_i^{n+2s}|G|^\gamma \,d\mu_{\tau,t} \right)^{\frac{\gamma-b_\tau}{\gamma}} \\ &\leq c\left(\fint_{\mathcal{Q}_{\rho_i}(z_i)} \left((\rho_i)^{s-\tau}|G|\right)^\gamma \,d\mu_{\tau,t} \right)^{\frac{\gamma-b_\tau}{\gamma}}. \end{aligned}$$

Using the above two estimates along with the third inequality in (5.28), we have

$$\mu_{\tau,t}\left(\mathcal{Q}_{\rho_i}(z_i)\right) \leq \frac{c}{(\delta\kappa\lambda)^{b_\tau}} \int_{\mathcal{Q}_{\rho_i}(z_i)} |G|^\gamma G_0 \,d\mu_{\tau,t}. \tag{5.30}$$

We combine (5.29) and (5.30) to see that

$$\mu_{\tau,t}\left(\mathcal{Q}_{\rho_i}(z_i)\right) \leq \frac{c}{(\kappa\lambda)^p} \int_{\mathcal{Q}_{\rho_i}(z_i)} |D^\tau d_s u|^p \,d\mu_{\tau,t} + \frac{c}{(\kappa\delta\lambda)^{\tilde{q}}} \int_{\mathcal{Q}_{\rho_i}(z_i)} |D^\tau d_0 f|^{\tilde{q}} \,d\mu_{\tau,t} + \frac{c}{(\delta\kappa\lambda)^{b_\tau}} \int_{\mathcal{Q}_{\rho_i}(z_i)} |G|^\gamma G_0 \,d\mu_{\tau,t}.$$

A suitable choice of the constants $\tilde{a}_u = \tilde{a}_u(n,s,L,q,\tau) \in \left(0,\frac{1}{8}\right]$, $\tilde{a}_f = \tilde{a}_f(n,s,L,q,\tau) \in (0,1]$ and $\tilde{a}_g = \tilde{a}_g(n,s,L,q,\tau) \in (0,1]$ yields

$$\begin{aligned} \mu_{\tau,t}\left(\mathcal{Q}_{\rho_i}(z_i)\right) \leq{}& \frac{c}{(\kappa\lambda)^p} \int_{\mathcal{Q}_{\rho_i}(z_i)\cap\{|D^\tau d_s u|>\tilde{a}_u\kappa\lambda\}} |D^\tau d_s u|^p \,d\mu_{\tau,t} + \frac{c}{(\kappa\delta\lambda)^{\tilde{q}}} \int_{\mathcal{Q}_{\rho_i}(z_i)\cap\{|D^\tau d_0 f|>\tilde{a}_f\kappa\delta\lambda\}} |D^\tau d_0 f|^{\tilde{q}} \,d\mu_{\tau,t} \\ &+ \frac{c}{(\kappa\delta\lambda)^{b_\tau}} \int_{\mathcal{Q}_{\rho_i}(z_i)\cap\{|G|^\gamma>(\tilde{a}_g\kappa\delta\lambda)^{b_\tau}G_0^{-1}\}} |G|^\gamma G_0 \,d\mu_{\tau,t}. \end{aligned} \tag{5.31}$$

Remark 6. We here remark on the second term appearing on the right-hand side of (5.11). We first note that this term is used to handle parabolic tail terms. Since $u \in C\left(-2^{2s},2^{2s};L^2(B_2)\right)$, there may exist some points $z_0 \in Q_{r_1}$ such that $\Theta_D(z_0,\rho_{z_0}) \geq \kappa\lambda$ and

$$\left(\fint_{\mathcal{Q}_{\rho_{z_0}}(z_0)} |D^\tau d_s u|^p \,d\mu_{\tau,t} \right)^{\frac{1}{p}} \leq \frac{\kappa\lambda}{2} \quad \text{with} \quad \tau^{\frac{1}{\gamma}} \left(\sup_{t\in\Lambda_{\rho_{z_0}}(t_0)} \fint_{B_{\rho_{z_0}}(x_0)} \frac{|u-(u)_{Q_{\rho_{z_0}}(z_0)}|^2}{\rho_{z_0}^{2s+2\tau}}\,dx \right)^{\frac{1}{2}} \geq \frac{\kappa\lambda}{2},$$

where ρ_{z_0} is the exit-radius of the point z_0. However, from energy estimates and rigorous tail estimates, we still have a sufficiently good bound on the measure of such cylinders as in (5.31), which is an essential ingredient to obtain L^q-regularity of $D^\tau d_s u$.

Step 2. Coverings for off-diagonal parts. We first note that for any $(x,y,t) \in \mathcal{Q}_{r_1}$ and $r \in \left(0, 5^{\frac{2}{s}} \times 2^{j_0+3}\mathcal{R}_{1,2}\right]$, we have

$$\mathcal{Q}_r(x,y,t) \subset \mathcal{Q}_{r_2},$$

where the parameter $\mathcal{R}_{1,2}$ is defined in (5.14). Let us define

$$\begin{aligned} E_{p,\tau}\left(u\,;\,Q_r(z_0)\right) = &\left(\fint_{\mathcal{Q}_r(z_0)} |D^\tau d_s u|^p \, d\mu_{\tau,t}\right)^{\frac{1}{p}} + \tau^{\frac{1}{\gamma}}\left(\sup_{t\in\Lambda_r(t_0)} \fint_{B_r(x_0)} \frac{|u-(u)_{Q_r(z_0)}|^2}{r^{2s+2\tau}}\, dx\right)^{\frac{1}{2}} \\ &+ \frac{1}{\delta}\left(\fint_{\mathcal{Q}_r(z_0)} |D^\tau d_0 f|^{\tilde{q}} \, d\mu_{\tau,t}\right)^{\frac{1}{\tilde{q}}}, \end{aligned} \tag{5.32}$$

for any $Q_r(z_0) \Subset Q_2$. Then we first note that for any $\rho < r$,

$$\sup_{t\in\Lambda_\rho}\left(\fint_{B_\rho} \frac{|u-(u)_{Q_\rho}|^2}{\rho^{2s+2\tau}}\, dx\right)^{\frac{1}{2}} \le 2\left(\frac{r}{\rho}\right)^{\frac{n}{2}+s+\tau} \sup_{t\in\Lambda_r}\left(\fint_{B_r} \frac{|u-(u)_{Q_r}|^2}{r^{2s+2\tau}}\, dx\right)^{\frac{1}{2}}.$$

We now intend to employ exit-time arguments, to this end, we first introduce a set function which is defined by

$$\mathfrak{A}\left(\mathcal{B}_r\right) = \begin{cases} \left(\frac{r}{\operatorname{dist}(B^1_r, B^2_r)}\right)^{s+\tau} & \text{if } \operatorname{dist}\left(B^1_r, B^2_r\right) \ge r, \\ 1 & \text{if } \operatorname{dist}\left(B^1_r, B^2_r\right) < r, \end{cases} \tag{5.33}$$

where $\mathcal{B}_r = B^1_r \times B^2_r$. We next define a functional

$$\Theta_{OD}\left(u; \mathcal{Q}_r(x_1,x_2,t_0)\right) = \left(\fint_{\mathcal{Q}_r(x_1,x_2,t_0)} |D^\tau d_s u|^p \, d\mu_{\tau,t}\right)^{\frac{1}{p}} + \mathfrak{A}\left(\mathcal{B}_r(x_1,x_2)\right)\sum_{d=1}^{2} E_{p,\tau}\left(u\,;\,Q_r(x_d,t_0)\right). \tag{5.34}$$

We then observe from (2.4) and (5.19) that for any $r \in \left[5^{-\frac{2}{s}}\mathcal{R}_{1,2}, 5^{\frac{2}{s}} \times 2^{j_0+3}\mathcal{R}_{1,2}\right]$,

$$\Theta_{OD}\left(u; \mathcal{Q}_r(x_1,x_2,t_0)\right) \le \lambda$$

holds by taking

$$\kappa = \frac{\tau^{\frac{1}{\gamma}}}{2^{4nj_0}5^{\frac{20n}{s}}(C_n+1)}, \tag{5.35}$$

where the constants C_n and j_0 are determined in (2.4) and (5.13), respectively. Let us define

$$OD_\lambda = \left\{(x_1,x_2,t_0) \in \mathcal{Q}_{r_1} \,:\, \sup_{0<\rho\le 5^{-\frac{2}{s}}\mathcal{R}_{1,2}} \Theta_{OD}\left(u; \mathcal{Q}_\rho(x_1,x_2,t_0)\right) \ge \lambda\right\}.$$

For each $(x_1,x_2,t_0) \in OD_\lambda$, there is an exit-time radius $\overline{r} \le 5^{-\frac{2}{s}}\mathcal{R}_{1,2}$ such that

$$\Theta_{OD}\left(u; \mathcal{Q}_\rho(x_1,x_2,t_0)\right) \le \lambda \quad \text{if } \rho \ge \overline{r} \quad \text{and} \quad \Theta_{OD}\left(u; \mathcal{Q}_{\overline{r}}(x_1,x_2,t_0)\right) \ge \lambda. \tag{5.36}$$

Using Vitali's covering lemma, we find a collection $\widetilde{\mathcal{A}} = \left\{\mathcal{Q}_{2^{j_0}\overline{r}_j}(x_{1,j},x_{2,j},t_{0,j})\right\}_{j\ge0}$ whose elements are mutually disjoint and satisfy

$$OD_\lambda \subset \bigcup_{j\ge0} \mathcal{Q}_{5^{\frac{1}{s}}2^{j_0}\overline{r}_j}(x_{1,j},x_{2,j},t_{0,j}),$$

where we denote by $\overline{r}_j$ the exit-time radius of the point $(x_{1,j},x_{2,j},t_{0,j})$. Therefore, we have

$$\{(x,y,t) \in \mathcal{Q}_{r_1} \,:\, |D^\tau d_s u(x,y,t)| \ge \lambda\} \subset \bigcup_{j\ge0} \mathcal{Q}_{5^{\frac{1}{s}}\widetilde{r}_j}(x_{1,j},x_{2,j},t_{0,j}), \tag{5.37}$$

where we denote

$$\widetilde{r}_j \equiv 2^{j_0}\overline{r}_j. \tag{5.38}$$

Remark 7. We give some remarks on the functional defined in (5.34). We first note that the second term in the right-hand side of (5.34) is needed to obtain a bound on the $L^{p\#}$-norm of $D^\tau d_s u$ (see (5.48), below). We now explain how to obtain a good upper bound on the measure of $\mathcal{Q}_{\widetilde{r}_j}(x_{1,j}, x_{2,j}, t_{0,j}) \in \widetilde{\mathcal{A}}$ which is an essential ingredient to get L^q-regularity of $D^\tau d_s u$. Indeed, by (2.5), (5.13) and (5.38), we observe

$$\mu_{\tau,t}\left(\mathcal{Q}_{\widetilde{r}_j}(x_{1,j}, x_{2,j}, t_{0,j})\right) \le c\mu_{\tau,t}\left(\mathcal{Q}_{\overline{r}_j}(x_{1,j}, x_{2,j}, t_{0,j})\right)$$

for some constant $c = c(n, s, \Lambda, q, \tau)$, which implies that it suffices to investigate a good upper bound on the measure of $\mathcal{Q}_{\overline{r}_j}(x_{1,j}, x_{2,j}, t_{0,j})$. Suppose that the selected cylinder $\mathcal{Q}_{\widetilde{r}_j}(x_{1,j}, x_{2,j}, t_{0,j})$ is close to the diagonal. In the next step, we will see that such a cylinder is indeed contained in a diagonal cylinder which we did choose in step 1. Therefore, the measure of $\mathcal{Q}_{\widetilde{r}_j}(x_{1,j}, x_{2,j}, t_{0,j})$ has a good upper bound as in (5.31).

On the other hand, if the selected cylinder $\mathcal{Q}_{\widetilde{r}_j}(x_{1,j}, x_{2,j}, t_{0,j}) \in \widetilde{\mathcal{A}}$ is not close to the diagonal, then from (5.36), we observe that there holds

$$\begin{aligned}\frac{3\lambda}{4} \le &\left(\frac{1}{\mu_{\tau,t}\left(\mathcal{Q}_{\overline{r}_j}(x_{1,j}, x_{2,j}, t_{0,j})\right)} \int_{\mathcal{Q}_{\overline{r}_j}(x_{1,j},x_{2,j},t_{0,j}) \cap \left\{|D^\tau d_s u| > \frac{\lambda}{16}\right\}} |D^\tau d_s u|^p \, d\mu_{\tau,t}\right)^{\frac{1}{p}} \\ &+ \mathfrak{A}\left(\mathcal{B}_{\overline{r}_j}(x_{1,j}, x_{2,j})\right)^{s+\tau} \left[\sum_{d=1}^{2} E_{p,\tau}\left(u\,;\, Q_{\overline{r}_j}(x_{d,j}, t_{0,j})\right)\right].\end{aligned} \tag{5.39}$$

Our approach to obtaining a good upper bound on the measure of such cylinders depends on the size of $E_{p,\tau}\left(u\,;\, Q_{\overline{r}_j}(x_{d,j}, t_{0,j})\right)$. Indeed, if $E_{p,\tau}\left(u\,;\, Q_{\overline{r}_j}(x_{d,j}, t_{0,j})\right) < \frac{\lambda}{16}$, then the second term on the right-hand side of (5.39) can be absorbed to the left-hand side. Thus we have a good upper bound on the measure of $\mathcal{Q}_{\widetilde{r}_j}(x_{1,j}, x_{2,j}, t_{0,j}) \in \widetilde{\mathcal{A}}$ (see (5.61) in step 6 below). We next suppose $E_{p,\tau}\left(u\,;\, Q_{\overline{r}_j}(x_{d,j}, t_{0,j})\right) \ge \frac{\lambda}{16}$. Then we first note that if

$$\lambda < a + b + c \Longrightarrow \left(\tfrac{\lambda}{3}\right)^p < a^p,\ \left(\tfrac{\lambda}{3}\right)^{\widetilde{q}} < b^{\widetilde{q}} \text{ or } \left(\tfrac{\lambda}{3}\right)^\gamma < c^\gamma, \tag{5.40}$$

where a, b and c are nonnegative constants. Applying this simple observation to (5.39), we get (5.64). Thus it remains to obtain a suitable upper bound on the second term in the right-hand side of (5.64). To this end, we find a suitable diagonal cylinder which was chosen in step 1 and contains $\mathcal{Q}_{5^{\frac{1}{s}}\widetilde{r}_j}(x_{d,j}, t_{0,j})$, and then we use some combinatorial arguments by taking advantage of the factor $\mathfrak{A}\left(\mathcal{B}_{\overline{r}_j}(x_{1,j}, x_{2,j})\right)$ (see step 7 for more details). As a result, we obtain a good upper bound on the measure of $\mathcal{Q}_{\widetilde{r}_j}(x_{1,j}, x_{2,j}, t_{0,j}) \in \widetilde{\mathcal{A}}$.

Step 3. First elimination of off-diagonal cylinders. We now prove that if $\mathcal{Q}_{\widetilde{r}_j}(x_{1,j}, x_{2,j}, t_{0,j}) \in \widetilde{\mathcal{A}}$ satisfies

$$\operatorname{dist}\left(B_{5^{\frac{1}{s}}\widetilde{r}_j}(x_{1,j}), B_{5^{\frac{1}{s}}\widetilde{r}_j}(x_{2,j})\right) < 5^{\frac{1}{s}}\widetilde{r}_j, \tag{5.41}$$

then

$$\mathcal{Q}_{5^{\frac{1}{s}}\widetilde{r}_j}(x_{1,j}, x_{2,j}, t_{0,j}) \subset \bigcup_i \mathcal{Q}_{5^{\frac{1}{s}}\rho_i}(z_i)\,. \tag{5.42}$$

By (5.36), one of the followings must hold:

$$\left(\fint_{\mathcal{Q}_{\overline{r}_j}(x_{1,j},x_{2,j},t_{0,j})} |D^\tau d_s u|^p \, d\mu_{\tau,t}\right)^{\frac{1}{p}} > \frac{\lambda}{3},\ E_{p,\tau}\left(u\,;\, Q_{\overline{r}_j}(x_{1,i}, t_{0,i})\right) > \frac{\lambda}{3},\ E_{p,\tau}\left(u\,;\, Q_{\overline{r}_j}(x_{2,i}, t_{0,i})\right) > \frac{\lambda}{3}. \tag{5.43}$$

Suppose that the first inequality in (5.43) holds. We now observe

$$\mathcal{B}_{5^{\frac{1}{s}}\widetilde{r}_j}(x_{1,j}, x_{2,j}) \subset \mathcal{B}_{5^{\frac{2}{s}}\widetilde{r}_j}(x_{1,j}).$$

Therefore, using (2.4), (5.12) and (5.35), we obtain

$$\left(\fint_{\mathcal{Q}_{5^{\frac{2}{s}}\widetilde{r}_j}(x_{1,j},t_{0,j})} |D^\tau d_s u|^p \, d\mu_{\tau,t}\right)^{\frac{1}{p}} \ge \left(\frac{\mu_{\tau,t}\left(\mathcal{Q}_{\overline{r}_j}(x_{1,j}, x_{2,j}, t_{0,j})\right)}{\mu_{\tau,t}\left(\mathcal{Q}_{5^{\frac{2}{s}}\widetilde{r}_j}(x_{1,j}, t_{0,j})\right)} \fint_{\mathcal{Q}_{\overline{r}_j}(x_{1,j},x_{2,j},t_{0,j})} |D^\tau d_s u|^p \, d\mu_{\tau,t}\right)^{\frac{1}{p}} \ge \kappa\lambda,$$

which implies that $\Theta_D\left((x_{1,j}, t_{0,j}), 5^{\frac{2}{s}}\widetilde{r}_j\right) > \kappa\lambda$. By the fact that $5^{\frac{2}{s}}\widetilde{r}_j \le \mathcal{R}_{1,2}$ and (5.20), (5.21) yields

$$\mathcal{Q}_{5^{\frac{2}{s}}\widetilde{r}_j}(x_{1,j}, t_{0,j}) \subset \bigcup_i \mathcal{Q}_{5^{\frac{1}{s}}\rho_i}(z_i)\,.$$

We next assume that the second inequality in (5.43) is true. By (5.35), we have $\Theta_D\left((x_{1,j},t_{0,j}),5^{\frac{2}{s}}\widetilde{r}_j\right) > \kappa\lambda$. Since $\mathcal{Q}_{5^{\frac{1}{s}}\widetilde{r}_j}(x_{1,j},x_{2,j},t_{0,j}) \subset \mathcal{Q}_{5^{\frac{2}{s}}\widetilde{r}_j}(x_{1,j},t_{0,j})$ which follows by (5.41), we have

$$\mathcal{Q}_{5^{\frac{1}{s}}\widetilde{r}_j}(x_{1,j},x_{2,j},t_{0,j}) \subset \mathcal{Q}_{5^{\frac{2}{s}}\widetilde{r}_j}(x_{1,j},t_{0,j}) \subset \bigcup_{i\geq 0}\mathcal{Q}_{5^{\frac{1}{s}}\rho_i}(z_i)\,.$$

Similarly, we get (5.42) if the third inequality in (5.43) holds. Thus, we now focus on the following subfamily of $\widetilde{\mathcal{A}}$:

$$\mathcal{A} = \left\{ \mathcal{Q}_{\widetilde{r}_j}(x_{1,j},x_{2,j},t_{0,j}) \;\middle|\; \begin{array}{l} \mathcal{Q}_{\widetilde{r}_j}(x_{1,j},x_{2,j},t_{0,j}) \in \widetilde{\mathcal{A}} \text{ and} \\ \mathcal{Q}_{5^{\frac{1}{s}}\widetilde{r}_j}(x_{1,j},x_{2,j},t_{0,j}) \not\subset \bigcup\limits_{i\geq 0}\mathcal{Q}_{5^{\frac{2}{s}}\rho_i}(z_i) \text{ for each } i \end{array} \right\} \tag{5.44}$$

Indeed, we take cylinders $\mathcal{Q}_{5^{\frac{2}{s}}\rho_i}(z_i)$ instead of the cylinder $\mathcal{Q}_{5^{\frac{1}{s}}\rho_i}(z_i)$, in order to eliminate other types of nearly diagonal cylinders (see (5.52) below).

Step 4. Off-diagonal estimates. We now obtain a bound on $L^{p\#}$-norm of $D^\tau d_s u$ and a reverse Hölder's inequality on cylinders which are far from the diagonal.

Lemma 5.2. *Let $\mathcal{Q} = \mathcal{Q}_r(x_1,x_2,t_0) \subset \mathcal{Q}_2$ be such that*

$$\operatorname{dist}(B_r(x_1),B_r(x_2)) \geq r. \tag{5.45}$$

Then there is a constant $c = c(n,s,p,\tau)$ such that

$$\left(\fint_{\mathcal{Q}} |D^\tau d_s u|^{p\#}\,d\mu_{\tau,t}\right)^{\frac{1}{p\#}} \leq c\,\Theta_{OD}(u;\mathcal{Q})\,.$$

Proof. We observe from (5.45) that

$$\operatorname{dist}(B_r(x_1),B_r(x_2)) \leq |x-y| \leq 5\operatorname{dist}(B_r(x_1),B_r(x_2)) \tag{5.46}$$

whenever $x\in B_r(x_1)$ and $y \in B_r(x_2)$. The above inequality and Jensen's inequality yield that

$$\frac{r^{2n}}{c(n)\operatorname{dist}(B_r(x_1),B_r(x_2))^{n-2\tau}} \leq \mu_\tau(\mathcal{B}_r(x_1,x_2)) \leq \frac{c(n)r^{2n}}{\operatorname{dist}(B_r(x_1),B_r(x_2))^{n-2\tau}} \tag{5.47}$$

and

$$\begin{aligned}
&\fint_{\mathcal{Q}_r(x_1,x_2,t_0)} |D^\tau d_s u|^{p\#}\,d\mu_{\tau,t} \\
&\leq \frac{c}{\operatorname{dist}(B_r(x_1),B_r(x_2))^{p\#(s+\tau)}}\fint_{Q_r(x_1,t_0)}\fint_{B_r(x_1)}|u(x,t)-u(y,t)|^{p\#}\,dz\,dt \\
&\leq \frac{cr^{p\#(s+\tau)}}{\operatorname{dist}(B_r(x_1),B_r(x_2))^{p\#(s+\tau)}}\sum_{d=1}^{2}\underbrace{\frac{1}{r^{p\#(s+\tau)}}\fint_{Q_r(x_d,t_0)}|u(x,t)-(u)_{B_r(x_d)}(t)|^{p\#}\,dz}_{=I_d} \\
&\quad + \frac{cr^{p\#(s+\tau)}}{\operatorname{dist}(B_r(x_1),B_r(x_2))^{p\#(s+\tau)}}\underbrace{\fint_{\Lambda_r(t_0)}\frac{|(u)_{B_r(x_1)}(t)-(u)_{B_r(x_2)}(t)|^{p\#}}{r^{p\#(s+\tau)}}\,dt}_{=J}
\end{aligned}$$

for some constant $c = c(n,s,p)$. We now further estimate I_1, I_2 and J as below.

Estimates of I_1 and I_2. Using (5.45) and Lemma 2.1, we estimate I_d as

$$I_d \leq c\left(\frac{1}{\tau}\fint_{\mathcal{Q}_r(x_d,t_0)}|D^\tau d_s u|^p\,d\mu_{\tau,t}\right)\left(\sup_{t\in\Lambda_r(t_0)}\fint_{B_r(x_d)}\frac{\left|u-(u)_{Q_r(x_d,t_0)}\right|^2}{r^{2s+2\tau}}\,dx\right)^{\frac{sp}{n}},$$

where we have used the fact that

$$\sup_{t\in\Lambda_r(t_0)}\fint_{B_r(x_d)}\frac{\left|u-(u)_{B_r(x_d)}(t)\right|^2}{r^{2s+2\tau}}\,dx \leq \sup_{t\in\Lambda_r(t_0)}\fint_{B_r(x_d)}\frac{\left|u-(u)_{Q_r(x_d,t_0)}\right|^2}{r^{2s+2\tau}}\,dx.$$

Applying Young's inequality to the above inequality, we see that there is a constant $c = c(\tau)$ such that

$$\frac{r^{p\#(s+\tau)}}{\operatorname{dist}(B_r(x_1),B_r(x_2))^{p\#(s+\tau)}}\sum_{d=1}^{2}I_d \leq \frac{cr^{p\#(s+\tau)}}{\operatorname{dist}(B_r(x_1),B_r(x_2))^{p\#(s+\tau)}}\left[\sum_{d=1}^{2}E_{p,\tau}(u\,;\,Q_r(x_d,t_0))\right]^{p\#}$$

Estimate of J. We first note that

$$J \leq c \sum_{d=1}^{2} \sup_{t\in\Lambda_r(t_0)} \left(\fint_{B_r(x_d)} \frac{|u-(u)_{Q_r(x_d,t_0)}|^2}{r^{2s+2\tau}}\,dx \right)^{\frac{p_\#}{2}} + \frac{|(u)_{Q_r(x_1,t_0)}-(u)_{Q_r(x_2,t_0)}|^{p_\#}}{r^{p_\#(s+\tau)}},$$

where we denote the last term by $\hat{J}$. In light of Jensen's inequality, (5.46) and (5.47), we estimate $\hat{J}$ as

$$\hat{J} \leq \left(\fint_{Q_r(x_1,t_0)} \fint_{B_r(x_2)} |u(x,t)-u(y,t)|^p\,dy\,dz \right)^{\frac{p_\#}{p}} \leq \frac{c\,\mathrm{dist}\,(B_r(x_1),B_r(x_2))^{p_\#(s+\tau)}}{r^{p_\#(s+\tau)}} \left(\fint_{\mathcal{Q}} |D^\tau d_s u|^p\,d\mu_{\tau,t} \right)^{\frac{p_\#}{p}}.$$

We finally combine all the estimates I_1, I_2 and J to get the desired result (5.48). □

Remark 8. Let $\mathcal{Q}_{\overline{r}_j}(x_{1,j},x_{2,j},t_{0,j}) \in \mathcal{A}$. By (5.41) and (5.44), we deduce $\mathcal{Q}_{5^{\frac{1}{s}}\widetilde{r}_j}(x_{1,j},x_{2,j},t_{0,j})$ satisfies (5.45). In light of (5.36) and Lemma 5.2, there is a constant $c_{od}=c_{od}(n,s,p,\tau)$ such that

$$\left(\fint_{\mathcal{Q}_{5^{\frac{1}{s}}\widetilde{r}_j}(x_{1,j},x_{2,j},t_{0,j})} |D^\tau d_s u|^{p_\#}\,d\mu_{\tau,t} \right)^{\frac{1}{p_\#}} \leq c_{od}\lambda. \tag{5.48}$$

Step 5. Decomposition of the family $\mathcal{A}$. We now decompose the family $\mathcal{A}$ into subfamilies $AD_\lambda = \bigcap_{d=1}^{2} AD_\lambda^d$ and $ND_\lambda = \bigcup_{d=1}^{2} ND_\lambda^d$ for

$$AD_\lambda^d = \left\{ \mathcal{Q} = B^1 \times B^2 \times I \in \mathcal{A} : \quad E_{p,\tau}\left(u\,;\,B^d\times I\right) \leq \frac{\lambda}{16} \right\} \tag{5.49}$$

and

$$ND_\lambda^d = \left\{ \mathcal{Q} = B^1 \times B^2 \times I \in \mathcal{A} : \quad E_{p,\tau}\left(u\,;\,B^d\times I\right) > \frac{\lambda}{16} \right\}.$$

Recalling (5.35), (5.32) and (5.20), we observe that for any $\mathcal{Q}_{\overline{r}_j}(x_{1,j},x_{2,j},t_{0,j}) \in ND_\lambda^d$, there is an exit-radius $\rho_{(x_{d,j},t_{0,j})} \geq \overline{r}_j$ such that (5.20) holds with $z_0=(x_{d,j},t_{0,j})$. Thus, there is a cylinder $Q_{\rho_i}(z_i)$ which is selected in (5.21) such that

$$Q_{2^{j_0}\overline{r}_j}(x_{d,j},t_{0,j}) = Q_{\widetilde{r}_j}(x_{d,j},t_{0,j}) \subset Q_{2^{j_0}\rho_{(x_d,t_0)}}(x_{d,j},t_{0,j}) \subset Q_{5^{\frac{1}{s}}\rho_i}(z_i) \quad \text{and} \quad \frac{\rho_i}{2} \leq 2^{j_0}\rho_{(x_d,t_0)} \leq 2\rho_i. \tag{5.50}$$

We have used (5.22) to obtain the second observation in (5.50). Therefore, the set

$$\mathcal{A}^d_{i,j,l} := \left\{ \mathcal{Q}_{\overline{r}_j}(x_{1,j},x_{2,j},t_{0,j}) \in ND_\lambda^d \ : \ \begin{array}{l} Q_{\widetilde{r}_j}(x_{d,j},t_{0,j}) \subset Q_{5^{\frac{1}{s}}\rho_i}(z_i), \text{ where } i \text{ is the smallest integer} \\ \text{satisfying (5.50) with } \rho_i \leq 2^l\overline{r}_j < 2\rho_i \end{array} \right\} \tag{5.51}$$

is either a singleton or an empty set for any $l \geq 0$. Then we will verify

$$\rho_i < \mathrm{dist}(B_{\overline{r}_j}(x_{1,j}), B_{\overline{r}_j}(x_{2,j})). \tag{5.52}$$

Suppose not, then we have

$$|x_{d',j}-x_i| \leq |x_{d',j}-x_{d,j}| + |x_{d,j}-x_i| \leq 5\rho_i + 5^{\frac{1}{s}}\rho_i < 2\times 5^{\frac{1}{s}}\rho_i,$$

where $d' \in \{1,2\}\setminus\{d\}$. Consequently, we have

$$\mathcal{Q}_{\overline{r}_j}(x_{1,j},x_{2,j},t_{0,j}) \subset \mathcal{Q}_{5^{\frac{2}{s}}\rho_i}(z_i),$$

which contradicts the definition of $\mathcal{A}$ defined in (5.44), and the claim follows.

We next define for any $k \geq 0$,

$$\mathcal{A}^d_{i,j,l,k} = \left\{ \mathcal{Q}_{\overline{r}_j}(x_{1,j},x_{2,j},t_{0,j}) \in \mathcal{A}^d_{i,j,l} \ : \ 2^k\rho_i \leq \mathrm{dist}(B_{\overline{r}_j}(x_{1,j}), B_{\overline{r}_j}(x_{2,j})) < 2^{k+1}\rho_i \right\} \tag{5.53}$$

to see that

$$ND_\lambda^d = \bigcup_i \bigcup_j \bigcup_{l\geq 0, k\geq 0} \mathcal{A}^d_{i,j,l,k}. \tag{5.54}$$

We next observe that, for every $\mathcal{Q}_{\overline{r}_j}(x_{1,j},x_{2,j},t_{0,j}) \in ND^d_\lambda$, there holds

$$\begin{aligned}\mathfrak{A}\left(\mathcal{B}_{\overline{r}_j}(x_{1,j},x_{2,j})\right)^{s+\tau} E_p(u;\mathcal{Q}_{\overline{r}_j}(x_{d,j},t_{0,j})) &\leq c\mathfrak{A}\left(\mathcal{B}_{\overline{r}_j}(x_{1,j},x_{2,j})\right)\left(\fint_{\mathcal{Q}_{\widetilde{r}_j}(x_{d,j},t_{0,j})} |D^\tau d_s u|^p \, d\mu_{\tau,t}\right)^{\frac{1}{p}} \\ &+ \frac{c\mathfrak{A}\left(\mathcal{B}_{\overline{r}_j}(x_{1,j},x_{2,j})\right)}{\delta}\left(\fint_{\mathcal{Q}_{\widetilde{r}_j}(x_{d,j},t_{0,j})} |D^\tau d_0 f|^{\tilde{q}} \, d\mu_{\tau,t}\right)^{\frac{1}{\tilde{q}}} \\ &+ \frac{c\mathfrak{A}\left(\mathcal{B}_{\overline{r}_j}(x_{1,j},x_{2,j})\right)}{\delta}\left(\fint_{\mathcal{Q}_{\widetilde{r}_j}(x_{d,j},t_{0,j})} (\widetilde{r}_j^{s-\tau}|G|)^\gamma \, d\mu_{\tau,t}\right)^{\frac{1}{\gamma}} + \frac{\lambda}{100}\end{aligned} \tag{5.55}$$

for some constant $c = c(n,s,L,q,\tau)$, where the constant $\widetilde{r}_j$ is determined in (5.38). Indeed, by following the same lines as in the proof of (5.27), and using (5.13), (5.36) along with the fact that

$$\mathfrak{A}\left(\mathcal{B}_{\overline{r}_j}(x_{1,j},x_{2,j})\right) \leq \mathfrak{A}\left(\mathcal{B}_{c\overline{r}_j}(x_{1,j},x_{2,j})\right)$$

for any $c \geq 1$, we get

$$\begin{aligned}&\mathfrak{A}\left(\mathcal{B}_{\overline{r}_j}(x_{1,j},x_{2,j})\right)\tau^{\frac{1}{\gamma}}\left(\sup_{t\in\Lambda_r(t_0)}\fint_{B_r(x_0)} \frac{|u-(u)_{Q_r(z_0)}|^2}{r^{2s+2\tau}}\,dx\right)^{\frac{1}{2}} \\ &\leq c\mathfrak{A}\left(\mathcal{B}_{\overline{r}_j}(x_{1,j},x_{2,j})\right)\left(\fint_{\mathcal{Q}_{2^{j_0}\overline{r}_j}(x_{d,j},t_{0,j})} |D^\tau d_s u|^p \, d\mu_{\tau,t}\right)^{\frac{1}{p}} \\ &\quad+ \frac{c\mathfrak{A}\left(\mathcal{B}_{\overline{r}_j}(x_{1,j},x_{2,j})\right)}{\delta}\left[\left(\fint_{\mathcal{Q}_{2^{j_0}\overline{r}_j}(x_{d,j},t_{0,j})} |D^\tau d_0 f|^{\tilde{q}} \, d\mu_{\tau,t}\right)^{\frac{1}{\tilde{q}}} + \left(\fint_{\mathcal{Q}_{2\overline{r}_j}(x_{d,j},t_{0,j})} (r_j^{s-\tau}|G|)^\gamma \, d\mu_{\tau,t}\right)^{\frac{1}{\gamma}}\right] + \frac{\lambda}{100}\end{aligned}$$

for some constant $c = c(n,s,L,q,\tau)$. Thus, using (5.32), (5.13) and (5.38), we get the desired estimate.

We end this step with the following lemma which is an essential ingredient for the next step.

Lemma 5.3. *Let us fix $i,l,k \geq 0$ and $d \in \{1,2\}$. Then there is a constant $c = c(n,s,L,q,\tau)$ such that*

$$\sum_{\mathcal{Q}_{\overline{r}_j}(x_{1,j},x_{2,j},t_{0,j}) \in \bigcup\limits_{\tilde{j}\geq 0} \mathcal{A}^d_{i,\tilde{j},l,k}} \int_{\mathcal{Q}_{\widetilde{r}_j}(x_{d,j},t_{0,j})} |D^\tau d_s u|^p \, d\mu_{\tau,t} \leq c2^{n(l+k)} \int_{\mathcal{Q}_{5^{\frac{1}{s}}\rho_i}(z_i)} |D^\tau d_s u|^p \, d\mu_{\tau,t}, \tag{5.56}$$

$$\sum_{\mathcal{Q}_{\overline{r}_j}(x_{1,j},x_{2,j},t_{0,j}) \in \bigcup\limits_{\tilde{j}\geq 0} \mathcal{A}^d_{i,\tilde{j},l,k}} \int_{\mathcal{Q}_{\widetilde{r}_j}(x_{d,j},t_{0,j})} |D^\tau d_0 f|^{\tilde{q}} \, d\mu_{\tau,t} \leq c2^{n(l+k)} \int_{\mathcal{Q}_{5^{\frac{1}{s}}\rho_i}(z_i)} |D^\tau d_0 f|^{\tilde{q}} \, d\mu_{\tau,t} \tag{5.57}$$

and

$$\sum_{\mathcal{Q}_{\overline{r}_j}(x_{1,j},x_{2,j},t_{0,j}) \in \bigcup\limits_{\tilde{j}\geq 0} \mathcal{A}^d_{i,\tilde{j},l,k}} \int_{\mathcal{Q}_{\widetilde{r}_j}(x_{d,j},t_{0,j})} (\widetilde{r}_j^{s-\tau}|G|)^\gamma \, d\mu_{\tau,t} \leq c2^{n(l+k)} \int_{\mathcal{Q}_{5^{\frac{1}{s}}\rho_i}(z_i)} (\rho_i^{s-\tau}|G|)^\gamma \, d\mu_{\tau,t}. \tag{5.58}$$

Proof. We will prove only for $d=1$ and we first claim that if

$$\bigcap_{\substack{j\in J \\ \mathcal{Q}_{\overline{r}_j}(x_{1,j},x_{2,j},t_{0,j}) \in \bigcup\limits_{\tilde{j}\geq 0} \mathcal{A}^1_{i,\tilde{j},l,k}}} \mathcal{Q}_{\widetilde{r}_j}(x_{1,j},t_{0,j}) \neq \emptyset \tag{5.59}$$

holds for some index set J, then $|J| \leq c2^{n(l+k)}$ for some constant c depending only on n,s,L,q and τ, where $|J|$ denotes the number of elements in the set J. To do this, suppose that $(x_1,t_0) \in \bigcap_{j\in J} Q_{\widetilde{r}_j}(x_{1,j},t_{0,j})$. Then, by the definition of the set $\mathcal{A}^1_{i,j,l,k}$ given in (5.53), we get

$$\operatorname{dist}(x_1, B_{\widetilde{r}_j}(x_{2,j})) < 2^{k+1}\rho_i + 2\widetilde{r}_j < c2^k\rho_i \tag{5.60}$$

for some constant $c = c(n,s,L,q,\tau)$, where we have used the fact that $\overline{r}_j \leq 2^{1-l}\rho_i$ by (5.51), (5.38) and (5.13). Since $\left\{\mathcal{Q}_{\widetilde{r}_j}(x_{1,j},x_{2,j},t_{0,j})\right\}_{j\in J}$ is a collection of mutually disjoint sets, we note from (5.59) that $\left\{B_{\widetilde{r}_j}(x_{2,j})\right\}_{j\in J}$ also consists of mutually disjoint sets. This along with (5.60) implies

$$|J| \leq \frac{\left|B_{c2^k\rho_i}\right|}{\left|B_{\widetilde{r}_j}\right|} \leq c2^{n(k+l)}$$

for some constant $c = c(n,s,L,q,\tau)$, where we have used (5.38), (5.13) and the relation between $\overline{r}_j$ and ρ_i given in (5.51). This proves (5.59). We are now ready to prove (5.56) using an inductive argument. We first note that the number of elements in $\bigcup_j \mathcal{A}^1_{i,j,l,k}$ is finite, as each cylinder in $\bigcup_j \mathcal{A}^1_{i,j,l,k}$ is of the form $\mathcal{Q} = \mathcal{Q}_r$, where $2^{-l}\rho_i < r \le 2^{-l+1}\rho_i$, is mutually disjoint and is contained in $\mathcal{Q}_{r_2}$. Let us denote j_i as the number of elements in $\bigcup_j \mathcal{A}^1_{i,j,l,k}$. For a clear notation, we assume $\mathcal{A}^1_{i,j,l,k} \neq \emptyset$ if $j \le j_i$ and $\mathcal{A}^1_{i,j,l,k} = \emptyset$ if $j \ge j_i + 1$. By (5.50), we note $\mathcal{Q}_{\widetilde{r}_j}(x_{1,j}, t_{0,j}) \subset Q_{5^{\frac{1}{s}}\rho_i}(z_i)$. We first define

$$\mathcal{D}_1 = \left\{ \mathcal{Q}_{5^{\frac{1}{s}}\rho_i}(z_i) \cap \mathcal{Q}_{\widetilde{r}_j}(x_{1,j}, t_{0,1}),\ \mathcal{Q}_{5^{\frac{1}{s}}\rho_i}(z_i) \setminus \mathcal{Q}_{\widetilde{r}_j}(x_{1,1}, t_{0,1}) \right\}.$$

Suppose $\mathcal{D}_k$ is determined, for some $k \ge 1$, then we define

$$\mathcal{D}_{k+1} = \bigcup_{\mathcal{F} \in \mathcal{D}_k} \left\{ \mathcal{F} \cap \mathcal{Q}_{\widetilde{r}_{k+1}}(x_{1,k+1}, t_{0,k+1}),\ \mathcal{F} \setminus \mathcal{Q}_{\widetilde{r}_{k+1}}(x_{1,k+1}, t_{0,k+1}) \right\}.$$

In this way, we obtain a collection $\mathcal{D}_{j_i}$ such that for any choice of two elements $\mathcal{F}$ and $\mathcal{F}'$ in $\mathcal{D}_{j_i}$, either $\mathcal{F} = \mathcal{F}'$ or $\mathcal{F} \cap \mathcal{F}' = \emptyset$. Thus, we can write

$$\mathcal{Q}_{5^{\frac{1}{s}}\rho_i}(z_i) = \bigcup_{\mathcal{F}_m \in \mathcal{D}_{j_i}} \mathcal{F}_m,$$

where $\mathcal{F}_m$'s are mutually disjoint, and for any m and $j \in [1, j_i]$, either

$$\mathcal{F}_m \subset \mathcal{Q}_{\widetilde{r}_j}(x_{1,j}, t_{0,j}) \quad \text{or} \quad \mathcal{F}_m \cap \mathcal{Q}_{\widetilde{r}_j}(x_{1,j}, t_{0,j}) = \emptyset.$$

Moreover, there are mutually disjoint elements $\mathcal{F}_m \in \mathcal{D}_{j_i}$ such that

$$\mathcal{Q}_{\widetilde{r}_j}(x_{1,j}, t_{0,j}) = \bigcup_{\mathcal{F}_m \subset \mathcal{Q}_{\widetilde{r}_j}(x_{1,j}, t_{0,j})} \mathcal{F}_m.$$

We note that for each $\mathcal{F}_m$, the number of elements in $J := \{j \in [1, j_i] : \mathcal{F}_m \subset \mathcal{Q}_{\widetilde{r}_j}(x_{1,j}, t_{0,j})\}$ is at most $c2^{n(k+l)}$, due to (5.59). As a result, by using the fact that $\widetilde{r}_j \le 2\rho_i$ (thanks to (5.50)), we have

$$\begin{aligned}
\sum_{\mathcal{Q}_{\widetilde{r}_j}(x_{1,j}, x_{2,j}, t_{0,j}) \in \bigcup_{\tilde{j}} \mathcal{A}^1_{i,\tilde{j},l,k}} \int_{\mathcal{Q}_{\widetilde{r}_j}(x_{1,j}, t_{0,j})} |D^\tau d_s u|^p \, d\mu_{\tau,t} &= \sum_{j=1}^{j_i} \sum_{\substack{m \\ \mathcal{F}_m \subset \mathcal{Q}_{\widetilde{r}_j}(x_{1,j}, t_{0,j})}} \int_{\mathcal{F}_m} |D^\tau d_s u|^p \, d\mu_{\tau,t} \\
&= \sum_m \sum_{\substack{j=1 \\ \mathcal{F}_m \subset \mathcal{Q}_{\widetilde{r}_j}(x_{1,j}, t_{0,j})}}^{j_i} \int_{\mathcal{F}_m} |D^\tau d_s u|^p \, d\mu_{\tau,t} \\
&\le c2^{n(l+k)} \sum_m \int_{\mathcal{F}_m} |D^\tau d_s u|^p \, d\mu_{\tau,t} \\
&\le c2^{n(l+k)} \int_{\mathcal{Q}_{5^{\frac{1}{s}}\rho_i}(z_i)} |D^\tau d_s u|^p \, d\mu_{\tau,t}
\end{aligned}$$

for some constant $c = c(n,s,L,q,\tau)$, which implies the desired result (5.56). By following the same lines as in the proof of (5.56) we can prove (5.57) and (5.58) by replacing $(D^\tau d_s u, p)$ with $(D^\tau d_0 f, \tilde{q})$ and $((\widetilde{r}_j^{\,s-\tau}|G|)^\gamma, 1)$, respectively. □

Step 6. Measure estimate of $\mathcal{Q} \in AD_\lambda$. We claim that for every $\mathcal{Q} \in AD_\lambda$, there holds

$$\mu_{\tau,t}(\mathcal{Q}) \le \frac{2^q}{\lambda^p} \int_{\mathcal{Q} \cap \left\{|D^\tau d_s u| > \frac{\lambda}{16}\right\}} |D^\tau d_s u|^p \, d\mu_{\tau,t}. \tag{5.61}$$

Indeed, on account of (5.36), we have

$$\lambda^p \le 2^p \left(\fint_{\mathcal{Q}} |D^\tau d_s u|^p \, d\mu_{\tau,t} \right) + 2^p \left(\mathfrak{A}(\mathcal{K})\right)^p \left[\sum_{d=1}^{2} E_{p,\tau}\left(u\,;\, B^d \times I\right) \right]^p. \tag{5.62}$$

Using (5.62), (5.49) and the fact that

$$\left(\fint_{\mathcal{Q}} |D^\tau d_s u|^p \, d\mu_{\tau,t} \right) \le \frac{\lambda^p}{8^p} + \frac{1}{\mu_{\tau,t}(\mathcal{Q})} \int_{\mathcal{Q} \cap \left\{|D^\tau d_s u| > \frac{\lambda}{4}\right\}} |D^\tau d_s u|^p \, d\mu_{\tau,t},$$

we deduce that

$$\frac{\lambda^p}{2^p} \le \frac{1}{\mu_{\tau,t}(\mathcal{Q})} \int_{\mathcal{Q} \cap \left\{|D^\tau d_s u| > \frac{\lambda}{16}\right\}} |D^\tau d_s u|^p \, d\mu_{\tau,t}.$$

This proves the claim of (5.61) as $p \le q$.
Step 7. Measure estimate of $\mathcal{Q} \in ND_\lambda$. Our next aim is to estimate $\mu_{\tau,t}(\mathcal{Q})$ for $\mathcal{Q} \in ND_\lambda$. With the aid of (5.61), (5.54) and Lemma 5.3, we prove the following result.

Lemma 5.4. *There exists a constant $c = c(n,s,L,q,\tau)$ such that*

$$\sum_{\mathcal{Q}\in ND_\lambda} \mu_{\tau,t}(\mathcal{Q}) \le \frac{c}{\lambda^p} \int_{\mathcal{Q}_{r_2}\cap\{|D^\tau d_s u| > \frac{\lambda}{16}\}} |D^\tau d_s u|^p \, d\mu_{\tau,t} + c \sum_i \mu_{\tau,t}\left(\mathcal{Q}_{\rho_i}(z_i)\right). \tag{5.63}$$

Proof. Let $\mathcal{Q} \equiv \mathcal{Q}_r(x_1,x_2,t_0) \in ND_\lambda$. For convenience in notation, we set

$$\mathbb{E}^d_u(\mathcal{Q}) = \mathfrak{A}\left(\mathcal{B}_r(x_1,x_2)\right)^{p(s+\tau)} \left(\fint_{\mathcal{Q}_{2^{j_0}r}(x_d,t_0)} |D^\tau d_s u|^p \, d\mu_{\tau,t} \right),$$

$$\mathbb{E}^d_f(\mathcal{Q}) = \frac{\mathfrak{A}\left(\mathcal{B}_r(x_1,x_2)\right)^{\tilde{q}(s+\tau)}}{\delta^{\tilde{q}}} \left(\fint_{2^{j_0}\mathcal{Q}_r(x_d,t_0)} |D^\tau d_0 f|^{\tilde{q}} \, d\mu_{\tau,t} \right)$$

and

$$\mathbb{E}^d_g(\mathcal{Q}) = \frac{\mathfrak{A}\left(\mathcal{B}_r(x_1,x_2)\right)^{\gamma(s+\tau)}}{\delta^{\gamma}} \left(\fint_{\mathcal{Q}_{2^{j_0}r}(x_d,t_0)} ((2^{j_0}r)^{s-\tau}|G|)^\gamma \, d\mu_{\tau,t} \right),$$

where the set function $\mathfrak{A}(\cdot)$ is defined in (5.33). We then note from (5.39), (5.49), (5.38) and (5.55) that there is a constant $c = c(n,s,L,q,\tau)$ such that

$$\frac{\lambda}{4} \le \left(\frac{1}{\mu_{\tau,t}(\mathcal{Q})} \int_{\mathcal{Q}\cap\left\{|D^\tau d_s u| > \frac{\lambda}{16}\right\}} |D^\tau d_s u|^p \, d\mu_{\tau,t} \right)^{\frac{1}{p}} + c \sum_{d=1}^{2} \left[\left(\mathbb{E}^d_u(\mathcal{Q})\right)^{\frac{1}{p}} + \left(\mathbb{E}^d_f(\mathcal{Q})\right)^{\frac{1}{\tilde{q}}} + \left(\mathbb{E}^d_g(\mathcal{Q})\right)^{\frac{1}{\gamma}} \right],$$

if $\mathcal{Q} \in ND^1_\lambda \cap ND^2_\lambda$ and

$$\frac{\lambda}{4} \le \left(\frac{1}{\mu_{\tau,t}(\mathcal{Q})} \int_{\mathcal{Q}\cap\left\{|D^\tau d_s u| > \frac{\lambda}{16}\right\}} |D^\tau d_s u|^p \, d\mu_{\tau,t} \right)^{\frac{1}{p}} + c \left[\left(\mathbb{E}^d_u(\mathcal{Q})\right)^{\frac{1}{p}} + \left(\mathbb{E}^d_f(\mathcal{Q})\right)^{\frac{1}{\tilde{q}}} + \left(\mathbb{E}^d_g(\mathcal{Q})\right)^{\frac{1}{\gamma}} \right],$$

if $\mathcal{Q} \in ND^d_\lambda \cap AD^{d'}_\lambda$, where we denote $d' \in \{1,2\} \setminus \{d\}$. Using (5.40) and then multiplying $\mu_{\tau,t}(\mathcal{Q})$ along with a few simple calculations, we obtain

$$\begin{aligned} \mu_{\tau,t}(\mathcal{Q}) &\le \frac{c}{\lambda^p} \int_{\mathcal{Q}\cap\left\{|D^\tau d_s u| > \frac{\lambda}{16}\right\}} |D^\tau d_s u|^p \, d\mu_{\tau,t} \\ &\quad + \mu_{\tau,t}(\mathcal{Q}) \left[\frac{c}{\lambda^p} \sum_{d=1}^{2} \mathbb{E}^d_u(\mathcal{Q}) + \frac{c}{\lambda^{\tilde{q}}} \sum_{d=1}^{2} \mathbb{E}^d_f(\mathcal{Q}) + \frac{c}{\lambda^\gamma} \sum_{d=1}^{2} \mathbb{E}^d_g(\mathcal{Q}) \right] \end{aligned} \tag{5.64}$$

if $\mathcal{Q} \in ND^1_\lambda \cap ND^2_\lambda$ and

$$\begin{aligned} \mu_{\tau,t}(\mathcal{Q}) &\le \frac{c}{\lambda^p} \int_{\mathcal{Q}\cap\left\{|D^\tau d_s u| > \frac{\lambda}{16}\right\}} |D^\tau d_s u|^p \, d\mu_{\tau,t} \\ &\quad + \mu_{\tau,t}(\mathcal{Q}) \left[\frac{c}{\lambda^p} \mathbb{E}^d_u(\mathcal{Q}) + \frac{c}{\lambda^{\tilde{q}}} \mathbb{E}^d_f(\mathcal{Q}) + \frac{c}{\lambda^\gamma} \mathbb{E}^d_g(\mathcal{Q}) \right] \end{aligned} \tag{5.65}$$

if $\mathcal{Q} \in ND^d_\lambda \cap AD^{d'}_\lambda$. We next observe from (5.13), (2.3), (2.4) and (5.47) that

$$\frac{\mu_{\tau,t}(\mathcal{Q})}{\mu_{\tau,t}\left(\mathcal{Q}_{2^{j_0}r}(x_d,t_0)\right)} \le c \frac{\mu_{\tau,t}(\mathcal{Q})}{\mu_{\tau,t}\left(\mathcal{Q}_r(x_d,t_0)\right)} \le c\tau \left(\frac{r}{\operatorname{dist}(B_r(x_1), B_r(x_2))} \right)^{n-2\tau} \tag{5.66}$$

for some constant $c = c(n,s,L,q,\tau)$. On account of (5.51), (5.53), (5.54), (5.66) and (5.56), we find that

$$\begin{aligned} \sum_{\mathcal{Q}\in ND^d_\lambda} \mu_{\tau,t}(\mathcal{Q}) \mathbb{E}^d_u(\mathcal{Q}) &\le c \sum_{i,j,k,l} \sum_{\mathcal{Q}\in \mathcal{A}^d_{i,j,l,k}} \left(\frac{r}{\operatorname{dist}(B_r(x_1), B_r(x_2))} \right)^{n-2\tau+ps+p\tau} \left(\int_{2^{j_0}P^d\mathcal{Q}} |D^\tau d_s u|^p \, d\mu_{\tau,t} \right) \\ &\le c \sum_{i,l,k\ge 0} \left(2^{-(l+k)} \right)^{ps} \left(\int_{\mathcal{Q}_{5^{\frac{1}{s}}\rho_i}(z_i)} |D^\tau d_s u|^p \, d\mu_{\tau,t} \right) \\ &\le c \sum_{i\ge 0} \left(\int_{\mathcal{Q}_{5^{\frac{1}{s}}\rho_i}(z_i)} |D^\tau d_s u|^p \, d\mu_{\tau,t} \right), \end{aligned} \tag{5.67}$$

where we denote $P^d\mathcal{Q} = \mathcal{Q}_r(x_d,t_0)$ for a given cylinder $\mathcal{Q} = \mathcal{Q}_r(x_1,x_2,t_0)$. For the last inequality, we have used the fact that

$$\sum_{i,j\geq 0}\left(2^{-(i+j)}\right)^a \leq \frac{2^a}{a\ln 2} \quad (a>0).$$

Therefore, using (5.20), we obtain

$$\sum_{\mathcal{Q}\in ND^d_\lambda}\mu_{\tau,t}(\mathcal{Q})\mathbb{E}^d_u(\mathcal{Q}) \leq c\lambda^p\sum_i \mu_{\tau,t}\left(\mathcal{Q}_{\rho_i}(z_i)\right). \tag{5.68}$$

As in (5.67) along with (5.57), (5.58) and (5.20), we have

$$\begin{aligned}\sum_{\mathcal{Q}\in ND^d_\lambda}\mu_{\tau,t}(\mathcal{Q})\mathbb{E}^d_f(\mathcal{Q}) &\leq c\sum_{i,l,k\geq 0}\left(2^{-(l+k)}\right)^{\tilde{q}s}\delta^{-\tilde{q}}\int_{\mathcal{Q}_{5^{\frac{1}{s}}\rho_i}(x_d,t_0)}|D^\tau d_0 f|^{\tilde{q}}\,d\mu_{\tau,t}\\ &\leq c\lambda^{\tilde{q}}\sum_i \mu_{\tau,t}\left(\mathcal{Q}_{\rho_i}(z_i)\right)\end{aligned}$$

and

$$\begin{aligned}\sum_{\mathcal{Q}\in ND^d_\lambda}\mu_{\tau,t}(\mathcal{Q})\mathbb{E}^d_f(\mathcal{Q}) &\leq c\sum_{i,l,k\geq 0}\left(2^{-(l+k)}\right)^{\gamma s}\delta^{-\gamma}\int_{\mathcal{Q}_{5^{\frac{1}{s}}\rho_i}(x_d,t_0)}\left((\rho_i)^{s-\tau}|G|\right)^\gamma d\mu_{\tau,t}\\ &\leq c\lambda^{\gamma}\sum_i \mu_{\tau,t}\left(\mathcal{Q}_{\rho_i}(z_i)\right)\end{aligned}$$

for some constant $c = c(n,s,L,q,\tau)$. Combining the above two inequalities, we get

$$\sum_{\mathcal{Q}\in ND^d_\lambda}\mu_{\tau,t}(\mathcal{Q})\left[\frac{1}{\lambda^{\tilde{q}}}\mathbb{E}^d_f(\mathcal{Q}) + \frac{1}{\lambda^\gamma}\mathbb{E}^d_g(\mathcal{Q})\right] \leq c\sum_i \mu_{\tau,t}\left(\mathcal{Q}_{\rho_i}(z_i)\right), \tag{5.69}$$

where we denote $\mathcal{Q} = B^1\times B^2\times\Lambda$. Plugging (5.68) and (5.69) into (5.64) and (5.65), we obtain (5.63), which completes the proof of the lemma. □

Step 8. Completion of the proof. Considering (5.18) and (5.35), the constant $\frac{M}{\kappa}$ depends only on n,s,L,q and τ. Therefore, if $\lambda\geq\lambda_0$ which is determined in (5.12), then we find two families of countable disjoint cylinders

$$\{\mathcal{Q}_{\rho_i}(z_i)\}_i \quad\text{and}\quad \left\{\mathcal{Q}_{\tilde{r}_j}(x_{1,j},x_{2,j},t_{0,j}\right\}_{\mathcal{Q}_{\tilde{r}_j}(x_{1,j},x_{2,j},t_{0,j})\in\mathcal{A}}$$

so that

$$\{(x,y,t)\in\mathcal{Q}_{r_1}\,:\,|D^\tau d_s u(x,y,t)|\geq\lambda\}\subset\bigcup_i \mathcal{Q}_{5^{\frac{2}{s}}\rho_i}(z_i)\cup\bigcup_{\substack{j\\ \mathcal{Q}_{\tilde{r}_j}(x_{1,j},x_{2,j},t_{0,j})\in\mathcal{A}}}\mathcal{Q}_{5^{\frac{1}{s}}\tilde{r}_j}(x_{1,j},x_{2,j},t_{0,j}),$$

which follows from the steps 2-3 along with (5.22), (5.37) and (5.44). In addition, using (5.48), (5.61), (5.63) and (5.31) along with the choice of the constants $a_u = \tilde{a}_u\kappa$, $a_f = \tilde{a}_f\kappa$ and $a_g = \tilde{a}_g\kappa$ given in (5.31), we get (5.10) and (5.8). This completes the proof. □

Before ending this section, we give some estimates which are useful in the context of the comparison Lemma 4.4.

Remark 9. Let $\mathcal{Q}_{\rho_i}(z_i)$ be the cylinder chosen in Lemma 5.1. Then we want to show that

$$\frac{1}{\tau}\fint_{Q_{4\times 5^{\frac{2}{s}}\rho_i}(z_i)}|D^\tau d_s u|^2\,d\mu_{\tau,t} + \mathrm{Tail}_{\infty,2s}\left(\frac{u-(u)_{B_{4\times 5^{\frac{2}{s}}\rho_i}(x_i)}(t)}{(4\times 5^{\frac{2}{s}}\rho_i)^{s+\tau}};Q_{4\times 5^{\frac{2}{s}}\rho_i}(z_i)\right)^2 \leq (c\lambda)^2 \tag{5.70}$$

and

$$\begin{aligned}&\left(\frac{1}{\tau}\fint_{Q_{4\times 5^{\frac{2}{s}}\rho_i}(z_i)}\left((4\times 5^{\frac{2}{s}}\rho_i)^{s-\tau}|G|\right)^\gamma d\mu_{\tau,t}\right)^{\frac{2}{\gamma}} + \frac{1}{\tau}\fint_{Q_{4\times 5^{\frac{2}{s}}\rho_i}(z_i)}|D^\tau d_0 f|^2\,d\mu_{\tau,t}\\ &\qquad + \mathrm{Tail}_{2,s}\left(\frac{f-(f)_{B_{4\times 5^{\frac{2}{s}}\rho_i}(x_i)}(t)}{(4\times 5^{\frac{2}{s}}\rho_i)^{\tau}};Q_{4\times 5^{\frac{2}{s}}\rho_i}(z_i)\right)^2 \leq (c\lambda\delta)^2\end{aligned} \tag{5.71}$$

for some constant $c = c(n, s, L, q, \tau)$. We first note from (5.14), (5.20) and (5.23) that there is a natural number l such that

$$5^{\frac{2}{s}} \times 2^{j_0+2}\mathcal{R}_{1,2} < 2^l \times \overline{\rho} \leq 5^{\frac{2}{s}} \times 2^{j_0+3}\mathcal{R}_{1,2},$$

where we denote $\overline{\rho} = 4 \times 5^{\frac{2}{s}}\rho_i$ for a clear notation. After a few modifications of the proof for (2.6) with $p = \infty$, we deduce that

$$\begin{aligned}
&\mathrm{Tail}_{\infty,2s}\left(\frac{u-(u)_{B_{\overline{\rho}}(x_i)}(t)}{\overline{\rho}^{s+\tau}}; Q_{\overline{\rho}}(z_i)\right) \\
&\leq c\left[\sum_{j=1}^{l} 2^{j(-s+\tau)} \sup_{t\in\Lambda_{2^j\overline{\rho}}(t_i)} \fint_{B_{2^j\overline{\rho}}(x_i)} \frac{|u-(u)_{B_{2^j\overline{\rho}}(x_i)}(t)|}{(2^j\overline{\rho})^{s+\tau}}\,dx + \frac{2^{-2sl+s+\tau}}{\overline{\rho}^{s+\tau}} \sup_{t\in\Lambda_2} \fint_{B_2} \frac{|u-(u)_{B_2}(t)|}{2^{s+\tau}}\,dx\right] \\
&\quad + \frac{c}{(2-r_1)^{n+2s}}\left(\frac{2}{\overline{\rho}}\right)^{-s+\tau} \mathrm{Tail}_{\infty,2s}\left(\frac{u-(u)_{B_2}(t)}{2^{s+\tau}}; Q_2\right),
\end{aligned}$$

where $c = c(n, s)$. Then using (5.7), (5.20) and the fact that

$$2^{-2sl}\left(\frac{2}{\overline{\rho}}\right)^{s+\tau} \leq \frac{c}{\mathcal{R}_{1,2}^{\frac{5n}{s}}} \quad \text{and} \quad \sum_{i=1}^{\infty} 2^{i(-s+\tau)} \leq c(n, s, \tau),$$

we estimate the above term as

$$\mathrm{Tail}_{\infty,2s}\left(\frac{u-(u)_{B_{\overline{\rho}}(x_i)}(t)}{\overline{\rho}^{s+\tau}}; Q_{\overline{\rho}}(z_i)\right) \leq c\lambda, \tag{5.72}$$

where $c = c(n, s, L, q, \tau)$. In addition, Hölder's inequality and (5.20) imply

$$\left(\frac{1}{\tau}\fint_{Q_{\overline{\rho}}(z_i)} |D^\tau d_s u|^2\,d\mu_{\tau,t}\right)^{\frac{1}{2}} \leq c\left(\fint_{Q_{\overline{\rho}}(z_i)} |D^\tau d_s u|^p\,d\mu_{\tau,t}\right)^{\frac{1}{p}} \leq c\lambda. \tag{5.73}$$

We combine the above two estimates to obtain (5.70). Similarly, with the aid of (1.7) and (2.6) along with (5.4), we get that

$$\begin{aligned}
&\mathrm{Tail}_{2,s}\left(\frac{f-(f)_{B_{\overline{\rho}}(x_i)}(t)}{(\overline{\rho})^{\tau}}\,;\,Q_{\overline{\rho}}(z_i)\right) \\
&\leq \mathrm{Tail}_{\tilde{q},s}\left(\frac{f-(f)_{B_{\overline{\rho}}(x_i)}(t)}{(\overline{\rho})^{\tau}}\,;\,Q_{\overline{\rho}}(z_i)\right) \\
&\leq c_q\sum_{j=1}^{l} 2^{i\left(-s+\tau+\frac{2s}{\tilde{q}}\right)}\left(\frac{1}{\tau}\fint_{\mathcal{Q}_{2^j\overline{\rho}}(z_i)} |D^\tau d_0 f|^{\tilde{q}}\,d\mu_{\tau,t}\right)^{\frac{1}{\tilde{q}}} + c_q 2^{-sl}\left(\frac{2}{\overline{\rho}}\right)^{\tau+s}\left(\frac{1}{\tau}\fint_{Q_2} |D^\tau d_0 f|^q\,d\mu_{\tau,t}\right)^{\frac{1}{q}} \\
&\quad + \frac{\tilde{c}}{(2-r_1)^{n+2s}}\left(\frac{2}{\overline{\rho}}\right)^{-s+\tau+\frac{2s}{\tilde{q}}} \mathrm{Tail}_{q,s}\left(\frac{f-(f)_{B_2}(t)}{2^{\tau}}; Q_2\right).
\end{aligned}$$

Hence, as in (5.72) and (5.73), there is a constant $c = c(n, s, L, q, \tau)$ such that

$$\left(\frac{1}{\tau}\fint_{Q_{\overline{\rho}}(z_i)} |D^\tau d_0 f|^2\,d\mu_{\tau,t}\right)^{\frac{1}{2}} + \mathrm{Tail}_{2,s}\left(\frac{f-(f)_{B_{\overline{\rho}}(x_i)}(t)}{(\overline{\rho})^{\tau}}\,;\,Q_{\overline{\rho}}(z_i)\right) \leq c\delta\lambda.$$

Additionally, (5.20) and (5.35) yield that

$$\left(\frac{1}{\tau}\fint_{Q_{4\times 5^{\frac{2}{s}}\rho_i}(z_i)} \left((4\times 5^{\frac{2}{s}}\rho_i\rho_i)^{s-\tau}|G|\right)^{\gamma}\,d\mu_{\tau,t}\right)^{\frac{2}{\gamma}} \leq c\delta\lambda.$$

We combine the above two inequality to show that (5.71) holds.

6. L^q-estimate of $d_s u$

In this section, we prove our main theorem. Since we have established comparison estimates and constructed coverings of upper level sets, we are able to obtain L^q-estimate of $d_s u$ with the estimate (1.9) via a bootstrap argument as in [13, Theorem 1.2]. Let us define

$$p_h = 2\left(1+\frac{s}{n}\right)^h, \qquad \text{for } h = 0, 1, 2, \ldots. \tag{6.1}$$

Then there is a positive integer h_q such that

$$p_{h_q-1} < q \leq p_{h_q}. \tag{6.2}$$

We now prove the following lemma which is an essential ingredient to use a boot strap argument.

Lemma 6.1. *Let $h \in \{0,1,\dots,h_q-1\}$. Suppose that u is a weak solution to (5.1) with*

$$f \in L^q\left(\Lambda_2; L^1_s(\mathbb{R}^n)\right) \quad \text{and} \quad \int_{\mathcal{Q}_2} |D^\tau d_s u|^{p_h} + |D^\tau d_0 f|^q + |G|^{\frac{q\gamma}{b_\tau}}\, d\mu_{\tau,t} < \infty.$$

Then there is a sufficiently small $\delta = \delta(n,s,L,q,\tau) \in (0,1]$ independent of h such that if A is $(\delta,2)$-vanishing in Q_2, then we have that

$$\begin{aligned}\left(\fint_{\mathcal{Q}_1} |D^\tau d_s u|^{\tilde p}\, d\mu_{\tau,t}\right)^{\frac{1}{\tilde p}} &\le c\left(\left(\fint_{\mathcal{Q}_2} |D^\tau d_s u|^{p_h}\, d\mu_{\tau,t}\right)^{\frac{1}{p_h}} + \mathrm{Tail}_{\infty,2s}\left(\frac{u-(u)_{B_2}(t)}{2^{s+\tau}}; Q_2\right)\right)\\ &\quad + c\left(\sup_{t\in\Lambda_2}\fint_{B_2}\frac{|u-(u)_{Q_2}|^2}{2^{2s+2\tau}}\,dx\right)^{\frac12} + c\left(\fint_{\mathcal{Q}_2}\left(2^{s-\tau}|G|\right)^{\frac{\tilde p\gamma}{b_\tau}}\, d\mu_{\tau,t}\right)^{\frac{b_\tau}{\tilde p\gamma}} \\ &\quad + c\left(\left(\fint_{\mathcal{Q}_2} |D^\tau d_0 f|^q\, d\mu_{\tau,t}\right)^{\frac1q} + \mathrm{Tail}_{q,s}\left(\frac{f-(f)_{B_2}(t)}{2^\tau}; Q_2\right)\right)\end{aligned} \tag{6.3}$$

for some constant $c = c(n,s,L,q,\tau)$, where the constant b_τ is defined in (5.9) and

$$\tilde p = \begin{cases} p_{h+1} & \text{if } h < h_q - 1,\\ q & \text{if } h = h_q - 1.\end{cases} \tag{6.4}$$

Proof. Let us first fix $1 \le r_1 < r_2 \le 2$ and $\epsilon > 0$. Then we select λ_0 as given in Lemma 5.1 with $p = p_h$ and select $\delta = \delta(n,s,L,\epsilon)$ determined in Lemma 4.4. For any $N \ge \lambda_0$, we define a function $\phi : [1,2] \to \mathbb{R}$ by

$$\phi_N(r) = \left(\fint_{\mathcal{Q}_r} |D^\tau d_s u|_N^{\tilde p}\, d\mu_{\tau,t}\right)^{\frac{1}{\tilde p}}, \tag{6.5}$$

where $|D^\tau d_s u|_N = \min\{|D^\tau d_s u|, N\}$. We now claim that for $N \ge \lambda_0$, the following holds

$$\phi_N(r_1) \le \frac{\phi_N(r_2)}{2} + c\lambda_0 + c\left(\fint_{\mathcal{Q}_2}\left(2^{s-\tau}|G|\right)^{\frac{\tilde p\gamma}{b_\tau}}\, d\mu_{\tau,t}\right)^{\frac{b_\tau}{\tilde p\gamma}} \tag{6.6}$$

for some constant $c = c(n,s,L,q,\tau)$. Using Fubini's theorem, we observe that

$$\begin{aligned}\int_{\mathcal{Q}_{r_1}} |D^\tau d_s u|_N^{\tilde p}\, d\mu_{\tau,t} &= \int_0^\infty \tilde p\lambda^{\tilde p-1}\mu_{\tau,t}\left(\{(x,y,t)\in\mathcal{Q}_{r_1} : |D^\tau d_s u|_N(x,y,t) > \lambda\}\right) d\lambda\\ &= \int_0^{\mathcal{M}\lambda_0} \tilde p\lambda^{\tilde p-1}\mu_{\tau,t}\left(\{(x,y,t)\in\mathcal{Q}_{r_1} : |D^\tau d_s u|_N(x,y,t) > \lambda\}\right) d\lambda\\ &\quad + \int_{\mathcal{M}\lambda_0}^N \tilde p\lambda^{\tilde p-1}\mu_{\tau,t}\left(\{(x,y,t)\in\mathcal{Q}_{r_1} : |D^\tau d_s u|_N(x,y,t) > \lambda\}\right) d\lambda =: I_1 + I_2,\end{aligned}$$

where $\mathcal{M} > 1$ is a constant which will be determined later and $N > \mathcal{M}\lambda_0$. We now estimate I_1 and I_2.

Estimate of I_1. A simple calculation yields that

$$I_1 \le \mu_{\tau,t}\left(\mathcal{Q}_{r_1}\right)(\mathcal{M}\lambda_0)^{\tilde p}.$$

Estimate of I_2. By a change of variable and (5.6) of Lemma 5.1 with $p = p_h$, we get that

$$\begin{aligned} I_2 &= \int_{\lambda_0}^{N\mathcal{M}^{-1}} \tilde p\mathcal{M}(\mathcal{M}\lambda)^{\tilde p-1}\mu_{\tau,t}\left(\{(x,y,t)\in\mathcal{Q}_{r_1} : |D^\tau d_s u|_N(x,y,t) > \mathcal{M}\lambda\}\right) d\lambda\\ &\le \sum_{i\ge0}\int_{\lambda_0}^{N\mathcal{M}^{-1}} \tilde p\mathcal{M}(\mathcal{M}\lambda)^{\tilde p-1}\mu_{\tau,t}\left(\left\{(x,y,t)\in\mathcal{Q}_{5^{\frac2s}\rho_i}(z_i) : |D^\tau d_s u|_N(x,y,t) > \mathcal{M}\lambda\right\}\right) d\lambda\\ &\quad + \sum_{j\ge0}\int_{\lambda_0}^{N\mathcal{M}^{-1}} \tilde p\mathcal{M}(\mathcal{M}\lambda)^{\tilde p-1}\mu_{\tau,t}\left(\left\{(x,y,t)\in\mathcal{Q}_{5^{\frac1s}\overline{r}_j}(x_{1,j},x_{2,j},t_{0,j}) : |D^\tau d_s u|_N(x,y,t) > \mathcal{M}\lambda\right\}\right) d\lambda\\ &=: I_{2,1} + I_{2,2},\end{aligned}$$

where we have used the fact that

$$\{(x,y,t)\in\mathcal{Q}_{r_1} : |D^\tau d_s u|_N(x,y,t) > \mathcal{M}\lambda\} \subset \{(x,y,t)\in\mathcal{Q}_{r_1} : |D^\tau d_s u|(x,y,t) > \lambda\}.$$

By assuming

$$\mathcal{M} > c_c c_d, \tag{6.7}$$

where the constants $c_c = c_c(n,s,L,\tau)$ and $c_d = c_d(n,s,L,q,\tau)$ are given in Lemma 4.4 with ρ_i replaced by $5^{\frac{2}{s}}\rho_i$ and Remark 9, respectively, we now estimate $I_{2,1}$ as

$$\begin{aligned}
I_{2,1} &\leq \sum_i \int_{\lambda_0}^{N\mathcal{M}^{-1}} \tilde{p}\mathcal{M}(\mathcal{M}\lambda)^{\tilde{p}-1}\mu_{\tau,t}\left(\left\{(x,y,t)\in\mathcal{Q}_{5^{\frac{2}{s}}\rho_i}(z_i) : |D^\tau d_s(u-v)|_N(x,y,t) > \mathcal{M}\lambda\right\}\right) d\lambda \\
&\leq \sum_i \int_{\lambda_0}^{N\mathcal{M}^{-1}} \tilde{p}\mathcal{M}(\mathcal{M}\lambda)^{\tilde{p}-3}\int_{\mathcal{Q}_{5^{\frac{2}{s}}\rho_i}(z_i)} |D^\tau d_s(u-v)|^2\, d\mu_{\tau,t}\, d\lambda \\
&\leq c\sum_i \int_{\lambda_0}^{N\mathcal{M}^{-1}} \mathcal{M}^{\tilde{p}-2}\lambda^{\tilde{p}-1}\epsilon^2\mu_{\tau,t}\left(\mathcal{Q}_{\rho_i}(z_i)\right) d\lambda
\end{aligned} \tag{6.8}$$

for some constant $c = c(n,s,L,q,\tau)$. In the above estimates, we have used weak 1-1 estimates, (4.11) and (2.3). On the other hand, using weak 1-1 estimates, (5.10), (2.1) and (6.1), we have that

$$\begin{aligned}
I_{2,2} &\leq \tilde{p}\lambda^{\tilde{p}}\sum_j \int_{\lambda_0}^{N\mathcal{M}^{-1}} \mathcal{M}(\mathcal{M}\lambda)^{-(p_h)_\#+\tilde{p}-1}\int_{\mathcal{Q}_{5^{\frac{1}{s}}\overline{r}_j}(x_{1,j},x_{2,j},t_{0,j})} |D^\tau d_s u|^{(p_h)_\#}\, d\mu_{\tau,t}\, d\lambda \\
&\leq \frac{\tilde{p}(c_{od,h})^{(p_h)_\#}\lambda^{\tilde{p}-1}}{\mathcal{M}^{\frac{2s}{n}}}\sum_j \int_{\lambda_0}^{N\mathcal{M}^{-1}} \mu_{\tau,t}\left(\mathcal{Q}_{\overline{r}_j}(x_{1,j},x_{2,j},t_{0,j})\right) d\lambda,
\end{aligned}$$

where the constant $c_{od,h} = c_{od,h}(n,s,p_h,\tau)$ is determined in (5.10). Note that the constant $c = \tilde{p}\times \max_{h=0,1,\ldots,h_q-1}(c_{od,h})^{q\#}$ depends only on n,s,L,q and τ and is bigger than the constant $\tilde{p}(c_{od,h})^{(p_h)_\#}$, as h_q depends only on n,s,q and τ (see (6.2)). Hence, we have that

$$I_{2,2} \leq \frac{c\lambda^{\tilde{p}-1}}{\mathcal{M}^{\frac{2s}{n}}}\sum_j \int_{\lambda_0}^{N\mathcal{M}^{-1}} \mu_{\tau,t}\left(\mathcal{Q}_{\overline{r}_j}(x_{1,j},x_{2,j},t_{0,j})\right) d\lambda \tag{6.9}$$

for some constant $c = c(n,s,L,q,\tau)$. Combine (5.8) with $p = p_h$, (6.8) and (6.9) to see that

$$\begin{aligned}
I_2 &\leq c\int_{\lambda_0}^{N\mathcal{M}^{-1}} \lambda^{\tilde{p}-1}\left(\mathcal{M}^{\tilde{p}-3}\epsilon^2 + \frac{1}{\mathcal{M}^{\frac{2s}{n}}}\right)\frac{c}{\lambda^{p_h}}\int_{\mathcal{Q}_{r_2}\cap\{|D^\tau d_s u|>a_u\lambda\}} |D^\tau d_s u|^{p_h}\, d\mu_{\tau,t}\, d\lambda \\
&\quad + c\int_{\lambda_0}^{N\mathcal{M}^{-1}} \lambda^{\tilde{p}-1}\left(\mathcal{M}^{\tilde{p}-3}\epsilon^2 + \frac{1}{\mathcal{M}^{\frac{2s}{n}}}\right)\frac{c}{(\delta\lambda)^{\tilde{q}}}\int_{\mathcal{Q}_{r_2}\cap\{|D^\tau d_0 f|>a_f\delta\lambda\}} |D^\tau d_0 f|^{\tilde{q}}\, d\mu_{\tau,t}\, d\lambda \\
&\quad + c\int_{\lambda_0}^{N\mathcal{M}^{-1}} \lambda^{\tilde{p}-1}\left(\mathcal{M}^{\tilde{p}-3}\epsilon^2 + \frac{1}{\mathcal{M}^{\frac{2s}{n}}}\right)\frac{c}{(\delta\lambda)^{b_\tau}}\int_{\mathcal{Q}_{r_2}\cap\{|G|^\gamma>(a_g\delta\lambda)^{b_\tau}G_0^{-1}\}} |G|^\gamma G_0\, d\mu_{\tau,t}\, d\lambda =: J_1 + J_2 + J_3
\end{aligned}$$

where the constants a_u, a_f and a_g are determined in (5.8), and the constants G_0 and $\tilde{q}$ are defined in (5.9) and (5.3), respectively. Using Fubini's theorem and taking $\mathcal{M} = \mathcal{M}(n,s,L,q,\tau) > 1$ which satisfies (6.7) and then choosing $\epsilon = \epsilon(n,s,L,q,\tau)\in(0,1)$, we have that

$$J_1 \leq \frac{1}{10^{5nq}}\int_{\mathcal{Q}_{r_2}} |D^\tau d_s u|_N^{\tilde{p}}\, d\mu_{\tau,t}.$$

On the other hand, if $\tilde{p}\leq\tilde{q}$, then we estimate

$$\begin{aligned}
J_2 &\leq \frac{c}{\lambda_0^{q-\tilde{p}}}\int_{\lambda_0}^{N\mathcal{M}^{-1}} \lambda^{q-1}\left(\mathcal{M}^{\tilde{p}-3}\epsilon^2 + \frac{1}{\mathcal{M}^{\frac{2s}{n}}}\right)\frac{c}{(\delta\lambda)^{\tilde{q}}}\int_{\mathcal{Q}_{r_2}\cap\{|D^\tau d_0 f|>a\lambda\}} |D^\tau d_0 f|^{\tilde{q}}\, d\mu_{\tau,t}\, d\lambda \\
&\leq \frac{c}{\lambda_0^{q-\tilde{p}}}\int_{\mathcal{Q}_{r_2}} |D^\tau d_0 f|^q\, d\mu_{\tau,t} \\
&\leq c\left(\int_{\mathcal{Q}_{r_2}} |D^\tau d_0 f|^q\, d\mu_{\tau,t}\right)^{\frac{\tilde{p}}{q}},
\end{aligned} \tag{6.10}$$

where we have used $1 \leq \left(\frac{\lambda}{\lambda_0}\right)^{q-\tilde{p}}$, Fubini's theorem (thanks to the relation $\tilde{q} < q$, from (5.5)) and the fact that $\lambda_0^{\tilde{p}-q} \leq c\left(\int_{\mathcal{Q}_{r_2}} |D^\tau d_0 f|^q\, d\mu_{\tau,t}\right)^{\frac{\tilde{p}-q}{q}}$. If $\tilde{p} > \tilde{q}$, Fubini's theorem and Hölder's inequality yield that

$$J_2 \le c \int_{\mathcal{Q}_{r_2}} |D^\tau d_0 f|^{\tilde{p}}\, d\mu_{\tau,t} \le c \left(\int_{\mathcal{Q}_{r_2}} |D^\tau d_0 f|^q \, d\mu_{\tau,t} \right)^{\frac{\tilde{p}}{q}}.$$

Similarly, we get that

$$J_3 \le c \int_{\mathcal{Q}_{r_2}} (|G|^\gamma G_0)^{\frac{\tilde{p}}{b_\tau}} \le c \left(\fint_{\mathcal{Q}_2} \left(2^{s-\tau}|G|\right)^{\frac{\tilde{p}\gamma}{b_\tau}} \, d\mu_{\tau,t} \right)^{\frac{b_\tau}{\gamma}}.$$

Consequently, combining all the estimates I_1 and I_2 and recalling (2.2) and (6.5), we obtain the desired result (6.6). We recall the definitions of λ_0 given in (5.7) and G_0 given in (5.9). Then using Lemma 2.4 and passing to the limit $N \to 0$, we get (6.3). □

With the aid of Lemma 6.1, we now prove our main theorem.

Proof of Theorem 1.1. Let us fix $Q_r(z_0) \Subset \Omega_T$ with $r \in (0, R]$. We now take

$$\tau = \frac{\sigma q}{q-2} < \min\left\{ s - \frac{2s}{q}, 1 - s \right\}. \tag{6.11}$$

Let us choose $\delta = \delta(n, s, L, q, \sigma)$ determined in Lemma 6.1. We claim that

$$\left(\fint_{\mathcal{Q}_{\frac{r}{2}}(z_0)} |D^\tau d_s u|^{\tilde{p}} \, d\mu_{\tau,t} \right)^{\frac{1}{\tilde{p}}} \le c\mathcal{R}(u, f, g, p_0, \tilde{p}, Q_r(z_0)) \tag{6.12}$$

for some constant $c = c(\mathsf{data})$, where $\tilde{p}$ is given in (6.4) with $h = 0$ and we denote

$$\begin{aligned} \mathcal{R}(u, f, g, p_0, \tilde{p}, Q_r(z_0)) = &\left(\fint_{\mathcal{Q}_r(z_0)} |D^\tau d_s u|^{p_0} \, d\mu_{\tau,t} \right)^{\frac{1}{p_0}} + \mathrm{Tail}_{\infty,2s}\left(\frac{u - (u)_{B_r(x_0)}(t)}{r^{s+\tau}} ; Q_r(z_0) \right) \\ &+ \left(\sup_{t \in \Lambda_r(t_0)} \fint_{B_r(x_0)} \frac{|u - (u)_{Q_r(z_0)}|^2}{r^{2s+2\tau}} \, dx \right)^{\frac{1}{2}} + \left(\fint_{\mathcal{Q}_r(z_0)} \left(r^{\frac{s-\tau}{\tilde{p}}} |G| \right)^{\frac{\tilde{p}\gamma}{b_\tau}} d\mu_{\tau,t} \right)^{\frac{b_\tau}{\tilde{p}\gamma}} \\ &+ \left(\fint_{\mathcal{Q}_r(z_0)} |D^\tau d_0 f|^q \, d\mu_{\tau,t} \right)^{\frac{1}{q}} + \mathrm{Tail}_{q,s}\left(\frac{f - (f)_{B_r(x_0)}(t)}{r^\tau} ; Q_r(z_0) \right). \end{aligned}$$

To this end, we define for any $x, y \in \mathbb{R}^n$, $t \in \Lambda_2$ and $\xi \in \mathbb{R}$,

$$\tilde{u}(x,t) = \frac{u\left(r_s x + x_1, r_s^{2s} t + t_1\right)}{r_s^{s+\tau}}, \quad \tilde{f}(x,t) = \frac{f\left(r_s x + x_1, r_s^{2s} t + t_1\right)}{r_s^\tau},$$

$$\tilde{g}(x,t) = r_s^{s-\tau} g\left(r_s x + x_1, r_s^{2s} t + t_1\right), \quad \tilde{A}(x,y,t) = A\left(r_s x + x_1, r_s y + x_1, r_s^{2s} t + t_1\right), \quad \tilde{\Phi}(\xi) = \frac{\Phi(r_s^\tau \xi)}{r_s^\tau}$$

where $z_1 \in \overline{Q}_{\frac{r}{2}}(z_0)$ and $r_s = \left(\frac{s(\sqrt{2}-1)}{4} \right)^{\frac{2}{s}} r$, in order to see that $\tilde{u}$ is a weak solution to (5.1) with $f = \tilde{f}$ $g = \tilde{g}$, $A = \tilde{A}$ and $\Phi = \tilde{\Phi}$. Moreover, we observe that

$$Q_{r_s}(z_1) \subset Q_{\frac{r}{\sqrt{2}}}(z_0). \tag{6.13}$$

We now apply Lemma 6.1 with $u = \tilde{u}$, $f = \tilde{f}$, $g = \tilde{g}$ and $h = 0$, and use change of the variables to get that

$$\left(\fint_{\mathcal{Q}_{\frac{r_s}{2}}(z_1)} |D^\tau d_s u|^{\tilde{p}} \, d\mu_{\tau,t} \right)^{\frac{1}{\tilde{p}}} \le c\mathcal{R}(u, f, g, p_0, \tilde{p}, Q_{r_s}(z_1)) \tag{6.14}$$

for some constant $c = c(\mathsf{data})$. After a few algebraic calculations along with Lemma 2.3, (2.2) and (6.13), the expression in (6.14) is estimated as

$$\left(\fint_{\mathcal{Q}_{\frac{r_s}{2}}(z_1)} |D^\tau d_s u|^{\tilde{p}} \, d\mu_{\tau,t} \right)^{\frac{1}{\tilde{p}}} \le c\mathcal{R}\left(u, f, g, p_0, \tilde{p}, Q_{\frac{r}{\sqrt{2}}}(z_0)\right) \tag{6.15}$$

On the other hand, for any $\mathcal{Q}_{\frac{r_s}{8\sqrt{n}}} \equiv \mathcal{B}_{\frac{r_s}{8\sqrt{n}}}(x_1, x_2) \times \Lambda_{\frac{r_s}{8\sqrt{n}}}(t_2)$ with $(x_1, t_2), (x_2, t_2) \in \overline{Q}_{\frac{r}{2}}(z_0)$ satisfying (5.45), we use Lemma 5.2 to obtain that

$$\left(\fint_{\mathcal{Q}_{\frac{r_s}{8\sqrt{n}}}} |D^\tau d_s u|^{\tilde{p}} \, d\mu_{\tau,t} \right)^{\frac{1}{\tilde{p}}} \le c \left(\fint_{\mathcal{Q}_{\frac{r_s}{8\sqrt{n}}}} |D^\tau d_s u|^{p_0} \, d\mu_{\tau,t} \right)^{\frac{1}{p_0}} + c \left[\sum_{d=1}^{2} E_{p,\tau}\left(u ; Q_{\frac{r_s}{8\sqrt{n}}}(x_d, t_2)\right) \right]$$

for some constant $c = c(\mathsf{data})$. As in (6.15) along with Lemma 2.3, (2.4) and (6.13), we deduce that

$$\left(\fint_{\mathcal{Q}_{\frac{r_s}{8\sqrt{n}}}} |D^\tau d_s u|^{\tilde{p}}\, d\mu_{\tau,t}\right)^{\frac{1}{\tilde{p}}} \le c\left(\fint_{Q_{\frac{r}{\sqrt{2}}}(z_0)} |D^\tau d_s u|^{p_0}\, d\mu_{\tau,t}\right)^{\frac{1}{p_0}} + c\left(\sup_{t\in\Lambda_{\frac{r}{\sqrt{2}}}(t_0)} \fint_{B_{\frac{r}{\sqrt{2}}}(x_0)} \frac{\left|u-(u)_{Q_{\frac{r}{\sqrt{2}}}(z_0)}\right|^2}{(\frac{r}{\sqrt{2}})^{2s+2\tau}}\, dx\right)^{\frac{1}{2}}. \tag{6.16}$$

Since $\overline{\mathcal{Q}}_{\frac{r}{2}}(z_0)$ is a compact set, there are finite mutually disjoint open sets $\mathcal{Q}_{r_s}(z_{1,i})$ and $\mathcal{Q}_{r_s}(x_{1,j}, x_{2,j}, t_{2,j})$ for some points $z_{1,i}, (x_{1,j}, t_{2,j}), (x_{2,j}, t_{2,j}) \in \overline{\mathcal{Q}}_{\frac{r}{2}}(z_0)$ such that

$$\overline{\mathcal{Q}}_{\frac{r}{2}}(z_0) \subset \left(\bigcup_i \mathcal{Q}_{r_s}(z_{1,i})\right) \bigcup \left(\bigcup_j \mathcal{Q}_{r_s}(x_{1,j}, x_{2,j}, t_{2,j})\right) \subset Q_{\frac{r}{\sqrt{2}}}(z_0),$$

and $\mathcal{Q}_{r_s}(x_{1,j}, x_{2,j}, t_{2,j})$ satisfies (5.45). Combine (6.12) and (6.16) to see that

$$\left(\fint_{\mathcal{Q}_{\frac{r}{2}}(z_0)} |D^\tau d_s u|^{\tilde{p}}\, d\mu_{\tau,t}\right)^{\frac{1}{\tilde{p}}} \le c\mathcal{R}\left(u, f, g, p_0, \tilde{p}, Q_{\frac{r}{\sqrt{2}}}(z_0)\right). \tag{6.17}$$

Consequently, after a few simple calculations with Lemma 2.3, we estimate the right-hand side of (6.17) to get that (6.12) holds. In addition, using the standard covering argument along with (6.16) and (6.17), we prove that $D^\tau d_s u \in L^{\tilde{p}}_{\mathrm{loc}}\left(d\mu_{\tau,t}; \Omega\times\Omega\times(0,T)\right)$. If $h = 0$, then by recalling (2.2), (3.5), (5.9), (6.2), (6.4) and (6.11), we obtain $d_s u \in L^q_{\mathrm{loc}}\left(d\mu_{\tau,t}; \Omega\times\Omega\times(0,T)\right)$ and the desired estimate (1.9). Let us assume that $h > 0$. We have shown that $D^\tau d_s u \in L^{p_1}_{\mathrm{loc}}\left(d\mu_{\tau,t}; \Omega\times\Omega\times(0,T)\right)$ and (6.16) and (6.17) with $\tilde{p} = p_1$. Thus, by following the same line as in the proof for (6.16) and (6.17) with p_0 replaced by p_1, we have that

$$\left(\fint_{\mathcal{Q}_{r_s}} |D^\tau d_s u|^{\tilde{p}_1}\, d\mu_{\tau,t}\right)^{\frac{1}{\tilde{p}_1}} \le c\left(\fint_{Q_{\frac{r}{\sqrt{2}}}(z_0)} |D^\tau d_s u|^{p_1}\, d\mu_{\tau,t}\right)^{\frac{1}{p_1}} + c\left(\sup_{t\in\Lambda_{\frac{r}{\sqrt{2}}}(t_0)} \fint_{B_{\frac{r}{\sqrt{2}}}(x_0)} \frac{\left|u-(u)_{Q_{\frac{r}{\sqrt{2}}}(z_0)}\right|^2}{(\frac{r}{\sqrt{2}})^{2s+2\tau}}\, dx\right)^{\frac{1}{2}}, \tag{6.18}$$

for any $\mathcal{Q}_{\frac{r_s}{8\sqrt{n}}} = \mathcal{B}_{\frac{r_s}{8\sqrt{n}}}(x_1, x_2) \times \Lambda_{\frac{r_s}{8\sqrt{n}}}(t_2)$ with $(x_1, t_2), (x_2, t_2) \in \overline{Q}_{\frac{r}{2}}(z_0)$ satisfying (5.45) and

$$\left(\fint_{\mathcal{Q}_{\frac{r}{2}}(z_0)} |D^\tau d_s u|^{\tilde{p}_1}\, d\mu_{\tau,t}\right)^{\frac{1}{\tilde{p}_1}} \le c\mathcal{R}\left(u, f, g, p_1, \tilde{p}_1, Q_{\frac{r}{\sqrt{2}}}(z_0)\right), \tag{6.19}$$

where $\tilde{p}_1$ is the constant defined in (6.4) with $h = 1$. As a result, plugging (6.17) to the first term in $\mathcal{R}\left(u, f, g, p_1, \tilde{p}_1, Q_{\frac{r}{\sqrt{2}}}(z_0)\right)$ and then after a few simple calculations with Lemma 2.3, we obtain (6.12) with $p_0 = p_1$ and $\tilde{p} = \tilde{p}_1$. In addition, using the standard covering argument along with (6.18) and (6.19), we prove that $D^\tau d_s u \in L^{\tilde{p}_1}_{\mathrm{loc}}\left(d\mu_{\tau,t}; \Omega\times\Omega_T\right)$. By iterating this procedure $l_q - 1$ times, we obtain $D^\tau d_s u \in L^q_{\mathrm{loc}}\left(d\mu_{\tau,t}; \Omega\times\Omega_T\right)$ with the estimate (6.12) with $\tilde{p} = q$. By recalling (2.2), (3.5), (5.9), (6.2), (6.4) and (6.11), we conclude that $d_s u \in L^q_{\mathrm{loc}}\left(d\mu_{\tau,t}; \Omega\times\Omega_T\right)$ with the desired estimate (1.9). □

We finish this section by proving Theorem 1.2.

Proof of Theorem 1.2. We note that in the proof of Lemma 6.1 and Theorem 1.1, we only use the condition (5.2) to apply Fubini's Theorem on the term $D^\tau d_0 f$ (see (6.10)). Taking into account Remark 5, if $f = 0$, then the constant τ can be chosen in $(0, \min\{s, 1-s\})$. Consequently, we allow for choosing $\sigma \in \left(0, \left(1 - \frac{2}{q}\right)\min\{s, 1-s\}\right)$ considering (6.11) and we are able to prove the Theorem 1.2 by following the same lines as in the proof of Theorem 1.1 with $f = 0$. □

Appendix A. self-improving property of nonlocal parabolic equations

In this appendix, we prove a self-improving property of a weak solution u to (1.1) with $f = g = 0$. Throughout this section, we take

$$\tau_0 = \min\left\{\frac{s}{2}, \frac{1-s}{2}\right\}. \tag{A.1}$$

Before proving Lemma 4.1, we are going to prove a reverse Hölder's inequality on diagonal parts and obtain another covering lemma.

With the aid of the gluing lemma, we first obtain the following inequality.

Lemma A.1. *Let u be a weak solution to (1.1) and let $Q_\rho(z_0) \Subset \Omega_T$. Then for any $\varsigma \in (0,1]$, we have*

$$\begin{aligned}
&\fint_{Q_\rho(z_0)} \frac{|u-(u)_{Q_\rho(z_0)}|^2}{\rho^{2s+2\tau}}\,dz \\
&\le \varsigma \sup_{t\in\Lambda_\rho(t_0)} \fint_{B_\rho(x_0)} \frac{|u-(u)_{B_\rho(x_0)}(t)|^2}{\rho^{2s+2\tau_0}}\,dx + \frac{c}{\varsigma^\beta}\left(\frac{1}{\tau_0}\fint_{\mathcal{Q}_\rho(z_0)} |D^{\tau_0} d_s u|^\gamma\,d\mu_{\tau_0,t}\right)^{\frac{2}{\gamma}} \\
&\quad + c\frac{1}{\tau_0}\fint_{\mathcal{Q}_\rho(z_0)} |D^{\tau_0} d_0 f|^2\,d\mu_{\tau_0,t} + c\,\mathrm{Tail}_{1,2s}\left(\frac{u-(u)_{B_\rho(x_0)}(t)}{\rho^{s+\tau_0}}; Q_\rho(z_0)\right)^2 \\
&\quad + c\,\mathrm{Tail}_{1,s}\left(\frac{f-(f)_{B_\rho(x_0)}(t)}{\rho^{\tau_0}}; Q_\rho(z_0)\right)^2 + c\left(\frac{1}{\tau_0}\fint_{\mathcal{Q}_{2\rho}(z_0)} \left((2\rho)^{s-\tau}|G|\right)^\gamma\,d\mu_{\tau_0,t}\right)^{\frac{2}{\gamma}}
\end{aligned} \tag{A.2}$$

for some constants $c = c(n,s,L)$ and $\beta = \beta(n,s)$, where the constant γ is defined in (3.2).

Proof. We may assume that $z_0 = 0$. Note that

$$\fint_{Q_\rho} |u-(u)_{Q_\rho}|^2\,dz \le c\fint_{Q_\rho} |u-(u)_{B_\rho}(t)|^2\,dz + c\fint_{Q_\rho} |(u)_{B_\rho}(t)-(u)_{Q_\rho}|^2\,dz =: I_1 + I_2.$$

In light of Lemma 2.1 with h, p and s, replaced by u, γ and $\tilde{s} := s + \tau_0\left(1-\frac{2}{\gamma}\right)$, respectively, we first estimate I_1 as

$$I_1 \le c\left(\fint_{Q_\rho} |u-(u)_{B_\rho}(t)|^{\tilde{\gamma}}\,dz\right)^{\frac{2}{\tilde{\gamma}}} \le c\left(\frac{\rho^{\gamma(s+\tau_0)}}{\tau_0}\fint_{\mathcal{Q}_\rho} |D^{\tau_0} d_s u|^\gamma\,d\mu_{\tau_0,t}\right)^{\frac{2}{\tilde{\gamma}}} \times \left(\sup_{t\in\Lambda_\rho}\fint_{B_\rho} |u-(u)_{B_\rho(t)}|^2\,dx\right)^{\frac{2\tilde{s}\gamma}{n\tilde{\gamma}}},$$

where $c = c(n,s)$ and $\tilde{\gamma} := \gamma\left(1+\frac{2\tilde{s}}{n}\right)$. We next apply Young's inequality along with the fact that $\frac{\gamma}{\tilde{\gamma}} + \frac{2\tilde{s}\gamma}{n\tilde{\gamma}} = 1$, in order to get

$$I_1 \le \vartheta\left(\frac{\rho^{\gamma(s+\tau_0)}}{\tau_0}\fint_{\mathcal{Q}_\rho} |D^{\tau_0} d_s u|^\gamma\,d\mu_{\tau_0,t}\right)^{\frac{2}{\gamma}} + \frac{c}{\vartheta^\beta}\sup_{t\in\Lambda_\rho}\fint_{B_\rho} |u-(u)_{B_\rho(t)}|^2\,dx$$

for any $\vartheta \in (0,1]$ and for some constant $\beta = \beta(n,s)$. We now estimate I_2 as

$$\begin{aligned}
I_2 &\le c\fint\!\!\fint_{\Lambda_\rho\times\Lambda_\rho} |(u)_{B_\rho}(t)-(u)_{B_\rho}(t')|^2\,dt\,dt' \\
&\le c\fint\!\!\fint_{\Lambda_\rho\times\Lambda_\rho} \left|(u)_{B_\rho}(t)-(u)^\psi_{B_\rho}(t)\right|^2\,dt\,dt' + c\fint\!\!\fint_{\Lambda_\rho\times\Lambda_\rho} \left|(u)^\psi_{B_\rho}(t)-(u)^\psi_{B_\rho}(t')\right|^2\,dt\,dt' \\
&\quad + c\fint\!\!\fint_{\Lambda_\rho\times\Lambda_\rho} \left|(u)^\psi_{B_\rho}(t')-(u)_{B_\rho}(t')\right|^2\,dt\,dt' \\
&\le c\fint_{\Lambda_\rho} \left|(u)_{B_\rho}(t)-(u)^\psi_{B_\rho}(t)\right|^2\,dt + c\sup_{t_1,t_2\in\Lambda_\rho} \left|(u)^\psi_{B_\rho}(t_1)-(u)^\psi_{B_\rho}(t_2)\right|^2 =: I_{2,1} + I_{2,2},
\end{aligned}$$

where ψ is the given function in Lemma 3.2 and

$$(u)^\psi_{B_\rho}(t) \equiv \frac{1}{\|\psi\|_{L^1}}\int_{B_\rho} u(x,t)\psi(x)\,dx.$$

Using the fact that $\|\psi\|_{L^1} \approx_n |B_\rho|$ and Hölder's inequality, we have

$$I_{2,1} \le c\fint_{\Lambda_\rho} \left|\frac{1}{\|\psi\|_{L^1}}\int_{B_\rho} \left(u-(u)_{B_\rho}(t)\right)\,dx\right|^2\,dt \le cI_1$$

for some constant $c = c(n, s)$. Using Lemma 3.2, we get that

$$I_{2,2}^{\frac{1}{2}} \le c\rho^{2s-1} \fint_{Q_\rho} \int_{B_\rho} \frac{|u(x,t) - u(y,t)|}{|x-y|^{n+2s-1}}\,dy\,dz + c\rho^{2s} \fint_{Q_\rho} \int_{\mathbb{R}^n \setminus B_\rho} \frac{|u(x,t) - u(y,t)|}{|y|^{n+2s}}\,dy\,dz$$
$$+ c\rho^{2s-1} \fint_{Q_\rho} \int_{B_\rho} \frac{|f(x,t) - f(y,t)|}{|x-y|^{n+s-1}}\,dy\,dz + c\rho^{2s} \fint_{Q_\rho} \int_{\mathbb{R}^n \setminus B_\rho} \frac{|f(x,t) - f(y,t)|}{|y|^{n+s}}\,dy\,dz$$
$$+ c\fint_{Q_\rho} \rho^{2s}|g|^\gamma\,dz =: \sum_{i=1}^{5} I_{2,2,i}$$

for some constant $c = c(n, s, L)$. We now estimate $I_{2,2,1}$ and $I_{2,2,2}$.
Estimate of $I_{2,2,1}$. Using Hölder's inequality and (2.2), we have

$$I_{2,2,1} \le c\rho^{2s-1} \left(\fint_{Q_\rho} \int_{B_\rho} \frac{|D^{\tau_0} d_s u|^\gamma}{|x-y|^{n-2\tau}}\,dx\,dy\,dt \right)^{\frac{1}{\gamma}} \left(\fint_{Q_\rho} \int_{B_\rho} \frac{dx\,dy\,dt}{|x-y|^{n+\gamma'\left(s-1+\tau_0\left(\frac{2}{\gamma}-1\right)\right)}} \right)^{\frac{1}{\gamma'}}$$
$$\le \frac{c\rho^{s+\tau_0}}{1-(s+\tau_0)} \left(\frac{1}{\tau_0} \fint_{\mathcal{Q}_\rho} |D^{\tau_0} d_s u|^\gamma\,d\mu_{\tau_0,t} \right)^{\frac{1}{\gamma}}$$

Estimate of $I_{2,2,2}$. A simple algebraic computation yields that

$$I_{2,2,2} \le c\rho^{2s} \fint_{Q_\rho} \int_{\mathbb{R}^n \setminus B_\rho} \frac{|u(x,t) - (u)_{B_\rho}(t)|}{|y|^{n+2s}}\,dy\,dz + c\rho^{2s} \fint_{Q_\rho} \int_{\mathbb{R}^n \setminus B_\rho} \frac{|(u)_{B_\rho}(t) - u(y,t)|}{|y|^{n+2s}}\,dy\,dz$$
$$\le cI_1^{\frac{1}{2}} + c\,\mathrm{Tail}_{1,s}\left(u - (u)_{B_\rho(x_0)}(t); Q_\rho(z_0)\right).$$

Similarly, we estimate $I_{2,2,3}$ and $I_{2,2,4}$ as

$$I_{2,2,3} + I_{2,2,4} \le c\frac{\rho^{s+\tau_0}}{1-s} \left(\frac{1}{\tau_0} \fint_{\mathcal{Q}_\rho(z_0)} |D^{\tau_0} d_0 f|^2\,d\mu_{\tau_0,t} \right)^{\frac{1}{2}} + c\rho^s\,\mathrm{Tail}_{1,\frac{s}{2}}\left(f - (f)_{B_\rho(x_0)}(t); Q_\rho(z_0)\right).$$

Using Hölder's inequality, (2.2) and (3.5), we estimate $I_{2,2,5}$ as

$$I_{2,2,5} \le c\left(\frac{1}{\tau_0} \fint_{\mathcal{Q}_{2\rho}(z_0)} \left((2\rho)^{s-\tau_0}|G|\right)^\gamma\,d\mu_{\tau_0,t} \right)^{\frac{2}{\gamma}}.$$

Take $\vartheta = \frac{\varsigma}{c}$ for some constant $c = c(n, s, L)$ and combine all the above estimates to get (A.2). □

Remark 10. By estimating I_1 in the proof of Lemma A.1 as

$$I_1 \le \frac{c}{\tau_0} \fint_{\mathcal{Q}_\rho(z_0)} |D^{\tau_0} d_s u|^2\,d\mu_{\tau_0,t},$$

we find that for any $\tau_0 \in (0, \min\{s, 1-s\})$, there holds

$$\fint_{Q_\rho(z_0)} \frac{|u - (u)_{Q_\rho(z_0)}|^2}{\rho^{2s+2\tau_0}}\,dz \le \frac{c}{\tau_0} \fint_{\mathcal{Q}_\rho(z_0)} |D^{\tau_0} d_s u|^2\,d\mu_{\tau_0,t} + c\,\mathrm{Tail}_{1,2s}\left(\frac{u - (u)_{B_\rho(x_0)}(t)}{\rho^{s+\tau_0}}; Q_\rho(z_0) \right)^2$$
$$+ c\frac{1}{\tau_0} \fint_{\mathcal{Q}_\rho(z_0)} |D^{\tau_0} d_0 f|^2\,d\mu_{\tau_0,t} + c\,\mathrm{Tail}_{1,s}\left(\frac{f - (f)_{B_\rho(x_0)}(t)}{\rho^{\tau_0}}; Q_\rho(z_0) \right)^2 \tag{A.3}$$
$$+ c\left(\frac{1}{\tau_0} \fint_{\mathcal{Q}_{2\rho}(z_0)} \left((2\rho)^{s-\tau_0}|G|\right)^\gamma\,d\mu_{\tau_0,t} \right)^{\frac{2}{\gamma}}$$

for some constant $c = c(n, s, L)$.

In light of Lemma A.1, we now prove the following reverse Hölder's inequality.

Lemma A.2. *Let u be a weak solution to (1.1) with $f = g = 0$ and let $Q_{2\rho}(z_0) \Subset \Omega_T$. Then we have*

$$E_{2,\tau_0}\left(u; Q_\rho(z_0)\right)^2 \le c_0 \left(\frac{1}{\tau_0} \fint_{\mathcal{Q}_{2\rho}(z_0)} |D^{\tau_0} d_s u|^\gamma\,d\mu_{\tau_0,t} \right)^{\frac{2}{\gamma}} + c_0\,\mathrm{Tail}_{\gamma,2s}\left(\frac{|u - (u)_{B_{2\rho}(x_0)}(t)|}{(2\rho)^{s+\tau_0}}; Q_{2\rho}(z_0) \right)^2,$$

where $c_0 = c_0(n, s, L)$ and $E_{2,\tau_0}(\cdot)$ is defined in (5.32).

Proof. From (2.2) and (3.1) with $k=(u)_{Q_r(z_0)}$, $f=0$ and $g=0$, we deduce that

$$E_{2,\tau_0}(u\,;Q_\rho)\leq \frac{cr^{n+2}}{(r-\rho)^{n+2}}\iint_{Q_r(z_0)}\frac{|u-(u)_{Q_r(z_0)}|^2}{r^{2s+2\tau_0}}\,dz+\frac{cr^{2n+4s}}{(r-\rho)^{2n+4s}}\,\mathrm{Tail}_{\gamma,2s}\left(\frac{u-(u)_{B_r(x_0)}(t)}{r^{s+\tau_0}};Q_r(z_0)\right)^2,$$

where $c=c(n,s,L)$. Using the estimate (A.2) with $\rho=r$, $f=0$ and $g=0$ into the first term in the right-hand side of the above inequality and Hölder's inequality, we get

$$\begin{aligned}E_{2,\tau_0}(u\,;Q_\rho)\leq{}&\frac{1}{4}E_{2,\tau_0}(u\,;Q_r)+\frac{cr^{n+2}}{(r-\rho)^{n+2}}\left(\frac{1}{\tau_0}\fint_{\mathcal{Q}_{2\rho}(z_0)}|D^{\tau_0}d_su|^\gamma\,d\mu_{\tau_0,t}\right)^{\frac{2}{\gamma}}\\&+c\left(\frac{r}{r-\rho}\right)^{2(n+2s)}\mathrm{Tail}_{\gamma,2s}\left(\frac{u-(u)_{B_{2\rho}(x_0)}(t)}{(2\rho)^{s+\tau_0}};Q_{2\rho}(z_0)\right)^2.\end{aligned}$$

by taking $\varsigma\in(0,1)$ sufficiently small depending only on n,s and L. Additionally, for the tail term, we have used Lemma 2.3. Finally, employing Lemma 2.4, we obtain the desired estimate (3.6). □

We next prove the following covering lemma.

Lemma A.3. *Let $1\leq r_1<r_2\leq 2$ and u be a weak solution to (5.1) with $f=g=0$. Then, there are two families of countable disjoint cylinders $\{\mathcal{Q}_{\rho_i}(z_i)\}_{i\geq 0}$ and $\left\{\mathcal{Q}_{\tilde{r}_j}(x_{1,j},x_{2,j},t_{0,j})\right\}_{j\geq 0}$, such that*

$$U_\lambda=\{(x,y,t)\in\mathcal{Q}_{r_1}\,:\,|D^{\tau_0}d_su(x,y,t)|\geq\lambda\}\subset\left(\bigcup_i\mathcal{Q}_{5^{\frac{2}{s}}\rho_i}(z_i)\right)\bigcup\left(\bigcup_j\mathcal{Q}_{5^{\frac{1}{s}}\tilde{r}_j}(x_{1,j},x_{2,j},t_{0,j})\right)\tag{A.4}$$

whenever $\lambda\geq\lambda_0$, where

$$\begin{aligned}\lambda_0:={}&\frac{c}{(r_2-r_1)^{\frac{5n}{s}}}\left[\left(\frac{1}{\tau_0}\fint_{\mathcal{Q}_2}|D^{\tau_0}d_su|^2\,d\mu_{\tau_0,t}\right)^{\frac{1}{2}}+\left(\sup_{t\in\Lambda_2}\fint_{B_2}\frac{|u-(u)_{Q_2}|^2}{2^{2s+2\tau_0}}\,dx\right)^{\frac{1}{2}}\right]\\&+\frac{c}{(r_2-r_1)^{\frac{5n}{s}}}\,\mathrm{Tail}_{\infty,2s}\left(\frac{u-(u)_{B_2}(t)}{2^{s+\tau_0}};Q_2\right)\end{aligned}\tag{A.5}$$

for some constant $c=c(n,s,L)$. In particular, we have

$$\sum_{i\geq 0}\mu_{\tau_0,t}\left(\mathcal{Q}_{\rho_i}(z_i)\right)\leq\frac{c}{\lambda^\gamma}\int_{\mathcal{Q}_{r_2}\cap\{|D^{\tau_0}d_su|>b_u\lambda\}}|D^{\tau_0}d_su|^\gamma\,d\mu_{\tau_0,t},\tag{A.6}$$

$$\sum_{j\geq 0}\mu_{\tau_0,t}\left(\mathcal{Q}_{\tilde{r}_j}(x_{1,j},x_{2,j},t_{0,j})\right)\leq\frac{c}{\lambda^2}\int_{\mathcal{Q}_{r_2}\cap\{|D^{\tau_0}d_su|>\frac{\lambda}{16}\}}|D^{\tau_0}d_su|^2\,d\mu_{\tau_0,t}++\sum_{i\geq 0}\mu_{\tau_0,t}\left(\mathcal{Q}_{\rho_i}(z_i)\right)\tag{A.7}$$

for some constant $b_u=b_u(n,s,L)\in(0,1]$, where the constant γ is defined in (3.2), and we also have

$$\left(\fint_{\mathcal{Q}_{5^{\frac{1}{s}}\tilde{r}_j}(x_{1,j},x_{2,j},t_{0,j})}|D^{\tau_0}d_su|^{2_\#}\,d\mu_\tau\right)^{\frac{1}{2_\#}}\leq c\lambda\quad\text{for any } j,\tag{A.8}$$

where the constant $2_\#$ is given in (2.1) with $p=2$.

Proof. We first define a functional

$$\Theta_D\left(z_0,r\right)=\left(\fint_{\mathcal{Q}_r(z_0)}|D^{\tau_0}d_su|^2\,d\mu_{\tau_0,t}\right)^{\frac{1}{2}}+\left(\tau_0\sup_{t\in\Lambda_r(t_0)}\fint_{B_r(x_0)}\frac{|u-(u)_{Q_r(z_0)}|^2}{r^{2s+2\tau_0}}\,dx\right)^{\frac{1}{2}}$$

for any $z_0\in Q_{r_1}$ and $r>0$ with $Q_r(z_0)\subset Q_2$. Let us take

$$\begin{aligned}\lambda_0={}&\frac{M\tau_0^{\frac{1}{2}}\kappa^{-1}}{(r_2-r_1)^{\frac{5n}{s}}}\left(\left(\frac{1}{\tau_0}\fint_{\mathcal{Q}_2}|D^{\tau_0}d_su|^2\,d\mu_{\tau_0,t}\right)^{\frac{1}{2}}+\mathrm{Tail}_{\infty,2s}\left(\frac{u-(u)_{B_2}(t)}{2^{s+\tau_0}};Q_2\right)\right)\\&+\frac{M\tau_0^{\frac{1}{2}}\kappa^{-1}}{(r_2-r_1)^{\frac{5n}{s}}}\left(\sup_{t\in\Lambda_2}\fint_{B_2}\frac{|u-(u)_{Q_2}|^2}{2^{2s+2\tau_0}}\,dx\right)^{\frac{1}{2}},\end{aligned}$$

where $M\geq 1$ and $\kappa\in(0,1]$ are free parameters which will be determined later. We next take a positive integer $j_0=j_0(n,s,\Lambda)$ such that

$$\frac{16(c_0+\tilde{c}+2c_2)^2}{1-2^{-s+\tau_0}}\leq 2^{j_0(s-\tau_0)},\tag{A.9}$$

where c_0 is the constant determined in Lemma A.2, and $\tilde{c}$ and c_2 are the constants determined in (2.6). We then note that for any $z_0 \in Q_{r_1}$,

$$Q_{5^{\frac{2}{s}}\times 2^{j_0+2}\mathcal{R}_{1,2}}(z_0) \subset Q_{r_2},$$

where $\mathcal{R}_{1,2}$ is defined in (5.14). Let us now define for $\lambda \geq \lambda_0$,

$$D_{\kappa\lambda} = \left\{ z_0 \in Q_{r_1} \,:\, \sup_{0<\rho\leq \mathcal{R}_{1,2}} \Theta_D\left(z_0,\rho\right) > \kappa\lambda \right\}.$$

We now take the constant M as in (5.18) with $q=2$. For any $z_0 \in Q_{r_1}$ and $r \in \left[\mathcal{R}_{1,2}, 5^{\frac{2}{s}}\times 2^{j_0+3}\mathcal{R}_{1,2}\right]$, we first note that

$$\Theta_D(z_0,r) \leq \kappa\lambda,$$

by following the same line as in the proof for (5.19) with $p=2$. Therefore, there is an exit radius ρ_z for each $z \in D_{\kappa\lambda}$ such that

$$\Theta_D\left(z,\rho_z\right) \geq \kappa\lambda \quad \text{and} \quad \Theta_D\left(z,\rho\right) \leq \kappa\lambda \quad \text{if } \rho_z \leq \rho \leq 5^{\frac{2}{s}}\times 2^{j_0+3}\mathcal{R}_{1,2}. \tag{A.10}$$

We now apply Vitali's covering lemma to find a family of mutually disjoint countable cylinders

$$\left\{Q_{2^{j_0}\rho_{z_i}}(z_i)\right\}_{i\geq 0} \text{ such that } D_{\kappa\lambda} \subset \bigcup_{i=0}^{\infty} Q_{5^{\frac{1}{s}}\times 2^{j_0}\rho_{z_i}}(z_i). \tag{A.11}$$

We now denote

$$\rho_i = 2^{j_0}\rho_{z_i} \quad \text{for each } i. \tag{A.12}$$

From (A.10), we have

$$\frac{\kappa\lambda}{\tau_0^{\frac{1}{2}}} \leq \left(\frac{1}{\tau_0}\fint_{\mathcal{Q}_{\rho_{z_i}}(z_i)} |D^{\tau_0}d_s u|^2\,d\mu_{\tau_0,t}\right)^{\frac{1}{2}} + \left(\sup_{t\in\Lambda_{\rho_{z_i}}(t_i)}\fint_{B_{\rho_{z_i}}(x_i)} \frac{|u-(u)_{Q_{\rho_{z_i}}(z_i)}|^2}{\rho_{z_i}^{2s+2\tau_0}}\,dx\right)^{\frac{1}{2}}. \tag{A.13}$$

On account of Lemma A.2, Lemma 2.6 and (A.12), we estimate the right-hand side of (A.13) as

$$\begin{aligned}\frac{\kappa\lambda}{\tau_0^{\frac{1}{2}}} &\leq c\left(\frac{1}{\tau_0}\fint_{\mathcal{Q}_{\rho_i}(z_i)} |D^{\tau_0}d_s u|^\gamma\,d\mu_{\tau_0,t}\right)^{\frac{1}{\gamma}} + c_0\tilde{c}\frac{1}{\frac{1}{\tau_0}^{\frac{1}{2}}}\sum_{j=j_0+1}^{l} 2^{i(-s+\tau_0)}\Theta_D\left(z_j,2^j\rho_{z_i}\right)\\ &\quad + c_0\tilde{c}\left(\frac{2}{2^{j_0}\mathcal{R}_{1,2}}\right)^{s+\tau_0}\left(\sup_{t\in\Lambda_2}\fint_{B_2}\frac{|u-(u)_{B_2}(t)|^2}{2^{2s+2\tau_0}}\,dx\right)^{\frac{1}{2}} + \frac{c_0\tilde{c}}{(r_2-r_1)^{5n}}\mathrm{Tail}_{\gamma,2s}\left(\frac{u-(u)_{B_2}(t)}{2^{s+\tau_0}};Q_2\right)\end{aligned}$$

where $c=c(n,s,L)$ and l is the positive integer such that $2^{j_0+1}\mathcal{R}_{1,2} \leq 2^l\rho_{z_i} < 2^{j_0+2}\mathcal{R}_{1,2}$. For the detailed calculations of the above inequality, we refer to (5.25) and (5.27) with $f=0$ and $g=0$. As a result, using (A.9) and (A.10), we find that

$$\frac{\kappa\lambda}{\tau_0^{\frac{1}{2}}} \leq c\left(\frac{1}{\tau_0}\fint_{\mathcal{Q}_{\rho_i}(z_i)} |D^{\tau_0}d_s u|^\gamma\,d\mu_{\tau_0,t}\right)^{\frac{1}{\gamma}}$$

for some constant $c=c(n,s,L)$. A suitable choice of the constant $\tilde{b}_u = \tilde{b}_u(n,s,L) \in \left(0,\frac{1}{8}\right]$ yields

$$\mu_{\tau_0,t}\left(\mathcal{Q}_{\rho_i}(z_i)\right) \leq \frac{c}{(\kappa\lambda)^\gamma}\int_{\mathcal{Q}_{\rho_i}(z_i)\cap\{|D^{\tau_0}d_s u|>\tilde{b}_u\kappa\lambda\}} |D^{\tau_0}d_s u|^\gamma\,d\mu_{\tau_0,t}, \tag{A.14}$$

as the constant τ_0 depends only on s (see (A.1)). With (A.10) and (A.11), we follow the same lines as in the proof of step 2 through step 8 given in Lemma 5.1 with $p=2$, $f=0$ and $g=0$. As a result, by taking κ as in (5.35) with $\tau^{\frac{1}{\gamma}}$ replaced by $\tau_0^{\frac{1}{2}}$, we find that there is a collection of countable disjoint cylinders $\{\mathcal{Q}\}_{\mathcal{Q}\in\mathcal{A}}$ such that

$$\{(x,y,t)\in\mathcal{Q}_{r_1} \,:\, |D^{\tau_0}d_s u(x,y,t)| \geq \lambda\} \subset \left(\bigcup_{i\geq 0}\mathcal{Q}_{5^{\frac{2}{s}}\rho_i}(z_i)\right)\bigcup\left(\bigcup_{j\geq 0}\mathcal{Q}_{5^{\frac{1}{s}}\tilde{r}_j}(x_{1,j},x_{2,j},t_{0,j})\right),$$

$$\sum_j \mu_{\tau_0,t}\left(\mathcal{Q}_{\tilde{r}_j}(x_{1,j},x_{2,j},t_{0,j})\right) \leq \frac{c}{\lambda^2}\int_{\mathcal{Q}_{r_2}\cap\left\{|D^{\tau_0}d_s u|>\frac{\lambda}{16}\right\}} |D^{\tau_0}d_s u|^2\,d\mu_{\tau_0,t} + c\sum_i \mu_{\tau_0,t}\left(\mathcal{Q}_{\rho_i}(z_i)\right)$$

and

$$\left(\fint_{\mathcal{Q}_{5^{\frac{1}{s}}\tilde{r}_j}(x_{1,j},x_{2,j},t_{0,j})} |D^{\tau_0} d_s u|^{2_\#} \, d\mu_\tau \right)^{\frac{1}{2_\#}} \leq c\lambda \quad \text{for any } j,$$

where $c = c(n,s,L)$ (see (5.63)and (5.48) for the second and the third inequalities, respectively). Let $b_u = \tilde{b}_u \kappa$. Then the above three observations and (A.14) yield the desired results as the constant τ_0 depends only on s. This completes the proof. □

We are now in the position to prove the following self-improving property for a weak solution to the corresponding homogeneous problem of (1.1).

Theorem A.1. *Let u be a weak solution to (1.1) with $f = 0$ and $g = 0$. Then there are constants $\epsilon = \epsilon(n,s,L) \in (0,1)$ and $c = c(n,s,L)$ such that*

$$\begin{aligned} \left(\frac{1}{\tau_0} \fint_{\mathcal{Q}_r(z_0)} |D^{\tau_0} d_s \tilde{u}|^{2+\epsilon} \, d\mu_{\tau_0,t} \right)^{\frac{1}{2+\epsilon}} &\leq c \left(\frac{1}{\tau_0} \fint_{\mathcal{Q}_{2r}(z_0)} |D^{\tau_0} d_s u|^2 d\mu_{\tau_0,t} \right)^{\frac{1}{2}} + c \sup_{t \in \Lambda_{2r}} \left(\fint_{B_{2r}} \frac{|u - (u)_{Q_{2r}(z_0)}|^2}{(2r)^{2s+2\tau}} \right)^{\frac{1}{2}} \\ &\quad + c \, \mathrm{Tail}_{\infty,2s} \left(\frac{u - (u)_{B_{2r}(x_0)}(t)}{(2r)^{s+\tau}} ; Q_{2r}(z_0) \right) \end{aligned} \tag{A.15}$$

whenever $Q_{2r}(z_0) \Subset \Omega_T$.

Proof. Let us fix $Q_r(z_0) \Subset \Omega_T$. We now define for any $x, y \in \mathbb{R}^n$, $t \in \Lambda_2$ and $\xi \in \mathbb{R}$,

$$\tilde{u}(x,t) = \frac{u\left(rx + x_0, r^{2s}t + t_0\right)}{r^{s+\tau}}, \ \tilde{A}(x,y,t) = A\left(rx + x_0, ry + x_0, r^{2s}t + t_0\right) \text{ and } \tilde{\Phi}(\xi) = \frac{\tilde{\Phi}\left(r^\tau \xi\right)}{r^\tau}$$

to see that $\tilde{u}$ is a weak solution to (5.1) with $f = g = 0$, $A = \tilde{A}$ and $\Phi = \tilde{\Phi}$. Let us take $\epsilon \in \left(0, \frac{2_\# - 2}{2}\right)$ which will be determined later. For each $N > 0$, we now define $\phi_N : [1,2] \to \mathbb{R}$ by

$$\phi_N(\rho) = \left(\fint_{\mathcal{Q}_\rho} |D^{\tau_0} d_s \tilde{u}|_N^{2+\epsilon} d\mu_{\tau_0,t} \right)^{\frac{1}{2+\epsilon}}.$$

For λ_0 as defined in (A.5) with $u = \tilde{u}$, we claim that if $N > \lambda_0$, then there is a constant $c = c(n,s,L)$ such that for any $1 \leq r_1 < r_2 \leq 2$,

$$\phi_N(r_1) \leq \frac{1}{2}\phi_N(r_2) + c\lambda_0. \tag{A.16}$$

By Fubini's theorem, we observe that

$$\begin{aligned} \int_{\mathcal{Q}_{r_1}} |D^{\tau_0} d_s \tilde{u}|_N^{2+\epsilon} d\mu_{\tau_0,t} &= \epsilon \int_0^{\mathcal{M}\lambda_0} \lambda^{\epsilon-1} \nu\left(\{\mathcal{Q}_{r_1} : |D^{\tau_0} d_s \tilde{u}|_N > \lambda\}\right) d\lambda \\ &\quad + \epsilon \int_{\mathcal{M}\lambda_0}^N \lambda^{\epsilon-1} \nu\left(\{\mathcal{Q}_{r_1} : |D^{\tau_0} d_s \tilde{u}|_N > \lambda\}\right) d\lambda =: I + J, \end{aligned}$$

where $d\nu = |D^{\tau_0} d_s \tilde{u}|^2 \, d\mu_{\tau_0,t}$ and $N > \mathcal{M}\lambda_0$ with $\mathcal{M} > 1$ to be determined later. We first estimate I as

$$I \leq c\mathcal{M}^\epsilon \lambda_0^\epsilon \int_{\mathcal{Q}_2} |D^{\tau_0} d_s \tilde{u}|^2 d\mu_{\tau_0,t} \leq c\mathcal{M}^\epsilon \lambda_0^{2+\epsilon} \mu_{\tau_0,t}\left(\mathcal{Q}_2\right),$$

where $c = c(n,s,L)$. We next estimate J as

$$\begin{aligned} J &= \epsilon \int_{\lambda_0}^{N\mathcal{M}^{-1}} \mathcal{M}^\epsilon \lambda^{\epsilon-1} \nu\left(\{(x,y,t) \in \mathcal{Q}_{r_1} : |D^{\tau_0} d_s \tilde{u}|_N > \mathcal{M}\lambda\}\right) d\lambda \\ &\leq \sum_{i \geq 0} \epsilon \int_{\lambda_0}^{N\mathcal{M}^{-1}} \mathcal{M}^\epsilon \lambda^{\epsilon-1} \nu\left(\mathcal{Q}_{5^{\frac{2}{s}}\rho_i}(z_i)\right) d\lambda \\ &\quad + \sum_{j \geq 0} \epsilon \int_{\lambda_0}^{N\mathcal{M}^{-1}} \mathcal{M}^\epsilon \lambda^{\epsilon-1} \nu\left(\left\{\mathcal{Q}_{5^{\frac{1}{s}}\tilde{r}_j}(x_{1,j}, x_{2,j}, t_{0,j}) : |D^{\tau_0} d_s \tilde{u}|_N > \mathcal{M}\lambda\right\}\right) d\lambda =: J_1 + J_2, \end{aligned}$$

where we have used the change of variables and (A.4). In light of the definition of the measure ν, (A.10), (2.3) and (A.6), we estimate J_1 as

$$\begin{aligned} J_1 &\leq \sum_i \epsilon \int_{\lambda_0}^{N\mathcal{M}^{-1}} \mathcal{M}^{\epsilon}\lambda^{\epsilon-1} \int_{\mathcal{Q}_{5^{\frac{2}{s}}\rho_i}(z_i)} |D^{\tau_0} d_s \tilde{u}|^2 d\mu_{\tau_0,t}\, d\lambda \\ &\leq c\sum_i \epsilon \int_{\lambda_0}^{N\mathcal{M}^{-1}} \mathcal{M}^{\epsilon}\lambda^{\epsilon-1}\lambda^2 \mu_{\tau_0,t}\left(\mathcal{Q}_{\rho_i}(z_i)\right) d\lambda \\ &\leq \frac{c\epsilon\mathcal{M}^{\epsilon}}{\lambda^{\gamma-(1+\epsilon)}} \int_{\lambda_0}^{N\mathcal{M}^{-1}} \int_{\mathcal{Q}_{r_2}\cap\{|D^{\tau_0} d_s\tilde{u}|\geq b_u\lambda\}} |D^{\tau_0} d_s \tilde{u}|^{\gamma} d\mu_{\tau_0,t}\, d\lambda, \end{aligned}$$

where $c = c(n, s, L)$ and the constant b_u is determined in Lemma A.3. Using Fubini's theorem, we get

$$J_1 \leq c\epsilon\mathcal{M}^{\epsilon} \int_{\mathcal{Q}_{r_2}} |D^{\tau_0} d_s \tilde{u}|^{2+\epsilon}_{N\mathcal{M}^{-1}} d\mu_{\tau_0,t},$$

where $c = c(n, s, L)$, as $\frac{1}{2-\gamma} > 0$ depends only on n and s. To estimate J_2, we first note from the weak 1-1 estimate and (A.8) that

$$\begin{aligned} \int_{\lambda_0}^{N\mathcal{M}^{-1}} \mathcal{M}^{\epsilon}\lambda^{\epsilon-1}\nu\left(\{\mathcal{Q} : |D^{\tau_0} d_s\tilde{u}|_N > \mathcal{M}\lambda\}\right) d\lambda &\leq c\int_{\lambda_0}^{N\mathcal{M}^{-1}} \mathcal{M}^{\epsilon}\lambda^{\epsilon-1} \int_{\mathcal{Q}} \frac{|D^{\tau} d_s \tilde{u}|_N^{2_\#-2}}{(\mathcal{M}\lambda)^{2_\#-2}} |D^{\tau_0} d_s u|^2 d\mu_{\tau_0,t}\, d\lambda \\ &\leq c\int_{\lambda_0}^{N\mathcal{M}^{-1}} \mathcal{M}^{\epsilon+2-2_\#}\lambda^{\epsilon+1}\mu_{\tau_0,t}\left(\mathcal{Q}\right) d\lambda, \end{aligned}$$

where we denote $\mathcal{Q} = \mathcal{Q}_{5^{\frac{1}{s}}\tilde{r}_j}(x_{1,j}, x_{2,j}, t_{0,j})$. We first note

$$\begin{aligned} \frac{1}{\lambda^{\gamma}} \int_{\lambda_0}^{N\mathcal{M}^{-1}} \int_{\mathcal{Q}_{r_2}\cap\{|D^{\tau_0} d_s\tilde{u}|\geq b_u\lambda\}} |D^{\tau_0} d_s \tilde{u}|^{\gamma} d\mu_{\tau_0,t}\, d\lambda &\leq \frac{\lambda_0^{2-\gamma}}{\lambda^2} \int_{\lambda_0}^{N\mathcal{M}^{-1}} \int_{\mathcal{Q}_{r_2}\cap\{|D^{\tau_0} d_s\tilde{u}|\geq b_u\lambda\}} |D^{\tau_0} d_s \tilde{u}|^{\gamma} d\mu_{\tau_0,t}\, d\lambda \\ &\leq \frac{c}{\lambda^2} \int_{\lambda_0}^{N\mathcal{M}^{-1}} \int_{\mathcal{Q}_{r_2}\cap\{|D^{\tau_0} d_s\tilde{u}|\geq b_u\lambda\}} |D^{\tau_0} d_s \tilde{u}|^{2} d\mu_{\tau_0,t}\, d\lambda \end{aligned}$$

for some constant $c = c(n, s, L)$ as $\gamma < 2$ and the constant b_u depends only on n, s and L. Considering the above two inequalities, (A.7) and the estimate J_1, we estimate J_2 as

$$\begin{aligned} J_2 &\leq c\epsilon \int_{\lambda_0}^{N\mathcal{M}^{-1}} \mathcal{M}^{\epsilon+2-2_\#}\lambda^{\epsilon-1} \int_{\mathcal{Q}_{r_2}\cap\{|D^{\tau_0} d_s\tilde{u}|\geq b_u\lambda\}} |D^{\tau_0} d_s \tilde{u}|^2 d\mu_{\tau_0,t}\, d\lambda \\ &\leq c\mathcal{M}^{\epsilon+2-2_\#} \int_{\mathcal{Q}_{r_2}} |D^{\tau_0} d_s \tilde{u}|^{2+\epsilon}_{N\mathcal{M}^{-1}} d\mu_{\tau_0,t} \\ &\leq \frac{1}{2^{10n}} \int_{\mathcal{Q}_{r_2}} |D^{\tau_0} d_s \tilde{u}|^{2+\epsilon}_{N\mathcal{M}^{-1}} d\mu_{\tau_0,t} \end{aligned}$$

by taking $\mathcal{M} = \mathcal{M}(n, s, L) > 1$ sufficiently large so that $c\mathcal{M}^{\frac{2-2_\#}{2}} \leq \frac{1}{2^{10n}}$ (thanks to $\mathcal{M}^{\epsilon+2-2_\#} \leq \mathcal{M}^{\frac{2-2_\#}{2}}$ as $\mathcal{M} \geq 1$ and $\epsilon < \frac{2_\#-2}{2}$). We next select $\epsilon = \epsilon(n, s, L) < 1$ such that

$$J_1 \leq \frac{1}{2^{10n}} \int_{\mathcal{Q}_{r_2}} |D^{\tau_0} d_s \tilde{u}|^{2+\epsilon}_{N\mathcal{M}^{-1}} d\mu_{\tau_0,t}.$$

Combining all the estimates of I and J, we observe that

$$\int_{\mathcal{Q}_{r_1}} |D^{\tau_0} d_s u|^{2+\epsilon}_{N\mathcal{M}^{-1}} d\mu_{\tau_0,t} \leq \int_{\mathcal{Q}_{r_1}} |D^{\tau_0} d_s u|^{2+\epsilon}_{N} d\mu_{\tau_0,t} \leq \frac{1}{2^{9n}} \int_{\mathcal{Q}_{r_2}} |D^{\tau_0} d_s u|^{2+\epsilon}_{N\mathcal{M}^{-1}} d\mu_{\tau_0,t} + c\lambda_0^{2+\epsilon}\mu_{\tau_0,t}\left(\mathcal{Q}_2\right)$$

holds if $N > \mathcal{M}\lambda_0$. After a few algebraic computations along with (2.2), we have (A.16). Applying Lemma 2.4 to (A.16), we obtain

$$\begin{aligned} \left(\frac{1}{\tau_0} \fint_{\mathcal{Q}_1} |D^{\tau_0} d_s \tilde{u}|^{2+\epsilon}_{N}\, d\mu_{\tau_0,t}\right)^{\frac{1}{2+\epsilon}} &\leq c\left(\frac{1}{\tau_0} \fint_{\mathcal{Q}_2} |D^{\tau_0} d_s \tilde{u}|^{2} d\mu_{\tau_0,t}\right)^{\frac{1}{2}} + c\,\mathrm{Tail}_{\infty,2s}\left(\frac{\tilde{u} - (\tilde{u})_{B_2}(t)}{2^{s+\tau}}; Q_2\right) \\ &\quad + c\sup_{t\in\Lambda_2}\left(\fint_{B_2} \frac{|\tilde{u} - (\tilde{u})_{Q_2}|^2}{2^{2s+2\tau}}\right)^{\frac{1}{2}} \end{aligned}$$

for some constant $c = c(n, s, L)$. By passing to the limit $N \to \infty$ and using the change of variables, we get the desired estimate (A.15). □

Appendix B. Existence and uniqueness

In this section, we present the existence result for the corresponding boundary value problem of (1.1) and the standard energy estimate. Before stating the result, we first note from [47, Proposition 1.2 in Chapter 3] with $V = W^{s,2}(\Omega)$ and $H = L^2(\Omega)$ that if $h \in L^2(0,T;W^{s,2}(\Omega))$ and $h_t \in \left(L^2(0,T;W^{s,2}(\Omega))\right)^*$, then $h \in C([0,T];L^2(\Omega))$ with the estimate

$$\sup_{t\in[0,T]} \|h(\cdot,t)\|_{L^2(\Omega)} \leq c\|h\|_{L^2(0,T;W^{s,2}(\Omega))} + c\|h_t\|_{(L^2(0,T;W^{s,2}(\Omega))^*} \tag{B.1}$$

for some constant $c = c(n,s)$.

Lemma B.1. *Let Ω' be an open and bounded set such that $\Omega \Subset \Omega'$. Suppose that*

$$\begin{aligned}
&f \in L^2\left(0,T;L^1_s(\mathbb{R}^n)\right) \quad \textit{with} \quad d_0 f \in L^2\left(\Omega'\times\Omega'\times(0,T)\,;\,\frac{dx\,dy\,dt}{|x-y|^n}\right),\\
&g \in L^{\frac{2(n+2s)}{n+4s}}(\Omega_T),\\
&h \in L^2(0,T;W^{s,2}(\Omega')) \cap L^\infty(0,T;L^1_{2s}(\mathbb{R}^n)) \quad \textit{and} \quad h_t \in \left(L^2(0,T;W^{s,2}(\Omega'))\right)^*.
\end{aligned} \tag{B.2}$$

Then there is a unique weak solution $u \in L^2(0,T;W^{s,2}(\Omega)) \cap L^\infty\left(0,T;L^1_{2s}(\mathbb{R}^n)\right) \cap C([0,T];L^2(\Omega))$ to

$$\begin{cases} u_t + \mathcal{L}^\Phi_A u = (-\Delta)^{\frac{s}{2}} f + g & \textit{in } \Omega\times(0,T),\\ \qquad u = h & \textit{in } \mathbb{R}^n\setminus\Omega\times[0,T],\\ \quad u(\cdot,0) = h(\cdot,0) & \textit{in } \Omega \end{cases} \tag{B.3}$$

with the estimate

$$\begin{aligned}
&\sup_{t\in(0,T)} \int_\Omega |u(x,t)|^2\,dx + \int_0^T\int_\Omega\int_\Omega |d_s u|^2 \frac{dx\,dy\,dt}{|x-y|^n}\\
&\leq c\int_0^T\int_{\Omega'}\int_{\Omega'} |d_0 f|^2\frac{dx\,dy\,dt}{|x-y|^n} + c\,\mathrm{Tail}_{2,s}(f-(f)_{\Omega'}(t);\Omega'_T)^2 + c\left(\int_0^T\int_\Omega |g|^{\frac{2(n+2s)}{n+4s}}\,dz\right)^{\frac{n+4s}{n+2s}}\\
&\quad + c\int_0^T\int_{\Omega'}\int_{\Omega'} |d_s h|^2\frac{dx\,dy}{|x-y|^n} + c\,\mathrm{Tail}_{2,2s}(h-(h)_{\Omega'}(t);\Omega'_T)^2 + c\|h_t\|^2_{(L^2(0,T;W^{s,2}(\Omega'))^*}
\end{aligned} \tag{B.4}$$

for some constant $c = c(n,s,L,T,\Omega,\Omega')$.

Proof. From [13, Lemma 2.7], we observe that

$$T_{f(\cdot,t)} : \phi \mapsto \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} (f(x,t)-f(y,t))(\phi(x)-\phi(y))\frac{dx\,dy}{|x-y|^{n+s}}, \quad \phi \in X^{s,2}_0(\Omega,\Omega')$$

is an element of the dual space of $X^{s,2}_0(\Omega,\Omega')$. This implies that

$$(-\Delta)^{\frac{s}{2}} f \in \left(L^2(0,T;W^{s,2}(\Omega'))\right)^*.$$

Therefore, combining [10, Thoerem A.3] and [11, Lemma A.1], we find a unique weak solution u to (B.3).

We are now in the position to prove (B.4). Since $u - h \in L^2\left(0,T;X^{s,2}_0(\Omega,\Omega')\right)$, using the standard approximation argument, we have

$$\begin{aligned}
&\sup_{t\in(0,T)} \int_\Omega |(u-h)(x,t)|^2\,dx + \int_0^T\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |d_s(u-h)|^2\frac{dx\,dy\,dt}{|x-y|^n}\\
&\leq c\left|\int_0^T \langle h_t, u-h\rangle_{X^{s,2}_0(\Omega,\Omega'),(X^{s,2}_0(\Omega,\Omega'))^*}\,dt\right| + c\int_0^T\int_\Omega |g(u-h)|\,dz\\
&\quad + c\int_0^T\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |d_0 f||d_s(u-h)|\frac{dx\,dy\,dt}{|x-y|^n} + c\int_0^T\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |d_s h||d_s(u-h)|\frac{dx\,dy\,dt}{|x-y|^n} =: I_1+I_2+I_3+I_4.
\end{aligned} \tag{B.5}$$

From (B.2) and Young's inequality, we first estimate I_1 as

$$I_1 \leq c\|h_t\|^2_{(L^2(0,T;W^{s,2}(\Omega'))^*} + \frac{1}{8}\int_0^T\int_{\Omega'}\int_{\Omega'} |d_s(u-h)|^2\frac{dx\,dy\,dt}{|x-y|^n}.$$

We next estimate I_2 with the help of (B.2) as below

$$I_2 \leq c\left(\int_0^T\int_\Omega |g|^{\frac{2(n+2s)}{n+4s}}\,dz\right)^{\frac{n+4s}{n+2s}} + \frac{1}{8}\sup_{t\in(0,T)}\int_\Omega |(u-h)(x,t)|^2\,dx + \frac{1}{8}\int_0^T\int_{\Omega'}\int_{\Omega'} |d_s(u-h)|^2\frac{dx\,dy\,dt}{|x-y|^n},$$

where we have used Lemma 2.1 and Young's inequality. For the estimate of I_3+I_4 , we follow the same line as in the estimate of J_1+J_2 in [13, Lemma 2.7] to see that

$$\begin{aligned} I_3+I_4 \leq \frac{1}{8}\int_0^T\int_{\Omega'}\int_{\Omega'} |d_s(u-h)|^2\frac{dx\,dy\,dt}{|x-y|^n} + c\int_0^T\int_{\Omega'}\int_{\Omega'} |d_0 f|^2\frac{dx\,dy\,dt}{|x-y|^n} + c\,\mathrm{Tail}_{2,s}(f-(f)_{\Omega'}(t);\Omega'_T)^2 \\ + c\int_0^T\int_{\Omega'}\int_{\Omega'} |d_s h|^2\frac{dx\,dy}{|x-y|^n} + c\,\mathrm{Tail}_{2,2s}(h-(h)_{\Omega'}(t);\Omega'_T)^2 \end{aligned}$$

for some constant $c=c(n,s,L,\Omega,\Omega',T)$. We now plug the estimates of I_1, I_2 and I_3 into (B.5) to see that

$$\begin{aligned} &\sup_{t\in(0,T)}\int_\Omega |(u-h)(x,t)|^2\,dx + \int_0^T\int_\Omega\int_\Omega |d_s(u-h)|^2\frac{dx\,dy\,dt}{|x-y|^n} \\ &\leq c\int_0^T\int_{\Omega'}\int_{\Omega'} |d_0 f|^2\frac{dx\,dy\,dt}{|x-y|^n} + c\,\mathrm{Tail}_{2,s}(f-(f)_{\Omega'}(t);\Omega'_T)^2 + c\left(\int_0^T\int_\Omega |g|^{\frac{2(n+2s)}{n+4s}}\,dz\right)^{\frac{n+4s}{n+2s}} \\ &\quad + c\int_0^T\int_{\Omega'}\int_{\Omega'} |d_s h|^2\frac{dx\,dy}{|x-y|^n} + c\,\mathrm{Tail}_{2,2s}(h-(h)_{\Omega'}(t);\Omega'_T)^2 + c\|h_t\|^2_{\left(L^2(0,T;W^{s,2}(\Omega'))\right)^*}. \end{aligned}$$

After a few simple calculations along with (B.1), we obtain the desired estimate (B.4). □

References

[1] B. Abdellaoui, A. J Fernández, T. Leonori, and A. Younes, *Global fractional Calderón-Zygmund regularity*, arXiv:2107.06535 (2021), 22pp.
[2] E. Acerbi and G. Mingione, *Gradient estimates for a class of parabolic systems*, Duke Math. J. **136** (2007), no. 2, 285–320.
[3] K. Adimurthi, H. Prasad, and V. Tewary, *Local Hölder regularity for nonlocal parabolic p-Laplace equations*, arXiv:2205.09695 (2022), 31pp.
[4] P. Auscher, S. Bortz, M. Egert, and O. Saari, *Nonlocal self-improving properties: a functional analytic approach*, Tunis. J. Math. **1** (2019), no. 2, 151–183.
[5] U. Biccari, M. Warma, and E. Zuazua, *Local elliptic regularity for the Dirichlet fractional Laplacian*, Adv. Nonlinear Stud. **17** (2017), no. 2, 387–409.
[6] U. Biccari, M. Warma, and E. Zuazua, *Local regularity for fractional heat equations*, "Recent advances in PDEs: analysis, numerics and control", SEMA SIMAI Springer Ser., vol. 17, Springer, Cham, 2018, pp. 233–249.
[7] V. Bögelein, *Global Calderón-Zygmund theory for nonlinear parabolic systems*, Calc. Var. Partial Differential Equations **51** (2014), no. 3-4, 555–596.
[8] L. Brasco and E. Lindgren, *Higher Sobolev regularity for the fractional p-Laplace equation in the superquadratic case*, Adv. Math. **304** (2017), 300–354.
[9] L. Brasco, E. Lindgren, and A. Schikorra, *Higher Hölder regularity for the fractional p-Laplacian in the superquadratic case*, Adv. Math. **338** (2018), 782–846.
[10] L. Brasco, E. Lindgren, and M. Strömqvist, *Continuity of solutions to a nonlinear fractional diffusion equation*, J. Evol. Equ. **21** (2021), no. 4, 4319–4381.
[11] S.-S. Byun, H. Kim, and K. Kim, *Higher Hölder regularity for nonlocal parabolic equations with irregular kernels*, J. Evol. Equ. **23** (2023), no. 3, Paper No. 53, 59.
[12] S.-S. Byun, H. Kim, and J. Ok, *Local Hölder continuity for fractional nonlocal equations with general growth*, Math. Ann. **387** (2023), no. 1-2, 807–846.
[13] S.-S. Byun and K. Kim, *L^q estimates for nonlocal p-Laplacian type equations with BMO kernel coefficients in divergence form*, arXiv:2303.08517 (2023), 48pp.
[14] S.-S. Byun, D. Kumar, and H.-S. Lee, *Global gradient estimates for the mixed local and nonlocal problems with measurable nonlinearities*, Calc. Var. Partial Differential Equations **63** (2024), no. 2, Paper No. 27, 48.
[15] S.-S. Byun, J. Ok, and K. Song, *Hölder regularity for weak solutions to nonlocal double phase problems*, J. Math. Pures Appl. (9) **168** (2022), 110–142.
[16] L. Caffarelli, C H Chan, and A Vasseur, *Regularity theory for parabolic nonlinear integral operators*, J. Amer. Math. Soc. **24** (2011), no. 3, 849–869.
[17] M. Cozzi, *Interior regularity of solutions of non-local equations in Sobolev and Nikol'skii spaces*, Ann. Mat. Pura Appl. (4) **196** (2017), no. 2, 555–578.
[18] A. Di Castro, T. Kuusi, and G. Palatucci, *Local behavior of fractional p-minimizers*, Ann. Inst. H. Poincaré Anal. Non Linéaire **33** (2016), no. 5, 1279–1299.
[19] L. Diening and S. Nowak, *Calderón-Zygmund estimates for the fractional p-Laplacian*, arXiv:2303.02116 (2023), 27pp.
[20] M. Ding, C. Zhang, and S. Zhou, *Local boundedness and Hölder continuity for the parabolic fractional p-Laplace equations*, Calc. Var. Partial Differential Equations **60** (2021), no. 1, Paper No. 38, 45.
[21] H. Dong, P. Jung, and D. Kim, *Boundedness of non-local operators with spatially dependent coefficients and L_p-estimates for non-local equations*, Calc. Var. Partial Differential Equations **62** (2023), no. 2, Paper No. 62, 28.

[22] H. Dong and D. Kim, *On L_p-estimates for a class of non-local elliptic equations*, J. Funct. Anal. **262** (2012), no. 3, 1166–1199.

[23] M. M. Fall, T. Mengesha, A. Schikorra, and S. Yeepo, *Calderón-Zygmund theory for non-convolution type nonlocal equations with continuous coefficient*, Partial Differ. Equ. Appl. **3** (2022), no. 2, Paper No. 24, 27.

[24] C. De Filippis and G. Mingione, *Gradient regularity in mixed local and nonlocal problems*, Math. Ann. **388** (2022), no. 1, 261–328.

[25] G. Giacomin, J.L. Lebowitz, and E. Presutti, *Deterministic and stochastic hydrodynamic equations arising from simple microscopic model systems*, "Stochastic partial differential equations: six perspectives", Math. Surveys Monogr., vol. 64, Amer. Math. Soc., Providence, RI, 1999, pp. 107–152.

[26] J. Giacomoni, D. Kumar, and K. Sreenadh, *Hölder regularity results for parabolic nonlocal double phase problems*, Adv. Differential Equations **29** (2024), no. 11-12, 899–950.

[27] G. Gilboa and S. Osher, *Nonlocal operators with applications to image processing*, Multiscale Model. Simul. **7** (2008), no. 3, 1005–1028.

[28] E. Giusti, "Direct methods in the calculus of variations", World Scientific Publishing Co., Inc., River Edge, NJ, 2003.

[29] G. Grubb, *Regularity in L_p Sobolev spaces of solutions to fractional heat equations*, J. Funct. Anal. **274** (2018), no. 9, 2634–2660.

[30] M. Kassmann, *A priori estimates for integro-differential operators with measurable kernels*, Calc. Var. Partial Differential Equations **34** (2009), no. 1, 1–21.

[31] M. Kassmann and M. Weidner, *The parabolic harnack inequality for nonlocal equations*, arXiv:2303.05975 (2023), 25pp.

[32] S. Khomrutai, A. Schikorra, A. Seesanea, and S. Yeepo, *Global tangential calderon-zygmund type estimates for the regional fractional laplacian*, arXiv:2209.07939 (2022), 22pp.

[33] T. Kuusi and G. Mingione, *Universal potential estimates*, J. Funct. Anal. **262** (2012), no. 10, 4205–4269.

[34] T. Kuusi, G. Mingione, and Y. Sire, *Nonlocal equations with measure data*, Comm. Math. Phys. **337** (2015), no. 3, 1317–1368.

[35] T. Kuusi, G. Mingione, and Y. Sire, *Nonlocal self-improving properties*, Anal. PDE **8** (2015), no. 1, 57–114.

[36] T. Kuusi, S. Nowak, and Y. Sire, *Gradient regularity and first-order potential estimates for a class of nonlocal equations*, arXiv:2212.01950 (2022), 56pp.

[37] T. Leonori, I. Peral, A. Primo, and F. Soria, *Basic estimates for solutions of a class of nonlocal elliptic and parabolic equations*, Discrete Contin. Dyn. Syst. **35** (2015), no. 12, 6031–6068.

[38] N. Liao, *Hölder regularity for parabolic fractional p-Laplacian*, Calc. Var. PDE**63**, (2024), 34pp.

[39] K. Mazowiecka and A. Schikorra, *Fractional div-curl quantities and applications to nonlocal geometric equations*, J. Funct. Anal. **275** (2018), no. 1, 1–44.

[40] T. Mengesha, A. Schikorra, and S. Yeepo, *Calderon-Zygmund type estimates for nonlocal PDE with Hölder continuous kernel*, Adv. Math. **383** (2021), Paper No. 107692, 64pp.

[41] S. Nowak, *$H^{s,p}$ regularity theory for a class of nonlocal elliptic equations*, Nonlinear Anal. **195** (2020), 111730, 28pp.

[42] S. Nowak, *Improved Sobolev regularity for linear nonlocal equations with VMO coefficients*, Math. Ann. **385** (2022), no. 3-4, 1323–1378.

[43] S. Nowak, *Regularity theory for nonlocal equations with VMO coefficients*, Ann. Inst. H. Poincaré C Anal. Non Linéaire **40** (2023), no. 1, 61–132.

[44] J. Ok, *Local Hölder regularity for nonlocal equations with variable powers*, Calc. Var. Partial Differential Equations **62** (2023), no. 1, Paper No. 32, 31.

[45] A. Schikorra, *Nonlinear commutators for the fractional p-Laplacian and applications*, Math. Ann. **366** (2016), no. 1-2, 695–720.

[46] J. M. Scott and T. Mengesha, *Self-improving inequalities for bounded weak solutions to nonlocal double phase equations*, Commun. Pure Appl. Anal. **21** (2022), no. 1, 183–212.

[47] R. E. Showalter, "Monotone operators in Banach space and nonlinear partial differential equations", Mathematical Surveys and Monographs, vol. 49, American Mathematical Society, Providence, RI, 1997.

[48] M. Strömqvist, *Harnack's inequality for parabolic nonlocal equations*, Ann. Inst. H. Poincaré Anal. Non Linéaire **36** (2019), no. 6, 1709–1745.

[49] A. Tavakoli, *A perturbative approach to Hölder continuity of solutions to a nonlocal p-parabolic equation*, J. Evol. Equ. **24**, (2024), 76pp.

[50] J. L. Vázquez, *Asymptotic behaviour for the fractional heat equation in the Euclidean space*, Complex Var. Elliptic Equ. **63** (2018), no. 7-8, 1216–1231.

[51] X. Zhang, *L^p-maximal regularity of nonlocal parabolic equations and applications*, Ann. Inst. H. Poincaré C Anal. Non Linéaire **30** (2013), no. 4, 573–614.

Sun-Sig Byun: Department of Mathematical Sciences and Research Institute of Mathematics, Seoul National University, Seoul 08826, Korea
Email address: byun@snu.ac.kr

Kyeongbae Kim: Department of Mathematical Sciences, Seoul National University, Seoul 08826, Korea
Email address: kkba6611@snu.ac.kr

Deepak Kumar: Research Institute of Mathematics, Seoul National University, Seoul 08826, Korea
Email address: deepak.kr0894@gmail.com