Borderline gradient continuity for fractional heat type operators

Vedansh Arya and Dharmendra Kumar
Tata Institute of Fundamental Research, Centre For Applicable Mathematics, Bangalore 560065, India (vedansh@tifrbng.res.in, dharmendra2020@tifrbng.res.in)

(Received 10 February 2022; accepted 8 September 2022)

In this paper, we establish gradient continuity for solutions to
\[(\partial_t - \text{div}(A(x)\nabla))^s u = f, \quad s \in (1/2, 1),\]
when \(f\) belongs to the scaling critical function space \(L\left(\frac{n+2}{2s-1}, 1\right)\). Our main results theorems 1.1 and 1.2 can be seen as a nonlocal generalization of a well-known result of Stein in the context of fractional heat type operators and sharpen some of the previous gradient continuity results which deal with \(f\) in subcritical spaces. Our proof is based on an appropriate adaptation of compactness arguments, which has its roots in a fundamental work of Caffarelli in [13].

Keywords: Extension problem; Fractional heat operator; Master equations; Borderline gradient continuity

2020 Mathematics subject classification: 35R11; 35B65; 35R09

1. Introduction and the statement of the main result

In this article we prove gradient continuity for the following nonlocal operators \((\partial_t - \text{div}(A(x)\nabla))^s\) which are modelled on the fractional heat operator \((\partial_t - \Delta)^s\) with critical scalar perturbations. To provide some context to our work, we note that the study of the fractional heat operator \((\partial_t - \Delta)^s\) was first proposed in M. Riesz’ visionary papers [37] and [38]. This nonlocal operator represents a basic model of the continuous time random walk (CTRW) introduced by Montroll and Weiss in [33]. We recall that a CTRW is a generalization of a random walk where the wandering particle waits for a random time between jumps. It is a stochastic jump process with arbitrary distributions of jump lengths and waiting times. In [32], Klafter and Metzler describe such processes by means of the nonlocal equation both in space and time

\[\eta(t, x) = \int_0^\infty \int_R \Psi(\tau, z) \eta(t - \tau, x - z) \, dz \, d\tau,\]

© The Author(s), 2022. Published by Cambridge University Press on behalf of The Royal Society of Edinburgh
which is an example of a master equation. Such equations were introduced in 1973 by
Kenkre, Montroll and Shlesinger in [24], and they are presently receiving increasing
attention by mathematicians also thanks to the work [16] of Caffarelli and
Silvestre in which the authors establish the Hölder continuity of viscosity solutions
of generalized master equations

\[ Lu(t, x) = \int_0^\infty \int_{\mathbb{R}^n} K(t, x; \tau, z)[u(t, x) - u(t - \tau, x - z)]\, dz\, d\tau = 0. \quad (1.1) \]

On the kernel \( K \) they assume that there exist \( 0 < s < 1 \) and \( \beta > 2s \) such that for
\( 0 < c_1 < c_2 \) one can find \( 0 < \lambda \leq \Lambda \) for which for a.e. \( (t, x) \in \mathbb{R} \times \mathbb{R}^n \) one has

\[
\begin{cases}
K(t, x; \tau, z) \geq \frac{\lambda}{|z|^{n+2s+\beta}} & \text{when } c_1|z|^{\beta} \leq \tau \leq c_2|z|^{\beta}, \\
K(t, x; \tau, z) \leq \frac{\Lambda}{|z|^{n+2s+\beta+\gamma}} & \text{when } 0 \leq \tau \leq c_1|z|^{\beta}.
\end{cases} \quad (1.2)
\]

The pseudo-differential operator \((\partial_t - \Delta)^s\) which is defined via the Bochner’s
subordination principle in the following way (see for instance [5, 8, 40]),

\[
(\partial_t - \Delta)^s u(t, x) = \frac{s}{\Gamma(1-s)} \int_0^\infty \int_{\mathbb{R}^n} \tau^{-s-1}G(\tau, z)[u(t, x) - u(t - \tau, x - z)]\, dz\, d\tau,
\]

\[
= \frac{s}{\Gamma(1-s)} \int_0^t \int_{\mathbb{R}^n} (t - \tau)^{-s-1}G(t - \tau, x - z)[u(t, x) - u(t, z)]\, dz\, d\tau, \quad (1.3)
\]

where \( G(\tau, z) = (4\pi\tau)^{-\frac{n}{2}} e^{-\frac{|z|^2}{4\tau}} \) is the standard heat kernel and \( \Gamma(z) \) indicates Euler
gamma function, is seen to be a special case of the master equation. We mention that
recently Nyström and Sande [35] and Stinga and Torrea [40] have independently adapted
to the fractional heat operator of the celebrated extension procedure of
Caffarelli and Silvestre in [15] which can be described as follows.

Given \( s \in (0, 1) \) we introduce the parameter

\[ a = 1 - 2s \in (-1, 1), \]

and indicate with \( U = U(t, X) \), where \( X = (x, y) \) is a point in \( \mathbb{R}^n \times (0, \infty) \), the
solution to the following extension problem

\[
\begin{aligned}
&y^a \frac{\partial U}{\partial t} = \text{div}_X(y^a \nabla X U), \quad (t, X) \in \mathbb{R} \times \mathbb{R}^n \times (0, \infty), \\
&U(t, x, 0) = u(t, x), \quad (t, x) \in \mathbb{R} \times \mathbb{R}^n.
\end{aligned} \quad (1.4)
\]

Using an appropriate Poisson representation of the extension problem it was
proved in [35, 40], see also § 3 in [8] for details, that one has in \( L^2(\mathbb{R} \times \mathbb{R}^n) \)

\[
- \frac{2^{2s-1}\Gamma(s)}{\Gamma(1-s)} \lim_{y \to 0^+} y^a \frac{\partial U}{\partial y}(t, x, y) = (\partial_t - \Delta)^s u(t, x). \quad (1.5)
\]

Such an extension problem has been generalized for fractional powers of variable
coefficient operators such as \((\partial_t - \text{div}(A(x)\nabla))^s\) in [11, 12]. We also refer to [9]
for a generalization of the extension problem in the subelliptic situation. We would
like to mention that the study of the fractional heat type operators as well as the related extension problem has received a lot of attention in recent times, see for instance [2–4, 6–8, 12, 19, 29, 30]. We would also like to mention that extension problem is prototype of the equations with general $A_2$ weight studied by Chiarenza and Serapioni in [17].

Now we will state our main result: Consider the following problem

$$
\begin{align*}
    y^a \partial_t U - \text{div}(y^a B(x) \nabla U) &= 0 \quad \text{in } Q_1 \times (0,1) \\
    -y^a U_y \big|_{y=0} &= f \quad \text{on } Q_1,
\end{align*}
$$

(1.6)

where $a = 1 - 2s$; for $1/2 < s < 1$,

$$
B = \begin{bmatrix} A & 0 \\ 0 & 1 \end{bmatrix}
$$

and $A$ is a uniformly elliptic matrix with Dini modulus of continuity $\omega_A$ and $f$ belongs to the scaling critical function space $L \left( \frac{n+2}{2s-1}, 1 \right)$.

**Theorem 1.1.** Let $U$ be a weak solution of (1.6) in $Q_1 \times (0,1)$. Then there exist modulus of continuity $K$ depending on ellipticity, $\omega_A$, $n$, $s$ and $f$ such that for all $(t_1, x_1, y_1), (t_2, x_2, y_2) \in Q_{1/2} \times (0,1/2)$

$$
|\nabla U(t_1, x_1, y_1) - \nabla U(t_2, x_2, y_2)| \leq CK(||(t_1, x_1, y_1) - (t_2, x_2, y_2)||) \quad (1.7)
$$

and

$$
|U(t_1, x, y) - U(t_2, x, y)| \leq CK(\sqrt{|t_1 - t_2|}) \sqrt{|t_1 - t_2|},
$$

where $C^2 = \int_{Q_1} U(t, x, 0)^2 \, dt \, dx + \int_{Q_1 \times (0,1)} U(t, x, y)^2 y^a \, dt \, dx \, dy$.

In view of the extension problem for $(\partial_t - \text{div}(A(x) \nabla))^s$ in [12] (see § 2 for relevant details), we obtain consequently that the following regularity result holds for the nonlocal fractional heat type problem.

**Theorem 1.2.** Let $s \in (1/2, 1)$ and let $u$ solve weakly to $(\partial_t - \text{div}(A(x) \nabla))^s u = f$ where $A(x)$ is uniformly elliptic with Dini coefficients and $f \in L \left( \frac{n+2}{2s-1}, 1 \right)$. Then $\nabla_x u$ is continuous.

Now to put our results in the right historical perspective, we note that in 1981, E. Stein in his work [39] showed the following ‘limiting’ case of Sobolev embedding theorem.

**Theorem 1.3.** Let $L(n, 1)$ denote the standard Lorentz space, then the following implication holds:

$$
\nabla v \in L(n, 1) \implies v \text{ is continuous.}
$$
The Lorentz space $L(n, 1)$ appearing in theorem 1.3 consists of those measurable functions $g$ satisfying the condition

$$\int_0^\infty |\{x : g(x) > t\}|^{1/n} \, dt < \infty.$$  

Theorem 1.3 can be regarded as the limiting case of Sobolev–Morrey embedding that asserts

$$\nabla v \in L^{n+\varepsilon} \implies v \in C^{0, \frac{n}{n+\varepsilon}}.$$  

Note that indeed $L^{n+\varepsilon} \subset L(n, 1) \subset L^n$ for any $\varepsilon > 0$ with all the inclusions being strict. Now theorem 1.3 coupled with the standard Calderon–Zygmund theory has the following interesting consequence.

**Theorem 1.4.** If $\Delta u \in L(n, 1)$ then this implies $\nabla u$ is continuous.

Similar result holds in the parabolic situation for more general variable coefficient operators when $f \in L(n+2, 1)$. The analogue of theorem 1.4 for general nonlinear and possibly degenerate elliptic and parabolic equations became accessible not so long ago through a rather sophisticated and powerful nonlinear potential theory (see for instance [21, 25, 27] and the references therein). The first breakthrough in this direction came up in the work of Kuusi and Mingione in [26] where they showed that the analogue of theorem 1.4 holds for operators modelled after the variational $p$-Laplacian. Such a result was subsequently generalized to $p$-Laplacian type systems by the same authors in [28].

Since then, there has been several generalizations of theorem 1.4 to operators with other kinds of nonlinearities and in context of fully nonlinear elliptic equations, the analogue of theorem 1.4 has been established by Daskalopoulos–Kuusi–Mingione in [18]. We also refer to [1] for the corresponding boundary regularity result and also to [10] for a similar result in the context of the game theoretic normalized $p$–Laplacian operator. Our main result theorem 1.2 can thus be thought of as a nonlocal generalization of the Stein’s theorem in the sense that $s \to 1$, it exactly reproduces the result theorem 1.3 for time-independent $f$. Moreover, theorem 1.2 is also seen to be the limiting case of theorem 1.2 (ii) in [12] which instead deals gradient Hölder continuity in case of subcritical scalar perturbations when $f \in L^p$ for $p > \frac{n+2}{2s-1}$. As the reader will see, our proof is based on a somewhat delicate adaptation of the Caffarelli style compactness arguments in [1] to the degenerate Neumann problem in (1.6).

The paper is organized as follows. In § 2, we introduce some basic notations and gather some preliminary results. In § 3, we prove our main result. Finally in the appendix, we give a self-contained proof of a basic existence result, used in our approximation lemma 3.1, and a self-contained proof of Companato type characterization, theorem 2.12.

## 2. Notations and preliminaries

We will denote generic point of thick space $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}$ by $(t, X) = (t, x, y)$. In general we will identify thin space $\mathbb{R} \times \mathbb{R}^n \times \{0\}$ by $\mathbb{R} \times \mathbb{R}^n$ and its generic point...
will be denoted by \((t, x)\). We will define cubes and balls in thin and thick space as following:

\[
Q_\rho(x_0, t_0) = (t_0 - \rho^2, t_0 + \rho^2) \times \{x \in \mathbb{R}^n : ||x - x_0|| < \rho\};
\]
\[
Q_\rho^*(x_0, t_0, y_0) = (t_0 - \rho^2, t_0 + \rho^2) \times \{x \in \mathbb{R}^n : ||x - x_0|| < \rho\} \times (y_0, y_0 + \rho);
\]
\[
Q_\rho(x_0, t_0, y_0) = (t_0 - \rho^2, t_0 + \rho^2) \times \{x \in \mathbb{R}^n : ||x - x_0|| < \rho\} \times (y_0 - \rho, y_0 + \rho).
\]

Similarly we define \(B_r, B_r^*\) and \(\mathcal{B}_r\). For \((t_1, X_1)\) and \((t_2, X_2)\), we let

\[
||(t_1, X_1) - (t_2, X_2)|| := \max\{\sqrt{|t_1 - t_2|}, ||X_1 - X_2||\},
\]

where \(||X_1 - X_2||\) denotes Euclidean norm. We denote \(Q_\rho(0, 0)\) by \(Q\) etc. For \(((t_1, t_2) \times \Omega) \subset \mathbb{R} \times \mathbb{R}^{n+1}\), \(\partial_p((t_1, t_2) \times \Omega)\) will denote the parabolic boundary of \((t_1, t_2) \times \Omega\) and defined as:

\[
\partial_p((t_1, t_2) \times \Omega) = ((t_1, t_2) \times \Omega) \cup ([t_1, t_2] \times \partial \Omega).
\]

We will denote \(L^2(Q_1^t)\) with measure \(g^a dt dX\) by \(L^2_{a}(Q_1^t)\). Similarly we define \(H^1_{a}(Q_1^t)\) or \(H^1_{0,a}(Q_1^t)\). Sometimes we will denote \(C([-1, 1]; L^2_{a}(B_1^t))\) by \(C(-1, 1; L^2_{a}(B_1^t))\) and other spaces in similar fashion. Let \(F = F(x, t) = (F_1, \ldots, F_n, F_{n+1})\) be an \(\mathbb{R}^{n+1}\)-valued vector field defined on \(Q_1^t\) such that

\[
F_{n+1} = 0, \quad |F| \in L^2(Q_1^t), \quad \text{and let } f = f(t, x) \in L^2(Q_1^t).
\]

We now recall the definition of the operator \((\partial_t - \text{div}(A(x)\nabla))^s\) from [11, 12].

Let \(\Omega \subset \mathbb{R}^n\) be a bounded Lipschitz domain. Define \(L = -\text{div}(A(x)\nabla)\). Then \(L\) has a discrete Dirichlet spectrum, that is \(L\) has a countable family of non-negative eigenvalues and eigenfunctions \((\lambda_k, \phi_k)\) for \(k = 0, 1, 2, \ldots\) such that the set \(\{\phi_k\}\) forms an orthonormal basis for \(L^2(\Omega)\). Hence \(u(t, x) \in L^2(\mathbb{R} \times \Omega)\) can be written as

\[
u(t, x) = \frac{1}{\sqrt{2\pi}} \int_0^\infty \sum_{k=0}^\infty \hat{u}_k(\rho) \phi_k(x) e^{i\nu \rho} d\rho
\]

where for almost every \(t \in \mathbb{R}\),

\[
u_k(t) = \int_\Omega u(t, x) \phi_k(x) dx
\]

and \(\hat{u}_k(\rho)\) is the Fourier transform of \(u_k(t)\) in \(t\) variable. We define the domain of the operator \((\partial_t - \text{div}(A(x)\nabla))^s, 0 \leq s \leq 1\), as following:

\[
\text{Dom}(H^s) = \left\{ u \in L^2(\mathbb{R} \times \Omega) : ||u||^2_{H^s} := \int_{\mathbb{R}} \int_\Omega \sum_{k=0}^\infty |i\rho + \lambda_k|^s |\phi_k(\rho)|^2 d\rho < \infty \right\}.
\]

Given \(u \in \text{Dom}(H^s)\), we define \(H^s u = (\partial_t - \text{div}(A(x)\nabla))^s u\) as a bounded linear functional on \(\text{Dom}(H^s)\) that acts on \(v \in \text{Dom}(H^s)\) by the following formula

\[
\langle H^s u, v \rangle := \int_{\mathbb{R}} \int_\Omega \sum_{k=0}^\infty (i\rho + \lambda_k)^s \hat{u}_k(\rho) \hat{\phi}_k(\rho) d\rho, \quad (2.1)
\]

with the understanding that we have chosen the principal branch of the complex function \(z \rightarrow z^s\) and \(\hat{v}_k(\rho)\) denotes the complex conjugate of \(\hat{v}_k(\rho)\).
Since the set \( \{ \phi_k \} \) is an orthonormal basis of \( L^2(\Omega) \), we can write the semigroup \( \{ e^{-\tau L} \}_{\tau \geq 0} \) generated by \( L \) as

\[
\langle e^{-\tau L} \phi, \psi \rangle = \sum_{k=0}^{\infty} e^{-\tau \lambda_k} \phi_k \psi_k
\]

for any \( \phi, \psi \in L^2(\Omega) \), where \( \phi_k = \int_\Omega \phi \phi_k \, dx \) and \( \psi_k = \int_\Omega \psi \phi_k \, dx \). Given \( u \in L^2(\mathbb{R} \times \Omega) \), we define \( e^{-\tau H} u \)

\[
e^{-\tau H} u(t, x) = e^{-\tau L}(u(t - \tau, \cdot))(x)
\]

in the sense that for any \( v \in L^2(\mathbb{R}) \),

\[
\langle e^{-\tau H} u, v \rangle_{L^2(\mathbb{R} \times \Omega)} = \int_{\mathbb{R}} \sum_{k=0}^{\infty} e^{-\tau \lambda_k} u_k(t - \tau) v_k(t) \, dt.
\]

The following is theorem 1.6 from [11].

**Theorem 2.1.** Let \( u \in \text{Dom}(H^s) \). For \( (t, x) \in \mathbb{R} \times \Omega \) and \( y > 0 \), we define

\[
U(t, x, y) = \frac{y^{2s}}{4^s \Gamma(s)} \int_0^\infty e^{-\frac{x^2}{4\tau}} e^{-\tau H} u(t, x) \, \frac{d\tau}{\tau^{1+s}}.
\]

Then,

\[
U \in C^\infty((0, \infty); L^2(\mathbb{R} \times \Omega)) \cap C([0, \infty); L^2(\mathbb{R} \times \Omega)) \cap L^2(\mathbb{R}; H^1_a(\Omega \times (0, \infty)))
\]

is a weak solution to the following problem

\[
\begin{cases}
\displaystyle y^a \partial_t U = \text{div}(y^a B(x) \nabla U) & \text{in } \mathbb{R} \times \Omega \times (0, \infty) \\
\displaystyle -\frac{2^{2s-1}\Gamma(s)}{\Gamma(1-s)} y^a U_y \big|_{y=0} = H^s u & \text{on } \mathbb{R} \times \Omega \\
\displaystyle U(t, x, 0) = u(t, x) & \text{on } \mathbb{R} \times \Omega \\
\displaystyle U(t, x, y) = 0 & \text{on } \mathbb{R} \times \partial\Omega \times (0, \infty).
\end{cases}
\]

We now recall the proposition 3.6 from [12].

**Proposition 2.2.** Let \( U \) be as defined in (2.2) and assume that \( f \overset{\text{def}}{=} H^s u \in L^2(\mathbb{R} \times \Omega) \). Then \( U_t \in L^2(\mathbb{R}; (H^1_a(\Omega \times (0, \infty))^*) \), where \( (H^1_a(\Omega \times (0, \infty))^* \) is the dual space of \( H^1_a(\Omega \times (0, \infty)) \). In particular, we have

\[
U \in C(\mathbb{R}; L^2(\Omega \times (0, \infty))) \cap L^2(\mathbb{R}; H^1_a(\Omega \times (0, \infty))).
\]

Moreover \( U \) satisfies

\[
\int_{B^2} y^a U \phi|_{t=t_1}^{t=t_2} \, dX - \int_{t_1}^{t_2} \int_{B^2} y^a U \partial_t \phi \, dt \, dX + \int_{t_1}^{t_2} \int_{B^2} y^a (B(x) \nabla U) \cdot \nabla \phi \, dt \, dX
\]
2.3. We say a constant to be universal if it depends only on ellipticity,\[\text{Remark}\]

\[
\phi \in H^1([-1, 1]; L^2(B_1^*) \cap L^2([-1, 1]; H^1(B_1^*)) \text{ such that } \phi = 0 \text{ on } \partial_p Q_1^* \setminus (Q_1 \times \{0\}).
\]

Now consider the following local problem

\[
y^a \partial_t U - \text{div}(y^a B(x) \nabla U) = -\text{div}(y^a F) \quad \text{in } Q_1^*
\]

\[-y^a U_y|_{y=0} = f \quad \text{on } Q_1.
\]

We say that \(U \in C([-1, 1]; L^2(B_1^*)) \cap L^2([-1, 1]; H^1(B_1^*))\) is a weak solution to (2.5) if for every \(-1 < t_1 < t_2 < 1\)

\[
\int_{B_1^*} y^a U \phi|_{t=t_1}^{t=t_2} dX - \int_{t_1}^{t_2} \int_{B_1^*} y^a U \partial_t \phi \, dt \, dX + \int_{t_1}^{t_2} \int_{B_1^*} y^a (B(x) \nabla U) \cdot \nabla \phi \, dt \, dX
\]

\[
= \int_{t_1}^{t_2} \int_{B_1} f(t, x) \phi(t, x, 0) \, dt \, dX + \int_{t_1}^{t_2} \int_{B_1^*} y^a F \cdot \nabla \phi \, dt \, dX
\]

holds for all \(\phi \in H^1([-1, 1]; L^2(B_1^*)) \cap L^2([-1, 1]; H^1(B_1^*))\) such that \(\phi = 0 \) on \(\partial_p Q_1^* \setminus (Q_1 \times \{0\})\). Before proceeding further, we make the following discursive remark.

**Remark 2.3.** We say a constant to be universal if it depends only on ellipticity, \(n, s, L\left(\frac{n+2}{2s-1}, 1\right)\) norm of \(f\) and the Dini character of \(\omega_A\).

We will need the following lemmas from [12].

**Lemma 2.4.** Assume that \(U\) is a weak solution to (2.5) with \(F\) as described above. Then, for each \(\varphi \in C_c^\infty(Q_1 \times [0, 1])\) and for each \(t_1, t_2 \in (-1, 1)\) with \(t_1 < t_2\),

\[
\sup_{t_1 < t < t_2} \int_{B_1^*} y^a U^2 \varphi^2 \, dX + \int_{t_1}^{t_2} \int_{B_1^*} y^a \varphi^2 |\nabla U|^2 \, dX \, dt
\]

\[
\leq C \left[ \int_{t_1}^{t_2} \int_{B_1^*} y^a \left(|\partial_t (\varphi^2)| + |\nabla \varphi|^2 U^2 + |F|^2 \varphi^2 \right) \, dX \, dt + \int_{t_1}^{t_2} \int_{B_1} (\varphi(t, x, 0))^2 |U(t, x, 0)||f(t, x)| \, dx \, dt + \int_{B_1^*} y^a U^2(t_1, X) \varphi^2(t_1, X) \, dX, \right]
\]

where \(C = C(n, s, \text{ellipticity}) > 0\).

**Lemma 2.5.** Let \(W\) be a weak solution to

\[
\begin{align*}
\{ & y^a \partial_t W = \text{div}(y^a \nabla W) \quad \text{in } Q_1^* \\
& -y^a W_y|_{y=0} = 0 \quad \text{on } Q_1.
\end{align*}
\]

Gradient Continuity etc. 1657
Then we have the following estimates:

(i) For \( k \in \mathbb{N} \cup \{0\} \), multi-index \( \alpha \) and \( Q_r(t_0, x_0) \subset Q_1 \), we have

\[
\sup_{Q^*_{2r}(t_0, x_0, 0)} |\partial^k_{D} D_2^\alpha W| \leq \frac{C(n, s)}{r^{k+|\alpha|}} \text{osc}_{Q_r(t_0, x_0, 0)} W.
\]

(ii) For \( Q_r(t_0, x_0) \subset Q_1 \),

\[
\max_{Q^*_{2r}(t_0, x_0, 0)} |W| \leq C(r, n, s) \|W\|_{L^2} a(Q^*_{1r}(t_0, x_0, 0)).
\]

(iii) For all \( 0 \leq y < \frac{1}{2} \),

\[
\sup_{(t, x) \in Q^*_{12}} |W_y(t, x, y)| \leq C(n, s) \|W\|_{L^2} a(Q^*_{1}(t_0, x_0)).
\]

**Theorem 2.6.** For every \( f \in L^2(Q_1) \) and \( F \in L^2(Q^*_1) \), there exists a \( u \in C(-1; 1; V(B^*_1)) \) unique weak solution to

\[
\begin{align*}
\partial^a D u - \text{div}(y^a B(x) \nabla u) &= \text{div}(y^a F) & \text{in } Q_1^* \\
y^a u_y |_{y=0} &= f & \text{on } Q_1 \\
u = 0 & \text{on } \partial Q^*_1 \setminus Q_1,
\end{align*}
\]

where \( V(B^*_1) = \{v \in H^1_a(B^*_1) : v = 0 \text{ on } \partial B^*_1 \setminus \{y = 0\}\} \).

We now recall the definition of modulus of continuity and provide a brief collection of basic results concerning to the modulus of continuity functions.

**Definition 2.7** Modulus of continuity. A function

\[ \Psi : [0, 1] \longrightarrow [0, \infty] \]

is said to be a modulus of continuity if the following conditions are satisfied:

(i) \( \Psi(s) \rightarrow 0 \) as \( s \searrow 0 \).

(ii) \( \Psi(s) \) is increasing as a function of \( s \).

(iii) \( \Psi(s_1 + s_2) \leq \Psi(s_1) + \Psi(s_2), \forall s_1, s_2 \in [0, 1] \).

(iv) \( \Psi(s) \) is continuous.

We now define the notion of Dini-continuity:

**Definition 2.8** Dini-continuity. Let \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) be a function. Define the following modulus of continuity:

\[ \omega_f(s) := \sup_{|x-y| \leq s} |f(x) - f(y)|. \]

We then say \( f \) is Dini-continuous if

\[ \int_0^1 \frac{\omega_f(s)}{s} \, ds < \infty. \]
We now recall the following result proved in [31, theorem 8]:

**Theorem 2.9.** For each modulus of continuity \( \psi(s) \) defined on \([0, 1]\), there is a concave modulus of continuity \( \tilde{\psi}(s) \) satisfying

\[
\psi(s) \leq \tilde{\psi}(s) \leq 2\psi(s) \quad \text{for all } s \in [0, 1].
\]

We will also need the following definition which captures a certain monotonicity property of the modulus of continuity.

**Definition 2.10.** Given \( \eta \in (0, 1] \), we say that \( \Psi : [0, 1) \rightarrow [0, \infty) \) is \( \eta \)-decreasing if the following condition holds:

\[
\frac{\Psi(s_1)}{s_1^\eta} \geq \frac{\Psi(s_2)}{s_2^\eta}, \quad \text{for all } s_1 \leq s_2.
\]

**Remark 2.11.** From [31, page 44], we have that any continuous, increasing function \( \Psi : [0, 1) \rightarrow [0, \infty) \) with \( \Psi(0) = 0 \) is a modulus of continuity if it is concave. More generally, it suffices to assume that \( \frac{\Psi(s)}{s} \) is decreasing instead of concavity for \( \Psi \).

Finally, we recall the Campanato type characterization.

**Theorem 2.12.** Suppose \( u \in L^2(Q_2) \) satisfies that for every \( (t_0, x_0) \in Q_1 \) there exists an affine function \( \ell_{(t_0,x_0)}(x) = a + b \cdot x \) such that

\[
\int_{Q_r(t_0,x_0)} |u(t, x) - \ell_{(t_0,x_0)}(x)|^2 \leq r^2 K^2(r),
\]

where \( K \) is a modulus of continuity. Then \( \nabla u \) exists with \( \nabla u(t_0, x_0) = \nabla \ell_{(t_0,x_0)} \) and there exists \( C = C(n) \) such that

\[
|\nabla u(t_1, x_1) - \nabla u(t_2, x_2)| \leq CK(2|t_1, x_1) - (t_2, x_2))
\]

and

\[
|u(t_1, x) - u(t_2, x)| \leq C\sqrt{|t_1 - t_2|}K(2\sqrt{|t_1 - t_2|}).
\]

We give a self-contained proof of theorem 2.12 in the appendix A. In the case when \( K(r) = r^\alpha \), we refer the reader to [12].

**3. Proof of main theorem**

Our proof is based on an appropriate adaptation of compactness arguments, which has its roots in a fundamental work of Caffarelli in [13]. This section is organized as follows. In § 3.1, we prove some real analysis estimates which are fractional analogues of the estimates in [18]. In § 3.2, we establish regularity at the boundary, i.e. at \( \{y = 0\} \). In § 3.3, we prove the required scaled version of the interior estimates. With such ingredients in hand, we finally give the proof of theorem 1.1.
3.1. Some real analysis estimates

We first introduce some notations. We define

$$\mathcal{I}_2^\{f\}((x_0, t_0), r) := \int_0^r \rho^{2s-2} \left( \int_{Q_{r, \rho}((x_0, t_0)} |f(x, t)|^2 \, dx \, dt \right)^{\frac{1}{2}} \, d\rho.$$ 

Given a positive function $g$, we denote $g^*: [0, \infty) \to [0, \infty)$ by the non-increasing rearrangement of $g$, that is $g^*(s) := \sup\{r \geq 0 : \{(t, x) \in \mathbb{R}^{n+1} : |g(t, x)| > r\} \mid r > s\}$.

The following estimate is a fractional analogue of [18, equation (3.4)].

**Estimate 1:** Let $r > 0$. For given $0 < \sigma < 1$, define

$$r_i := \frac{\sigma^i r}{2}$$

for $i = 0, 1, 2, 3, \ldots$. We have

$$\sum_{i=0}^\infty r_i^{2s-1} \left( \int_{Q_{r_i, \rho}((x_0, t_0)} |f(x, t)|^2 \, dx \, dt \right)^{\frac{1}{2}} \leq c \int_0^r \rho^{2s-2} \left( \int_{Q_{r, \rho}((x_0, t_0)} |f(x, t)|^2 \, dx \, dt \right)^{\frac{1}{2}} \, d\rho \quad (3.1)$$

for some $c$ depends on $s, \sigma$ and $n$.

**Proof.** In this proof we will denote $Q_{r_i, \rho}(x_0, t_0)$ by $Q_i$. We will assume $|Q_1| = 1$.

$$\sum_{i=0}^\infty r_i^{2s-1} \left( \int_{Q_i} |f(x, t)|^2 \, dx \, dt \right)^{\frac{1}{2}} = r_0^{2s-1} \left( \int_{Q_0} |f(x, t)|^2 \, dx \, dt \right)^{\frac{1}{2}} + \sum_{i=1}^\infty r_i^{2s-1} \left( \int_{Q_i} |f(x, t)|^2 \, dx \, dt \right)^{\frac{1}{2}}$$

$$= \frac{2s - 1}{2^{2s-1} - 1} \int_{r/2}^r \rho^{2s-2} \, d\rho \left( \int_{Q_0} |f(x, t)|^2 \, dx \, dt \right)^{\frac{1}{2}}$$

$$+ \sum_{i=1}^\infty \frac{\sigma^{2s-1}(2s - 1)}{1 - \sigma^{2s}} \int_{Q_i} |f(x, t)|^2 \, dx \, dt \right)^{\frac{1}{2}}$$

$$= \frac{2s - 1}{2^{2s-1} - 1} \int_{r/2}^r \rho^{2s-2} \, d\rho \left( \int_{Q_0} |f(x, t)|^2 \, dx \, dt \right)^{\frac{1}{2}}$$

$$+ \sum_{i=1}^\infty \frac{\sigma^{2s-1}(2s - 1)}{1 - \sigma^{2s}} \int_{Q_i} |f(x, t)|^2 \, dx \, dt \right)^{\frac{1}{2}}$$

$$\leq \frac{(2s - 1)2^{(n+2)/2}}{(2^{2s-1} - 1)} \int_{r/2}^r \rho^{2s-2} \, d\rho \left( \int_{Q_0} |f(x, t)|^2 \, dx \, dt \right)^{\frac{1}{2}}$$
From Hardy–Littlewood inequality

\[ \text{Proof.} \]

Then, given any \( \epsilon > 0 \), there exists a \( \delta = \delta(\epsilon, n, s, \text{ellipticity}) > 0 \) such that if

\[ \int_{Q_1} f^2 \, dx \leq \delta^2, \quad \int_{Q_1} y^a |F|^2 \, dx \leq \delta^2, \quad \text{and} \quad \omega_A(1) \leq \delta^2, \]

then there exists \( V \) which solves weakly

\[
\begin{cases}
\text{div} (y^a \nabla V) = y^a \partial_t V & \text{in } (-3/4, 3/4) \times B_{1/2}^1 \\
\lim_{y \to 0} y^a V_y = 0 & \text{on } (-3/4, 3/4) \times B_{1/2} \times \{0\}
\end{cases}
\]
such that

\[ \int_{Q_{1/2}} y^a |U - V|^2 \, dt \, dX < \epsilon^2 \quad \text{and} \quad \int_{Q_{1/2}} |U - V|^2 \, dt \, dx < \epsilon^2. \]

Proof. Consider the following problem:

\[
\begin{aligned}
\langle \text{div}(y^a \nabla V) = y^a \partial_t V \quad &\text{in} \ (-1/2, 3/4) \times B^*_1/2 \\
V = U \quad &\text{on} \ \partial_p((-1/2, 3/4) \times B^*_1/2) \setminus \{(-1/2, 3/4) \times B_{1/2} \times \{0\}\} \\
\lim_{y \to 0} y^a V_y = 0 \quad &\text{on} \ (-1/2, 3/4) \times B_{1/2} \times \{0\}. 
\end{aligned}
\]

(3.6)

Note that from theorem 2.6, we have the unique solution, which we will denote by \( V \), of (3.6). Now using the regularity of \( V \) (lemma 2.5) and the properties of Steklov average, we have, the Steklov average of \( V \), i.e., \( V_h(t, x, y) = \frac{1}{t} \int_0^t V(\tau, x, y) \, d\tau \), \( h < 1/4 \), solves

\[
\begin{aligned}
\langle \text{div}(y^a \nabla V_h) = y^a \partial_t V_h \quad &\text{in} \ Q^*_1/2 \\
V_h = U_h \quad &\text{on} \ \partial_p Q^*_1/2 \setminus \{Q_{1/2} \cup \{t = -1/2\}\} \\
\lim_{y \to 0} y^a (V_h)_y = 0 \quad &\text{on} \ Q_{1/2},
\end{aligned}
\]

(3.7)

where \( U_h(t, x, y) = \frac{1}{t} \int_0^t U(\tau, x, y) \, d\tau \), \( h < 1/4 \), is the Steklov average of \( U \). Now, we multiply (3.7) by \( (U_h - V_h) \) and integrate by parts to get

\[
- \int_{Q^*_1/2} y^a \nabla V_h \nabla (U_h - V_h) \, dt \, dX - \int_{Q^*_1/2} y^a \partial_t V_h (U_h - V_h) \, dt \, dX = 0. \tag{3.8}
\]

By standard arguments, we know \( U \) can be replaced by \( U_h \) in (2.6). Also \( V_h \) solves (3.7) so we can put \( \phi = (U_h - V_h)\eta(t) \in H^1([-1,1], L^2_a(B_1^*) \cap L^2([-1,1]; H^2_a(B_1^*)) \text{ in (2.6), where } \eta(t) \text{ is chosen such that } \eta(t) = 1 \text{ in } (-1/2, 1/2) \text{ and compactly supported in } (-5/8, 5/8). \text{ Hence, after integrating by parts in (2.6) with respect to } t \text{ we get}

\[
\int_{Q^*_1/2} y^a \partial_t U_h (U_h - V_h) \, dt \, dX + \int_{Q^*_1/2} y^a B(x) \nabla U_h \nabla (U_h - V_h) \, dt \, dX
\]

\[
= \int_{Q_{1/2}} f_h (U_h - V_h) \, dt \, dx + \int_{Q^*_1/2} y^a F_h \nabla (U_h - V_h) \, dt \, dX, \tag{3.9}
\]

where \( f_h \) and \( F_h \) denote the Steklov average of \( f \) and \( F \) respectively. On adding (3.8) and (3.9), we get

\[
\int_{Q^*_1/2} y^a \partial_t (U_h - V_h) (U_h - V_h) \, dX \, dt
\]

\[
+ \int_{Q^*_1/2} y^a (B(x) - I) \nabla U_h \nabla (U_h - V_h) \, dX \, dt
\]
\[ + \int_{Q_{1/2}^*} y^a |\nabla (U_h - V_h)|^2 \, dX \, dt \]
\[ = \int_{Q_{1/2}^*} f_h (U_h - V_h) \, dx \, dt + \int_{Q_{1/2}^*} y^a F_h \nabla (U_h - V_h) \, dX \, dt. \tag{3.10} \]

Using Fundamental theorem of calculus in \( t \) -variable and the Young’s inequality with \( \varepsilon \) we get,

\[ I_h(0) + \int_{Q_{1/2}^*} y^a |\nabla (U_h - V_h)|^2 \, dX \, dt \]
\[ \leq \frac{1}{2} \int_{Q_{1/2}^*} y^a |\nabla (U_h - V_h)|^2 \, dX \, dt \]
\[ + \frac{1}{2} \int_{Q_{1/2}^*} y^a |B(x) - I|^2 |\nabla U_h|^2 \, dX \, dt \]
\[ + \frac{1}{4\varepsilon} \int_{Q_{1/2}^*} f_h^2 \, dx \, dt + \varepsilon \int_{Q_{1/2}^*} (U_h - V_h)^2 \, dx \, dt \]
\[ + \int_{Q_{1/2}^*} y^a F_h^2 \, dX \, dt + \frac{1}{4} \int_{Q_{1/2}^*} y^a |\nabla (U_h - V_h)|^2 \, dX \, dt, \tag{3.11} \]

where \( I_h(0) = \int_{B_1^*} y^a (U_h - V_h)(0) \, dX \). We rewrite above equation as

\[ I_h(0) + \frac{1}{4} \int_{Q_{1/2}^*} y^a |\nabla (U_h - V_h)|^2 \, dX \, dt \]
\[ \leq \frac{1}{2} \int_{Q_{1/2}^*} y^a |B(x) - I|^2 |\nabla U_h|^2 \, dX \, dt + \varepsilon \int_{Q_{1/2}^*} (U_h - V_h)^2 \, dx \, dt \]
\[ + \frac{1}{4\varepsilon} \int_{Q_{1/2}^*} f_h^2 \, dx \, dt + \int_{Q_{1/2}^*} y^a F_h^2 \, dX \, dt. \tag{3.12} \]

Note that we can make third and fourth term small by taking \( \delta \) small enough. For the first term, take \( \varphi \) compactly supported in \( Q_{3/4}^* \) such that \( \varphi = 1 \) in \( Q_{1/2}^* \) and use lemma 2.4 to get

\[ \frac{1}{2} \int_{Q_{1/2}^*} y^a |B(x) - I|^2 |\nabla U_h|^2 \, dX \, dt \]
\[ \leq \frac{\omega_A(1)}{2} \int_{Q_{1/2}^*} y^a |\nabla U_h|^2 \, dX \, dt \]
\[ \leq C \frac{\omega_A(1)}{2} \left[ \int_{B_1^*} y^a (U_h^2 + F_h^2) \, dx \, dt + \int_{B_1} |U_h(t, x, 0)| |f_h(t, x)| \, dx \, dt \right] \]
\[ \leq C \frac{\omega_A(1)}{2} \left[ \int_{-1/2}^{1/2} y^a (U_h^2 + F_h^2) \, dx \, dt + \frac{1}{2} \int_{-1/2}^{1/2} |U_h(t, x, 0)|^2 \, dx \, dt \right] \]
\[\frac{1}{2} \int_{-1/2}^{1/2} \int_{B_1} |f_h(t,x)|^2 \, dx \, dt \leq C \frac{\omega_A(1)}{2} \left[ \int_{Q_1^*} y^a (U^2 + F^2) \, dX \, dt + \frac{1}{2} \int_{Q_1} |U(t,x,0)|^2 \, dx \, dt \right.
\]
\[\left. + \frac{1}{2} \int_{Q_1} |f(t,x)|^2 \, dx \, dt \right] \leq C \omega_A(1) (1 + \delta^2),\]

where \( C = C(n, s, \text{ellipticity}) > 0 \). Also by trace theorem [34], we have
\[\int_{Q_{1/2}} (U_h - V_h)^2 \, dx \, dt \leq C_T \int_{Q_{1/2}^*} y^a (U_h - V_h)^2 \, dX \, dt \]
\[+ C_T \int_{Q_{1/2}^*} y^a |\nabla (U_h - V_h)|^2 \, dX \, dt \tag{3.13}\]
where \( C_T(n, s) \) is constant from trace theorem. Hence, (3.12) becomes
\[I_h(0) + \frac{1}{4} \int_{Q_{1/2}^*} y^a |\nabla (U_h - V_h)|^2 \, dX \, dt \]
\[\leq C \omega_A(1) (1 + \delta^2) + \varepsilon C_T \int_{Q_{1/2}^*} y^a (U_h - V_h)^2 \, dX \, dt \]
\[+ \varepsilon C_T \int_{Q_{1/2}^*} y^a |\nabla (U_h - V_h)|^2 \, dX \, dt + \frac{\delta^2}{4\varepsilon} + \delta^2. \tag{3.14}\]

Also, for each time label we will apply the Poincaré inequality [22] to get
\[I_h(0) + ((1/4 - \varepsilon C_T) C_P - \varepsilon C_T) \int_{Q_{1/2}^*} y^a (U_h - V_h)^2 \, dX \, dt \]
\[\leq C \omega_A(1) (1 + \delta^2) + \frac{\delta^2}{4\varepsilon} + \delta^2\]
where \( C_P(n, s) \) is constant in the Poincaré inequality.

Now, first choose \( \varepsilon > 0 \) small enough such that \((1/4 - \varepsilon C_T) C_P - \varepsilon C_T > 0\) then choose \( \delta \) to get
\[I_h(0) + \int_{Q_{1/2}^*} y^a (U_h - V_h)^2 \, dX \, dt \leq \epsilon^2.\]

Also, from (3.13) and (3.14) we have
\[\int_{Q_{1/2}} (U_h - V_h)^2 \, dx \, dt \leq M \epsilon^2 + M I_h(0)\]
where \( M \) is a universal constant. Since \( U \in C([-1, 1], L^2_a(B_1^*)) \cap L^2([-1, 1], H^1_a(B_1^*)) \), \( U_h - V_h \to U - V \), \( \nabla (U_h - V_h) \to \nabla (U - V) \) in \( L^2_a(Q_{1/2}^*) \) and using (3.6), we
have $I_h(0) \to 0$. Hence by trace theorem $(U_h - V_h)(t, x, 0) \to (U - V)(t, x, 0)$ in $L^2(Q_{1/2})$. This completes the proof of the lemma.

We now prove the closeness of solution $U$ to (2.5) by an affine function at some fixed scale.

**Lemma 3.2.** There exist $0 < \delta, \lambda < 1$ (depending on $n$, $s$ and ellipticity), a linear function $\ell(x) = A + B \cdot x$ and constant $C = C(n, s) > 0$ such that for any solution $U$ of (2.5) which satisfies (3.4) and (3.5),

$$\frac{1}{\lambda^{n+2}} \int_{Q_\lambda} |U(t, x, 0) - \ell(x)|^2 \, dt \, dx + \frac{1}{\lambda^{n+3+a}} \int_{Q_\lambda} y^a |U - \ell(x)|^2 \, dt \, dX < \lambda^3$$

and $|A| + |B| \leq C$.

**Proof.** Let $0 < \epsilon < 1$ be any real number. From lemma 3.1, there exists a $\delta(\epsilon) > 0$ and a solution $V$ to (3.6) such that if (3.5) holds, then

$$\int_{Q_{1/2}} y^a |V - V|^2 \, dt \, dX < \epsilon^2 \quad \text{and} \quad \int_{Q_{1/2}} |U - V|^2 \, dt \, dx < \epsilon^2 \quad (3.15)$$

where $V$ solves (3.6). Now by lemma 2.5, $V$ is smooth, we define

$$\ell(x) = V(0, 0, 0) + \nabla_x V(0, 0, 0) \cdot x = A + B \cdot x.$$ 

Also by lemma 2.5, there exists a constant $\tilde{C} = \tilde{C}(n, s)$ such that

$$|V(0, 0, 0)| + |
abla_x V(0, 0, 0)| \leq \tilde{C}||V||_{L^2(Q_{1/2} \times [0, 1/2])}.$$

Now, using triangle inequality, $(a + b)^2 \leq 2a^2 + 2b^2$, (3.15) and (3.4), we get

$$\int_{Q_{1/2}} y^a |V|^2 \, dt \, dX \leq 2 \int_{Q_{1/2}} y^a |U - V|^2 \, dt \, dX$$

$$+ 2 \int_{Q_{1/2}} y^a |U|^2 \, dt \, dX \leq 2 \epsilon^2 + 2 \leq 4.$$ 

Hence we get

$$|V(0, 0, 0)| + |
abla_x V(0, 0, 0)| \leq C,$$

where $C = 4\tilde{C}$.

Using triangle inequality, we get

$$\frac{1}{\lambda^{n+2}} \int_{Q_\lambda} |U(t, x, 0) - \ell(x)|^2 \, dt \, dx + \frac{1}{\lambda^{n+3+a}} \int_{Q_\lambda} y^a |U - \ell(x)|^2 \, dt \, dX$$

$$\leq \frac{2}{\lambda^{n+3+a}} \int_{Q_\lambda} y^a |U - V|^2 \, dt \, dX + \frac{2}{\lambda^{n+3+a}} \int_{Q_\lambda} y^a |V - \ell(x)|^2 \, dt \, dX$$

$$+ \frac{2}{\lambda^{n+2}} \int_{Q_\lambda} |U - V|^2 \, dt \, dx + \frac{2}{\lambda^{n+2}} \int_{Q_\lambda} |V - \ell(x)|^2 \, dt \, dx.$$
For any \((t, x, y) \in Q^{1/4}_1\), using Mean value theorem, Taylor’s theorem and lemma 2.5, we have a universal constant \(D\) such that
\[
|V(t, x, y) - \ell(x)| \leq |V(t, x, y) - V(t, x, 0)| + |V(t, x, 0) - V(0, x, 0)|
+ |V(0, x, 0) - V(0, 0, 0) - \nabla_x V(0, 0, 0) \cdot x|
\leq |V_y(t, x, \xi_1)|y + |V_t(\xi_2, x, 0)|t + D|x|^2
\text{(for some } \xi_1 \in (0, y) \text{ and } \xi_2 \in (0, t))
\leq D\xi_1 y + Dt + D|x|^2 \leq D(|X|^2 + t).
\]
Hence, using (3.15) we get
\[
\frac{1}{\lambda^{n+2}} \int_{Q^\lambda} |U(t, x, 0) - \ell(x)|^2 dt \, dx + \frac{1}{\lambda^{n+3+a}} \int_{Q^\lambda} y^a |U - \ell(x)|^2 dt \, dx
\leq \frac{2\epsilon^2}{\lambda^{n+3+a}} + \frac{8D^2}{\lambda^{n+3+a}} \int_{Q^\lambda} y^a |X|^4
\text{+ } |t|^2 \, dt \, dx + \frac{2\epsilon^2}{\lambda^{n+2}} + \frac{4D^2}{\lambda^{n+2}} \int_{Q^\lambda} (|X|^4 + |t|^2) \, dt \, dx
\leq \frac{2\epsilon^2}{\lambda^{n+3+a}} + \frac{8D^2}{1 + a} \lambda^4 + \frac{2\epsilon^2}{\lambda^{n+2}} + 8D^2 \lambda^4.
\]
Now first choose \(0 < \lambda < 1/4\) small enough such that
\[
\frac{8D^2}{1 + a} \lambda^4 + 8D^2 \lambda^4 \leq \lambda^3 / 2,
\]
then take \(\epsilon\) such that
\[
\epsilon^2 = \frac{\lambda^{n+6+a}}{8},
\]
which in turn fixes \(\delta\). \qed

Now we will define some functions as in [1]:

- Define
  \[
  \tilde{\omega}_1(r) := \max\{\omega_A(\gamma r) / \tilde{\delta}, r\},
  \]
  where \(\gamma\) and \(\tilde{\delta}\) will be fixed later. By theorem 2.9, we can assume that \(\tilde{\omega}_1\) is concave. Without loss of generality we can assume \(\tilde{\omega}_1(1) = 1\). Finally we define
  \[
  \omega_1(r) := \tilde{\omega}_1(\sqrt{r}).
  \]
  Then, \(\omega_1(r)\) is \(1/2\)-decreasing function in the sense of definition 2.10. Also, by application of change of variables we have \(\omega_1(r)\) is Dini continuous.

- We define
  \[
  \omega_2(r) := \max\left\{\gamma I(\gamma r) / \tilde{\delta}, r\right\},
  \]
  where \(I(r) := r^{2s-1} \left(\int_{Q^r} |f(x, t)|^2 dt \, dx\right)^{1/2}\).
• We define
\[ \omega_3(\lambda^k) := \sum_{i=0}^{k} \omega_1(\lambda^{k-i})\omega_2(\lambda^i), \]
where \( \lambda \) is from lemma 3.2.

• Finally, we define
\[ \omega(\lambda^k) = \max\{\omega_3(\lambda^k), \lambda^{k/2}\}. \]

**Lemma 3.3.** There exists a universal constant \( C_{sum} \) such that
\[ \sum_{i=0}^{\infty} \omega(\lambda^i) \leq C_{sum}. \]

**Proof.** From the definition of \( \omega \), we have
\[
\sum_{i=0}^{\infty} \omega(\lambda^i) \leq \sum_{i=0}^{\infty} \omega_3(\lambda^i) + \frac{1}{1 - \sqrt{\lambda}}.
\]

An application of Fubini’s theorem and reindexing gives \( \sum_{i=0}^{\infty} \omega_3(\lambda^i) \leq \sum_{i=0}^{\infty} \omega_1(\lambda^i) \sum_{i=0}^{\infty} \omega_2(\lambda^i) \). Consequently, we have
\[
\sum_{i=0}^{\infty} \omega(\lambda^i) \leq \sum_{i=0}^{\infty} \omega_1(\lambda^i) \sum_{i=0}^{\infty} \omega_2(\lambda^i) + \frac{1}{1 - \sqrt{\lambda}}.
\]

Using the Dini continuity, increasing nature of \( \omega_A \) and change of variables, we have
\[
\sum_{i=0}^{\infty} \omega_1(\lambda^i) \leq \sum_{i=0}^{\infty} \frac{\omega_A(\gamma \lambda^{i/2})}{\delta} + \frac{1}{1 - \sqrt{\lambda}}
\]
\[
\leq \frac{\omega_A(\gamma)}{\delta} + \frac{1}{(\log \sqrt{\lambda})\delta} \int_0^{\gamma} \frac{\omega_A(s)}{s} ds + \frac{1}{1 - \sqrt{\lambda}}.
\]

Choose \( \gamma < \tilde{\delta} \) such that
\[
\sum_{i=0}^{\infty} \omega_1(\lambda^i) \leq 1 + \frac{1}{1 - \sqrt{\lambda}}.
\]

Note that from (3.1), we have
\[
\sum_{i=0}^{\infty} \omega_2(\lambda^i) \leq \frac{C_\gamma}{\delta} I_2((0,0),\gamma) + \frac{1}{1 - \lambda}.
\]

Using (3.2) we have
\[
\sum_{i=0}^{\infty} \omega_2(\lambda^i) \leq \frac{C_\gamma}{\delta(n + 2)} C_{n+1}^{(2s-1)/(n+2)} \int_0^{C_{n+1}^{\gamma n+2}} \frac{\rho^{2s-1}}{\rho^{n+2}} (g^{**}(\rho)) \frac{1}{\rho} d\rho + \frac{1}{1 - \lambda},
\]
where $g = f^2$. We have given that $f \in L\left(\frac{n+2}{2s-1}, 1\right)$, which gives us $g \in L\left(\frac{n+2}{2(2s-1)}, \frac{1}{2}\right)$. Since $2(2s - 1) < n + 2$, therefore from [36, equation (6.8)], we get
\[
\int_0^\infty \rho^{\frac{2s-1}{n+2}} (g^{**}(\rho))^{\frac{1}{2}} \frac{d\rho}{\rho} < \infty.
\]
Hence we get a universal bound on $\sum_{i=0}^\infty \omega_2(\lambda^i)$ as $\lambda$ is a universal constant. This completes the proof of the lemma. \(\square\)

We now prove the closeness of solution $U$ to (2.5) by affine function at each dyadic scale and closeness of affine functions as well. Given the solution $U(t, X)$ of (2.5), we define
\[
U_\gamma(t, X) = \left(\int_{Q_1} U(t, x, 0)^2 \, dt \, dx + \int_{Q_1^*} U(t, X)^2 \, dt \, dX + 1\right)^{-1} U(\gamma^2 t, \gamma X).
\]
Then $U_\gamma$ solves (2.5) with modulus of continuity for $A$ is $\omega_A(\gamma r)$,
\[
\int_{Q_1} U_\gamma(t, x, 0)^2 \, dt \, dx + \int_{Q_1^*} U_\gamma(t, X)^2 \, dt \, dX \leq 1
\]
and $r^{2s-1}(\int f_{Q_r} |f(x, t)|^2 \, dt \, dx)^{1/2} \leq \gamma I(\gamma r)$.

**Lemma 3.4.** There exist a sequence of linear functions $\ell_k(x) = a_k + b_k \cdot x$ and a constant $C = C(n, s) > 0$ such that for any solution $U$ of (2.5) satisfies
\[
\frac{1}{\lambda^{k(n+2)}} \int_{Q_{\lambda^k}} |U_\gamma(t, x, 0) - \ell_k(x)|^2 \, dt \, dx
\]
\[
+ \frac{1}{\lambda^{k(n+3+a)}} \int_{Q_{\lambda^k}} y^a |U_\gamma - \ell_k(x)|^2 \, dt \, dX < \lambda^{2k} \omega^2(\lambda^k), \tag{3.16}
\]
\[
|a_{k+1} - a_k| \leq C \lambda^k \omega(\lambda^k) \quad \text{and} \quad |b_{k+1} - b_k| \leq C \omega(\lambda^k). \tag{3.17}
\]

**Proof.** For notational ease we shall denote $U_\gamma(t, X)$ by $U(t, X)$, $\omega_A(\gamma r)$ by $\omega_A(r)$ and $\gamma I(\gamma r)$ by $I(r)$. We will prove the lemma by induction on $k$. For $k = 0$, take $\ell_k(x) = 0$ and we get
\[
\int_{Q_1} |U(t, x, 0)|^2 \, dt \, dx + \int_{Q_1^*} y^a |U|^2 \, dt \, dX \leq 1 \leq \omega^2(1).
\]
Thus we are done for $k = 0$. Let us assume result is true for $k = 0, 1, \ldots, i$. We will prove it for $k = i + 1$. Define
\[
\tilde{U}(t, X) := \frac{U(\lambda^2 t, \lambda^i X) - \ell_i(\lambda^i x)}{\lambda^i \omega(\lambda^i)},
\]
then $\tilde{U}$ is a weak solution to
\[
\begin{cases}
y^a \partial_t \tilde{U} - \div(y^a B(\lambda^i X) \nabla \tilde{U}) = -\div(y^a F) \quad \text{in} \ Q_1^* \\
y^a \tilde{U}_y |_{y=0} = \tilde{f} \quad \text{on} \ Q_1,
\end{cases}
\]
where  \( \tilde{f}(t, x) = \frac{\lambda^{-ia}}{\omega^2(\lambda^i)} f(\lambda^i t, \lambda^i x) \) and the vector field \( F \) is given by \( \lambda^i \omega(\lambda^i) F = ((I - A(\lambda^i x)) \nabla_x \ell(\lambda^i x), 0) \). Now we estimate the following:

**Bound for \( f \):**

\[
\int_{Q_1} \tilde{f}^2 \, dt \, dx = \frac{\lambda^{-2ia}}{\omega^2(\lambda^i)} \int_{Q_1} f^2(\lambda^i t, \lambda^i x) \, dt \, dx = \frac{\lambda^{-2ia}}{\omega^2(\lambda^i)} \int_{Q_{\lambda^i}} f^2(t, x) \, dt \, dx
\]

\[
\leq \frac{\omega^2(\lambda^i)}{\omega^2(\lambda^i)} \leq \tilde{\delta}^2 \frac{\omega^2(\lambda^i)}{\omega^2(\lambda^i)} = \tilde{\delta}^2.
\]

**Bound for \( F \):** Note that by induction hypothesis

\[
|b_i| \leq \sum_{j=1}^{i} |b_j - b_{j-1}| \leq \sum_{j=0}^{i-1} C\omega(\lambda^i) \leq CC_{\text{sum}} =: C_1.
\]

Now

\[
\int_{Q_1} y^a |F|^2 \, dt \, dX = \lambda^{-2i} \omega^{-2}(\lambda^i) \int_{Q_1} y^a |(I - A(\lambda^i x)) \nabla_x \ell(\lambda^i x)|^2 \, dt \, dX.
\]

We now integrate in \( y \)-variable and then use change of variable to obtain

\[
\int_{Q_1} y^a |F|^2 \, dt \, dX = 2\omega^{-2}(\lambda^i) \int_{B_{\lambda^i}} |(I - A(x)) b_i|^2 \, dt \, dx
\]

\[
\leq 2\omega^{-2}(\lambda^i) \omega^2_A(\lambda^i) C_1^2
\]

\[
\leq 2C_1^2 \frac{\omega^2_A(\lambda^i)}{\omega^2(\lambda^i)} \leq 2C_1^2 \delta^2 \omega^2_A(\lambda^i) = 2C_1^2 \delta^2.
\]

Also, note that

\[
|B(\lambda^i x) - I|^2 = |A(\lambda^i x) - I|^2 \leq \omega^2_A(\lambda^i) \leq \omega^2_A(1) \leq \tilde{\delta}^2.
\]

Choose \( \tilde{\delta} < \delta \) such that \( 2C_1^2 \delta^2 < \delta^2 \). We rewrite \((3.16)\) for \( k = i \) as

\[
\frac{1}{\lambda^i(n+2)} \int_{Q_{\lambda^i}} |U(t, x, 0) - \ell_i(x)|^2 \, dt \, dx + \frac{1}{\lambda^i(n+3+a)} \int_{Q_{\lambda^i}} y^a |U(t, x, y) - \ell_i(x)|^2 \, dt \, dX < 1.
\]

Now after changing the variables, we get

\[
\int_{Q_1} |\tilde{U}(t, x, 0)|^2 \, dt \, dx + \int_{Q_1} y^a |\tilde{U}(t, x, y)|^2 \, dt \, dX \leq 1.
\]

By lemma 3.2 there exists a \( l(x) = a + b \cdot x \) such that

\[
\frac{1}{\lambda^{n+2}} \int_{Q_\lambda} |\tilde{U}(t, x, 0) - l(x)|^2 \, dt \, dx + \frac{1}{\lambda^{n+3+a}} \int_{Q_\lambda} y^a |\tilde{U} - l(x)|^2 \, dt \, dX < \lambda^3
\]
and \(|a| + |b| \leq C\). After putting the value of \(\tilde{U}\) and changing the variables, we get
\[
\frac{1}{\lambda^{(i+1)(n+2)}} \int_{Q_{x,1}} |U(t, x, 0) - \ell_{i+1}(x)|^2 \, dt \, dx
\]
\[
+ \frac{1}{\lambda^{(i+1)(n+3+a)}} \int_{Q_{x,1}} y^a |U - \ell_{i+1}(x)|^2 \, dt \, dX < \lambda^{2(i+1)} \omega^2(\lambda^{i+1})
\]
where \(\ell_{i+1}(x) = \ell_i(x) + \lambda_i \omega(\lambda^i) f(\lambda^{-i} x)\). By putting \(x = 0\) we get \(|a_{i+1} - a_i| \leq C \lambda_i \omega(\lambda^i)\). Take gradient to get \(|b_{i+1} - b_i| \leq C \omega(\lambda^i)\).

With lemmas 3.3 and 3.4 in hand, we get the approximation of \(U\) by an affine function.

**Lemma 3.5.** There exists a linear function \(\ell_{t_0, x_0}(x) = a_\infty + b_\infty \cdot x\) such that for any solution \(U\) of (2.5) with (3.4) satisfies
\[
\frac{1}{r^{n+3+a}} \int_{Q_{x}^r(t_0, x_0, 0)} |U(t, x, y) - \ell_{(t_0, x_0)}(x)\|^2 y^a \, dt \, dX \leq C_1 r^2 K^2(r)
\]
for all \(0 < r < 1/2\), where \(C_1\) is a universal constant and \(K\) is 1/2–decreasing concave modulus of continuity.

**Proof.** Without loss of generality we will assume that \((t_0, x_0) = (0, 0)\). First we will estimate for \(U_\gamma\) with small \(r\) then will do it for \(U\). Note that \(a_k\) and \(b_k\) are convergent as
\[
\sum_{i=0}^{\infty} \omega(\lambda^i) < \infty.
\]
We define \(a_\infty\) and \(b_\infty\) as limit of \(a_k\) and \(b_k\) respectively. We will denote \(\ell_{(0,0)}\) by \(\ell_0\).

Take \(r\) such that \(\lambda^{k+1} \leq r \leq \lambda^k\) for some \(k \in \mathbb{N}\). Now using triangle inequality, (3.16) and (3.17) we have
\[
\int_{Q_r} |U(t, x, 0) - \ell_0(x)|^2 \, dt \, dx
\]
\[
\leq 2 \int_{Q_r} |U(t, x, 0) - \ell_k(x)|^2 \, dt \, dx + 2 \int_{Q_r} |\ell_k(x) - \ell_0(x)|^2 \, dt \, dx
\]
\[
\leq 2\lambda^{-n-2} \lambda^{2k} \omega^2(\lambda^k) + 2 \int_{Q_r} \sum_{i=k}^{\infty} |\ell_i(x) - \ell_{i+1}(x)|^2 \, dt \, dx
\]
\[
\leq 2\lambda^{-n-2} \lambda^{2k} \omega^2(\lambda^k) + 4 \sum_{i=k}^{\infty} |a_i - a_{i+1}|^2 + 4r^2 \sum_{i=k}^{\infty} |b_i - b_{i+1}|^2
\]
\[
\leq 2\lambda^{-n-2} \lambda^{2k} \omega^2(\lambda^k) + 4C \lambda^{2k} \sum_{i=k}^{\infty} \omega^2(\lambda^i) + 4C \lambda^{2k} \sum_{i=k}^{\infty} \omega^2(\lambda^i)
\]
\[
\leq (2\lambda^{-n-2} + 8C) \lambda^{2k} \left( \sum_{i=k}^{\infty} \omega(\lambda^i) \right)^2.
\]
Note that from lemma 3.2, $\lambda$ depends only on $n, s$, ellipticity therefore $(2\lambda^{-n-2} + 8C)$ is a universal constant. We now follow the same lines of proof as in [1] from (4.30) to (4.35) to get

$$\frac{1}{r^{n+3+a}} \int_{Q_r^\gamma} |U(t, x, y) - \ell_0(x)|^2 y^a \, dt \, dX \leq C_t r^2 K^2(r),$$

(3.18)

where $K = K_1(r) + K_2(r) + K_3(r)$ and

$$K_1(r) := \sup_{a \geq 0} \int_{a}^{a+\sqrt{7}} \frac{\omega_1(t)}{t} \, dt, \quad K_2(r) := \sqrt{r}$$

and

$$K_3(r) = \sup_{a \geq 0} \int_{a}^{a+C_{n+1}r} \frac{2^s-1}{\rho^{n+2}} (g^{**} (\rho))^\frac{1}{2} \frac{d\rho}{\rho}.$$\n
Now put the value of $U_\gamma$ and do change of variables to get

$$\int_{Q_{r^\gamma}} |U(t, x, y) - \ell_0(\gamma^{-1} x)|^2 y^a \, dt \, dX \leq \gamma^{-2} (1 + ||U(t, x, 0)||_L^2(Q_{r^\gamma})) + ||U(t, X)||_L^2(Q_{r^\gamma}) C_t(r\gamma)^2 K^2(\gamma r / \gamma).$$

(3.19)

Call $K(r/\gamma)$ by $K(r)$, $\ell_0(\gamma^{-1} x)$ by $\ell_0(x)$ and use the fact $\gamma$ is a universal constant to get

$$\int_{Q_{r^\gamma}} |U(t, x, y) - \ell_0(x)|^2 y^a \, dt \, dX \leq \tilde{C}_t r^2 K^2(r),$$

(3.20)

for $0 < r < \gamma \lambda$. For $\gamma \lambda < r < 1/2$, it will be done by replacing $\tilde{C}_t$ to $C \tilde{C}_t / (\gamma^{n+6} \lambda K(\gamma \lambda))$, call it again by $C_t$. \hfill \Box

**Lemma 3.6.** There exists a universal constant $C_{bdr}$ such that

$$|\nabla \ell_{(t_1, x_1)} - \nabla \ell_{(t_2, x_2)}| \leq C_{bdr} K(||(t_1, x_1) - (t_2, x_2)||),$$

and

$$|U(t_1, x, 0) - U(t_2, x, 0)| \leq C_{bdr} K(\sqrt{|t_1 - t_2|}) |t_1 - t_2|^{1/2}$$

for $(t_i, x_i) \in Q_{1/2}$, where $\ell_{(t_i, x_i)}$ denotes the linear function constructed in lemma 3.5.

**Proof.** After proceeding as in the proof of lemma 3.5 we get a linear function $\ell_{t_0, x_0}(x) = a_\infty + b_\infty \cdot x$ such that

$$\int_{Q_r(t_0, x_0, 0)} |U(t, x, 0) - \ell_{(t_0, x_0)}(x)|^2 \, dt \, dx \leq C_r^2 K^2(r).$$

Now apply theorem 2.12 (Companato type characterization) to get the result with modulus of continuity $K(2r)$, which we again call as $K(r)$. \hfill \Box
3.3. Interior estimate

To prove theorem 1.1, we have to combine the above boundary estimates with known interior estimates in [14]. In order to do this we need the following rescaled version of interior estimate.

**Theorem 3.7.** Let \( u \) be a weak solution of
\[
\frac{\partial u}{\partial t} - \text{div}(B(x)\nabla u) - b \cdot \nabla u = \text{div}(g) \quad \text{in} \ Q_1, \tag{3.21}
\]
with \( \|u\|_{L^2(Q_1)} \leq 1 \), where \( B, g \) are Dini continuous with modulus of continuity \( \Psi_B \) and \( \Psi_g \) respectively and \( b \) is smooth. Then there exists a constant \( C_{ir} > 0 \) depending on \( n, \text{ellipticity}, \|b\|_1 \ (C^1\text{-norm of} \ b), \|g\|_\infty, \) and \( \int_0^1 \frac{\Psi_B(t)}{t} \int_0^1 \frac{\Psi_g(t)}{t} \) such that for all \( (t_1, x_1), (t_2, x_2) \in Q_{1/2} \), we have
\[
|\nabla u(t_1, x_1) - \nabla u(t_2, x_2)| \leq C_{ir}(\Psi(\sqrt{|t_1 - t_2|}) + |t_1 - t_2|^{1/4} \sqrt{|t_1 - t_2|}),
\]
where
\[
\Psi(r) = \sup_{a \geq 0} \int_a^{a+\sqrt{r}} \frac{\Psi_B(s)}{s} \, ds + \sup_{a \geq 0} \int_a^{a+\sqrt{r}} \frac{\Psi_g(s)}{s} \, ds
\]
and
\[
|u(t_1, x) - u(t_2, x)| \leq C_{ir}(\Psi(\sqrt{|t_1 - t_2|}) + |t_1 - t_2|^{1/4} \sqrt{|t_1 - t_2|}).
\]
Furthermore for all \( (t, x) \in Q_{1/2}, \)
\[
|\nabla u(t, x)| \leq C_{ir}(1 + \Psi(1)). \tag{3.22}
\]

The proof of the theorem 3.7 will be done by compactness method. We will sketch it as most of the details are similar to boundary estimate. We need the following lemmas.

**Lemma 3.8.** Let \( u \) be a weak solution of (3.21) with \( \|u\|_{L^2(Q_1)} \leq 1 \). Then, given any \( \epsilon > 0 \), there exists a \( \delta = \delta(\epsilon, n, \text{ellipticity}, \|b\|_1, \|g\|_\infty) > 0 \) such that if
\[
\omega_B(1) \leq \delta^2 \quad \text{and} \quad \omega_g(1) \leq \delta^2, \tag{3.23}
\]
then there exists \( v \) which solves weakly
\[
\begin{cases}
\frac{\partial v}{\partial t} - \Delta v + b \cdot \nabla v = 0 & \text{in} \ Q_{1/2} \\
v = u & \text{on} \ \partial_p Q_{1/2}
\end{cases} \tag{3.24}
\]
such that
\[
\int_{Q_{1/2}} |u - v|^2 \, dt \, dx < \epsilon^2.
\]

**Proof.** We have existence of solution of (3.24) by remark A.1. We now proceed along the similar lines as in the proof of lemma 3.1. Also, for the required energy estimate, we refer to lemma 1.2.3 ([14]). \( \square \)
Lemma 3.9. There exist $0 < \delta_{\text{int}}, \lambda_{\text{int}} < 1$ (depending on $n$, ellipticity, $\|b\|_1$), a linear function $\ell(x) = A_{\text{int}} + B_{\text{int}} \cdot x$ and constant $C_{\text{int}} = C(n, \|b\|_1) > 0$ such that for any solution $u$ of (3.21) with $\|u\|_{L^2(Q_1)}$ and satisfies (3.23),

$$\frac{1}{\lambda^{n+2}} \int_{Q_{\lambda}} |u - \ell(x)|^2 \, dt \, dx < \lambda^3$$

and $|A_{\text{int}}| + |B_{\text{int}}| \leq C_{\text{int}}$.

Proof. Since $v$, solution of (3.24), is smooth therefore we can proceed like the proof of the lemma 3.2.

Lemma 3.10. There exist a sequence of linear functions $\ell^k_{\text{int}}(x) = a^k_{\text{int}} + b^k_{\text{int}} \cdot x$ and a constant $C_{\text{int}} = C(n, \|b\|_1) > 0$ such that for any weak solution $u$ of (3.21) with $\|u\|_{L^2(Q_1)}$,

$$\frac{1}{\lambda^{n+2}} \int_{Q_{\lambda}^k_{\text{int}}} |u - \ell^k_{\text{int}}(x)|^2 \, dt \, dx < \lambda^{2k} \psi^2(\lambda^k_{\text{int}})$$

(3.25)

and

$$|a^{k+1}_{\text{int}} - a^k_{\text{int}}| \leq C_{\text{int}} \lambda^k_{\text{int}} \psi(\lambda^k_{\text{int}}) \quad \text{and} \quad |b^{k+1}_{\text{int}} - b^k_{\text{int}}| \leq C_{\text{int}} \psi(\lambda^k_{\text{int}})$$

(3.26)

Proof. Note that $u_\gamma(t, x) = u(\gamma^2 t, \gamma x)$ will satisfy (3.21) with $\Psi_B(\gamma r)$ and $\gamma g(\gamma x)$.

Now we will define some functions:

Define

$$\hat{\psi}_1(r) := \max\{\Psi_B(\gamma_{\text{int}} r)/\delta_{\text{int}}, r\},$$

where $\gamma_{\text{int}}$ and $\delta_{\text{int}}$ will be fixed later. By theorem 2.9, we can assume $\hat{\Psi}_1$ is concave. Without loss of generality we can assume $\hat{\Psi}_1(1) = 1$. Finally we define

$$\psi_1(r) := \hat{\psi}_1(\sqrt{r}).$$

Then, $\psi_1(r)$ is $1/2$-decreasing function. Also, using change of variables we have $\psi_1(r)$ is Dini continuous. Similarly, define $\psi_2(r)$ corresponding to Dini continuity of $g$, i.e., $\omega_g$. Finally define

$$\psi := \max\{\psi_3(r), \sqrt{r}\}$$

where $\psi_3(\lambda^k) := \sum_{i=0}^k \psi_1(\lambda^{k-i}) \psi_2(\lambda^i)$. Now proceed as in the proof of lemma 3.4.

Sketch of proof of theorem 3.7. First we will get the estimate as in lemma 3.5 and then use theorem 2.12 (Companato type characterization) to get the implications of the theorem.

Proof of theorem 1.1. Let $(t_1, x_1, y_1)$ and $(t_2, x_2, y_2)$ be in $Q^*_1/2$. Without loss of generality we shall assume $y_1 \leq y_2$. We will do proof in two cases:

(i) $|(t_1, X_1) - (t_2, X_2)| \leq \frac{y_1}{4}$
\( \left( \ell(t_1, x_1) - \ell(t_2, x_2) \right) \geq \frac{y_1}{t}. \)

In the first case, on using lemma 3.5 for \( r = y_1/2 \), we have

\[
\int_{Q^*_{y_1}(t_1, x_1, y_1)} |U(t, x, y) - \ell(t_1, x_1)(x)|^2 y^a \, dt \, dX \leq C_n C_l y_1^{n+5+a} K^2(y_1),
\]

where \( C_n = 2^{n+5+a} \). Consider

\[
\tilde{W}(t, x, y) := U(t, x, y) - \ell(t_1, x_1)(x).
\]

Now observe that the following rescaled function

\[
W(t, x, y) = \tilde{W}(t_1 + y_1^2 t, x_1 + y_1 x, y_1 y)/\sqrt{C_n C_l K(y_1) y_1}
\]

solves

\[
W_t - \text{div}(B(y_1 X) \nabla W) - (a/y)W_y = -\frac{\text{div}((B(y_1 X) - I) \cdot \nabla \ell)}{\sqrt{C_n C_l K(y_1)}}
\]

in \( Q_{1/2}(0, 0, 1) \) and

\[
||W||_{L^2(Q_{1/2}(0, 0, 1))} \leq 1.
\]

Note that for \( g = (B(y_1 X) - I) \cdot \nabla \ell/\sqrt{C_n C_l K(y_1)} \), we have \( \int_0^1 \frac{\Psi(t)}{t} \, dt \) and \( ||g||_{\infty} \) is bounded by a universal constant. We now apply theorem 3.7 to get

\[
|\nabla W(t, X) - \nabla W(0, 0, 1)| \leq C_{ir}(K(y_1 |(t, X) - (0, 0, 1)|) / K(y_1)
\]

\[
+ |(t, X) - (0, 0, 1)|^{1/2}
\]

(3.27)

for all \((t, X) \in Q_{1/2}(0, 0, 1)\). Note that condition of case(i) implies that

\[
(\bar{t}, \bar{X}) = \left( \frac{t_2 - t_1}{y_1^2}, \frac{x_2 - x_1}{y_1}, \frac{y_2}{y_1} \right) \in Q_{1/2}(0, 0, 1).
\]

Therefore we put \((\bar{t}, \bar{X})\) in (3.27). Subsequently, we use \( \ell(t_1, x_1) \) is linear and re-scale back to find

\[
|\nabla U(t_1, X_1) - \nabla U(t_2, X_2)| \leq C(K(|(t_1, X_1) - (t_2, X_2)|)
\]

\[
+ \frac{K(y_1)}{y_1^{3/2}} |(t_1, X_1) - (t_2, X_2)|^{1/2}.
\]

Since \( K \) is 1/2-decreasing, we obtain

\[
|\nabla U(t_1, X_1) - \nabla U(t_2, X_2)| \leq C K(|(t_1, X_1) - (t_2, X_2)|).
\]

Similarly, we get

\[
|U(t_1, X) - U(t_2, X)| \leq C K(|t_1 - t_2|^{1/2})|t_1 - t_2|^{1/2}.
\]

This proves the first case.
To prove the case(ii) first note that using triangle inequality we get

\[ |y_2| = |(t_2, X_2) - (t_2, x_2, 0)| \leq |(t_2, X_2) - (t_1, X_1)| + |(t_1, X_1) - (t_1, x_1, 0)| \]

\[ + |(t_1, x_1, 0) - (t_2, x_2, 0)| \leq |(t_2, X_2) - (t_1, X_1)| + |y_1| + |(t_2, X_2) - (t_1, X_1)| \leq 6|(t_2, X_2) - (t_1, X_1)|. \]

Use lemma 3.6 and (3.22) of theorem 3.7 to get the following estimate

\[ |\nabla U(t_1, X_1) - \nabla U(t_2, X_2)| \leq |\nabla U(t_1, X_1) - \nabla \ell(t_1, x_1)| + |\nabla \ell(t_1, x_1) - \nabla \ell(t_2, x_2)| \]

\[ + |\nabla U(t_2, X_2) - \nabla \ell(t_2, x_2)| \leq CK(y_1) + CK(|(t_1, x_1) - (t_2, x_2)|) + CK(y_2) \]

\[ \leq CK(6)|X_1 - X_2|. \]

Now we will get estimate for continuity in \(t\) variable:

\[ |U(t_1, X) - U(t_2, X)| \leq |U(t_1, x, y) - U(t_1, x, 0) + U(t_2, x, 0) - U(t_2, x, y)| \]

\[ + |U(t_1, x, 0) - U(t_2, x, 0)| \leq |U_y(t_1, x, \xi_1) - U_y(t_2, x, \xi_2)||y| + CK(|t_1 - t_2|^{1/2})|t_1 - t_2|^{1/2} \]

\[ \leq C(K(|(t_1, x, \xi_1) - (t_1, x, \xi_2)|) + K(|t_1 - t_2|^{1/2}))|t_1 - t_2|^{1/2}, \]

where \(\xi_1, \xi_2\) lies between 0 and \(y\). Since \(\xi_1, \xi_2\) lies between 0 and \(y\) therefore \(|(t_1, x, \xi_1) - (t_1, x, \xi_2)| \leq y\). We now use the condition of case(ii) to find \(|(t_1, x, \xi_1) - (t_1, x, \xi_2)| \leq 4|(t_1, X) - (t_2, X)|. Hence we get

\[ |U(t_1, X) - U(t_2, X)| \leq CK(4|t_1 - t_2|^{1/2})|t_1 - t_2|^{1/2}. \]

This completes the proof of case(ii). \(\square\)

**Acknowledgments**

We are thankful to Agnid Banerjee for various helpful discussions and suggestions. We would also like to thank the editor for the kind handling of our paper and the reviewer for various comments and suggestions which helped in improving the presentation of the manuscript.

**Appendix A. Appendix**

**Proof of theorem 2.6.** Consider \(f \in L^2(Q_1)\) and \(F \in L_a^2(Q_1^*)\). Now, using the density of smooth function, we have \(f_k \in C_c^\infty(Q_1)\) and \(F_k \in C_c^\infty(Q_1^*)\) such that \(f_k \rightarrow f\) in \(L^2(Q_1)\) and \(F_k \rightarrow F\) in \(L_a^2(Q_1^*)\). Let \(V^*\) be the dual space of \(V\). Define \(g_k :\)
$[-1, 1] \to V^*$ as following: Fix $t$, for $v \in V$, 
\[
g_k(t)v = \int_{B_1} f_k(t, x)v(x) \, dx + \int_{B_1^*} y^a F_k(t, X) \cdot \nabla v(X) \, dX.
\]
Using triangle inequality, Cauchy–Schwarz inequality and trace theorem (in the first term), we have $g_k(t) \in V^*$ with
\[
\|g_k(t)\|_{V^*} \leq \left( \int_{B_1} |f_k(t, x)|^2 \, dx \right)^{1/2} + \left( \int_{B_1^*} y^a |F_k(t, X)|^2 \, dX \right)^{1/2}.
\]
Use $f_k \in L^2(Q_1)$ and $F_k \in L^2_0(Q_1^*)$ to get $g_k \in L^2(-1, 1; V^*)$. Also, define $\hat{g}_k : [-1, 1] \to V^*$ as following: Fix $t$, for $v \in V$, 
\[
\hat{g}_k(t)v = \int_{B_1} (f_k)_t(t, x)v(x) \, dx + \int_{B_1^*} y^a (F_k)_t(t, X)v(X) \, dX,
\]
then $g'_k = \hat{g}_k \in L^2(-1, 1; V^*)$. For $u, v \in V$, define 
\[
a(u, v) = \int_{B_1^*} y^a \nabla u \cdot \nabla v \, dX.
\]
Note that by Poincaré inequality, for all $u \in V$ we have
\[
a(u, u) \geq C\|u\|^2.
\]
Thus we have verified conditions of theorem 5.1 of [20] for $u_0 = 0$ and $\chi = 0$. Hence we get $u_k \in L^2(-1, 1; V)$ and $(u_k)_t \in L^2(-1, 1; V)$ such that for almost all $t$ and for all $v \in V$ we have 
\[
\int_{B_1^*} (u_k)_t v \, dX + \int_{B_1^*} y^a \nabla u_k \cdot \nabla v \, dX = \int_{B_1} f_k(t, x)v(x) \, dx 
+ \int_{B_1^*} y^a F_k(t, X) \cdot \nabla v(X) \, dX. \tag{A.1}
\]
Put $v = u_k(t)$ in (A.1) and integrate with respect to $t$ to get
\[
\frac{1}{2} \int_{B_1^*} u_k^2(1) \, dX \, dt + \int_{B_1^*} y^a \nabla u_k \cdot \nabla u_k \, dX \, dt 
= \int_{B_1} f_k u_k \, dx \, dt + \int_{B_1^*} y^a F_k \cdot \nabla u_k \, dX \, dt.
\]
Since $t \to \|u_k(t)\|_V$ is a absolutely continuous function, we get 
\[
\frac{1}{2} \int_{B_1^*} u_k^2(1) \, dX \, dt + \int_{B_1^*} y^a \nabla u_k \cdot \nabla u_k \, dX \, dt 
= \int_{B_1} f_k u_k \, dx \, dt + \int_{B_1^*} y^a F_k \cdot \nabla u_k \, dX \, dt.
\]
On applying Hölder’s inequality, trace theorem and AM-GM inequality in the first term of right-hand side and Hölder’s inequality and AM-GM inequality in the second term we get

\[
\int_{-1}^{1} \int_{B_{1}^{*}} y^a \nabla u_k \cdot \nabla u_k \, dX \, dt \leq 2 C_{Tr} \int_{-1}^{1} \int_{B_{1}} f_k^2 \, dx \, dt + 2 \int_{-1}^{1} \int_{B_{1}^{*}} y^a |F_k|^2 \, dX \, dt.
\] (A.2)

Now write (A.1) as

\[
\int_{B_{1}^{*}} (u_k)_t v \, dX = \int_{B_{1}} f_k(t, x)v(x) \, dx + \int_{B_{1}^{*}} y^a F_k(t, X) \cdot \nabla v(X) \, dX
\]

\[- \int_{B_{1}^{*}} y^a \nabla u_k \cdot \nabla v \, dX. \] (A.3)

After applying Hölder’s inequality and trace theorem in first term of right-hand side and Hölder’s inequality in second term and third term of right-hand side we get, for almost all \(t\)

\[
|| (u_k)_t (t) ||_{V^*} \leq \left( \int_{B_{1}} |f_k(t, x)|^2 \, dx \right)^{1/2} + \left( \int_{B_{1}^{*}} y^a |F_k(t, X)|^2 \, dX \right)^{1/2}
\]

\[+ \left( \int_{B_{1}^{*}} y^a |\nabla u_k|^2 \, dX \right)^{1/2}. \]

After squaring both side and applying AM-GM inequality we get

\[
\int_{-1}^{1} || (u_k)_t (t) ||_{V^*}^2 \, dt \leq 4 \int_{-1}^{1} \int_{B_{1}} |f_k(t, x)|^2 \, dx \, dt + 4 \int_{-1}^{1} \int_{B_{1}^{*}} y^a |F_k(t, X)|^2 \, dX \, dt
\]

\[+ 4 \int_{-1}^{1} \int_{B_{1}^{*}} y^a |\nabla u_k|^2 \, dX \, dt. \]

Using (A.2) and boundedness of ||\(f_k||_{L^2}|| \) and ||\(F_k||_{L^2}|| \), we get \(u_k\) and \((u_k)_t\) is bounded in \(L^2(-1, 1; V)\) and \(L^2(-1, 1; V^*)\) respectively. Hence we will get a weak convergent subsequence \(u_k\) such that \(u_k \rightharpoonup u\) in \(L^2(-1, 1; V)\) and \((u_k)_t \rightharpoonup u_t\) in \(L^2(-1, 1; V^*)\). Hence, for all \(\phi \in H_{1}^1(-1, 1; V)\) we have

\[
- \int_{-1}^{1} \int_{B_{1}^{*}} u \phi_t \, dX \, dt + \int_{-1}^{1} \int_{B_{1}^{*}} y^a \nabla u \cdot \nabla \phi \, dX \, dt
\]

\[= \int_{-1}^{1} \int_{B_{1}} f(t, x)\phi \, dx \, dt + \int_{-1}^{1} \int_{B_{1}^{*}} y^a F_k(t, X) \cdot \nabla \phi(t, X) \, dX \, dt. \] (A.4)

Also \(u \in L^2(-1, 1; V)\) and \((u_k)_t \in L^2(-1, 1; V^*)\) implies \(u \in C(-1, 1; V)\). Since \(u_k(0) = 0\) for all \(k\), \(u(0) = 0\). Note that (A.4) is equivalent to following: for all
Lebesgue differentiation theorem gives $u \in C^1$. For every $\phi \in H^1(\Omega)$ we have

$$
\int_{B_1^1} u \phi(t_1) dX - \int_{B_1^1} u \phi(t_2) dX dt - \int_{t_1}^{t_2} \int_{B_1^1} u \phi_t dX dt + \int_{t_1}^{t_2} \int_{B_1^1} y^a \nabla u \cdot \nabla \phi dX dt
$$

$$
= \int_{t_1}^{t_2} \int_{B_1^1} f(t,x) \phi dx dt + \int_{t_1}^{t_2} \int_{B_1^1} y^a F(t,X) \cdot \nabla \phi(t,X) dX dt.
$$

Now we will prove uniqueness. Assume there are two solutions $u_1$ and $u_2$ in $C([-1,1]; V) \cap L^2(-1,1; V)$. By standard argument Steklov average of $u^1, u^2$ belongs to $H^1(-1,1; V)$ and satisfy

$$
\int_{t_1}^{t_2} \int_{B_1^1} (u^1_1)^t \phi dX dt + \int_{t_1}^{t_2} \int_{B_1^1} y^a \nabla u^1_1 \cdot \nabla \phi dX dt
$$

$$
= \int_{t_1}^{t_2} \int_{B_1^1} f_h(t,x) \phi dx dt + \int_{t_1}^{t_2} \int_{B_1^1} y^a F_h(t,X) \cdot \nabla \phi(t,X) dX dt.
$$

On putting $\phi = u^2_1 - u^1_1$ and subtracting two equations we get

$$
\int_{t_1}^{t_2} \int_{B_1^1} (u^2_1 - u^1_1)^t (u^2_1 - u^1_1) dX dt
$$

$$
+ \int_{t_1}^{t_2} \int_{B_1^1} y^a \nabla (u^2_1 - u^1_1) \cdot \nabla (u^2_1 - u^1_1) dX dt = 0.
$$

On passing limit as $h \to 0$ we get

$$
\int_{B_1^1} (u^2 - u^1)^2(t_2) dX - \int_{B_1^1} (u^2 - u^1)^2(t_1) dX \leq 0
$$

Take $t_1 \to -1$ and conclude $u^2 = u^1$. This completes the proof of uniqueness.

**Remark A.1.** For every $F \in L^2(Q_1)$, there exists a unique weak solution $U \in C([-1,1; V(B_1))$ to

$$
\begin{cases}
\partial_t u - \text{div}(B(x) \nabla u) = \text{div}(F) & \text{in } Q_1 \\
u = 0 & \text{on } \partial_p Q_1.
\end{cases}
$$

This can be obtained by taking $V = H^1_0(B_1)$ and measure $dx dt$ instead of $y^a dx dt$.

**Proof of theorem 2.12.** We first prove the existence of $\nabla u$. Observe that the Lebesgue differentiation theorem gives $u(t_0, x_0) = \ell_{(t_0, x_0)}(x_0)$. Consequently, we have

$$
|u(t_0, x_0 + re_i) - u(t_0, x_0) - re_i \cdot \nabla \ell_{(t_0, x_0)}|
$$

$$
= \frac{|\ell_{(t_0, x_0 + re_i)}(x_0 + re_i) - \ell_{(t_0, x_0)}(x_0) - re_i \cdot \nabla \ell_{(t_0, x_0)}|}{r},
$$
where \( e_i \in \mathbb{R}^n \) denote the usual \( i^{th} \) basis element. Since \( \ell(t_0, x_0)(x_0 + e_i r) = \ell(t_0, x_0)(x_0) + re_i \cdot \nabla \ell(t_0, x_0) \), therefore we get
\[
\frac{|u(t_0, x_0 + re_i) - u(t_0, x_0) - re_i \cdot \nabla \ell(t_0, x_0)|}{r} = \frac{|\ell(t_0, x_0 + re_i)(x_0 + re_i) - \ell(t_0, x_0)(x_0 + re_i)|}{r}.
\] (A.6)

Now we use \( \int_{Q_r(x_0 + re_i)} (x - x_0 - re_i) \, dx = 0 \) to obtain
\[
|\ell(t_0, x_0 + re_i)(x_0 + re_i) - \ell(t_0, x_0)(x_0 + re_i)|^2 \leq \int_{Q_r(t_0, x_0 + re_i)} |\ell(t_0, x_0 + re_i)(x) - \ell(t_0, x_0)(x)|^2 \, dt \, dx,
\]
here for simplicity we have assumed \(|Q_1| = 1\). We now add and subtract \( u(t, x) \) and use (A.6) to find
\[
|u(t_0, x_0 + re_i) - u(t_0, x_0) - re_i \cdot \nabla \ell(t_0, x_0)|^2 \leq \frac{2}{r^2} \int_{Q_r(t_0, x_0 + re_i)} |\ell(t_0, x_0 + re_i)(x) - u(t, x)|^2 \, dt \, dx
\]
\[
+ \frac{2}{r^2} \int_{Q_r(t_0, x_0 + re_i)} |u(t, x) - \ell(t_0, x_0)(x)|^2 \, dt \, dx.
\]
It is easy to see for \( r < 1/2 \), \( Q_r(t_0, x_0 + re_i) \subset Q_{2r}(t_0, x_0) \). Hence using (2.9) we get
\[
|u(t_0, x_0 + re_i) - u(t_0, x_0) - re_i \cdot \nabla \ell(t_0, x_0)|^2 \leq 2K^2(r) + 8K^2(2r) \leq 10K^2(2r).
\]

Now take \( r \to 0 \) to obtain \( \partial_x u(t_0, x_0) = e_i \cdot \nabla \ell(t_0, x_0) \). This proves the existence of \( \nabla u \).

We now prove the continuity of \( \nabla u \). Let \((t_1, x_1), (t_2, x_2) \in Q_1 \) with \( r = |(t_1, x_1) - (t_2, x_2)| \). Since \( \int_{Q_r} x_i \, dx = 0 \) for all \( i \) and for \( i \neq j \) we have \( \int_{Q_r} (e_i \cdot x)(e_j \cdot x) = 0 \).
Therefore, we find
\[
|e_i \cdot \nabla (\ell(t_1, x_1) - \ell(t_2, x_2))|^2 \leq \frac{1}{r^2} \int_{Q_r(t_1, x_1)} |\ell(t_1, x_1)(x) - \ell(t_2, x_2)(x)|^2 \, dt \, dx
\]
\[
\leq \frac{2}{r^2} \int_{Q_r(t_1, x_1)} |\ell(t_1, x_1)(x) - u(t, x)|^2 \, dt \, dx
\]
\[
+ \frac{2}{r^2} \int_{Q_{2r}(t_2, x_2)} |u(t, x) - \ell(t_2, x_2)(x)|^2 \, dt \, dx,
\]
where we have used \( Q_r(t_1, x_1) \subset Q_{2r}(t_2, x_2) \). Hence using (2.9) we obtain
\[
|e_i \cdot \nabla (\ell(t_1, x_1) - \ell(t_2, x_2))|^2 \leq 2K^2(r) + 8K^2(2r) \leq 10K^2(2r).
\]
This completes the proof of continuity of \( \nabla u \).
We now prove continuity in \( t \)-variable. Let \((t_1, x_0)\) and \((t_2, x_0)\) with \( r^2 = |t_1 - t_2| \).

Using \( u(t, x) = \ell_{(t,x)}(x) \) we get

\[
|u(t_1, x_0) - u(t_2, x_0)|^2 = |\ell_{(t_1,x_0)}(x_0) - \ell_{(t_2,x_0)}(x_0)|^2.
\]

On using \( \int_{Q_r(t_1,x_0)} (x-x_0) = 0 \), we find

\[
|u(t_1, x_0) - u(t_2, x_0)|^2 \leq 2 \int_{Q_r(t_1,x_0)} |\ell_{(t_1,x_0)}(x) - \ell_{(t_2,x_0)}(x)|^2 \, dt \, dx
\]

\[
\leq 2 \int_{Q_r(t_1,x_0)} |\ell_{(t_1,x_0)}(x) - u(t, x)|^2 \, dt \, dx + 2 \int_{Q_r(t_1,x_0)} |u(t, x) - \ell_{(t_2,x_0)}(x)|^2 \, dt \, dx
\]

\[
\leq 2 \int_{Q_r(t_1,x_0)} |\ell_{(t_1,x_0)}(x) - u(t, x)|^2 \, dt \, dx + 2 \int_{Q_{2r}(t_2,x_0)} |u(t, x) - \ell_{(t_2,x_0)}(x)|^2 \, dt \, dx,
\]

where last inequality is a consequence of \( Q_r(t_1, x_0) \subset Q_{2r}(t_2, x_0) \). We now use (2.9) to get

\[
|u(t_1, x_0) - u(t_2, x_0)|^2 \leq 2r^2 K^2(r) + 8r^2 K^2(2r) \leq 10r^2 K^2(2r).
\]

This completes the proof of the lemma. \( \square \)

References

1. K. Adimurthi and A. Banerjee. Borderline regularity for fully nonlinear equations in Dini domains. To appear in Adv. Calc. Var. arXiv:1806.07652v2.
2. I. Athanasopoulos, L. Caffarelli and E. Milakis. On the regularity of the non-dynamic parabolic fractional obstacle problem. J. Differ. Equ. 265 (2018), 2614–2647.
3. A. Audrito and S. Terracini. On the nodal set of solutions to a class of nonlocal parabolic reaction-diffusion equations. arXiv:1807.10135.
4. A. Audrito. On the existence and Hölder regularity of solutions to some nonlinear Cauchy-Neumann problems. arXiv:2107.03308.
5. A. Balakrishnan. Fractional powers of closed operators and the semigroups generated by them. Pac. J. Math. 10 (1960), 419–437.
6. A. Banerjee, D. Danielli, N. Garofalo and A. Petrosyan. The regular free boundary in the thin obstacle problem for degenerate parabolic equations. Algebra Anal. 32 (2020), 84–126.
7. A. Banerjee, D. Danielli, N. Garofalo and A. Petrosyan. The structure of the singular set in the thin obstacle problem for degenerate parabolic equations. Calc. Var. Partial Differ. Equ. 60 (2021), 91.
8. A. Banerjee and N. Garofalo. Monotonicity of generalized frequencies and the strong unique continuation property for fractional parabolic equations. Adv. Math. 336 (2018), 149–241.
9. A. Banerjee, N. Garofalo, I. Munive and D. Nhieu. The Harnack inequality for a class of nonlocal parabolic equations. Commun. Contemp. Math. 23 (2021), 2050050.
10. A. Banerjee and I. Munive. Gradient continuity estimates for the normalized p-Poisson equation. Commun. Contemp. Math. 22 (2020), 1950069.
11. A. Biswas, M. De Leon-Contreras and P. Stinga. Harnack inequalities and Hölder estimates for master equations. SIAM J. Math. Anal. 53 (2021), 2319–2348.
Gradient Continuity etc.

12 A. Biswas and P. R. Stinga. Regularity estimates for nonlocal space-time master equations in bounded domains. *J. Evol. Equ.* 21 (2021), 503–565.

13 L. Caffarelli. Interior estimates for fully nonlinear equations. *Ann. Math.* 130 (1989), 189–213.

14 L. Caffarelli and C. E. Kenig. Gradient estimates for variable coefficient parabolic equations and singular perturbation problems. *Am. J. Math.* 120 (1998), 391–439.

15 L. Caffarelli and L. Silvestre. An extension problem related to the fractional Laplacian. *Commun. Partial Differ. Equ.* 32 (2007), 1245–1260.

16 L. Caffarelli and L. Silvestre. Hölder regularity for generalized master equations with rough kernels. Advances in Analysis: The Legacy of Elias M. Stein, Princeton Math. Ser., vol. 50, pp. 63–83 (Princeton, NJ: Princeton University Press, 2014).

17 F. Chiarenza and R. Serapioni. A remark on a Harnack inequality for degenerate parabolic equations. *Rend. Sem. Mat. Univ. Padova* 73 (1985), 179–190.

18 P. Daskalopoulos, T. Kuusi and G. Mingione. Borderline estimates for fully nonlinear elliptic equations. *Calc. Var. Partial Differ. Equ.* 39 (2014), 574–590.

19 H. Dong and T. Phan. Regularity for parabolic equations with singular or degenerate coefficients. *Calc. Var. Partial Differ. Equ.* 60 (2021), 44.

20 G. Duzaar and G. Mingione. Gradient estimates via non-linear potentials. *Arch. Ration. Mech. Anal.* 207 (2013), 215–246.

21 T. Kuusi and G. Mingione. Universal potential estimates. *J. Funct. Anal.* 262 (2012), 4205–4269.

22 T. Kuusi and G. Mingione. Linear potentials in nonlinear potential theory. *Arch. Ration. Mech. Anal.* 207 (2013), 1–82.

23 T. Kuusi and G. Mingione. Guide to nonlinear potential estimates. *Bull. Math. Sci.* 4 (2014), 1–86.

24 R. Lai, Y. Lin and A. Ruland. The Calderón problem for a space-time fractional parabolic equation. *SIAM J. Math. Anal.* 52 (2020), 2655–2688.

25 M. Litsgard and K. Nyström. Fractional powers of parabolic operators with time-dependent measurable coefficients. *arXiv:2104.07313*.

26 G. G. Lorentz. *Approximation of functions* (New York-Chicago, Ill.-Toronto, ON: Holt, Rinehart and Winston, 1966).

27 R. Metzler and J. Klafter. The random walk’s guide to anomalous diffusion: a fractional dynamics approach. *Phys. Rep.* 339 (2000), 1–77.

28 E. W. Montroll and G. H. Weiss. Random walks on lattices. II. *J. Math. Phys.* 6 (1965), 167–181.

29 M. Riesz. Intégrales de Riemann-Liouville et potentiels. *Acta Sci. Math. Szeged* 9 (1938), 1–42.

30 M. Riesz. L’intégrale de Riemann-Liouville et le problème de Cauchy. *Acta Math.* 81 (1949), 1–223.
E. M. Stein. Editor’s note: the differentiability of functions in $\mathbb{R}^n$. *Ann. Math.* **113** (1981), 383–385.

P. R. Stinga and J. L. Torrea. Regularity theory and extension problem for fractional nonlocal parabolic equations and the master equation. *SIAM J. Math. Anal.* **49** (2017), 3893–3924.