PARABOLIC Lp DIRICHLET BOUNDARY VALUE PROBLEM AND VMO-TYPE TIME-VARYING DOMAINS

Citation for published version:
Dindos, M, Dyer, L & Hwang, S 2019, 'PARABOLIC Lp DIRICHLET BOUNDARY VALUE PROBLEM AND VMO-TYPE TIME-VARYING DOMAINS' Analysis & PDE.

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Analysis & PDE

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
PARABOLIC $L^p$ DIRICHLET BOUNDARY VALUE PROBLEM
AND VMO-TYPE TIME-VARYING DOMAINS

MARTIN DINDOŠ, LUKE DYER, AND SUKJUNG HWANG

Abstract. We prove the solvability of the parabolic $L^p$ Dirichlet boundary value problem for $1 < p \leq \infty$ for a PDE of the form $u_t = \text{div}(A\nabla u) + B \cdot \nabla u$ on time-varying domains where the coefficients $A = [a_{ij}(X, t)]$ and $B = [b_i]$ satisfy a certain natural small Carleson condition. This result brings the state of affairs in the parabolic setting up to the elliptic standard.

Furthermore, we establish that if the coefficients of the operator $A, B$ satisfy a vanishing Carleson condition and the time-varying domain is of VMO type then the parabolic $L^p$ Dirichlet boundary value problem is solvable for all $1 < p \leq \infty$. This result is related to results in papers by Maz’ya, Mitrea and Shaposhnikova, and Hofmann, Mitrea and Taylor where the fact that boundary of domain has normal in VMO or near VMO implies invertibility of certain boundary operators in $L^p$ for all $1 < p \leq \infty$ which then (using the method of layer potentials) implies solvability of the $L^p$ boundary value problem in the same range for certain elliptic PDEs.

Our result does not use the method of layer potentials since the coefficients we consider are too rough to use this technique, but remarkably we recover $L^p$ solvability in the full range of $p$’s as the two papers mentioned above.

1. Introduction

Let us consider a parabolic differential equation on a time-varying domain $\Omega$ of the form

\[
\begin{cases}
    u_t = \text{div}(A\nabla u) + B \cdot \nabla u & \text{in } \Omega \subset \mathbb{R}^{n+1}, \\
    u = f & \text{on } \partial \Omega,
\end{cases}
\]

where $A = [a_{ij}(X, t)]$ is a $n \times n$ matrix satisfying the uniform ellipticity condition with $X \in \mathbb{R}^n$, $t \in \mathbb{R}$. That is, there exist positive constants $\lambda$ and $\Lambda$ such that

\[
\lambda|\xi|^2 \leq \sum_{i,j} a_{ij}(X, t)\xi_i \xi_j \leq \Lambda|\xi|^2
\]

for almost every $(X, t) \in \Omega$ and all $\xi \in \mathbb{R}^n$. In addition, we assume that the coefficients of $A$ and $B$ satisfy a natural, minimal smoothness condition (1.6) and we do not assume any symmetry on $A$.

It has been observed via the method of layer potentials that when the domain on which we consider certain boundary value problems for elliptic or parabolic PDEs is sufficiently smooth the question of $L^p$ invertibility of certain boundary operator can be resolved using the Fredholm theory since this operator is just a compact perturbation of the identity. This observation then implies invertibility of this boundary operator for all $1 < p \leq \infty$ and hence solvability of the corresponding $L^p$ boundary value problem in this range.

Luke Dyer was supported by The Maxwell Institute Graduate School in Analysis and its Applications, a Centre for Doctoral Training funded by the UK Engineering and Physical Sciences Research Council (grant EP/L016508/01), the Scottish Funding Council, the University of Edinburgh, and Heriot-Watt University.

Sukjung Hwang was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2017R1D1A1B03035152).
The notion of how smooth the domain has to be for the above observation to hold has evolved. Initial results for constant coefficient elliptic PDEs required domains of at least $C^{1,\alpha}$ type. This was reduced to $C^{1}$ domains in an important paper of Fabes, Jodeit, and Rivière [FJR78]. Later the method of layer potentials was adapted to variable coefficient settings, and the results were extended to elliptic PDEs with variable coefficients [Din08] on $C^{1}$ domains.

Further progress was made after advancements in singular integrals theory on sets that are not necessary of graph-type [Sem91, HMT10]. It turns out that compactness of the mentioned boundary operator only requires that the normal (which must be well defined at almost every boundary point) belongs to VMO.

This observation for the Stokes system was made in [MMS09] where boundary value problems for domains whose normal belongs to VMO (or is near to VMO in the BMO norm) were considered. In [HMT15] symbol calculus for operators of layer potential type on surfaces with VMO normals was developed and applied to various elliptic PDEs including elliptic systems.

So far we have only mentioned elliptic results. One of the first results for the heat equation in Lipschitz cylinders is by Brown [Bro89]. Here the domain considered is time independent and Fourier methods in the time variable are used. Domains of a time-varying type for the heat operator were first considered the papers [LM95, HL96] and again the method of layer potentials was used to establish $L^2$ solvability. The question of solvability of various boundary value problems for parabolic PDEs on time-varying, domains has a long history. Recall, that in the elliptic setting [Dah77] has shown in a Lipschitz domain that the harmonic measure and surface measure are mutually absolutely continuous and that the elliptic Dirichlet problem is solvable with data in $L^2$ with respect to the surface measure. R. Hunt then asked whether Dahlberg’s result holds for the heat equation in domains whose boundaries are given locally as functions $\phi(x,t)$, Lipschitz in the spatial variable. It was conjectured (due to the natural parabolic scaling) that the correct regularity of $\phi(x,t)$ should be a Hölder condition of order $1/2$ in the time variable $t$ and Lipschitz in $x$. It turns out that under this assumption the parabolic measure associated with the equation (1.1) is doubling [Nys97].

However, to answer R. Hunt’s question positively, one has to consider more regular classes of domains than the one just described above. This follows from the counterexample of [KW88] where it was shown that under just the $\text{Lip}(1,1/2)$ condition on the domain $\Omega$ the associated caloric measure (that is the measure associated with the operator $\partial_t - \Delta$) might not be mutually absolutely continuous with the natural surface measure. The issue was resolved in [LM95] where it was established that mutual absolute continuity of caloric measure and a certain parabolic analog of the surface measure holds when $\phi$ has $1/2$ of a time derivative in the parabolic $\text{BMO}(\mathbb{R}^n)$ space, which is a slightly stronger condition than $\text{Lip}(1,1/2)$. We shall call such domains to be of Lewis-Murray type. [HL96] subsequently showed that this condition is sharp. We thoroughly discuss these domains in section 2.1.

Further work was done by [HL01, Riv03, Riv14] in graph domains and time-varying cylinders satisfying the Lewis-Murray condition where they proved the $L^p$ Dirichlet problem was solvable for all $p > p'$ for some potentially very large $p'$ (due to the technique used there is no control on the size of $p'$). Finally, [DH18] has established $L^p$ solvability $2 \leq p \leq \infty$ in domains that are local of Lewis-Murray type under a small Carleson condition.

While researching literature on domains of Lewis-Murray type and ways this concept can be localized (in the time variable the half-derivative is a nonlocal operator, and hence any condition imposed on it is difficult to localize) we have realized that important results we have planned to rely on have issues (either in
their proofs or even worse are simply false, see in particular remark 2.7 in the next section). This has prompted us to write section 2.1 on parabolic domains in substantially more detail we originally intended to. This sets the literature record straight and more importantly in detail explains the concept of localized domains of Lewis-Murray type. For readability of the paper and this section, we have moved long proofs into an appendix.

In this paper, we establish $L^p$ solvability results for parabolic PDEs on time-varying cylinders satisfying locally the Lewis-Murray condition in the full range $1 < p \leq \infty$ improving the solvability range from [DH18] as well as older results such as [HL96], where only $p = 2$ was considered. The coefficients we consider are very rough and, in particular, the method of layer potentials cannot be used. Despite this, we recover (in the parabolic setting) an analog of [MMS09] and [HMT15]. When the domain $\Omega$, on which the parabolic PDE is considered, is of VMO type (that is certain derivatives both in temporal and spatial variables will be in VMO) and the coefficients of the operator satisfy a vanishing Carleson condition the $L^p$ solvability can be established for all $1 < p \leq \infty$. Remarkably this is the full range of solvability that holds for smooth coefficients (via the layer potential method).

Our proof is however completely different from the layer potential method, for example at no point is compactness used. The proof is also substantially different than the case $2 \leq p \leq \infty$ of [DH18] in the following way. We were inspired by [DPP07] and have used a so-called $p$-adapted square function to prove $L^p$ solvability. However, due to the presence of the parabolic term, a second square function type object will arise, namely

$$
\int_{\Omega} |u_t(X,t)|^2|u(X,t)|^{p-2}\delta(X,t)^3 \, dX \, dt, \tag{1.3}
$$

where $\delta(X,t)$ is the parabolic distance to the boundary defined as

$$
\delta(X,t) = \inf_{(Y,\tau) \in \partial\Omega} \left( |X-Y|^2 + |t-\tau| \right)^{1/2}.
$$

When $p = 2$ such object was called the “area function” and in [DH18] it was shown that it the usual square function can dominate it. It turns out however that the case $1 < p < 2$ is substantially more complicated and we were only able to establish required bounds for (1.3) for non-negative $u$ after a substantial effort.

There is also an issue of whether the $p$-adapted square function is actually well-defined and locally finite (as the exponent on $|u|$ is negative). We prove that when $u$ is a solution of a parabolic PDE the $p$-adapted square function is indeed well defined by adapting a recent regularity result [DP16]. The paper [DP16] deals with complex coefficient elliptic PDEs but the method used there can be adapted to the parabolic setting; see Theorem 4.1 for details.

Many results in the parabolic setting, are motivated by previous results in the elliptic setting and ours is not different. Let us, therefore, overview the major elliptic results related to our main theorem.

The papers [KKPT00] and [KP01] started the study of non-symmetric divergence elliptic operators with bounded and measurable coefficients. [KP01] used [KKPT00] to show that the elliptic measure of operators satisfying a type of Carleson measure condition is in $A_\infty$ and hence the $L^p$ Dirichlet problem is solvable for some, potentially large, $p$. In [DPP07], the authors improved the result of [KP01] in the following way. They showed that if

$$
\delta(X)^{-1} \left( \operatorname{osc}_{B_a(X)/2(X)} a_{ij} \right)^2 \quad \text{and} \quad \delta(X) \left( \sup_{B_a(X)/2(X)} b_i \right)^2 \tag{1.4}
$$
are densities of Carleson measures with vanishing Carleson norms then the $L^p$ Dirichlet problem is solvable for all $1 < p \leq \infty$. A similar result for the elliptic Neumann and regularity boundary value problem was established in [DPR17].

The parabolic analogue of the elliptic Carleson condition (1.4) is that

$$\delta(X,t)^{-1} \sup_{i,j} \left( \frac{\text{osc}}{B_{a/2}(X,t)} a_{ij} \right)^2 + \delta(X,t) \sup_{B_a(X,t)} b_i$$

is the density of a Carleson measure on $\Omega$ with a small Carleson norm and $\delta(X,t)$ is the parabolic distance of a point $(X,t)$ to the boundary $\partial \Omega$.

The condition (1.5) arises naturally as follows. Let $\Omega = \{ (x_0, t) : x_0 > \phi(x,t) \}$ for a function $\phi$ which satisfies the Lewis-Murray condition above. Let $\rho : U \to \Omega$ be a mapping from the upper half space $U$ to $\Omega$. Consider $v = u \circ \rho$. It will follow that if $u$ solves (1.1) in $\Omega$ then $v$ will be a solution to a parabolic PDE similar to (1.1) in $U$. In particular, if $\rho$ is chosen to be the mapping in (2.26) then the coefficients of the new PDE for $v$ will satisfy a Carleson condition like (1.5), c.f. Lemma 2.18, provided the original coefficients (for $u$) were either smooth or constant.

Furthermore, if we do not insist on control over the size of the Carleson norm, then we can still infer solvability of the $L^p$ Dirichlet problem for large $p$, as in [HL01, Riv03, Riv14].

Finally, we are ready to state our main result; some notions used here are defined in detail in section 2.

**Theorem 1.1.** Let $\Omega$ be a domain as in definition 2.10 with character $(\ell, \eta, N, d)$ and let $A$ be bounded and elliptic (1.2), and $B$ be measurable. Consider any $1 < p \leq \infty$ and assume that either:

(1) $d\mu_1 = \left[ \delta(X,t)^{-1} \sup_{i,j} \left( \frac{\text{osc}}{B_{a/2}(X,t)} a_{ij} \right)^2 + \delta(X,t) \sup_{B_a(X,t)} |B|^2 \right] dX dt$ is a Carleson measure on $\Omega$ with Carleson norm $\| \mu_1 \|_C$.

(2) Or assume in addition that $\nabla A, \partial_t A$ are well defined a.e. point in $\Omega$ and

$$d\mu_2 = (\delta(X,t)|\nabla A|^2 + \delta(X,t)\partial_t |A|^2 + \delta(X,t)|B|^2) \, dX dt$$

is a Carleson measure on $\Omega$ with Carleson norm $\| \mu_2 \|_C$ and

$$\delta(X,t)|\nabla A| + \delta(X,t)^2 |\partial_t A| + \delta(X,t)|B| \leq \| \mu_2 \|_C^{1/2}. \tag{1.8}$$

Then there exists $K = K(\lambda, \Lambda, \ell, n, p) > 0$ such that if for some $r_0 > 0$

$$\max|\eta|, \| \mu_1 \|_{C, r_0} < K \quad \text{or} \quad \max|\eta|, \| \mu_2 \|_{C, r_0} < K$$

the $L^p$ Dirichlet boundary value problem (1.1) is solvable (c.f. definition 2.26) and the following estimate holds for all continuous boundary data $f \in C_0(\partial \Omega)$

$$\| N(u) \|_{L^p(\partial \Omega, d\sigma)} \lesssim \| f \|_{L^p(\partial \Omega, d\sigma)},$$

where the implied constant depends only on the operator, $n$, $p$ and character $(\ell, \eta, N, d)$, and $N(u)$ is the non-tangential maximal function of $u$.

**Corollary 1.2.** In particular, if $\Omega$ is of VMO-type ($\eta$ in the character $(\ell, \eta, N, d)$ can be taken arbitrary small), and the Carleson measure $\mu$ from Theorem 1.1 is a vanishing Carleson measure then the $L^p$ Dirichlet boundary value problem (1.1) is solvable for all $1 < p \leq \infty$.

2. Preliminaries

Here and throughout we consistently use $\nabla u$ to denote the gradient in the spatial variables and $u_t$ or $\partial_t u$ the gradient in the time variable.
2.1. Parabolic Domains. In this subsection, we define a class of time-varying domains whose boundaries are given locally as functions $\phi(x,t)$, Lipschitz in the spatial variable and satisfying the Lewis-Murray condition in the time variable. At each time $t \in \mathbb{R}$ the set of points in $\Omega$ with fixed time $t = \tau$, that is $\Omega_\tau = \Omega \cap \{t = \tau\}$, is a non-empty bounded Lipschitz domain in $\mathbb{R}^n$. We start with a discussion of the Lewis-Murray condition, give a summary and clarification of the results in the literature, and introduce some new equivalent definitions.

Let $Q_r(x,t) = \{(y,s) \in \mathbb{R}^{n-1} \times \mathbb{R} : |x_i - y_i| < r$ for all $1 \leq i \leq n - 1, |t - s|^{1/2} < r \}$. Let $J_r \subset \mathbb{R}^{n-1}$ be a spatial cube of radius $r$. For a given $f : \mathbb{R}^n \to \mathbb{R}$ let

$$f_{Q_r} = \frac{1}{|Q_r|} \int_{Q_r} f(x,t) \, dx \, dt.$$ 

When we write $f \in \text{BMO}(\mathbb{R}^n)$ we mean that $f$ belongs to the parabolic version of the usual BMO space with the norm $\|f\|_\ast$ where

$$\|f\|_\ast = \sup_{Q_r} \frac{1}{|Q_r|} \int_{Q_r} |f - f_{Q_r}| \, dx \, dt < \infty. \quad (2.1)$$

Recall that the Lewis-Murray condition imposed that a half derivative in time of $\phi(x,t)$ belongs to parabolic BMO. There are a few different ways one can define half derivatives, and BMO-Sobolev spaces and there are also some erroneous results in the literature which we correct here. To bring clarity, we start by discussing the various definitions in the global setting of a graph domain $\Omega = \{(x_0, x, t) : x_0 > \phi(x,t)\}$, where $\phi : \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R}$. We follow the standard notation of [HL96].

If $g \in C^\infty_c(\mathbb{R})$ and $0 < \alpha < 2$ then the one-dimensional fractional differentiation operators $D_\alpha$ are defined on the Fourier side by

$$\hat{D_\alpha}g(\tau) = |\tau|^\alpha \hat{g}(\tau).$$

If $0 < \alpha < 1$ then by standard results

$$D_\alpha g(t) = c \int_\mathbb{R} \frac{g(t) - g(s)}{|t - s|^{1+\alpha}} \, ds.$$ 

Therefore, we define the \textit{pointwise half derivative in time} of $\phi : \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R}$ to be

$$D^\prime_{1/2} \phi(x,t) = c_n \int_\mathbb{R} \phi(x,s) - \phi(x,t) \frac{|s - t|^{1/2}}{|s - t|^{1/2}} \, ds, \quad (2.2)$$

for a properly chosen constant $c_n$ (c.f. [HL96]). In order for the Fourier transform to be well-defined, $\phi$ should be a tempered distribution modulo first degree polynomials in $x$ (c.f. [HL96] and [Str80]).

However, this definition ignores the spatial coordinates. Instead by following [FR67] we may define the \textit{parabolic half derivative in time} of $\phi : \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R}$ to be

$$\widehat{D_\alpha} \phi(\xi, \tau) = \frac{\tau}{\|\xi, \tau\|^\alpha} \hat{\phi}(\xi, \tau), \quad (2.3)$$

where $\xi$ and $\tau$ denote the spatial and temporal variables on the Fourier side respectively, and $\|x(t)\| = |x| + |t|^{1/2}$ denotes the parabolic norm. In addition we define the \textit{parabolic derivative} (in space and time) of $\phi : \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R}$ to be

$$\widehat{D_\alpha} \phi(\xi, \tau) = \|\xi, \tau\|^\alpha \hat{\phi}(\xi, \tau). \quad (2.4)$$

$\mathbb{D}^{-1}$ is the parabolic Riesz potential. Again, we assume here that $\phi$ is a tempered distribution modulo first degree polynomials in $x$. One can also represent $\mathbb{D}$ as

$$\mathbb{D} = \sum_{j=1}^n R_j \mathbb{D}_j, \quad (2.5)$$
where $\mathbb{D}_j = \partial_j$ for $1 \leq j \leq n - 1$, $\mathbb{D}_n$ is defined above and $R_j$ are the parabolic Riesz transforms defined on the Fourier side as
\[
\hat{R}_j(\xi, \tau) = \frac{i\xi_j}{\|\xi, \tau\|} \quad \text{for } 1 \leq j \leq n - 1 \quad \text{and} \quad \hat{R}_n(\xi, \tau) = \frac{\tau}{\|\xi, \tau\|}. \tag{2.6}
\]

Furthermore the kernels of $R_j$ have average zero on (parabolically weighted) spheres around the origin, obey the standard Calderón-Zygmund kernel and therefore by standard Calderón-Zygmund theory each $R_j$ defines a bounded operator on $L^p(\mathbb{R}^n)$ for $1 < p < \infty$ and is bounded on $\text{BMO}(\mathbb{R}^n)$ [Pee66, FR66, FR67, HL96].

We say that $\phi : \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R}$ is $\text{Lip}(1,1/2)$ with Lipschitz constant $\ell$ if $\phi$ is Lipschitz in the spatial variables and Hölder continuous of order $1/2$ in the temporal variable. That is
\[
|\phi_j(x,t) - \phi_j(y,t)| \leq \ell \left( |x-y| + |t-s|^{1/2} \right). \tag{2.7}
\]

The Lewis-Murray condition on the domain $\Omega$, for which they proved the mutual absolute continuity of the caloric measure and the natural surface measure, is $\phi \in \text{Lip}(1,1/2)$ and $\|D_{1/2}\phi\|_* \leq \eta$; note this BMO norm is taken over $\mathbb{R}^n$.

It is worth remarking that neither the operators $D_{1/2}$, $\mathbb{D}_n$ or $\mathbb{D}$ easily lend themselves to being localized to a function $\phi : Q_d \to \mathbb{R}$ due to their non-local natures. However, our goal is to provide a theory where the domain is locally given by graphs satisfying the Lewis-Murray condition. The parabolic nature of the PDE (especially time irreversibility and exponential decay of solutions with vanishing boundary data) suggest we should expect to need only local conditions on the functions describing the boundary.

To this end, we state the following theorems where we show some equivalent statements to the Lewis-Murray condition for a global function $\phi : \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R}$. Furthermore, the final conditions admit themselves to both being localized easily as well as amiable to an extension, see Theorem 2.8 later for details on an extension.

The equivalence of conditions (1) and (2) below is shown in [HL96] with an equivalence of norms in the small and large sense, see [HL96, (2.10) and Theorem 7.4] for precise details, c.f. (2.5) and (2.6).

**Theorem 2.1.** Let $\phi : \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R}$ and $\phi \in \text{Lip}(1,1/2)$ then the following conditions are equivalent:

1. $D_{1/2}\phi \in \text{BMO}(\mathbb{R}^n)$
2. $\mathbb{D}_n\phi \in \text{BMO}(\mathbb{R}^n)$
3. $D\phi \in \text{BMO}(\mathbb{R}^n)$.

We also note that since $\mathbb{D}_n\phi = R_nD\phi$ we have $\|\mathbb{D}_n\phi\|_* \lesssim \|D\phi\|_*$ by the boundedness of $R_n$ on $\text{BMO}(\mathbb{R}^n)$.

We now extend this theorem by adding three more equivalent statements. To motivate condition (6) of Theorem 2.3 below we first recall a characterisation of $\text{BMO}$ from [Str80, p. 546]. Let $M(f, Q) = \frac{1}{|Q|} \int_Q f$ denote the average of $f$ over a cube $Q$, and let $\bar{Q}_\rho(x)$ be the cube of radius $\rho$ with $x$ in the upper right corner.

**Lemma 2.2** ([Str80]). $f \in \text{BMO}(\mathbb{R}^n)$ is equivalent to
\[
\sup_{Q, r} \sum_{k=1}^n \frac{1}{|Q|} \int_{\bar{Q}_r} \int_0^r \left| M(f, \bar{Q}_\rho(x)) - M(f, \bar{Q}_\rho(x - \rho e_k)) \right|^2 \frac{d\rho}{\rho} \, dx = B < \infty, \tag{2.8}
\]
where $e_k$ are the usual unit vectors in $\mathbb{R}^n$, and $\|f\|_*^2 \sim B$. 

The equivalence of conditions (3) and (4) in theorem below is a generalisation of [Str80] to the parabolic setting that is stated in [Riv03], c.f. [FS72, CT75, CT77]. We have some question marks over the proof given in [Riv03]; however, the argument we give for condition (5) also works for condition (4) and hence the claim in [Riv03] is correct.

**Theorem 2.3.** Let $\phi : \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R}$ and $\phi \in \text{Lip}(1,1/2)$ then the following conditions are equivalent:

1. $\mathbb{D}\phi \in \text{BMO}(\mathbb{R}^n)$
2. $\sup_{Q_r} \frac{1}{|Q_r|} \int_{Q_r} \frac{|\phi(x+y,t+s) - 2\phi(x,t) + \phi(x-y,t-s)|^2}{\|(y,s)\|^{n+3}} dy \, ds \, dt = B_{(4)} < \infty,$ \hspace{1cm} (2.9)

3. $\sup_{Q_r} \frac{1}{|Q_r|} \int_{Q_r} \frac{|\phi(x+y,t) - 2\phi(x,t) + \phi(x-y,t)|^2}{|y|^{n+1}} dy \, dx \, dt = B_{(5,a)} < \infty,$ \hspace{1cm} (2.10)

4. $\sup_{Q_{r,t}} \frac{1}{|Q_{r,t}|} \int_{Q_{r,t}} \frac{|\phi(x,t) - \phi(x,s)|^2}{|t-s|^2} ds \, dt \, dx = B_{(5,b)} < \infty.$ \hspace{1cm} (2.11)

5. Let $a = (a', a_n)$ and $e_n$ be the unit vector in the time direction. For $k = 1, \ldots, n - 1$ let

$$A_k = \int_0^1 \rho u' \cdot \left( M \left( \nabla \phi, \tilde{Q}_{\rho}(x + \lambda \rho u', t) \right) - M \left( \nabla \phi, \tilde{Q}_{\rho}(x + \lambda \rho u' - \rho e_k, t) \right) \right) d\lambda,$$

$$A_n = \int_0^1 \rho u' \cdot \left( M \left( \nabla \phi, \tilde{Q}_{\rho}(x + \lambda \rho u', t) \right) - M \left( \nabla \phi, \tilde{Q}_{\rho}(x + \lambda \rho u', t - \rho^2) \right) \right) d\lambda.$$

6. a) $\sup_{Q_r} \sum_{k=1}^n \frac{1}{|Q_r|} \int_{Q_r} \int_{u \in \mathbb{S}^{n-1}} \int_0^r \frac{|A_k|^2}{\rho^2} d\rho \, du \, dx \, dt = B_{(6,a)} < \infty,$ \hspace{1cm} (2.12)

b) $\sup_{Q_{r,t}} \frac{1}{|Q_{r,t}|} \int_{Q_{r,t}} \frac{|\phi(x,t) - \phi(x,s)|^2}{|t-s|^2} ds \, dt \, dx = B_{(6,b)} < \infty.$ \hspace{1cm} (2.11)

Furthermore we have equivalence of the norms

$$\|\mathbb{D}\phi\|_p^2 \sim B_{(4)} \sim B_{(5,a)} + B_{(5,b)} \sim B_{(6,a)} + B_{(6,b)}.$$ \hspace{1cm} (2.13)

We give a proof of this result in the appendix at the end of the paper.

**Remark 2.4.** Condition (6,a) does not immediately look too similar to its supposed motivation, (2.8) in Lemma 2.2. However, if we move back into Cartesian coordinates and undo the mean value theorem, then we obtain something much similar to a combination of (2.8) and an endpoint version of [Str80, (3.1)]. The reason why we can obtain the endpoint, whereas [Str80, (3.1)] can only be used for a fractional derivative smaller than 1, is due to additional integrability and cancellation coming from (8.1). Consider
then condition (6.a) is equivalent to
\[

t_{n} \sim \|D \phi \|_{\text{BMO}(\mathbb{R}^{n-1})} < \infty.
\]  

(2.14)

**Proposition 2.5.** Let \( \phi : \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R} \) and \( \phi \in \text{Lip}(1, 1/2) \). Let \( B_{(5.a)} \) and \( B_{(6.a)} \) be as in Theorem 2.3. Then we have
\[
B_{(6.a)} \lesssim \|\nabla \phi\|^{2}, \quad B_{(5.a)} \lesssim \sup_{t} \|\nabla \phi(\cdot, t)\|^{2}_{\text{BMO}(\mathbb{R}^{n-1})},
\]
where \( \text{BMO}(\mathbb{R}^{n-1}) \) denotes the \( \text{BMO} \) norm in the spatial variables only.

**Proof.** The statement \( \nabla \phi \in \text{BMO}(\mathbb{R}^{n-1}) \) implies condition (5.a) follows from [Str80, Theorem 3.3]. In order to establish the second claim for the ease of notation let us fix \( Q_{e} \) and \( k \) in \( 1 \leq k \leq n - 1 \). Then since \( |u| \leq 1 \) after changing the order of integration (and the substitution \( y = x + \lambda \rho u' \in Q_{2r} \)) we get that \( B_{(6.a)} \) defined by (2.12) is bounded by
\[
\int_{0}^{1} \int_{0}^{r} \int_{Q_{e}} \frac{1}{|Q_{e}|} \int_{Q_{2r}} \left| \left( M (\nabla u, \hat{\phi}_{r}(y, t)) - M (\nabla u, \hat{\phi}_{r}(y - \rho e_{k}, t)) \right) \right|^{2} dy \, dt \, d\rho \, du \, d\lambda.
\]

Then by Lemma 2.2 the two interior integrals are bounded by \( C\|\nabla \phi\|^{2} \). Therefore \( B_{(6.a)} \) is controlled by \( C\|\nabla \phi\|^{2} \).

It’s not immediately obvious whether the opposite implications is true or false due to the highly singular nature of Riesz potentials, c.f. (2.5) and (2.6).

**Corollary 2.6.** Let \( \phi : \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R} \) and \( \phi \in \text{Lip}(1, 1/2) \). If \( \|\nabla \phi\|_{\ast} \lesssim \eta \), and
\[
B_{(5.a)} \lesssim \eta^{2} \text{ then } \|\nabla \phi\|_{\ast} \lesssim \eta.
\]

Here we have replaced conditions (5.a) or (6.a) by slightly stronger but easier to verify condition \( \|\nabla \phi\|_{\ast} \lesssim \eta \). We believe that, without too much extra work, one could formulate our main theorem and associated lemmas with a local version of condition (5.a) in place of \( \|\nabla \phi\|_{\ast} \).

**Remark 2.7.** In [Riv03, Lemma 2.1], it is stated that another condition is equivalent to those given in Theorems 2.1 and 2.3; however this claim is not correct and only one of the stated implications holds.

A result of Strichartz [Str80, Theorem 3.3] states that in the one dimensional setting \( D_{1/2}^{\frac{1}{2}} \phi(t) \in \text{BMO}(\mathbb{R}) \) is equivalent to the one dimensional version of conditions (5.b) and (6.b)
\[
\sup_{t' \subseteq \mathbb{R}} \left( \frac{1}{|t'|} \int_{t'} \int_{t'} |\phi(t) - \phi(s)|^{2} dt \, ds \right)^{1/2} \leq B,
\]
with \( B \sim \|D_{1/2}^{\frac{1}{2}} \phi(t)\|_{\text{BMO}(\mathbb{R})} \).

In [Riv03, Lemma 2.1] it is claimed that given \( \phi : \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R} \) and \( \phi \in \text{Lip}(1, 1/2) \) the pointwise \( n \)-dimensional analogue of (2.15)
\[
\sup_{x \in \mathbb{R}^{n-1}} \sup_{t' \subseteq \mathbb{R}} \left( \frac{1}{|t'|} \int_{t'} \int_{t'} |\phi(x, t) - \phi(x, s)|^{2} dt \, ds \right)^{1/2} \leq B
\]
(2.16)
is equivalent to $D\phi \in \text{BMO}(\mathbb{R}^n)$ with $B \sim \|D\phi\|_{\text{BMO}(\mathbb{R}^n)}$. This is does not appear to be correct. The paper [Riv03] does not give proof and provides instead a reference to [Str80] that is irrelevant for the claim. By [Str80] (2.16) is equivalent to $D^{1/2}_{t} \phi(x, \cdot) \in \text{BMO}(\mathbb{R})$ pointwise for a.e. $x$. After some tedious and technical calculations we were able to show $\sup D^{1/2}_{t} \phi(x, \cdot) \in \text{BMO}(\mathbb{R})$ implies $D^{1/2}_{t} \phi \in \text{BMO}(\mathbb{R}^n)$ and hence $D\phi \in \text{BMO}(\mathbb{R}^n)$ via condition (4) of Theorem 2.3. However, whether the converse is not clear even if we assume more structure for the function $\phi(x, t)$. This is due to the fact that there is “no reasonable Fubini theorem relating $\text{BMO}(\mathbb{R}^n)$ to $\text{BMO}(\mathbb{R})$” [Str80, p. 558].

Fortunately, the lack of a converse implication does not cast doubt over the subsequent results of [Riv03] since the author only uses the claimed equivalence in the correct direction — that (2.16) implies $D\phi \in \text{BMO}(\mathbb{R}^n)$.

**Localisation.** After the comprehensive review of the Lewis-Murray condition for a graph domain $\Omega$ we continue in our aim to construct a time-varying domain which is locally described by local graphs $\phi_j$.

For a vector $x \in \mathbb{R}^{n-1}$ we denote consider the norm $|x|_{\infty} = \sup_i |x_i|$.

Consider $\phi : Q_{bd} \to \mathbb{R}^{n-1} \times \mathbb{R}$. The localised version of (2.11) from Theorem 2.3 is simply

$$
\sup_{Q_r = I_r \times I_r, Q_r \subset Q_{bd}} \frac{1}{|Q_r|} \int_{I_r} \int_{I_r} |\phi(x, t) - \phi(x, s)|^2 \frac{dt}{s^2} dx < \infty. \quad (2.17)
$$

We write $\|f\|_{*,d}$ to be the BMO norm of $f$ where the supremum in the BMO norm, c.f. (2.1), is taken over all cubes $Q_r$ with $r \leq d$. For a function $f : J \times I \to \mathbb{R}$, where $J \subset \mathbb{R}^{n-1}$ and $I \subset \mathbb{R}$ are closed bounded cubes we consider the norm $\|f\|_{*,J \times I}$ defined as above where the supremum is taken over all parabolic cubes $Q_r$ contained in $J \times I$. The norm $\|f\|_{*,J \times I,d}$ is where the supremum is taken over all parabolic cubes $Q_r$ with $r \leq d$ contained in $J \times I$. If the context is clear we suppress the $J \times I$ and write $\|f\|_*$ or $\|f\|_{*,d}$.

Recall that $\text{VMO}(\mathbb{R}^n)$ is defined as the closure of all bounded uniformly continuous functions (which be denote by $C_{b,u}(\mathbb{R}^n)$) in the BMO norm or equivalently BMO functions $f$ such that $\|f\|_{*,d} \to 0$ as $d \to 0$. Alternatively, if we define

$$
d(f, \text{VMO}) := \inf_{h \in C_{b,u}(\mathbb{R}^n)} \|f - h\|_*
$$

then $f \in \text{VMO}$ if and only if $d(f, \text{VMO}) = 0$; for $f \in \text{BMO}$ this measures the distance of $f$ to VMO. In our case, the boundary of the parabolic domains we consider can be locally described as a graph of a continuous function. However, as our domain is unbounded in time, we may potentially require an infinite family of local graphs $\{\phi_i\}$. Therefore we need to measure the distance to VMO uniformly across this infinite family.

Let $\delta : \mathbb{R}_+ \to \mathbb{R}_+, \delta(0) = 0$ and $\delta$ be continuous at 0 then we define $C_\delta$ to be the set of bounded continuous functions with the same modulus of continuity $\delta$. That is

$$
C_\delta = \{g : \mathbb{R}^n \to \mathbb{R} : |g(x) - g(y)| \leq \delta(|x - y|) \text{ for all } x, y, \text{ & } g \text{ is bounded}\}. \quad (2.18)
$$

Note that every family of bounded equicontinuous functions is a subset of $C_\delta$ for some modulus of continuity $\delta$. Also $C_{b,u} = \bigcup \delta C_\delta$. For $f : Q_{bd} \to \mathbb{R}$ we define $d(f, C_\delta)$ as

$$
d(f, C_\delta) = \inf_{h \in C_\delta} \|f - h\|_{*,Q_{bd}}.
$$

We are now ready to state and prove the result on extensibility of $\phi : Q_{bd} \to \mathbb{R}$ to a global function.
Theorem 2.8. Let $\phi : Q_{sd} \subset \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R}$ be Lip$(1,1/2)$ with Lipschitz constant $\ell$. If there exist a scale $r_1$, a constant $\eta > 0$ and a modulus of continuity $\delta$ such that
\[
\sup_{Q_s = J_s \times I_s} \frac{1}{|Q_s|} \int_{Q_s} \int_{I_s} \frac{|\phi(x,t) - \phi(x,\tau)|^2}{|t - \tau|^2} \, dt \, dx \leq \eta^2
\tag{2.19}
\]
and
\[
d(\nabla \phi, C_8) \leq \eta
\tag{2.20}
\]
then there exists a scale $d' \leq d$, that only depends on $d$, $\delta$, $\eta$, and $r_1$ and not $\phi$, such that for all $Q_s \subset Q_{sd}$ with $r \leq d'$ there exists a global Lip$(1,1/2)$ function $\Phi : \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R}$ with the following properties for all $0 < \varepsilon < 1$:

(i) $\Phi|_{Q_r} = \phi|_{Q_r}$,

(ii) the Lip$(1,1/2)$ constant of $\Phi$ is $\ell$,

(iii) $\|\nabla \Phi\|_* \lesssim \eta^{1-\varepsilon} + \eta \ell$, and

(iv) $\sup_{Q_s = J_s \times I_s} \frac{1}{|Q_s|} \int_{Q_s} \int_{I_s} \frac{|\Phi(x,t) - \Phi(x,\tau)|^2}{|t - \tau|^2} \, dt \, dx \lesssim \eta^2$.

Therefore by Corollary 2.6, $\|D\Phi\|_* \lesssim \eta^{1-\varepsilon} + \eta \ell$.

We again give the proof of this result in the appendix. We are now ready to define the class of parabolic domains on which we will work. Motivated by the usual definition of a Lipschitz domain we have:

Definition 2.9. $Z \subset \mathbb{R}^n \times \mathbb{R}$ is an $\ell$-cylinder of diameter $d$ if there exists a coordinate system $(x_0, x, t) \in \mathbb{R} \times \mathbb{R}^{n-1} \times \mathbb{R}$ obtained from the original coordinate system by translation in spatial and time variables, and rotation only in the spatial variables such that
\[
Z = \{(x_0, x, t) : |x| \leq d, |t|^{1/2} \leq d, |x_0| \leq (\ell + 1)d\}
\]
and for $s > 0$
\[
sZ := \{(x_0, x, t) : |x| < sd, |t|^{1/2} \leq sd, |x_0| \leq (\ell + 1)sd\}.
\]

Definition 2.10. $\Omega \subset \mathbb{R}^n \times \mathbb{R}$ is an admissible parabolic domain with character $(\ell, \eta, N, d)$ if there exists a positive scale $r_1$, and a modulus of continuity $\delta$ such that for any time $\tau \in \mathbb{R}$ there are at most $N$ $\ell$-cylinders $\{Z_j\}_{j=1}^N$ of diameter $d$ satisfying the following conditions:

1. $\partial \Omega \cap \{|t - \tau| \leq d^2\} = \bigcup_j (Z_j \cap \partial \Omega)$.

2. In the coordinate system $(x_0, x, t)$ of the $\ell$-cylinder $Z_j$ the set $Z_j \cap \Omega \supset \{(x_0, x, t) \in \Omega : |x| < d, |t| < d^2, \delta(x_0, x, t) \leq d/2\}$.

3. $sZ_j \cap \partial \Omega$ is the graph $\{(x_0 = \phi_j(x, t))\}$ of a function $\phi_j : Q_{sd} \to \mathbb{R}$, with $Q_{sd} \subset \mathbb{R}^{n-1} \times \mathbb{R}$, such that
\[
|\phi_j(x, t) - \phi_j(y, s)| \leq \ell \left(|x - y| + |t - s|^{1/2}\right) \quad \text{and} \quad \phi_j(0, 0) = 0.
\tag{2.21}
\]

4. $d(\nabla \phi_j, C_8) \leq \eta$
\tag{2.22}
and
\[
\sup_{Q_s = J_s \times I_s} \frac{1}{|Q_s|} \int_{Q_s} \int_{I_s} \frac{|\phi_j(x, t) - \phi_j(x, \tau)|^2}{|t - \tau|^2} \, dt \, dx \leq \eta^2.
\tag{2.23}
\]

Here and throughout $\delta(x_0, x, t) := \text{dist}((x_0, x, t), \partial \Omega)$ and $\text{dist}$ is the parabolic distance $\text{dist}([X, t], [Y, s]) = |X - Y| + |t - s|^{1/2}$.

We say that $\Omega$ is of VMO type if $\eta$ in the character $\ell, \eta, N, d$ can be taken arbitrarily small (at the expense of a potentially smaller $d$ and $r_1$, and larger $N$).
Remark 2.11. When (2.22) holds for small or vanishing $\eta$ it follows that for a fixed time $\tau$ the normal $\nu$ to the fixed-time spatial domain $\Omega_\tau = \Omega \cap \{t = \tau\}$ can be written in local coordinates as
\[
\nu = \frac{1}{\|(-1, \nabla \phi_j)\|}(-1, \nabla \phi_j)
\]
and hence $d(\nu, \text{VMO}) \lesssim \eta$. Therefore $\Omega_\tau$ is similar to the domains considered in the papers [MMS09] and [HMT15] which have dealt with the elliptic problems on domains with normal in or near VMO.

Remark 2.12. It follows from this definition that for each $\tau \in \mathbb{R}$ the time-slice $\Omega_\tau$ of an admissible parabolic domain $\Omega \subset \mathbb{R}^n \times \mathbb{R}$ is a bounded Lipschitz domain in $\mathbb{R}^n$ and they all have a uniformly bounded diameter. That is
\[
\inf_{\tau \in \mathbb{R}} \text{diam}(\Omega_\tau) \sim d \sim \sup_{\tau \in \mathbb{R}} \text{diam}(\Omega_\tau),
\]
where $d$ is the scale from definition 2.10 and the implied constants only depends on $N$. In particular, if $\mathcal{O} \subset \mathbb{R}^n$ is a bounded Lipschitz domain then the parabolic cylinder $\Omega = \mathcal{O} \times \mathbb{R}$ is an example of a domain satisfying definition 2.10.

Definition 2.13. Let $\Omega \subset \mathbb{R}^n \times \mathbb{R}$ be an admissible parabolic domain with character $(\ell, \eta, N, d)$. The measure $\sigma$ defined on sets $A \subset \partial \Omega$ is
\[
\sigma(A) = \int_{-\infty}^{\infty} \mathscr{H}^{n-1}(A \cap \{(X, t) \in \partial \Omega\}) \, dt,
\]
where $\mathscr{H}^{n-1}$ is the $n-1$ dimensional Hausdorff measure on the Lipschitz boundary $\partial \Omega$.

We consider solvability of the $L^p$ Dirichlet boundary value problem with respect to this measure $\sigma$. The measure $\sigma$ may not be comparable to the usual surface measure on $\partial \Omega$: in the $t$-direction the functions $\phi_j$ from definition 2.10 are only $1/2$ Lipschitz and hence the standard surface measure might not be locally finite. Our definition assures that for any $A \subset 8Z_j$, where $Z_j$ is an $\ell$-cylinder, we have
\[
\mathscr{H}^n(A) \sim \sigma \left( \{(\phi_j(x, t), x, t) : (x, t) \in A\} \right),
\]
where the constants in (2.25), by which these measures are comparable, only depend on $\ell$ of the character $(\ell, \eta, N, d)$ of the domain $\Omega$. If $\Omega$ has a smoother boundary, such as Lipschitz (in all variables) or better, then the measure $\sigma$ is comparable to the usual $n$-dimensional Hausdorff measure $\mathscr{H}^n$. In particular, this holds for a parabolic cylinder $\Omega = \mathcal{O} \times \mathbb{R}$.

Corollary 2.14. Let $\Omega$ be defined as in definition 2.10 by a family of functions $\{\phi_j\}, \phi_j : Q_{\delta d} \to \mathbb{R}$. Then there exists an extended family $\{\Phi_j\}, \Phi_j : \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R}$, such that
\[
(i) \left\{ \Phi_j|_{Q_{\delta d}} \right\} \text{ still describes } \Omega, \text{ as in definition 2.10, but with character } (\ell, \eta, \bar{N}, r)
\]
\[
\text{ instead of } (\ell, \eta, N, d), \text{ where } \bar{N} \geq N \text{ and } r \leq r_1 \leq d \text{ an by Theorem 2.8;}
\]
\[
(ii) \|\nabla \Phi_j\|_\ast \lesssim \epsilon \eta^{1-\varepsilon} + nf, \text{ and}
\]
\[
(iii) \|D\Phi_j\|_\ast \lesssim \epsilon \eta^{1-\varepsilon} + nf.
\]
Here, $\bar{N}, r$ depend on the original character variables $\ell, \eta, N, d$, the modulus of continuity $\delta$ and the dimension $n$.

Proof. This follows from Theorem 2.8 and by tiling the support of each $\phi_j$ into parabolic cubes of size $8r$ with enough overlap.

Corollary 2.15. If $\Omega$ is a VMO type domain then we may take $\eta$ arbitrarily small in Corollary 2.14, or in (2.22) and (2.23) of definition 2.10, by reducing $r$. 

□
2.2. Pullback Transformation and Carleson Condition. We now briefly recall the pullback mapping of Dahlberg-Kenig-Nečas-Stein on the upper half-space $U \rho : U \to \Omega$ (c.f. [HL96, HL01]) in the setting of parabolic equations defined by

$$\rho(x_0, x, t) = (x_0 + P_{x_0} \phi(x, t), x, t). \quad (2.26)$$

For simplicity, assume

$$\Omega = \{(x_0, x, t) \in \mathbb{R} \times \mathbb{R}^{n-1} \times \mathbb{R} : x_0 > \phi(x, t)\} \quad (2.27)$$

where $\phi(x, t) : \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R}$ and satisfies conditions (3) and (4) of definition 2.10. This transformation maps the upper half-space $U = \{(x_0, x, t) : x_0 > 0, x \in \mathbb{R}^{n-1}, t \in \mathbb{R}\} \quad (2.28)$

into $\Omega$ and allows us to consider the $L^p$ solvability of the PDE (1.1) in the upper half-space instead of in the original domain $\Omega$.

To complete the definition of the mapping $\rho$ we define a parabolic approximation to the identity $P$ to be an even non-negative function $P(x, t) \in C^\infty_0(Q_1(0, 0))$, for $(x, t) \in \mathbb{R}^{n-1} \times \mathbb{R}$, with $\int P(x, t) \, dx \, dt = 1$ and set

$$P_\lambda(x) := \lambda^{-(n+1)} P \left( \frac{x}{\lambda^2} \right).$$

Let $P_\lambda \phi$ be the convolution operator

$$P_\lambda \phi(x, t) := \int_{\mathbb{R}^{n-1} \times \mathbb{R}} P_\lambda(x - y, t - s) \phi(y, s) \, dy \, ds$$

then $P$ satisfies for constants $\gamma$

$$\lim_{(y_0, y, s) \to (0, x, t)} P_\gamma y_0 \phi(y, s) = \phi(x, t)$$

and $\rho$ defined in (2.26) extends continuously to $\rho : \overline{U} \to \overline{\Omega}$. The usual surface measure on $\partial U$ is comparable with the measure $\sigma$ defined by (2.24) on $\partial \Omega$.

Suppose that $v = u \circ \rho$ and $f^v = f \circ \rho$ then (1.1) transforms to a new PDE for the variable $v$

\[
\begin{align*}
  v_t &= \text{div}(A^v \nabla v) + B^v \cdot \nabla v \quad \text{in } U, \\
  v &= f^v \quad \text{on } \partial U,
\end{align*}
\]

where $A^v = [a^v_{ij}(X, t)]$, $B^v = [b^v_i(X, t)]$ are $(n \times n)$ and $(1 \times n)$ matrices.

The precise relations between the original coefficients $A$ and $B$ and the new coefficients $A^v$ and $B^v$ are detailed in [Riv14, p. 448]. We note that if the constant $\gamma > 0$ is chosen small enough then the coefficients $a^v_{ij}, b^v_i : U \to \mathbb{R}$ are Lebesgue measurable and $A^v$ satisfies the standard uniform ellipticity condition with constants $\lambda^v$ and $\Lambda^v$, since the original matrix $A$ did.

**Definition 2.16.** Let $\Omega$ be a parabolic domain from definition 2.10. For $(Y, s) \in \partial \Omega$, $(X, t), (Z, \tau) \in \Omega$ and $r > 0$ we write:

$$B_r(X, t) = \{(Z, \tau) \in \mathbb{R}^n \times \mathbb{R} : \text{dist}([X, t], (Z, \tau]) < r\},$$

$$Q_r(X, t) = \{(Z, \tau) \in \mathbb{R}^n \times \mathbb{R} : |x_i - z_i| < r \text{ for all } 0 \leq i \leq n - 1, |t - \tau|^{1/2} < r\},$$

$$\Delta_r(Y, s) = \partial \Omega \cap B_r(Y, s), \quad T(\Delta_r) = \Omega \cap B_r(Y, s),$$

$$\delta(X, t) = \inf_{(Y, s) \in \partial \Omega} \text{dist}([X, t], (Y, s)].$$

**Definition 2.17** (Carleson measure). A measure $\mu : \Omega \to \mathbb{R}^+$ is a Carleson measure if there exists a constant $C = C(d)$ such that for all $r \leq d$ and all surface balls $\Delta_r$,

$$\mu(T(\Delta_r)) \leq C \sigma(\Delta_r). \quad (2.30)$$
The best possible constant $C$ is called the Carleson norm and is denoted by $\|\mu\|_{C,d}$. Occasionally, for brevity, we drop the $d$ and just write $\|\mu\|_C$ if the context is clear. We say that $\mu$ is a vanishing Carleson measure if $\|\mu\|_{C,d} \to 0$ as $d \to 0+$.

When $\partial \Omega$ is locally given as a graph of a function $x_0 = \phi(x,t)$ in the coordinate system $(x_0, x,t)$ and $\mu$ is a measure supported on $\{x_0 > \phi(x,t)\}$ we can reformulate the Carleson condition locally using the parabolic boundary cubes $Q_r$ and corresponding Carleson regions $T(Q_r)$. The Carleson condition (2.30) then becomes

$$\mu(T(Q_r)) \leq C|Q_r| = Cr^{n+1}. \quad (2.31)$$

Note that the Carleson norms induced from (2.30) and (2.31) are not equal but are comparable.

We now return to the pullback transformation and investigate the Carleson condition on the coefficients of $A$ and $B$. The following result comes directly from a careful reading of the proofs of Lemma 2.8 and Theorem 7.4 in [HL96] combined with Theorems 2.1 and 2.3.

**Lemma 2.18.** Let $\sigma$ and $\theta$ be non-negative integers, $\alpha = (\alpha_1, \ldots, \alpha_n - 1)$ a multi-index with $l = \sigma + |\alpha| + \theta$, $d$ a scale and fix $\gamma$. If $\phi: \mathbb{R}^{n+1} \times \mathbb{R} \to \mathbb{R}$ satisfies for all $x, y \in \mathbb{R}^{n+1}$, $t, s \in \mathbb{R}$ and for some positive constants $\ell$ and $\eta$

$$|\phi(x, t) - \phi(y, s)| \leq \ell \left( |x - y| + |t - s|^{1/2} \right),$$

$$\|D\phi\|_\sigma \leq \eta \quad (2.32)$$

then the measure $\nu$ defined at $(x_0, x, t)$ by

$$d\nu = \left( \frac{\partial \phi_{x_0} \phi}{\partial x^\alpha} \frac{\partial \phi_{x_0} \phi}{\partial \gamma} \right)^2 \frac{2^\ell}{\alpha + 1} dx dt dx_0$$

is a Carleson measure on cubes of diameter $\leq d/4$ whenever either $\sigma + \theta \geq 1$ or $|\alpha| \geq 2$, with

$$\nu \left[ (0, r) \times Q_r(x, t) \right] \lesssim \eta |Q_r(x, t)|,$$

where $r \leq d/4$. Moreover, if $l \geq 1$ then at $(x_0, x, t)$, with $x_0 \leq d/4$,

$$\left| \frac{\partial \phi_{x_0} \phi}{\partial x^\alpha} \frac{\partial \phi_{x_0} \phi}{\partial \gamma} \right| \lesssim \eta (1 + \ell) x_0^{1-l-\theta}, \quad (2.33)$$

where the implicit constants depend on $d, l, n$.

The drift term $B^\nu$ from the pullback transformation in (2.29) includes the term

$$\frac{\partial}{\partial t} \phi_{x_0} \phi u_{x_0}.$$

From Lemma 2.18 with $\sigma = |\alpha| = 0$, $\theta = 1$, we see that

$$x_0 \left[ \frac{\partial}{\partial t} \phi_{x_0} \phi(x_0, t) \right]^2 dX dt$$

is a Carleson measure in $U$. Thus it is natural to expect that

$$d\mu_1(X, t) = x_0 |B^\nu|^2(X, t) dX dt \quad (2.34)$$

is a Carleson measure in $U$ and $B^\nu$ satisfies

$$x_0 |B^\nu(X, t)| \leq \Lambda_B < \|\mu_1\|_C^{1/2}. \quad (2.35)$$

Indeed, this is the case provided the original vector $B$ satisfies the assumption that

$$d\mu(X, t) = \delta(X, t) \left[ \sup_{B_{4l(X, t)}/2(X, t)} \|B\| \right]^2 dX dt \quad (2.36)$$
is a Carleson measure in $\Omega$. Here $\|\mu_1\|_C$ depends on $\eta$ and the Carleson norm of (2.36).

Similarly, for the matrix $A^\nu$ if we apply Lemma 2.18 and use the calculations in [Riv14, §6] then

$$d\mu_2(X, t) = (x_0|\nabla A^\nu|^2 + x_0^2|A^\nu|^2)(X, t) \, dX \, dt$$

(2.37)

is a Carleson measure in $U$ and $A^\nu$ satisfies

$$(x_0|\nabla A^\nu| + x_0^2|A^\nu|)(X, t) \leq \|\mu_2\|_C^{1/2}$$

(2.38)

for almost everywhere $(X, t) \in U$ provided the original matrix $A$ satisfies that

$$d\mu(X, t) = \left( \frac{\delta(X, t)}{\sup_{B_{s(X, t)}(X, t)} |\nabla A|} \right)^2 + \frac{\delta(X, t)^3}{\sup_{B_{s(X, t)}(X, t)} |\partial_1 A|^2} \right)^2 \, dX \, dt$$

(2.39)

is a Carleson measure in $\Omega$.

We note that if both $\|\mu\|_{C, r}$ and $\eta$ are small then too are the Carleson norms $\|\mu_1\|_{C, r}$ and $\|\mu_2\|_{C, r}$ of the matrix $A^\nu$ and vector $B^\nu$, at least if we restrict ourselves to small Carleson regions $r \leq d$; this comes from Theorem 2.8 and Corollaries 2.14 and 2.15. Then by Lemma 2.18 we see that $\|\mu_1\|_{C, r}$ and $\|\mu_2\|_{C, r}$ only depend on $\eta$ and $\|\mu\|_{C, r}$ on Carleson regions of size $r \leq d$. In particular, they are small if both $\eta$ and $\|\mu\|_{C, r}$ are small. It further follows by Corollary 2.15 that we can make $\|\mu_1\|_{C, r}$ and $\|\mu_2\|_{C, r}$ as small as we like if $\mu$ is a vanishing Carleson norm and the domain $\Omega$ is of VMO type.

Observe that condition (2.39) is slightly stronger than (1.6), which we claimed to assume in Theorem 1.1. We replace condition (2.39) by the weaker condition (1.6) later via perturbation results of [Swe98].

**Definition 2.19.** We define $\rho_j : U \rightarrow \mathbb{R} Z_j$ to be the local pullback mapping in $8 Z_j$ associated to the function $\Phi_j$ in Theorem 2.8, the extension of $\phi_j$ from definition 2.10.

**Remark 2.20.** By [BZ17] and its adaptation to the setting of admissible domains in [DH18, §2.3], one may construct a ‘proper generalised distance’ globally when $\eta$ in the character of the domain is small. The smallness of $\eta$ in the character of the domain is used to guarantee that overlapping coordinate charts, generated by a local construction, are almost parallel. We may then use the result of [BZ17, Theorem 5.1] to show there exists a domain $\Omega^F$ of class $C^\infty$, a homeomorphism $f^\nu : \Omega \rightarrow \Omega^F$ such that $f^\nu(\partial \Omega) = \partial \Omega^F$ and $f^\nu : \Omega \rightarrow \Omega^F$ is a $C^\infty$ diffeomorphism.

2.3. Parabolic Non-tangential Cones, Maximal Functions and $p$-adapted Square and Area Functions. We proceed with the definition of parabolic non-tangential cones and define the cones in a (local) coordinate system where $\Omega = \{(x_0, x, t) : x_0 > \phi(x, t)\}$, which also applies to the upper half-space $U$.

**Definition 2.21.** For a constant $a > 0$, we define the parabolic non-tangential cone at a point $(x_0, x, t) \in \partial \Omega$ as follows

$$\Gamma_a(x_0, x, t) = \left\{ (y_0, y, s) \in \Omega : |y - x| + |s - t|^{1/2} \leq a(y_0 - x_0), \ x_0 < y_0 \right\}.$$ 

We occasionally truncate the cone $\Gamma$ at the height $r$

$$\Gamma_a^*(x_0, x, t) = \left\{ (y_0, y, s) \in \Omega : |y - x| + |s - t|^{1/2} < a(y_0 - x_0), \ x_0 < y_0 < x_0 + r \right\}.$$

**Definition 2.22** (Non-tangential maximal function). For a function $u : \Omega \rightarrow \mathbb{R}$, the non-tangential maximal function $N_a(u) : \partial \Omega \rightarrow \mathbb{R}$ and its truncated version at a
height $r$ are defined as
\begin{align*}
N_a(u)(x_0, x, t) &= \sup_{(y_0, y, s) \in \Gamma_a(x_0, x, t)} |u(y_0, y, s)|, \\
N'_a(u)(x_0, x, t) &= \sup_{(y_0, y, s) \in \Gamma'_a(x_0, x, t)} |u(y_0, y, s)| \quad \text{for} \ (x_0, x, t) \in \partial\Omega. \quad (2.40)
\end{align*}

The following $p$-adapted square function was introduced in [DPP07] and has been modified appropriately for the parabolic setting. It is used to control the spatial derivatives of the solution. When $p = 2$ it is equivalent to the usual square function and when $p < 2$ we use the convention that the expression $|\nabla u|^2|u|^{p-2}$ is zero whenever $\nabla u$ vanishes.

**Definition 2.23** ($p$-adapted square function). For a function $u : \Omega \to \mathbb{R}$, the $p$-adapted square function $S_{p,a}(u) : \partial\Omega \to \mathbb{R}$ and its truncated version at a height $r$ are defined as
\begin{align*}
S_{p,a}(u)(Y, s) &= \left( \int_{\Gamma_a(Y, s)} |\nabla u(X, t)|^2|u(X, t)|^{p-2}\delta(X, t)^{-n} \, dX \, dt \right)^{1/p}, \\
S'_{p,a}(u)(Y, s) &= \left( \int_{\Gamma'_a(Y, s)} |\nabla u(X, t)|^2|u(X, t)|^{p-2}\delta(X, t)^{-n} \, dX \, dt \right)^{1/p}. \quad (2.41)
\end{align*}

By applying Fubini we also have
\begin{align*}
\|S_{p,a}(u)\|_{L^p(\partial\Omega)}^p \sim \int_U |\nabla u|^2|u|^{p-2}x \, dx \, dt. \quad (2.42)
\end{align*}

It is not known a priori if these integrals are locally integrable even for $p > 2$. However, Theorem 4.1 shows that these expressions make sense and are finite for solutions to (1.1).

We also need a $p$-adapted version of an object called the area function which was introduced in [DH18] and is used to control the solution in the time variable. Again when $p = 2$ this is just the area function of [DH18].

**Definition 2.24** ($p$-adapted area function). For a function $u : \Omega \to \mathbb{R}$, the $p$-adapted area function $A_{p,a}(u) : \partial\Omega \to \mathbb{R}$ and its truncated version at a height $r$ are defined as
\begin{align*}
A_{p,a}(u)(Y, s) &= \left( \int_{\Gamma_a(Y, s)} |u_t|^2|u(X, t)|^{p-2}\delta(X, t)^{2-n} \, dX \, dt \right)^{1/p}, \\
A'_{p,a}(u)(Y, s) &= \left( \int_{\Gamma'_a(Y, s)} |u_t|^2|u(X, t)|^{p-2}\delta(X, t)^{2-n} \, dX \, dt \right)^{1/p}. \quad (2.43)
\end{align*}

Also by Fubini
\begin{align*}
\|A_{p,a}(u)\|_{L^p(\partial\Omega)}^p \sim \int_U |u_t|^2|u|^{p-2}x_0^2 \, dx_0 \, dt. \quad (2.44)
\end{align*}

As before, it is not known a priori if these expressions are finite for solutions to (1.1) but in Lemma 4.5 we establish control of $A_{p,a}$ by $S_{p,2a}$ and use the finiteness of $S_{p,a}$ from Theorem 4.1.

2.4. The $L^p$ Solvability of the Dirichlet Boundary Value Problem. We are now in the position to define the $L^p$ Dirichlet boundary value problem and our main results.
Definition 2.25 ([Aro68]). We say that \( u \) is a weak solution to a parabolic operator of the form (1.1) in \( \Omega \) if \( u, \nabla u \in L^2_{loc}(\Omega) \), \( \sup \|u(\cdot, t)\|L^2_{loc}(\Omega_t) < \infty \) and

\[
\int_\Omega (-u\phi_t + A\nabla u \cdot \nabla \phi - \phi B \cdot \nabla u) \, dX \, dt = 0
\]

for all \( \phi \in C^\infty_0(\Omega) \).

Definition 2.26. We say that the \( L^p \) Dirichlet problem with boundary data in \( L^p(\partial \Omega, d\sigma) \) is solvable if the unique solution \( u \) to (1.1) for any continuous boundary data \( f \) decaying to 0 as \( t \to \pm \infty \) satisfies the following non-tangential maximum function estimate

\[
\|N(u)\|_{L^p(\partial \Omega, d\sigma)} \lesssim \|f\|_{L^p(\partial \Omega, d\sigma)},
\]

with the implied constant depending only on the operator, \( n, p \), and \( \Omega \).

Remark 2.27. Since the space \( C_0(\Omega) \) is dense in \( L^p(\Omega) \) for \( p < \infty \) it follows that the solution operator \( f \mapsto u \) has a unique extension onto the whole space \( L^p(\Omega) \) with the bound (2.45) being satisfied for such \( u \). Hence we can assign to every boundary datum \( f \in L^p(\Omega) \) a unique solution \( u \) such that (2.45) holds.

3. Basic Results and Interior Estimates

In this section, we now recall some foundational estimates that will be used later. First, we state interior estimates of a weak solution of the parabolic operator

\[
u_t = \text{div} (A \nabla u) + B \cdot \nabla u.
\]

Definition 3.1 ([Aro68]). We say that \( u \) is a weak solution to a parabolic operator of the form (3.1) in \( \Omega \) if \( u, \nabla u \in L^2_{loc}(\Omega) \), \( \sup \|u(\cdot, t)\|L^2_{loc}(\Omega_t) < \infty \) and

\[
\int_\Omega (-u\phi_t + A\nabla u \cdot \nabla \phi - \phi B \cdot \nabla u) \, dX \, dt = 0
\]

for all \( \phi \in C^\infty_0(\Omega) \).

Lemma 3.2 (A Cacciopoli inequality, see [Aro68]). Let \( A \) and \( B \) satisfy (1.2) and (2.35) and suppose that \( u \) is a weak solution of (3.1) in \( Q_r(X,t) \) with \( 0 < r < \delta(X,t)/8 \). Then there exists a constant \( C = C(\Lambda, \Lambda, n) \) such that

\[
r^2 \left( \sup_{Q_{r/2}(X,t)} u \right)^2 \leq C \sup_{t-r^2 \leq s \leq t+r^2} \int_{Q_r(X,t) \cap \{t=s\}} u^2(Y,s) \, dY + C \int_{Q_r(X,t)} |\nabla u|^2 \, dY \, ds
\]

\[
\leq \frac{C^2}{r^2} \int_{Q_r(X,t)} u^2(Y,s) \, dY \, ds.
\]

Lemmas 3.4 and 3.5 in [HL01] give the following estimates for weak solutions of parabolic.

Lemma 3.3 (Interior Hölder continuity). Let \( A \) and \( B \) satisfy (1.2) and (2.35) and suppose that \( u \) is a weak solution of (3.1) in \( Q_r(X,t) \) with \( 0 < r < \delta(X,t)/8 \). Then for any \( (Y, s), (Z, \tau) \in Q_{2r}(X,t) \)

\[
|u(Y,s) - u(Z,\tau)| \leq C \left( \frac{|Y-Z| + |s-\tau|^{1/2}}{r} \right)^{\alpha} \sup_{Q_r(X,t)} |u|,
\]

where \( C = C(\Lambda, \Lambda, n), \alpha = \alpha(\Lambda, \Lambda, n), \) and \( 0 < \alpha < 1 \).

Lemma 3.4 (Harnack inequality). Let \( A \) and \( B \) satisfy (1.2) and (2.35) and suppose that \( u \) is a weak non-negative solution of (3.1) in \( Q_r(X,t) \), with \( 0 < r < \delta(X,t)/8 \).
Suppose that \((Y,s), (Z,\tau) \in Q_{2r}(X,t)\) then there exists \(C = C(\lambda, \Lambda, n)\) such that, for \(\tau < s,\)
\[
u(Z,\tau) \leq u(Y,s) \exp \left[ C \left( \frac{|Y-Z|^2}{|s-\tau|} + 1 \right) \right].
\]

We state a version of the maximum principle from [DH18] that is a modification of [HL01, Lemma 3.38].

**Lemma 3.5 (Maximum Principle).** Let \(A\) and \(B\) satisfy (1.2) and (2.35), and let \(u\) and \(v\) be bounded continuous weak solutions to (3.1) in \(\Omega\). If \(|u|, |v| \to 0\) uniformly as \(t \to -\infty\) and
\[
\limsup_{(Y,s) \to (X,t)} (u-v)(Y,s) \leq 0
\]
for all \((X,t) \in \partial\Omega\), then \(u \leq v\) in \(\Omega\).

**Remark 3.6 ([DH18]).** The proof of Lemma 3.38 from [HL01] works given the assumption that \(|u|, |v| \to 0\) uniformly as \(t \to -\infty\). Even with this additional assumption, the lemma as stated is sufficient for our purposes. We shall mostly use it when \(u \leq v\) on the boundary of \(\Omega\) \(\{t \geq \tau\}\) for a given time \(\tau\). Obviously then the assumption that \(|u|, |v| \to 0\) uniformly as \(t \to -\infty\) is not necessary. Another case when the lemma as stated here applies is when \(u|_{\partial\Omega}, v|_{\partial\Omega} \in C_0(\partial\Omega)\), where \(C_0(\partial\Omega)\) denotes the class of continuous functions decaying to zero as \(t \to \pm\infty\). This class is dense in any \(L^p(\partial\Omega, dr)\), \(1 < p < \infty\) allowing us to consider an extension of the solution operator from \(C_0(\partial\Omega)\) to \(L^p\).

The following result is from [DH18], which was adapted from the elliptic result in [Din02].

**Lemma 3.7.** Let \(r > 0\) and \(0 < a < b\). Consider the non-tangential maximal functions defined using two set of cones cones \(\Gamma^r_a\) and \(\Gamma^r_b\). Then for any \(p > 0\) there exists a constant \(C_p > 0\) such that for all \(u : U \to \mathbb{R}\)
\[
N_p^r(u) \leq N_p^r(u) \quad \text{and} \quad \|N_p^r(u)\|_{L^p(\partial U)} \leq C_p\|N_p^r(u)\|_{L^p(\partial U)}.
\]

## 4. Improved Regularity for \(p\)-adapted square function

Here we extend recent work of [DP16] for complex coefficient elliptic equations to the real parabolic setting. The goal is to obtain an improved regularity result for weak solutions of (1.1) implying that \(|\nabla u|^2|u|^{p-2}\) belongs to \(L^1_{\text{loc}}(\Omega)\) when \(1 < p < 2\). Having this it follows that the \(p\)-adapted square function \(S_{p,a}\) is well defined at almost every boundary point.

**Theorem 4.1 (c.f. [DP16, Theorem 1.1]).** Suppose \(u \in W^{1,2}_{\text{loc}}(\Omega)\) is a weak solution to \(\mathcal{L}u = u_t\), where \(\mathcal{L}u = \text{div}(A\nabla u) + B\nabla u\), \(A\) is bounded and elliptic, \(B\) is locally bounded and satisfies
\[
\delta(X,t)|B(X,t)| \leq K
\]
for some uniform constant \(K > 0\). Then for any parabolic ball \(B_{2r}(X,t) \subset \Omega\) and \(p,q \in (1,\infty)\) we have the following improvement in regularity
\[
\left( \frac{1}{r^2} \int_{B_{r}(X,t)} |u|^p \right)^{1/p} \leq C_\varepsilon \left( \frac{1}{q} \int_{B_{2r}(X,t)} |u|^q \right)^{1/q} + \varepsilon \left( \frac{1}{r^2} \int_{B_{2r}(X,t)} |u|^2 \right)^{1/2}. \tag{4.2}
\]

Here the constant \(C_\varepsilon\) only depends on \(p, q, \varepsilon, n, \lambda, \Lambda, \) and \(K\) but not on \(u, (X,t)\) or \(r\). In addition, for all \(1 < p < \infty\)
\[
r^2 \int_{B_{r}(X,t)} |\nabla u|^2|u|^{p-2} \leq C_\varepsilon \int_{B_{2r}(X,t)} |u|^p + \varepsilon \left( \frac{1}{r^2} \int_{B_{2r}(X,t)} |u|^2 \right)^{p/2}. \tag{4.3}
\]
where again the constant only depends on ε, p, n, the ellipticity constants of A, and K. This also shows that |u|^{(p-2)/2}∇u ∈ L^2_{loc}(Ω).

**Remark 4.2.** If q ≥ 2 in (4.2) or if p ≥ 2 in (4.3) then one can take ε = 0 because the L^2 averages can be controlled by the first term on the right hand side of these inequalities.

We focus only on the case 1 < p < 2 as the p ≥ 2 result above follows from the Cacciopoli inequality, Lemma 3.2. We shall establish the following lemma for the 1 < p < 2 case which concludes the proof of Theorem 4.1.

**Lemma 4.3** (c.f. [DP16, Lemma 2.7]). Let u be a weak solution to Lu = u_t in Ω for A elliptic and bounded, and B bounded satisfying (4.1). Then for any p < 2, any ball B_r(X,t) with r < δ(X,t)/4, and any ε > 0

\[ p^2 \int_{B_r(X,t)} |\nabla u|^2 |u|^{p-2} \leq C_\varepsilon \int_{B_{2r}(X,t)} |u|^p + \varepsilon \left( \int_{B_{2r}(X,t)} |u|^2 \right)^{p/2} \quad (4.4) \]

and

\[ \left( \int_{B_r(X,t)} |u|^2 \right)^{1/2} \leq C_\varepsilon \left( \int_{B_{2r}(X,t)} |u|^p \right)^{1/p} + \varepsilon \left( \int_{B_{2r}(X,t)} |u|^2 \right)^{1/2} \],\quad (4.5) \]

where the constants only depend on n, ε, λ, Λ, and K. In particular, |u|^{(p-2)/2}∇u ∈ L^2_{loc}(Ω).

**Proof.** We start by assuming that A and B are smooth then the solution u to Lu = u_t is smooth. We prove the above inequalities with constants that do not depend on the smoothness of A or B. It follows then that the smoothness assumption can be removed by a limiting argument; that is A and B are approximated by sequences of smooth functions for which (4.4) and (4.5) hold uniformly. This is done in detail in the elliptic setting in [DP16, Lemma 2.7] and a similar argument in the parabolic case is shown in [HL01]. We skip further details as the argument is fairly standard.

To simplify notation, we suppress the argument of the ball B_r(X,t). Let

\[ ρ_δ(s) = \begin{cases} \delta^{(p-2)/2} & 0 \leq s \leq \delta \\ \sigma^{(p-2)/2} & s > \delta. \end{cases} \quad (4.6) \]

The choice of cut off function ρ_δ in this proof is inspired by [Lan99, p. 311], [CM05, p. 1088]. We multiply Lu = u_t by ρ_δ^2(|u|)u and integrate by parts to obtain

\[ \int_{B_r} \nabla (ρ_δ^2(|u|)u) A∇u = \int_{B_r} ρ_δ^2(|u|)uu_t + \int_{B_r} ρ_δ^2(|u|)B · ∇u \]

\[ + \int_{B_r} (ρ_δ^2(|u|))ν · A∇u \mathrm{d}σ(y,s), \quad (4.7) \]

where ν is the outer unit normal to B_r. Consider E_δ = \{u > \delta\} then the left hand side of (4.7) is

\[ \int_{B_r} \nabla (ρ_δ^2(|u|)u) A∇u = \delta^{p-2} \int_{B_r \setminus E_δ} ∇u · A∇u + \int_{B_r \cap E_δ} A∇u · ∇(|u|^{p-2}u) \quad (4.8) \]

and by ellipticity of A on the open set B_r \cap E_δ we have for some λ' > 0

\[ λ' \int_{B_r \cap E_δ} |u|^{p-2} |∇u|^2 \leq \int_{B_r \cap E_δ} A∇u · ∇(|u|^{p-2}u). \quad (4.9) \]

Our strategy is to let δ → 0 and show all the integrals involving B_r \setminus E_δ tend to 0.
First, we use the following result from [Lan99]. They proved if \( u \in C^2(B_r) \) and \( u = 0 \) on \( \partial B_r \) then for \( q > -1 \)
\[
\lim_{\delta \to 0} \delta^q \int_{B_r \setminus E_\delta} |\nabla u|^2 = 0. \tag{4.10}
\]

To deal with the boundary integral in (4.7) we note that equations (4.7) to (4.9) remain valid for any enlarged ball \( B_{ar} \) for \( 1 \leq \alpha \leq 5/4 \). We write (4.7) for every \( B_{ar} \) and then average in \( \alpha \) over the interval \([1, 5/4]\). The last term in (4.7) then turns into a solid integral over \( B_{5r/4} \setminus B_r \). Therefore,
\[
\lambda' \int_{B_r \cap E_\delta} |u|^{p-2} |\nabla u|^2 \leq \sup_{\alpha \in [1, 5/4]} \int_{B_{ar}} \rho_\delta^2 (|u|) uu_t + \sup_{\alpha \in [1, 5/4]} \int_{B_{ar}} \rho_\delta^2 (|u|) uB \cdot \nabla u + r^{-1} \int_{B_{5r/4} \setminus B_r} \left( |u| \right) \left( uu + A \nabla u \right) + o(1)
\]
where \( o(1) \) contains the integral over \( B_{ar} \setminus E_\delta \), which tends to 0 as \( \delta \to 0 \). We bound \( II \) and \( III \) as [DP16]
\[
II + III \leq C_\varepsilon r^{-2} \int_{B_{5r/4}} |u|^p + \varepsilon r^{-2} \int_{B_{5r/4}} |\nabla u|^p + o(1).
\]

Now we turn to \( I \) and use the same idea as the proof of (4.10) in [Lan99, (3.3)] to show \( I \) converges as expected. By splitting the integral with the set \( E_\delta \), using the fact \( \delta^{p-2} \leq |u|^{p-2} \) on \( B_{ar} \setminus E_\delta \) (since \( p < 2 \)), and the smoothness of \( u \), which implies \( |u|^{p-2} uu_t, \in L^1(B_{ar}) \), we obtain
\[
\int_{B_{ar}} \rho_\delta^2 (|u|) uu_t = \int_{B_{ar} \cap E_\delta} |u|^{p-2} uu_t + \delta^{p-2} \int_{B_{ar} \setminus E_\delta} uu_t \\
\leq \int_{B_{ar} \cap E_\delta} |u|^{p-2} uu_t + \int_{B_{ar} \setminus E_\delta} |u|^{p-2} uu_t \\
\leq \int_{B_{ar}} |u|^{p-1} |u_t| < \infty.
\]
Therefore by the dominated convergence theorem
\[
\int_{B_{ar}} \rho_\delta^2 (|u|) uu_t \to \int_{B_{ar}} |u|^{p-2} uu_t. \tag{4.11}
\]

We change from working with balls to integrating over parabolic cubes \( Q_{ar} \) and denote by \( Q_{ar} |_{t = s} \) the cube \( Q_{ar} \) restricted to the hypersurface \( \{ t = s \} \). Using the fundamental theorem of calculus, we obtain in the limit that
\[
\int_{B_{ar}} |u|^{p-2} uu_t \sim \int_{B_{ar}} \frac{\partial}{\partial t} (|u|^p) \, dt \, dX \\
\leq \int_{Q_{ar}} \frac{\partial}{\partial t} (|u|^p) \, dt \, dX = \int_{t_0 + (ar)^2}^{t_0} \frac{d}{dt} \int_{Q_{ar}|_{t=s}} |u|^p \, dX \, ds \tag{4.12}
\]
\[
\leq ||u||_{L^\infty_x (Q_{ar}|_{t_0 + (ar)^2})}^p + ||u||_{L^p_x (Q_{ar}|_{t_0 - (ar)^2})}^p
\]
Observe that (4.12) holds for all time restricted cubes \( Q_{ar}|_{t_0 \pm (ar)^2} \) with \( \alpha \in [1, 1.1] \). Once again we average over these cubes to show
\[
\int_{B_{ar}} |u|^{p-2} uu_t \leq \frac{1}{r^2} \int_{Q_{1, ar}} |u|^p \, dX \, dt.
\]
Since \( Q_{1,10r} \subset B_{2r} \), in the limit as \( \delta \to 0 \)
\[
I \lesssim \frac{1}{r^2} \int_{B_{2r}} |u|^p \, dX \, dt.
\]

Therefore grouping the estimates we have the following bound
\[
\lambda' \int_{B_{2r}} |u|^p \frac{\nabla u^2}{\delta^2} \lesssim C r^{-2} \int_{B_{2r}} |u|^p + \varepsilon r^{p-2} \int_{B_{3r/4}} |\nabla u|^p + o(1). \tag{4.13}
\]
We let \( \delta \to 0 \) and proceed as [DP16] to obtain (4.4) and (4.5) for smooth \( A \) and \( B \).
Finally, since no constants depend on the smoothness of \( A \) or \( B \), we can remove the smoothness assumption by the same argument as in [HL01]. We suppose \( A \) is just elliptic and bounded, and \( B \) satisfies (1.1) then we approximate \( A \) and \( B \) by smooth matrices and vectors respectively. For each smooth approximation, we have (4.4) and (4.5) and then passing to the limit we obtain analogous estimates for \( W_{1,2}^{1,2} \) solutions \( u \) of \( Lu = u \) with the constants having the same dependence as before.

It follows that the \( p \)-adapted square function \( S_{p,a} \) is well defined. [DH18] also considered an area function and established [DH18, Lemma 5.2] that the usual square function can control this area function. The case \( 1 < p < 2 \) is significantly more complicated so for this reason we focus only on non-negative solutions \( u \).

We fix a boundary point \( (Y, s) \in \partial \Omega \) and consider \( A_{p,a}(Y, s) \). Clearly, the non-tangential cone \( \Gamma_a(Y, s) \) can be covered by a non-overlapping collection of Whitney cubes \( \{ Q_i \} \) with the following properties:
\[
\Gamma_a(Y, s) \subset \bigcup_i Q_i \subset \Gamma_{2a}(Y, s), \quad r_i := \text{diam}(Q_i) \sim \text{dist}(Q_i, \partial \Omega), \quad 4Q_i \subset \Omega, \tag{4.14}
\]
and the cubes \( \{ 2Q_i \} \) having only finite overlap. It follows that
\[
[A_{p,a}(Y, s)]^p \lesssim \sum_i (r_i)^{-2n} \int_{Q_i} |u|^p u^{p-2} \, dX \, dt \tag{4.15}
\]
\[
\lesssim \sum_i (r_i)^{-2n} \int_{Q_i} |\nabla u^2|^p u^{p-2} + (|A|^2 + |B|^2) |\nabla u|^2 u^{p-2} \, dX \, dt.
\]

We need the following estimate on each \( Q_i \).

**Lemma 4.4.** Assume the ellipticity condition (1.2) and that the coefficients \( A \) and \( B \) of (1.1) satisfy the conditions
\[
|\nabla A(X, t)| \leq K/\delta(X, t) \quad \text{and} \quad |B(X, t)| \leq K/\delta(X, t),
\]
for some uniform constant \( K > 0 \). Then for all non-negative solutions \( u \) of (1.1) and any parabolic cube \( Q \) such that \( 4Q \subset \Omega \) we have the following estimate
\[
\int_Q |\nabla u|^2 u^{p-2} \, dX \, dt \lesssim r^{-2} \int_{2Q} |\nabla u|^2 u^{p-2} \, dX \, dt, \tag{4.16}
\]
where \( r = \text{diam}(Q) \).

**Proof.** Since we assume differentiability of the matrix \( A \) in the spatial variables, we may also assume that \( A \) is symmetric. Let us denote by \( W = (w_k) \), where \( w_k = \delta_k u \) for \( k = 0, 1, \ldots, n - 1 \). Differentiating (1.1) we obtain the following PDE for each \( w_k \)
\[
(w_k)_t - \text{div}(A \nabla w_k) = \text{div}((\delta_k A) W + \delta_k (B \cdot W)) \tag{4.17}
\]
We multiply (4.17) by \( w_k u^{p-2} \zeta^2 \), integrate over \( 2Q \) and integrate by parts. Here \( 0 \leq \zeta \leq 1 \) is a smooth cut-off function equal to 1 on \( Q \), vanishing outside \( 2Q \) and
satisfying \( r|\nabla \zeta| + r^2|\zeta_t| \leq C \) for some \( C > 0 \) independent of \( Q \). This gives
\[
\int_{2Q} (w_k) w_k u^{p^2-2} \zeta^2 \, dX \, dt + \int_{2Q} a_{ij} (\partial_j w_k) \partial_i (w_k u^{p^2-2} \zeta^2) \, dX \, dt
\]
\[=- \int_{2Q} (\partial_i a_{ij}) w_j \partial_i (w_k u^{p^2-2} \zeta^2) \, dX \, dt - \int_{2Q} b_i w_i \partial_k (w_k u^{p^2-2} \zeta^2) \, dX \, dt. \tag{4.18}
\]
We rearrange and group similar terms together
\[
\frac{1}{2} \int_{2Q} \left[ (w_k u^{p^2-1})^2 \right] \, dX \, dt - \frac{p}{2} \int_{2Q} w_k^2 u^{p^2-3} u_t \zeta^2 \, dX \, dt
\]
\[+ \int_{2Q} A \left( \nabla (w_k \zeta) u^{p^2-1} \right) \cdot \left( \nabla (w_k \zeta) u^{p^2-1} \right) \, dX \, dt
\]
\[+ (p - 2) \int_{2Q} A \left( \nabla (w_k \zeta) u^{p^2-1} \right) \cdot \left( (\nabla u) w_k u^{p^2-2} \zeta \right) \, dX \, dt
\]
\[= \int_{2Q} |w_k|^2 u^{p^2-2} \zeta^2 \, dX \, dt + \int_{2Q} |w_k|^2 u^{p^2-2} A \nabla \zeta \cdot \nabla \zeta \, dX \, dt
\]
\[- \int_{2Q} b_i w_i \partial_k (w_k \zeta) u^{p^2-2} \zeta \, dX \, dt
\]
\[- (p - 2) \int_{2Q} b_i w_i \left( \partial_k w_k u^{p^2-2} \zeta \right) u^{p^2-1} \zeta \, dX \, dt
\]
\[- \int_{2Q} b_i w_i w_k u^{p^2-2} \zeta_k \, dX \, dt - \int_{2Q} (\partial_k a_{ij}) w_j w_k u^{p^2-2} \zeta_i \, dX \, dt
\]
\[- \int_{2Q} (\partial_k a_{ij}) w_j (\partial_k w_k \zeta) u^{p^2-2} \zeta \, dX \, dt
\]
\[- (p - 2) \int_{2Q} (\partial_k a_{ij}) w_j \left( \partial_l (w_k \zeta) \right) u^{p^2-2} \zeta \, dX \, dt. \tag{4.19}
\]
All the terms after the equal sign are “error” terms since they either contain a derivative of \( \zeta \), or coefficients \( \nabla A \) or \( B \). These will be handled using the Cauchy-Schwarz inequality and the estimates for \( |\nabla A|, |B| \leq K/r \). The four main terms are on the left-hand side of (4.19). The term that needs further work is the second term, and we use the PDE (1.1) for \( u_t \). This gives
\[
- \frac{p}{2} \int_{2Q} w_k^2 u^{p^2-3} u_t \zeta^2 \, dX \, dt = - \frac{p}{2} \int_{2Q} w_k^2 u^{p^2-3} \text{div}(A \nabla u) \zeta^2 \, dX \, dt
\]
\[- \frac{p}{2} \int_{2Q} w_k^2 u^{p^2-3} B \cdot W \zeta^2 \, dX \, dt. \tag{4.20}
\]
Again the second term will be an “error” term. For the first term, we observe the equality
\[u^{p^2-3} \text{div}(A \nabla u) = \text{div}(A(\nabla u) u^{p^2-3}) - (p - 3) A((\nabla u) u^{p^2-2}) \cdot ((\nabla u) u^{p^2-2}).\]

It follows (by integrating by parts)
\[- \frac{p}{2} \int_{2Q} w_k^2 u^{p^2-3} \text{div}(A \nabla u) \zeta^2 \, dX \, dt
\]\[- \frac{(p - 2)}{2} \int_{2Q} A((\nabla u) w_k u^{p^2-2} \zeta) \cdot ((\nabla u) w_k u^{p^2-2} \zeta) \, dX \, dt. \tag{4.21}
\]
We now group all main terms together; these are the first, second and fourth terms on the left-hand side of (4.19) and the terms of (4.21). This gives

\[
\text{LHS of (4.19)} = \frac{1}{2} \int_{2Q} \left[ \left( w_k u^{p/2-1} \right)^2 \right]_t \, dX \, dt \\
+ \int_{2Q} A \left( \nabla (w_k \zeta) u^{p/2-1} \right) \cdot \left( \nabla (w_k \zeta) u^{p/2-1} \right) \, dX \, dt \\
+ 2(p - 2) \int_{2Q} A \left( \nabla (w_k \zeta) u^{p/2-1} \right) \cdot \left( \nabla u \right) w_k u^{p/2-2} \zeta \, dX \, dt \\
+ \frac{(2 - p)(3 - p)}{2} \int_{2Q} A \left( \nabla u \right) w_k u^{p/2-2} \zeta \cdot \left( \nabla u \right) w_k u^{p/2-2} \zeta \, dX \, dt \\
= \frac{1}{2} \int_{2Q} \left[ \left( w_k u^{p/2-1} \right)^2 \right]_t \, dX \, dt \\
+ \left( 1 - \frac{2(2 - p)}{3 - p} \right) \int_{2Q} A \left( \nabla (w_k \zeta) u^{p/2-1} \right) \cdot \left( \nabla (w_k \zeta) u^{p/2-1} \right) \, dX \, dt \\
+ \int_{2Q} A \left( \sqrt{\frac{2(2 - p)}{3 - p}} \left[ \nabla (w_k \zeta) u^{p/2-1} \right] - \sqrt{\frac{(2 - p)(3 - p)}{2}} \left[ \nabla u \right] w_k u^{p/2-2} \zeta \right) \\
\cdot \left( \sqrt{\frac{2(2 - p)}{3 - p}} \left[ \nabla (w_k \zeta) u^{p/2-1} \right] - \sqrt{\frac{(2 - p)(3 - p)}{2}} \left[ \nabla u \right] w_k u^{p/2-2} \zeta \right) \, dX \, dt \\
\geq \frac{1}{2} \int_{2Q} \left[ \left( w_k u^{p/2-1} \right)^2 \right]_t \, dX \, dt + \frac{(p - 1)\lambda}{3 - p} \int_{2Q} \left| \nabla (w_k \zeta) u^{p/2-1} \right|^2 \, dX \, dt. \tag{4.22}
\]

Here we have first completed the square (using the symmetry of \( A \)), and then used the ellipticity of the matrix \( A \). The important point is that for all \( 1 < p < 2 \) the coefficient \( \frac{(p - 1)\lambda}{3 - p} \) is positive.

We also note that we could have completed the square differently and, instead of (4.22), obtain the following estimate:

\[
\text{LHS of (4.19)} \geq \frac{1}{2} \int_{2Q} \left[ \left( w_k u^{p/2-1} \right)^2 \right]_t \, dX \, dt \\
+ \frac{(p - 1)(2 - p)\lambda}{2} \int_{2Q} \left| \nabla (w_k u^{p/2-2} \zeta) \right|^2 \, dX \, dt. \tag{4.23}
\]

It follows that we could average (4.22) and (4.23) and have both

\[
\int_{2Q} \left| \nabla (w_k \zeta) u^{p/2-1} \right|^2 \, dX \, dt \quad \text{and} \quad \int_{2Q} \left| \nabla (w_k u^{p/2-2} \zeta) \right|^2 \, dX \, dt
\]
in the estimate with small positive constants.

Now we briefly mention how all the error terms of (4.19), (4.20) and (4.22) can be handled. Some can be immediately estimated from above by

\[
r^{-2} \int_{2Q} |W|^2 u^{p-2} \, dX \, dt,
\]

where the scaling factor \( r^{-2} \) comes from the estimates on \( \nabla \zeta, \zeta_t, |\nabla A| \) and \( |\mathcal{B}| \). For other terms (for example the first term of fifth line of (4.19) or the second term of the same line) we use Cauchy-Schwarz. One of the terms in the product will be

\[
\left( r^{-2} \int_{2Q} |W|^2 u^{p-2} \, dX \, dt \right)^{1/2},
\]

while the other term is one of

\[
\left( \int_{2Q} \left| \nabla (w_k \zeta) u^{p/2-1} \right|^2 \, dX \, dt \right)^{1/2} \quad \text{or} \quad \left( \int_{2Q} \left| \nabla (w_k u^{p/2-2} \zeta) \right|^2 \, dX \, dt \right)^{1/2}.
\]
It follows using the $\varepsilon$-Cauchy-Schwarz inequality that we can hide these on the left-hand side of (4.19). Finally, we put everything together by summing over all $k$ and recalling that $W = \nabla u$. This gives for some constant $c = c(p, \lambda, n) > 0$ with $\varepsilon \to 0$ as $p \to 1$.

\[
\sup_{\tau} \int_{Q \cap \{t = \tau\}} |\nabla u|^2 u^{p-2} \, dx + \varepsilon \int_{Q} |\nabla^2 u|^2 u^{p-2} \, dx + \varepsilon \int_{Q} |\nabla u|^4 u^{p-4} \, dx \, dt \\
\leq C r^{-2} \int_{2Q} |\nabla u|^2 u^{p-2} \, dx. \tag{4.24}
\]

In particular (4.16) holds. \qed

After using (4.16) in (4.15) we can conclude the following.

**Lemma 4.5.** Let $u$ be a non-negative solution of (1.1) with matrix $A$ satisfying the ellipticity hypothesis and the coefficients satisfying the bound $|\nabla A|, |B| \leq K/\delta$. Then given $\alpha > 0$ there exists a constant $C = (\Lambda, \lambda, \alpha, K, p, n)$ such that

\[
A_{p,a}(u)(X, t) \leq C S_{p,2a}(u)(X, t). \tag{4.25}
\]

From this we have the global estimate

\[
\|A_{p,a}(u)\|_{L^p(\partial\Omega)}^p \leq C_2 \|S_{p,a}(u)\|_{L^p(\partial\Omega)}^p. \tag{4.26}
\]

As far as the proof goes, the calculations above clearly work for solutions $u$ with a uniform bound $u \geq \varepsilon > 0$. Hence considering $v_\varepsilon = u + \varepsilon$ and then taking the limit $\varepsilon \to 0^+$ using Fatou’s lemma yields (4.25) for all non-negative $u$, where we have used the convention that $|\nabla u|^2 u^{p-2} = 0$ whenever $u = 0$ and $\nabla u = 0$ with a similar convention for the second gradient in $A_{p,a}$.

5. **Bounding the $p$-adapted square function by the non-tangential maximum function**

We slightly abuse notation and only work on a Carleson region $T(\Delta_r)$ in the upper half space $\mathcal{U}$ even though we formulate the following lemmas on any admissible domain $\Omega$. The equivalence of these formulations via the pullback map $\rho$ is discussed in section 2.2 and [DH18], and hence we omit the details. We start with a local bound of the $p$-adapted square function by the non-tangential maximal function.

**Lemma 5.1.** Let $\Omega$ be an admissible domain from definition 2.10 with character $(\ell, \eta, N, d)$. Let $1 < p < 2$ and $u$ be a non-negative solution of (1.1), with the Carleson conditions (1.7) and (1.8) on the coefficients $A$ and $B$. Then there exists a constant $C = (\Lambda, \lambda, N, C_0)$ such that for any solution $u$ with boundary data $f$ on any ball $\Delta_r \subset \partial\Omega$ with $r \leq \min\{d/4, d/(4C_0)\}$ we have

\[
\int_{T(\Delta_r)} |\nabla u|^2 |u|^{p-2} x_0 \, dx_0 \, dx \leq C(1 + \|f\|)(1 + \ell^2) \int_{\partial\Omega} (N^{2r})^2 (u) \, dx_0. \tag{5.1}
\]

In addition, we have the following global result.

**Lemma 5.2.** Let $\Omega$ be an admissible domain with smooth boundary $\partial\Omega$. Let $1 < p < 2$ and $u$ be a weak non-negative solution of (1.1) satisfying (2.34), (2.35), (2.37) and (2.38) with Dirichlet boundary data $f \in L^p(\partial\Omega)$. Then there exists positive constants $C_1$ and $C_2$ independent of $u$ such that for small $r_0 > 0$ we have

\[
C_1 \int_0^{r_0/2} \int_{\partial\Omega} |\nabla u|^2 |u|^{p-2} x_0 \, dx_0 \, dt \, dx_0 + \frac{2}{r_0} \int_0^{r_0} \int_{\partial\Omega} u^p(x_0, x, t) \, dx_0 \, dx_0 \\
\leq \int_{\partial\Omega} u^p(r_0, x, t) \, dx_0 \, dt + \int_{\partial\Omega} u^p(0, x, t) \, dx_0 \, dt \\
+ C_2 \left( \|\mu_1\|_{C, 2r} + \|\mu_2\|_{C, 2r} + \|\mu_2\|_{C, 2r}^{1/2} \right) \int_{\partial\Omega} (N^{2r}(u))^p \, dx_0. \tag{5.2}
\]
Proof of Lemmas 5.1 and 5.2. Let \( Q_r(y, s) \) be a parabolic cube on the boundary with \( r < d \) and let \( \zeta \) be a smooth cut off function independent of the \( x_0 \) variable. As long as there is no ambiguity we suppress the argument of \( Q_r \) and extensively use the Einstein summation convention. Let \( \zeta \) be supported in \( Q_{2r} \), equal 1 in \( Q_r \) and satisfy the estimate \( r|\nabla \zeta| + r^2|\zeta| \leq C \) for some constant \( C \) and.

We start by estimating

\[
\int_0^r \int_{Q_{2r}} |u|^{p-2} \frac{a_{ij}}{a_{00}} (\partial_i u)(\partial_j u)\zeta^2 x_0 \, dx \, dt \, dx_0,
\]

where by ellipticity we have

\[
\frac{\lambda}{X} \int_0^r \int_{Q_r} |\nabla u|^2 |u|^{p-2} x_0 \, dx \, dt \, dx_0 \leq \int_0^r \int_{Q_{2r}} |u|^{p-2} \frac{a_{ij}}{a_{00}} (\partial_i u)(\partial_j u)\zeta^2 x_0 \, dx \, dt \, dx_0.
\]

Now we integrate by parts whilst noting that \( \nu = (1, 0, 0, \ldots, 0) \) since the domain is \( \{x_0 > 0\} \).

\[
\int_0^r \int_{Q_{2r}} |u|^{p-2} \frac{a_{ij}}{a_{00}} (\partial_i u)(\partial_j u)\zeta^2 x_0 \, dx \, dt \, dx_0
= \frac{1}{p} \int_{Q_{2r}} \frac{a_{0i}}{a_{00}} \partial_j \left(|u(r, x, t)|^p\right) r \zeta^2 \, dx \, dt
- \int_0^r \int_{Q_{2r}} \frac{1}{a_{00}} |u|^{p-2} u \partial_j (a_{ij} \partial_i u)\zeta^2 x_0 \, dx \, dt \, dx_0
- \int_0^r \int_{Q_{2r}} \partial_i \left(\frac{1}{a_{00}}\right) |u|^{p-2} u \partial_j u \zeta^2 x_0 \, dx \, dt \, dx_0
- \frac{2}{a_{00}} \int_0^r \int_{Q_{2r}} |u|^{p-2} u (\partial_j u)\zeta\partial_i x_0 \, dx \, dt \, dx_0
- \int_0^r \int_{Q_{2r}} \frac{a_{ij}}{a_{00}} \partial_i \left(|u|^{p-2} u(\partial_j u)\zeta^2 x_0 \right) \, dx \, dt \, dx_0
= I + II + III + IV + V + VI
\]

Our strategy is to further estimate all these terms and then group similar terms together. First consider \( II \), we use that \( u \) is a solution to (1.1)

\[
II = - \int_0^r \int_{Q_{2r}} \frac{1}{a_{00}} |u|^{p-2} u u_i \zeta^2 x_0 \, dx \, dt \, dx_0 + \int_0^r \int_{Q_{2r}} \frac{1}{a_{00}} |u|^{p-2} u_0 \partial_i u \zeta^2 x_0 \, dx \, dt \, dx_0
= II_1 + II_2.
\]

Using the identity \( 2x_0 = \partial_0 x_0^2 \) we integrate by parts in \( x_0 \) to obtain

\[
II_1 = -\frac{1}{2} \int_0^r \int_{Q_{2r}} \frac{1}{a_{00}} |u|^{p-2} u u_i \zeta^2 \partial_0 x_0^2 \, dx \, dt \, dx_0
= -\frac{1}{2} \int_0^r \int_{Q_{2r}} \frac{1}{a_{00}} |u(r, x, t)|^{p-2} u(r, x, t) u_i(r, x, t) \zeta^2 r^2 \, dx \, dt
+ \frac{1}{2} \int_0^r \int_{Q_{2r}} \partial_0 \left(\frac{1}{a_{00}}\right) |u|^{p-2} u u_i \zeta^2 x_0^2 \, dx \, dt \, dx_0
+ \frac{p-1}{2} \int_0^r \int_{Q_{2r}} \frac{1}{a_{00}} |u|^{p-2} \partial_0 u u_i \zeta^2 x_0^2 \, dx \, dt \, dx_0
+ \frac{1}{2} \int_0^r \int_{Q_{2r}} \frac{1}{a_{00}} |u|^{p-2} u \partial_0 \partial_i u \zeta^2 x_0^2 \, dx \, dt \, dx_0
= II_{11} + II_{12} + II_{13} + II_{14}.
\]
Consider the boundary term $II_{11}$ and we integrate by parts in $t$

$$II_{11} = -\frac{1}{4} \int_{Q_2} \left( \frac{1}{a_0} \right) u(r,x,t)^p - \partial_t \left( u^2(r,x,t) \right) \zeta^2 r^2 \, dx \, dt$$

$$= \frac{1}{4} \int_{Q_2} \partial_t \left( \frac{1}{a_0} \right) u(r,x,t)^p - \partial_t \left( u^2(r,x,t) \zeta^2 r^2 \right) \, dx \, dt$$

$$\leq \frac{1}{2} \int_{Q_2} \partial_t \left( \frac{1}{a_0} \right) u(r,x,t)^p \zeta^2 r^2 \, dx \, dt$$

$$+ \frac{p-2}{4} \int_{Q_2} \left( \frac{1}{a_0} \right) u(r,x,t)^p - \partial_t \left( u^2(r,x,t) \right) \zeta^2 r^2 \, dx \, dt$$

$$= II_{111} + II_{112} + II_{113}.$$

Since $p < 2$, so $p - 2 < 0$, we can absorb $II_{113}$ into $II_{11}$ and save $II_{112}$ to bound later on.

Considering $II_{14}$, we swap the order of differentiation on $\partial_0 \partial_1 u$ and integrate by parts in $t$ to show

$$II_{14} = 1/2 \int_0^r \int_{Q_{2r}} \left( \frac{1}{a_0} \right) u(r,x,t)^p - \partial_0 \partial_1 u \zeta^2 x_0^2 \, dx \, dt \, dx_0$$

$$= -1/2 \int_0^r \int_{Q_{2r}} \partial_0 \left( \frac{1}{a_0} \right) u(r,x,t)^p - \partial_1 u \zeta^2 x_0^2 \, dx \, dt \, dx_0$$

$$- \frac{p-1}{2} \int_0^r \int_{Q_{2r}} \left( \frac{1}{a_0} \right) u(r,x,t)^p - \partial_1 u \zeta^2 x_0^2 \, dx \, dt \, dx_0$$

$$= II_{141} + II_{142} + II_{143}.$$

Observe that $II_{142} = -II_{13}$ so these terms cancel. We bound $II_{141}$ by

$$II_{141} = 1/2 \int_0^r \int_{Q_{2r}} \partial_0 \left( \frac{1}{a_0} \right) u(r,x,t)^p - \partial_1 u \zeta^2 x_0^2 \, dx \, dt \, dx_0$$

$$\leq \left( \int_0^r \int_{Q_{2r}} |A|^2 |u(r,x,t)^p \zeta^2 | \, dx \, dt \, dx_0 \right)^{1/2} \left( \int_0^r \int_{Q_{2r}} |\nabla u|^2 |u(r,x,t)^p \zeta^2 | \, dx \, dt \, dx_0 \right)^{1/2}.$$

Two parts of $II_1$ we have left to bound are $II_{112}$ and $II_{143}$. Both of these integrals involve $\zeta \zeta$ and therefore if $\zeta$ is a partition of unity when we sum over that partition these terms sum to 0.

The terms $II_2$ and $III$ are simply dealt with by

$$II_2 \leq \left( \int_0^r \int_{Q_{2r}} |B|^2 |u(r,x,t)^p \zeta^2 | \, dx \, dt \, dx_0 \right)^{1/2} \left( \int_0^r \int_{Q_{2r}} |\nabla u|^2 |u(r,x,t)^p \zeta^2 | \, dx \, dt \, dx_0 \right)^{1/2}$$

and

$$III \leq \left( \int_0^r \int_{Q_{2r}} |\nabla A|^2 |u(r,x,t)^p \zeta^2 | \, dx \, dt \, dx_0 \right)^{1/2} \left( \int_0^r \int_{Q_{2r}} |\nabla u|^2 |u(r,x,t)^p \zeta^2 | \, dx \, dt \, dx_0 \right)^{1/2}.$$

The integral in the term $IV$ contains the terms $\zeta \partial_t \zeta$ and as before if $\zeta$ is a partition of unity then after summing this term cancels out. Therefore the terms that we have yet to estimate are $I$, $V$, and $VI$. 
We consider \( V \) in the two cases \( j = 0 \) and \( j \neq 0 \) separately. Since \( \zeta \) is independent of \( x_0 \) by the fundamental theorem of calculus

\[
V_{(j=0)} = - \frac{1}{p} \int_0^r \int_{Q_{2r}} |u|^{p-2} u (\partial_0 u) \zeta^2 \, dx \, dt \, dx_0 = - \frac{1}{p} \frac{1}{2} \int_0^r \int_{Q_{2r}} \partial_0 (|u|^p \zeta^2) \, dx \, dt \, dx_0 \\
= \frac{1}{p} \int_0^r \int_{Q_{2r}} |u(0, x, t)|^p \zeta^2 \, dx \, dt - \frac{1}{p} \int_0^r \int_{Q_{2r}} |u(r, x, t)|^p \zeta^2 \, dx \, dt.
\]

For the \( j \neq 0 \) case we use that \( \partial_0 x_0 = 1 \) and integrate this case by parts in \( x_0 \)

\[
V_{(j \neq 0)} = - \frac{1}{p} \int_0^r \int_{Q_{2r}} \frac{a_{0j}}{a_{00}} \partial_j (|u|^p) \zeta^2 \, dx \, dt \, dx_0 \\
= - \frac{1}{p} \int_0^r \int_{Q_{2r}} \frac{a_{0j}}{a_{00}} \partial_j (|u|^p) \zeta^2 \partial_0 x_0 \, dx \, dt \, dx_0 \\
= - \frac{1}{p} \int_0^r \int_{Q_{2r}} \frac{a_{0j}}{a_{00}} \partial_j (|u(r, x, t)|^p) \zeta^2 x_0 \, dx \, dt + \frac{1}{p} \int_0^r \int_{Q_{2r}} \frac{a_{0j}}{a_{00}} \partial_j \partial_0 (|u|^p) \zeta^2 x_0 \, dx \, dt \\
+ \frac{1}{p} \int_0^r \int_{Q_{2r}} \partial_0 \left( \frac{a_{0j}}{a_{00}} \right) \partial_j (|u|^p) \zeta^2 x_0 \, dx \, dt \\
= V_1 + V_2 + V_3.
\]

The term \( V_1 = -I_{(j \neq 0)} \) so they cancel out. For \( V_2 \) we integrate by parts in \( x_j \)

\[
V_2 = - \sum_{j \neq 0} \frac{1}{p} \int_0^r \int_{Q_{2r}} \frac{a_{0j}}{a_{00}} \partial_j (|u(r, x, t)|^p) \zeta^2 x_0 \, dx \, dt \\
- \frac{1}{p} \int_0^r \int_{Q_{2r}} \partial_j \left( \frac{a_{0j}}{a_{00}} \partial_0 (|u|^p) \zeta^2 x_0 \right) \, dx \, dt \\
- \frac{2}{p} \int_0^r \int_{Q_{2r}} \partial_0 \left( |u|^p \right) \zeta \partial_j x_0 \, dx \, dt \\
= V_{21} + V_{22} + V_{23}.
\]

\( V_{22} \) and \( V_{33} \) are of the same type and can be estimated as \( III \) by

\[
\left| \int_0^r \int_{Q_{2r}} \nabla \left( \frac{a_{0j}}{a_{00}} \right) \nabla (|u|^p) \zeta^2 x_0 \, dx \, dt \, dx_0 \right| \lesssim \int_0^r \int_{Q_{2r}} |u|^{p-1} |\nabla u| |\nabla A| |\zeta^2 x_0| \, dx \, dt \, dx_0 \\
\lesssim \left( \int_0^r \int_{Q_{2r}} |\nabla A|^2 |u|^p \zeta^2 x_0 \, dx \, dt \, dx_0 \right)^{1/2} \left( \int_0^r \int_{Q_{2r}} |\nabla u|^2 |u|^{p-2} \zeta^2 x_0 \, dx \, dt \, dx_0 \right)^{1/2}.
\]

The final term from (5.4) to estimate is \( VI \)

\[
VI = - \int_0^r \int_{Q_{2r}} \frac{a_{0j}}{a_{00}} \partial_j (|u|^{p-2}) u (\partial_j u) \zeta^2 \, dx \, dt \, dx_0
\]

\[
= (2 - p) \int_0^r \int_{Q_{2r}} |u|^{p-2} (\partial_j u) (\partial_j u) \zeta^2 \, dx \, dt \, dx_0
\]

and since \( 2 - p < 1 \) we can hide \( VI \) in the left hand side of (5.4).

We are now at the stage where we can group all the similar terms and estimate them. There are four different types of terms:

\[
J_1 = I_{(j=0)} + II_{111} + V_{(j=0)} + V_{21}
\]

\[
J_2 = II_{12}
\]

\[
J_3 = II_{141} + II_{12} + III + \sum_{j \neq 0} V_{22} + \sum_{j \neq 0} V_{33}
\]

\[
J_4 = II_{112} + II_{143} + IV + \sum_{j \neq 0} V_{23}.
\]
We shall use the following standard result multiple times to deal with terms containing $|\nabla A|^2$, $|A_1|$ or $|B|$; a reference for this is [Ste93, p. 59]. Let $\mu$ be a Carleson measure and $U$ the upper half space then for any function $u$ we have

$$\int_{U}\lvert u \rvert^p \,d\mu \leq \lVert \mu \lVert \lVert N(u) \rVert^p_{L^p(U^\ast)},$$

with a local version holding on Carleson boxes as well.

First we consider $J_1$, which consists of boundary terms at $(0, x, t)$ and $(r, x, t)$.

$$J_1 = \frac{1}{p} \int_{Q_{2r}} \partial_0 (\lvert u(r, x, t) \rvert^p) \zeta^2 r \,dx \,dt - \frac{1}{4} \int_{Q_{2r}} \frac{\partial a_{00}}{a_{00}} \lvert u(r, x, t) \rvert^{p-2} u^2(r, x, t) \zeta^2 r^2 \,dx \,dt$$

$$+ \frac{1}{p} \int_{Q_{2r}} \lvert u(0, x, t) \rvert^p \zeta^2 \,dx \,dt - \frac{1}{p} \int_{Q_{2r}} \lvert u(r, x, t) \rvert^p \zeta^2 \,dx \,dt$$

$$- \sum_{j \neq 0} \frac{1}{p} \int_{Q_{2r}} \frac{a_{0j}}{a_{00}} \partial_j (\lvert u(r, x, t) \rvert^p) \zeta^2 r \,dx \,dt.$$

The second term in $J_1$, originating from $II_{111}$, has the bound

$$II_{111} = \frac{1}{4} \int_{Q_{2r}} \frac{\partial a_{00}}{a_{00}} \lvert u(r, x, t) \rvert^{p-2} u^2(r, x, t) \zeta^2 r^2 \,dx \,dt$$

$$\leq \frac{1}{4\lambda^2} \int_{Q_{2r}} \lvert A_1 \lvert \lvert u(r, x, t) \rvert^p \zeta^2 r^2 \,dx \,dt \leq \frac{\lVert \mu_2 \lVert_{C^{1,2}}^{1/2}}{\lambda^2} \lVert N^\ast (u) \rVert^p_{L^p(Q_{2r})}.$$

$$J_2 = \frac{1}{2} \int_0^r \int_{Q_{2r}} \partial_0 \left( \frac{1}{a_{00}} \right) \lvert u \rvert^{p-2} u u_0 \zeta^2 x_0^2 \,dx \,dt \,dx_0$$

$$\leq \frac{1}{2\lambda^2} \left( \int_0^r \int_{Q_{2r}} |\nabla A|^2 \lvert u \rvert x_0^2 \zeta^2 \,dx \,dt \,dx_0 \right)^{1/2} \left( \int_0^r \int_{Q_{2r}} |u_0|^2 \lvert u \rvert^{p-2} x_0^3 \zeta^2 \,dx \,dt \,dx_0 \right)^{1/2}$$

$$\leq \frac{1}{\lambda^2} \left( \lVert \mu_2 \lVert_{C^{1,2}} \lVert N^\ast (u) \rVert^p_{L^p(Q_{2r})} \right)^{1/2} \left( \int_0^r \int_{Q_{2r}} |u_0|^2 \lvert u \rvert^{p-2} x_0^3 \zeta^2 \,dx \,dt \,dx_0 \right)^{1/2}.$$

With a constant $C_3 = C_3(\lambda, \Lambda, n)$ we can bound $J_3$ by

$$J_3 \leq C_3 \left( \int_0^r \int_{Q_{2r}} (x_0 |\nabla A|^2 + x_0 |B|^2 + x_0^3 |A_1|^2) \lvert u \rvert^p \zeta^2 \,dx \,dt \,dx_0 \right)^{1/2}$$

$$\times \left( \int_0^r \int_{Q_{2r}} |\nabla u|^2 \lvert u \rvert^{p-2} x_0^3 \zeta^2 \,dx \,dt \,dx_0 \right)^{1/2}$$

$$\leq C_3 \left( \lVert \mu_1 \lVert_{C^{1,2}} + \lVert \mu_2 \lVert_{C^{1,2}} \lVert N^\ast (u) \rVert^p_{L^p(Q_{2r})} \right)^{1/2} \left( \int_0^r \int_{Q_{2r}} |\nabla u|^2 \lvert u \rvert^{p-2} x_0^3 \zeta^2 \,dx \,dt \,dx_0 \right)^{1/2}.$$

Finally, $J_4$ consists of terms of the type $\zeta \partial_0 \zeta$ or $\zeta \partial_i \zeta$. Later we take $\zeta$ to be a partition of unity and so when we sum up over the partition all the terms in $J_4$ sum to 0.
Therefore after all these calculations
\[
\int_0^r \int_{Q_{2r}} |u|^{p-2} \frac{\partial_j u}{a_{00}} (\partial_i u)(\partial_j u) \xi^2 x_0 \, dx \, dx_0 = J_1 + J_2 + J_3 + J_4
\]
\[
\leq \frac{nA}{\lambda} \int_{Q_{2r}} \partial_0 (|u(r, x, t)|^p) \xi^2 x_0 \, dx \, dx_0 + \int_{Q_{2r}} |u(0, x, t)|^p \xi^2 \, dx \, dt
\]
\[- \int_{Q_{2r}} |u(r, x, t)|^p \xi^2 \, dx \, dt + \frac{\|\mu_2\|_{L^p(\Omega)}^{1/2}}{2^2} \ ||N^r(u)\|_{L^p(\Omega)}^p
\]
\[
+ \frac{1}{\lambda^2} \left( \|\mu_2\|_{C,2r} ||N^r(u)\|_{L^p(\Omega)}^p \right)^{1/2} \left( \int_0^r \int_{Q_{2r}} |u|^2 |u|^{p-2} x_0^3 \xi^2 \, dx \, dx_0 \right)^{1/2}
\]
\[+ C_3 \left( \|\mu_1\|_{C,2r} + \|\mu_2\|_{C,2r} ||N^r(u)\|_{L^p(\Omega)}^p \right)^{1/2}
\]
\[\times \left( \int_0^r \int_{Q_{2r}} |\nabla u|^2 |u|^{p-2} x_0^3 \xi^2 \, dx \, dx_0 \right)^{1/2} + J_4.
\]

By assuming that \( \Omega \) is smooth as well as an admissible domain (definition 2.10) there exists a collar neighbourhood \( V \) of \( \partial \Omega \) in \( \mathbb{R}^{n+1} \) such that \( \Omega \cap V \) can be globally parametrised by \( (0, r) \times \partial \Omega \) for some small \( r > 0 \), see remark 2.20 and [DH18] for details. Using definition 2.10, there is a collection of charts covering \( \partial \Omega \) with bounded overlap, say by \( M \). We consider a partition of unity of these charts \( \xi_j \), with \( \xi_j \) having the same definition, support and estimates as \( \xi \) before, and \( \sum_j \xi_j = 1 \) everywhere. Therefore, when we sum (5.6) over this partition of unity the term on the left-hand side is bounded below by
\[
\frac{1}{\lambda} \int_0^r \int_{\partial \Omega} |u|^{p-2} (A \nabla u \cdot \nabla u) x_0 \, dx \, dt \, dx_0,
\]
which is comparable to the truncated \( p \)-adapted square function \( \|S^p_p(u)\|_{L^p(\partial \Omega)}^p \).

Therefore, remembering that after summing \( J_4 = 0 \), for any \( \varepsilon > 0 \) we have
\[
\frac{\lambda}{\lambda} \|S^p_p(u)\|_{L^p(\partial \Omega)}^p \sim \frac{\lambda}{\lambda} \int_0^r \int_{\partial \Omega} |u|^{p-2} |\nabla u|^2 x_0 \, dx \, dt \, dx_0
\]
\[
\leq \frac{nA}{\lambda} \int_{\partial \Omega} \partial_0 (|u(r, x, t)|^p) r \, dx \, dt + \int_{\partial \Omega} |u(0, x, t)|^p \, dx \, dt - \int_{\partial \Omega} |u(r, x, t)|^p \, dx \, dt
\]
\[+ \frac{M \|\mu_2\|_{L^p(\Omega)}^{1/2}}{\lambda^2} ||N^r(u)\|_{L^p(\partial \Omega)}^p + \frac{\|\mu_2\|_{C,2r}}{4 \lambda} ||N^r(u)\|_{L^p(\partial \Omega)}^p
\]
\[+ \varepsilon \int_0^r \int_{\partial \Omega} |u|^2 |u|^{p-2} x_0^3 \xi^2 \, dx \, dx_0 + C_3 \left( \|\mu_1\|_{C,2r} + \|\mu_2\|_{C,2r} ||N^r(u)\|_{L^p(\partial \Omega)}^p \right)
\]
\[\times \left( \int_0^r \int_{\partial \Omega} |\nabla u|^2 |u|^{p-2} x_0^3 \xi^2 \, dx \, dx_0 \right)^{1/2} + J_4.
\]

By applying Lemma 4.5 to the \( p \)-adapted area function in (5.7) we see that the \( p \)-adapted square function on the right-hand side of (5.7) is always multiplied by \( \varepsilon \).

By choosing \( \varepsilon \) small enough we can absorb this \( p \)-adapted square function into the left-hand side yielding
\[
C_1 \|S^p_p(u)\|_{L^p(\partial \Omega)}^p \leq \int_{\partial \Omega} \partial_0 (|u(r, x, t)|^p) r \, dx \, dt
\]
\[+ \int_{\partial \Omega} |u(0, x, t)|^p \, dx \, dt - \int_{\partial \Omega} |u(r, x, t)|^p \, dx \, dt
\]
\[+ C_2 \left( \|\mu_1\|_{C,2r} + \|\mu_2\|_{C,2r} + \|\mu_2\|_{1/2} \right) ||N^r(u)\|_{L^p(\partial \Omega)}^p.
\]
We integrate (5.8) in the $r$ variable, average over $[0, r_0]$ and use the identity $(\partial_0 |u|^p)x_0 = \partial_0( |u|^p x_0) - |u|^p$ to give
\begin{align}
C_1 \int_0^{r_0} \int_{\Omega} \left( x_0 - \frac{r^2}{r_0} \right) |\nabla u|^2 |u|^{p-2} dx dt dx_0 \\
+ \frac{2}{r_0} \int_0^{r_0} \int_{\Omega} |u(x_0, x, t)|^p dx dt dx_0 \\
\leq \int_{\partial \Omega} |u(r_0, x, t)|^p dx dt + \int_{\Omega} |u(0, x, t)|^p dx dt \\
+ C_2 \left( \|\mu_1\|_{C, 2r} + \|\mu_2\|_{C, 2r} + \|\mu_2\|_{C, 2r}^{1/2} \right) \|N^r(u)\|_{L^p(\partial \Omega)}.
\end{align}

Finally truncating the first integral on the left hand side to $[0, r_0/2]$ gives
\begin{align}
C_1 \frac{1}{2} \int_0^{r_0/2} \int_{\partial \Omega} |\nabla u|^2 |u|^{p-2} x_0 dx dt dx_0 + \frac{2}{r_0} \int_0^{r_0} \int_{\partial \Omega} |u(x_0, x, t)|^p dx dt dx_0 \\
\leq \int_{\partial \Omega} |u(r_0, x, t)|^p dx dt + \int_{\Omega} |u(0, x, t)|^p dx dt \\
+ C_2 \left( \|\mu_1\|_{C, 2r} + \|\mu_2\|_{C, 2r} + \|\mu_2\|_{C, 2r}^{1/2} \right) \|N^r(u)\|_{L^p(\partial \Omega)}.
\end{align}

The local estimate for Lemma 5.1 is obtained (exactly as in [DH18]) if we do not sum over all the coordinate patches but instead use the estimates derived for a single boundary cube $Q_r$ in (5.6).

We need to control the first integral on the right-hand side of (5.2) to achieve our goal of controlling the $p$-adapted square function. Thankfully this has already been done for us in the proof of [DH18, Cor. 5.3] which we encapsulate below.

**Lemma 5.3.** Let $\Omega$ be as in Lemma 5.2 and $u$ be a non-negative solution to (1.1). For a small $r_0 > 0$ depending on the geometry the domain $\Omega$ there exists a constant $C$ such that for $\varepsilon = \|\mu_1\|_{C, 2r} + \|\mu_2\|_{C, 2r} + \|\mu_2\|_{C, 2r}^{1/2}$
\begin{align}
\int_{\partial \Omega} u(r_0, x, t)^p dx dt \leq \frac{2}{r_0} \int_0^{r_0} \int_{\partial \Omega} u(x_0, x, t)^p dx dt dx_0 + C\varepsilon \|N^r(u)\|_{L^p(\partial \Omega)}.
\end{align}

Combining Lemmas 5.2 and 5.3 gives us the desired result.

**Corollary 5.4.** Let $\Omega$ be as in Lemma 5.2 and $u$ be a non-negative solution to (1.1). For a small $r_0 > 0$ depending on the geometry the domain $\Omega$ there exists constants $C_1, C_2 > 0$ such that for $\varepsilon = \|\mu_1\|_{C, 2r} + \|\mu_2\|_{C, 2r} + \|\mu_2\|_{C, 2r}^{1/2}$
\begin{align}
\|S_{p}^{r_0/2}(u)\|_{L^p(\partial \Omega)} \leq \int_0^{r_0/2} \int_{\partial \Omega} |\nabla u|^2 |u|^{p-2} x_0 dx dt dx_0 \\
\leq C_1 \int_{\partial \Omega} |u(0, x, t)|^p dx dt + C_2 \varepsilon \|N^r(u)\|_{L^p(\partial \Omega)}.
\end{align}

6. Bounding the non-tangential maximum function by the $p$-adapted square function

Our goal in this section has been vastly simplified due to [Riv03] proving a local good-$\lambda$ inequality. We use this to bound the non-tangential maximum function by the $p$-adapted square function. We first bound the non-tangential maximum function by the usual $L^2$ based square function $S_2(u)$ but a simple argument from [DPP07, (3.41)] shows that for $1 < p < 2$ and any $\varepsilon > 0$ we have
\begin{align}
\|S_2(u)\|_{L^p(\partial \Omega)} \leq C_\varepsilon \|S_2'(u)\|_{L^p(\partial \Omega)} + \varepsilon \|N^r(u)\|_{L^p(\partial \Omega)},
\end{align}
with a local version of this statement holding as well.

The good-$\lambda$ inequality from [Riv03, p. 508] is expressed in the following lemma.
Lemma 6.1. Let $v$ be a solution to (2.29) and $v(X, t) = 0$ for some point $(X, t) \in Q_r$. Let $E = \{(0, x, t) \in Q_r : S_{2, \alpha}(v) \leq \lambda \}$ and $q > 2$ then

$$|\{(0, x, t) \in Q_r : N_q(v) > \lambda\}| \lesssim |\{(0, x, t) \in Q_r : S_{2, \alpha}(v) > \lambda\}|$$

$$+ \frac{1}{\lambda^{q-1}} \int_E S_{2, \alpha}(v)^q \, dx \, dt.$$  \hspace{1cm} (6.2)

If $p \geq 2$ then the following lemma is immediate from [DH18, Lemma 6.1], which is an adaptation of [Riv03, Theorem 1.3 and Proposition 5.3].

Lemma 6.2. Let $v$ be a solution to (2.29) in $U$ and the coefficients of (2.29) satisfy the Carleson estimates (2.34), (2.35), (2.37) and (2.38) on all parabolic balls of size $\leq r_0$. Then there exists a constant $C$ such that for any $r \in (0, r_0/8)$

$$\int_{Q_r} N_{a_{1/2}}(v)^p \, dx \, dt \leq C \left( \int_{Q_{2r}} A_{2, \alpha}(v)^p \, dx \, dt + \int_{Q_{2r}} S_{2, \alpha}(v)^p \, dx \, dt \right)$$

$$+ \lambda^{q+1} |\{A_{u_{2r}}\}|^p,$$

where $A_{2r}$ is a corkscrew point of the boundary ball $\Delta_r$. That is a point $2r^2$ later in time than the center of $\Delta_r$ and at a distance comparable to $r$ from the boundary and $\lambda$ from the center of the ball $\Delta_r$.

Proof. We first assume that $v(X, t) = 0$ for some $(X, t) \in Q_r$ and then we have the good-$\lambda$ inequality (6.2). The passage from this good-$\lambda$ inequality to a local $L^p$ estimate is standard in the spirit of [FS72]. We remove the assumption $v(X, t) = 0$ for the cost of adding the $\lambda^{q+1} |\{A_{u_{2r}}\}|^p$ term in the same way as [Riv03, DH18].

From this local estimate, we can obtain the following global $L^p$ estimate by the same proof as the global $L^2$ estimate from [DH18, Theorem 6.3].

Theorem 6.3. Let $u$ be a solution to (1.1) and the coefficients of (1.1) satisfy the Carleson estimates (2.36) and (2.39) then

$$\|N^+(u)\|_{L^p(\partial \Omega)} \lesssim \|S_{\alpha}^+(u)\|_{L^p(\partial \Omega)} + \|u\|_{L^p(\partial \Omega)}$$  \hspace{1cm} (6.4)

and by (6.1)

$$\|N^-(u)\|_{L^p(\partial \Omega)} \lesssim \|S_{\alpha}^-(u)\|_{L^p(\partial \Omega)} + \|u\|_{L^p(\partial \Omega)}.$$  \hspace{1cm} (6.5)

7. Proof of Theorem 1.1

We only consider the case $1 < p < 2$ and use interpolation to obtain solvability for $p \geq 2$. First assume either stronger Carleson condition of (2.39), or (1.7) and (1.8) holds. Therefore the Carleson conditions on the pullback coefficients (2.34), (2.35), (2.37) and (2.38) hold.

Without loss of generality, by remark 2.20, we may assume that our domain is smooth. Consider $f^+ = \max\{0, f\}$ and $f^- = \max\{0, -f\}$ where $f \in C_{0}(\partial \Omega)$ and denote the corresponding solutions with these boundary data $u^+$ and $u^-$ respectively. Hence we may apply the Corollary 5.4 separately to $u^+$ and $u^-$. By the maximum principle, these two solutions are non-negative. It follows that for any such non-negative $u$ we have

$$\|S_{\alpha}^+(u)\|_{L^p(\partial \Omega)} \leq C \|f\|_{L^p(\partial \Omega)} + C(\|\mu\|_{C}^{1/2} + \|\mu\|_{C}) \|N^2(u)\|_{L^p(\partial \Omega)}^p$$

and Theorem 6.3 gives

$$\|N^+(u)\|_{L^p(\partial \Omega)} \leq C \|f\|_{L^p(\partial \Omega)}^p + C \|S_{\alpha}^2(u)\|_{L^p(\partial \Omega)}^p.$$
(depending on the domain $\Omega$) such that the Carleson norm after the pullback is said only twice the original Carleson norm of the coefficients over all balls of size $\leq r_0$.

Since we are assuming $\|\mu\|_C$ is small, apparently we also have $\|\mu\|_C \leq C\|\mu\|_C^{1/2}$. By rearranging these two inequalities and combining estimates for $u^+$ and $u^-$, we obtain, for $0 < r \leq r_0/8$, $$
\|N^r(u)\|_{L^p(\partial \Omega)} \leq C\|f\|_{L^p(\partial \Omega)} + C\|\mu\|_C^{1/2}\|N^{4r}(u)\|_{L^p(\partial \Omega)}.
$$ By a simple geometric argument in [DH18] involving cones of different apertures, Lemmas 3.4 and 3.7 show there exists a constant $M$ such that $$
\|N^{4r}(u)\|_{L^p(\partial \Omega)} \leq M\|N^r(u)\|_{L^p(\partial \Omega)}. 
$$ It follows that if $CM\|\mu\|_C^{1/2} < 1/2$ by combining the last two inequalities we obtain $$
\|N^r(u)\|_{L^p(\partial \Omega)} \leq 2C\|f\|_{L^p(\partial \Omega)},
$$ which is the desired estimate (for truncated version of non-tangential maximum function). The result with the non-truncated version of the non-tangential maximum function $N(u)$ follows as our domain is bounded in space and hence (7.1) can be iterated finitely many times until the non-tangential cones have sufficient height to cover the whole domain.

Finally, we comment on how the Carleson condition (2.39) can be relaxed to the weaker condition (1.6). The idea is the same as [DH18, Theorem 3.1]. As shown there, if the operator $L$ satisfies the weaker condition (1.6), then it is possible (via mollification of coefficients) to find another operator $L_1$ which is a small perturbation of the operator $L$ and $L_1$ satisfies (2.39). The solvability of the $L^p$ Dirichlet problem in the range $1 < p < 2$ for $L_1$ follows by our previous arguments. However, as $L$ is a small perturbation of the operator $L_1$ we have by the perturbation argument of [Swe98] $L^p$ solvability of $L$ as well.

Finally, for larger values of $p$ we use the maximum principle and interpolation to obtain solvability results in the full range $1 < p < \infty$.

\section{Appendix — Proofs of Results from Section 2}

\textbf{Proof of Theorem 2.3.} We begin by proving the equivalence of conditions (3) and (6) using ideas from [Str80] and write $F = \mathcal{D}\hat{\phi}$ where $F$ is a tempered distribution. Let $$
\varphi^k = \chi_{Q_1(0,0)} - \chi_{Q_1(c_k)}
$$ then for $1 \leq k \leq n - 1$
\begin{align*}
\hat{\varphi}^k(\xi, \tau) &= \frac{2\sin^2(\xi_k/2)}{\xi_k} \frac{1 - e^{-i\tau}}{i\tau} \prod_{j \neq k} \frac{1 - e^{-i\xi_j}}{i\xi_j}, \\
\varphi^k(\xi, \tau) &= \frac{2\sin^2(\tau/2)}{\tau} \prod_{j = 1}^{n-1} \frac{1 - e^{-i\xi_j}}{i\xi_j}, \tag{8.1}
\end{align*}
with $\hat{\varphi}^k(\xi, \tau) \sim \xi_k$ for small $\xi_k$ and $1 \leq k \leq n - 1$. We let $$\hat{\psi}_u = \frac{e^{i(\xi_0)\cdot u} - 1}{\|\xi, \tau\|}$$ and denote by $\psi^u_p(x, t)$ the usual parabolic dilation by $\rho$, that is $$\psi^u_p(x, t) = \rho^{-(n+1)}\psi_u(x/\rho, t/\rho^2).$$
It is worth noting that \((ψ^k * ϕ_u)_ρ = ϕ^k_ρ * ϕ_u\). Therefore we may rewrite condition (6.a), by remark 2.4, as
\[
\sup_{Q_r} \sum_{k=1}^{n-1} \frac{1}{|Q_r|} \int_{Q_r} \int_{u \in \mathbb{S}^{n-1}} \int_{0}^{r} (ψ^k_ρ * ϕ^k_ρ * F)^2 \frac{dρ}{ρ} \, du \, dx \, dt \sim B_{(6.a)}.
\]
(8.2)

Similarly if we let
\[
\tilde{ψ}^u_n = \frac{ψ((0, τ) - 1)}{∥(ξ, τ)∥}
\]
then we may rewrite condition (6.b) as
\[
\sup_{Q_r} \frac{1}{|Q_r|} \int_{Q_r} \int_{u \in \mathbb{S}^{n-1}} \int_{0}^{r} (ψ^u_n ρ * F)^2 \frac{dρ}{ρ} \, du \, dx \, dt \sim B_{(6.b)}.
\]
(8.4)

The functions \(ψ^k \ast ϕ_u\) and \(ψ^u_n\) all satisfy the following conditions for some \(ε_i > 0\)
\[
\int ψ \, dx \, dt = 0,
\]
\[
|ψ(x, t)| \lesssim ∥(x, t)∥^{n-1-ε_1} \text{ for } ∥(x, t)∥ \geq a > 0,
\]
\[
|ϕ(ξ, τ)| \lesssim ∥(ξ, τ)∥^{ε_2} \text{ for } ∥(ξ, τ)∥ \leq 1,
\]
\[
|ϕ(ξ, τ)| \lesssim ∥(ξ, τ)∥^{-ε_2} \text{ for } ∥(ξ, τ)∥ \geq 1.
\]
(8.5)

Therefore if \(Dϕ = F \in \text{BMO}(\mathbb{R}^n)\) then \(B_{(6.a)} \lesssim ||Dϕ||^2\) and \(B_{(6.b)} \lesssim ||Dϕ||^2\) by \[Str80, Theorem 2.1\]; this shows condition (3) implies condition (6).

For the converse, we proceed via an analog of the proof of \[Str80, Theorem 2.6\].

Consider
\[
\hat{θ}(ξ, τ) = ∥(ξ, τ)∥\hat{ϕ}(ξ, τ),
\]
where \(ϕ \in C_0^∞(\mathbb{R})\). Let \(H_{00}\) be the dense subclass of continuous \(H^1\) functions \(ϕ\) such that \(g\) and all its derivatives decay rapidly, see \[Ste70, p. 225\]. Via an analogue of \[FS72, Theorem 3\], \[Str80, Lemma 2.7\] by assuming conditions (6.a) and (6.b) if \(g \in H_{00}(\mathbb{R}^n)\) then for each \(1 \leq k \leq n - 1\)
\[
\left| \int_{\mathbb{S}^{n-1}} \int_{0}^{∞} \int_{\mathbb{R}^{n-1} \times \mathbb{R}} ψ^u_n \ast ϕ^k_ρ \ast F(x, t) ϕ_ρ \ast g(x, t) \, dx \, dt \, \frac{dρ}{ρ} \, du \right| \lesssim B_{(6.a)}^{1/2} ||g||_{H^1},
\]
(8.6)

and
\[
\left| \int_{\mathbb{S}^{n-1}} \int_{0}^{∞} \int_{\mathbb{R}^{n-1} \times \mathbb{R}} ψ^u_n ρ \ast F(x, t) ϕ_ρ \ast g(x, t) \, dx \, dt \, \frac{dρ}{ρ} \, du \right| \lesssim B_{(6.b)}^{1/2} ||g||_{H^1}.
\]
(8.7)

For \(1 \leq k \leq n - 1\) let
\[
m_k(ξ, τ) = \int_{\mathbb{S}^{n-1}} \int_{0}^{∞} \hat{ϕ}(−ρξ, −ρ^2 τ) \hat{ϕ}^k(−ρξ, −ρ^2 τ) ||(ξ, τ)|| (ρ ||(ξ, τ)||) \, dρ \, du,
\]
\[
m_u(ξ, τ) = \int_{\mathbb{S}^{n-1}} \int_{0}^{∞} \hat{ϕ}(−ρξ, −ρ^2 τ) ||(ξ, τ)|| (ρ ||(ξ, τ)||) \, dρ \, du.
\]
(8.8)

All of these functions \(m_i\) are homogeneous of degree zero, smooth away from the origin and the associated Fourier multipliers \(M_k\), for \(1 \leq k \leq n\), are Caldarón-Zygmund operators that preserve the class \(H^1_{00}\) and are bounded on \(H^1\).

The non-degeneracy condition from \[CT75\] on the family of functions \(\{m_k\}_{k=1}^n\) holds — that is the property that \(\sum_k |m_k(rξ, r^2 τ)|^2\) does not vanish identically in \(r\).
for \((\xi, \tau) \neq (0,0)\). Therefore by [CT75, CT77] we can find smooth homogeneous of degree zero functions \(u_{k,j}(\xi, \tau)\) and positive numbers \(r_j\) such that for all \((\xi, \tau) \neq (0,0)\)

\[
\sum_{k=1}^{n} \sum_{j=1}^{j_0} m_{k,r_j}(\xi, \tau) u_{k,j}(\xi, \tau) = 1,
\]

where \(m_{k,r_j}\) are as \(m_k\) but with \(r_j\rho\) replacing \(\rho\) in the arguments of \(\hat{\psi}^u, \hat{\varphi}^k\) and \(\hat{\psi}^u_2\) in (8.8) (but not \(\zeta\)).

Let \(M_{k,j}\) and \(U_{k,j}\) be the associated Fourier multiplier operators to their respective multipliers \(m_{k,r_j}\) and \(u_{k,j}\) then \(\sum \sum m_{k,j} U_{k,j} g = g\) for all \(g \in H^0_{\delta_0}\). By [FS72, Theorem 3], [Str80, Lemma 2.7] there exists \(h_{k,j} \in \text{BMO}(\mathbb{R}^n)\) such that \(\|h_{k,j}\|_{\text{BMO}}^2 \lesssim B(6.6)\) or \(B(6.6)\), and \((h_{k,j}, g) = (F, M_{k,j} g)\) for all \(g \in H^0_{\delta_0}\). If we replace \(g\) by \(U_{k,j} g \in H^0_{\delta_0}\) in the previous identity and sum over \(j\) and \(k\) we obtain \((h, g) = (F, g)\) for all \(g \in H^0_{\delta_0}\) where \(h = \sum \sum U_{k,j}^{*} h_{k,j}\); furthermore by the BMO condition on \(h_{k,j}\), \(\|h\|_{\text{BMO}}^2 \lesssim B(6.6) + B(6.6)\). The identity (8.9) does not need to hold at the origin therefore \(h \sim \tilde{F}\) may be supported at the origin and hence \(F = h + p\) where \(p\) is a polynomial. Due to the assumption \(\phi \in \text{Lip}(1,1/2)\) clearly \(F\) must be a tempered distribution. Hence as in [Str80] we may conclude \(F = h \in \text{BMO}(\mathbb{R}^n)\). This implies equivalence of conditions (3) and (6).

Similarly we may prove the equivalence of conditions (4) and (5) to condition (3). The changes needed are outlined below.

We first look at condition (5) \(\iff\) condition (3). In this instance we replace the convolutions \(\varphi^k \ast \psi^u\) by

\[
\hat{\psi}^u_i(\xi, \tau) = \frac{\mathrm{e}^{i(\xi \cdot 0)\cdot u} - 2 - \mathrm{e}^{-i(\xi \cdot 0)\cdot u}}{\|\xi, \tau\|},
\]

which corresponds to condition (5.a), and we keep the convolution \(\hat{\psi}^u\) as it is in (8.3). The same proof then goes through to give that condition (5) holds if and only if condition (3) holds with equivalent norms, as in (2.13).

Condition (4) \(\iff\) condition (3). This case is stated in [Riv03, Proposition 3.2]. Again the proof proceeds as above with one convolution

\[
\hat{\psi}^u(\xi, \tau) = \frac{\mathrm{e}^{i(\xi \cdot 0)\cdot u} - 2 - \mathrm{e}^{-i(\xi \cdot 0)\cdot u}}{\|\xi, \tau\|}.
\]

**Proof of Theorem 2.8.** Without loss of generality, we only consider the case \(\eta < 1\). When \(\eta \geq 1\) the existence of a extension with \(\|B\Phi\| \lesssim \eta + \ell\) requires a much simpler argument.

By (2.20) there exists \(f \in C_\delta\) such that \(\|\nabla \phi - f\|_{Q_{2R}} \leq 2\eta\) and a scale \(0 < r_0 = r_0(\delta) \leq d\) such that

\[
\|f\|_{Q_{2R}, r_0} \leq 2\eta.
\]

Let \(d' = \eta \min(r_0, r_1)/2\) and consider some \(r \leq d'\) and \(Q_r \subset Q_{4d}\). Find a natural number \(k\) such that \(R = 2^k r\) and \(R_\eta/2 < r \leq R_\eta\). By our choice of \(d'\) the cube \(Q_{2R}\) which is an enlargement of \(Q_r\) by a factor \(2^{k+1}\) is still contained in the original cube \(Q_{8d}\).

It follows that

\[
\sup_{Q_{2R} = J_{2R} \times I_{2R}} \int_{Q_{2R}} \frac{1}{|Q_{2R}|} \int_{I_{2R}} \int_{I_{2R}} \frac{|\phi(x,t) - \phi(x,\tau)|^2}{|\tau - \tau'|^2} \mathrm{d}\tau \mathrm{d}x \mathrm{d}t \leq \eta^2.
\]

Without loss of generality, we may now assume that the cube \(Q_{2R}\) is centered at the origin \((0,0)\) and that \(\phi(0,0) = 0\), since the BMO norm is invariant under
translation and ignores constants. We first define \( \tilde{\phi} \) as an extension in time via reflection and tiling of the cube \( Q_r \subset \mathbb{R}^2 \):

\[
\tilde{\phi}(x, t) = \begin{cases} 
\phi(x, t) & t \in [-r^2, r^2] + 4kr^2, \\
\phi(x, 2r^2 - t) & t \in [r^2, 3r^2] + 4kr^2, \ k \in \mathbb{Z}.
\end{cases}
\] (8.10)

See figure 1 on page 34 for an illustration of this. Clearly \( \tilde{\phi} \) coincides with \( \phi \) on \( Q_r \).

It follows that \( \tilde{\phi} \) is a function \( \tilde{\phi} : \{|x|_\infty < 2R\} \times \mathbb{R} \to \mathbb{R} \) and \( (\nabla \tilde{\phi})_{Q_r} = (\nabla \phi)_{Q_r} \).

Consider a cut off function \( \rho \) such that

\[
\rho(x) = \begin{cases} 
1 & \text{if } |x|_\infty < r \\
0 & \text{if } |x|_\infty > 2R,
\end{cases}
\]

and \( |\nabla \rho| \lesssim 1/R \lesssim \eta/r \). Finally define

\[
\Phi = \tilde{\phi}\rho + (1 - \rho)(x \cdot (\nabla \tilde{\phi})_{Q_r}).
\] (8.11)

Clearly \( \Phi \) is well defined on \( \mathbb{R}^{n-1} \times \mathbb{R} \) as \( \rho = 0 \) outside the support of \( \tilde{\phi} \). We claim that \( \Phi \) satisfies \( (i)-(iv) \) of Theorem 2.8 which we establish in a sequence of lemmas below. Observe also that from our definition of \( \Phi \) we have

\[
\nabla \Phi = (\nabla \tilde{\phi} - (\nabla \tilde{\phi})_{Q_r}) \rho + \nabla \rho(x \cdot (\nabla \tilde{\phi})_{Q_r}) + (\nabla \tilde{\phi})_{Q_r}.
\] (8.12)

We start with a couple of lemmas that allows us to reduce our claim to the dyadic case; this is to make the geometry easier to handle.

**Lemma 8.1** ([Jon80, Lemma 2.3], c.f. [Str80, Theorem 2.8]). Let \( f \) be defined on \( \mathbb{R}^n \) and

\[
\sup_Q \frac{1}{|Q|} \int_Q |f - f_Q| \leq c(\eta),
\] (8.13)
where the supremum is taken over all dyadic cubes \( Q \subset \mathbb{R}^n \). Further, assume that

\[
\sup_{Q_1, Q_2} |f_{Q_1} - f_{Q_2}| \leq c(\eta),
\]

(8.14)

where the supremum is taken over all dyadic cubes \( Q_1, Q_2 \) of equal edge length with a touching edge. Then

\[
\|f\|_* \lesssim c(\eta).
\]

Below \( l(Q_s) = s \) denotes the radius of a parabolic cube.

**Lemma 8.2** ([Jon80, Lemma 2.1 and pp. 44-45]). Let \( f \in BMO(Q) \) and \( Q_0 \subset Q_1 \subset Q \) then

\[
|f_{Q_0} - f_{Q_1}| \lesssim \log \left( 2 + \frac{l(Q_1)}{l(Q_0)} \right) \|f\|_{*, Q}.
\]

(8.15)

Furthermore, the same proof in [Jon80] gives the following slightly stronger result

\[
\frac{1}{|Q_0|} \int_Q |f - f_{Q_1}| \lesssim \log \left( 2 + \frac{l(Q_1)}{l(Q_0)} \right) \|f\|_{*, Q}.
\]

(8.16)

If \( Q_0, Q_1 \subset Q \) and \( l(Q_0) \leq l(Q_1) \) but they are not necessarily nested then

\[
|f_{Q_0} - f_{Q_1}| \lesssim \left( \log \frac{2 + l(Q_1)}{l(Q_0)} \right) + \log \left[ 2 + \frac{\text{dist}(Q_0, Q_1)}{l(Q_1)} \right] \|f\|_{*, Q}.
\]

(8.17)

If the cubes \( Q_0, Q_1 \) and \( Q \) are dyadic then we may replace BMO by dyadic BMO.

There is a typo at the top of [Jon80, p. 45]. It should read \( l(Q_4) \leq l(Q_3) \) (it currently reads the converse).

**Claim 8.3.** Let \( \tilde{\phi} \) be defined as in (8.10), \( \|\nabla \tilde{\phi}\|_{*, Q} \lesssim \eta \), and let \( Q \) be dyadic with \( r \leq l(Q) \leq 2R \) then

\[
\frac{1}{|Q|} \int_Q |\nabla \tilde{\phi} - \nabla \phi_{Q}\prime| \lesssim \eta^{1-\varepsilon}.
\]

(8.18)

**Proof of claim.** Let \( N \in \mathbb{N} \) be such that \( l(Q) = 2^N l(Q_1) \). Let \( \{Q^i\} \) be the \( 2^{N(n-1)} \) dyadic cubes that are translations of \( Q_r \) and partition \( Q \cap \{t \leq r^2 \} \). Then by Lemma 8.2

\[
\frac{1}{|Q|} \int_Q |\nabla \tilde{\phi} - \nabla \phi_{Q}\prime| = \sum_i \frac{2^{2N} |Q^i|}{|Q|} \frac{1}{|Q^i|} \int_{Q^i} |\nabla \tilde{\phi} - \nabla \phi_{Q}\prime| \\
\leq \sum_i \frac{2^{2N} |Q^i|}{|Q|} \left( \frac{1}{|Q^i|} \int_{Q^i} |\nabla \phi - \nabla \phi_{Q}\prime| + |\nabla \phi_{Q}\prime - \nabla \phi_{Q}\prime| \right) \\
\lesssim (\eta + \eta \log(2 + R/r)) \lesssim \eta + \eta \log(1 + 1/\eta) \lesssim \eta^{1-\varepsilon}. \quad \Box
\]

**Lemma 8.4** ([Ste76]). Let \( g, h \in L^1_{loc} \) then

\[
\frac{1}{|Q|} \int_Q |gh - (gh)_Q| \leq \frac{2}{|Q|} \int_Q |g(h - h_Q)| + \frac{|h_Q|}{|Q|} \int_Q |g - g_Q|. \tag{8.19}
\]

**Proof.** This small reduction is from [Ste76, p. 582]. First observe

\[
gh - (gh)_Q = g(h - h_Q) + h_Q(g - g_Q) + g_Qh_Q - (gh)_Q
\]

and

\[
|g_Qh_Q - (gh)_Q| = \left| \frac{1}{|Q|} \int_Q gh - \frac{1}{|Q|} \int_Q gh \right| \leq \frac{1}{|Q|} \int_Q |g(h - h_Q)|
\]

hence

\[
\frac{1}{|Q|} \int_Q |gh - (gh)_Q| - \frac{|h_Q|}{|Q|} \int_Q |g - g_Q| \leq 2 \frac{1}{|Q|} \int_Q |g(h - h_Q)|. \tag{8.20}
\]

\[\Box\]
We can now prove property (iii) of Theorem 2.8.

**Lemma 8.5.** Let $\Phi: \mathbb{R}^n \to \mathbb{R}$ be defined as in (8.11) with $\|\nabla \phi\|_{*,Q_{2R}} \lesssim \eta$ then $\nabla \Phi \in \text{BMO}(\mathbb{R}^n)$ and for all $0 < \varepsilon < 1$

$$\|\nabla \Phi\|_* \lesssim \eta^{1-\varepsilon} + \eta^\ell. \quad (8.21)$$

**Proof.** Recall $\nabla \Phi = (\nabla \tilde{\phi} - (\nabla \tilde{\phi})_{Q_r}) \rho + \nabla \rho(\tilde{\phi} - x \cdot (\nabla \tilde{\phi})_{Q_r}) + (\nabla \tilde{\phi})_{Q_r}$; we can ignore the constant term as the BMO norm doesn’t see it. Let $\psi = \nabla \tilde{\phi} - (\nabla \tilde{\phi})_{Q_r}$ and $\theta = \tilde{\phi} - x \cdot (\nabla \tilde{\phi})_{Q_r}$. We want to bound $\|\rho \psi\|_*$ and $\|\nabla \rho \theta\|_*$. We first tackle the term $\|\rho \psi\|_*$.

**Step 1:** $(8.14)$ holds: $\sup_{Q_1, Q_2} |(\rho \psi)_{Q_1} - (\rho \psi)_{Q_2}| \leq c(\eta)$ for $Q_1, Q_2$ dyadic cubes of equal side length and with a touching edge.

Since $\tilde{\phi}$ is the extension in the time direction by reflection and tiling (c.f. (8.10)), and $Q_1, Q_2$ and $Q_3$ are all dyadic cubes we may assume that if $l(Q_1) \leq r$ then $Q_1, Q_2 \subset \{ |t| < r^2 \}$, and if $l(Q_1) > r$ then $\{ |t| < r^2 \} \subset Q_1$.

If $Q_1, Q_2 \subset Q_{2R}$ then $|(\rho \psi)_{Q_1} - (\rho \psi)_{Q_2}| \lesssim \|\rho \psi\|_{*,\text{dyadic}, Q_{2R}}$. Therefore, if we show (8.13) for $f = \rho \psi$ then by Lemmas 8.2 and 8.4 clearly

$$|(\rho \psi)_{Q_1} - (\rho \psi)_{Q_2}| \lesssim \|\rho \psi\|_{*,\text{dyadic}, Q_{2R}} \leq \|\nabla \tilde{\phi}\|_{*,\text{dyadic}, Q_{2R}} \lesssim \eta.$$

Now look at the other cases: $Q_1 \subset Q_{2R}$ and $Q_2 \cap Q_{2R} = \emptyset$, or $Q_2 \subset Q_1$ and $Q_2 \cap Q_{2R} = \emptyset$. In both cases, we wish to control $|(\rho \psi)_{Q_1}|$.

**Step 1.a:** Case $Q_1 \subset Q_{2R}, Q_2 \cap Q_{2R} = \emptyset$ and $l(Q_1) \lesssim \frac{R}{\eta^2}$.

$Q_1$ is small here and touches the boundary of $Q_{2R}$. This means that $\|\rho\|_{L^\infty(Q_1)} \lesssim \frac{l(Q_1)}{R}$ since $\rho$ is 0 outside $Q_{2R}$. Therefore we apply the trivial bound

$$|(\rho \psi)_{Q_1}| \leq \|\rho\|_{L^\infty(Q_1)} \|\psi\|_{L^\infty(Q_1)} \lesssim \frac{l(Q_1)}{R} \lesssim \eta.$$

**Step 1.b:** Case $Q_1 \subset Q_{2R}, Q_2 \cap Q_{2R} = \emptyset$ and $l(Q_1) \lesssim 2R$.

Since $Q_1 \subset Q_{2R}$ we have $\frac{R}{\eta^2} \lesssim l(Q_1) \leq 2R$. $Q_1$ is dyadic so there exists $N \in \mathbb{Z}$ such that $l(Q_1) = 2^N l(Q_r)$.

**Step 1.b.1:** $N \leq 0$.

This means that $l(Q_1) \leq l(Q_r)$ and so by the reflection and tiling in time, (8.10), we may assume $Q_1 \subset \{ |t| \leq r^2 \}$ and by Lemma 8.2

$$|(\rho \psi)_{Q_1}| \leq |\psi|_{Q_1} = \frac{1}{|Q_1|} \int_{Q_1} |\nabla \tilde{\phi} - \nabla \phi_{Q_1}| \leq \frac{1}{|Q_1|} \int_{Q_1} |\nabla \tilde{\phi} - \nabla \phi_{Q_1}| + |\nabla \phi_{Q_1} - \nabla \phi_{Q_2}| \lesssim \eta + \eta \log (1 + \ell) + \eta \log (1 + 1/\eta) \lesssim \eta^{1-\varepsilon}.$$

**Step 1.b.ii:** $N > 0$.

By Claim 8.3 we obtain

$$|(\rho \psi)_{Q_1}| \leq |\psi|_{Q_1} = \frac{1}{|Q_1|} \int_{Q_1} |\nabla \tilde{\phi} - \nabla \phi_{Q_1}| \lesssim \eta^{1-\varepsilon}.$$

**Step 1.c:** Case $Q_{2R} \subset Q_1, Q_2 \cap Q_{2R} = \emptyset$ so $l(Q_1) \geq 2R$.

Let $N$ satisfy $l(Q_1) = 2^N l(Q_{2R})$, the number of dyadic generations separating $Q_1$ and $Q_{2R}$. Then $Q_1$ overlaps $Q_{2R}$ (and its dyadic translates in time) exactly $2^N$ times. Therefore by claim 8.3,

$$|(\rho \psi)_{Q_1}| \leq |\psi|_{Q_1} \leq \frac{2^N}{|Q_1|} \int_{Q_{2R}} |\nabla \tilde{\phi} - \nabla \phi_{Q_2}| \leq \frac{2^N}{2^{N(1+\varepsilon)}} \frac{1}{|Q_{2R}|} \int_{Q_{2R}} |\nabla \tilde{\phi} - \nabla \phi_{Q_2}| \lesssim \eta^{1-\varepsilon}.$$

Hence, modulo the unproved statement $\|\rho \psi\|_{*,\text{dyadic}, Q_{2R}} \lesssim \eta$ we have shown

$$|(\rho \psi)_{Q_1} - (\rho \psi)_{Q_2}| \lesssim \eta^{1-\varepsilon} + \eta \log (1 + \ell).$$
Step 2: (8.13) holds, that is: \( \|\rho \phi\|_{L^\infty(Q)} \lesssim c(\eta) \).

To apply Lemma 8.4 we need to control two terms

\[
\sup_{Q \text{ dyadic}} \|\rho\|_{L^\infty(Q)} \frac{1}{|Q|} \int_Q |\psi - \psi_Q| 
\]

and

\[
\sup_{Q \text{ dyadic}} \frac{\psi_Q}{|Q|} \int_Q |\rho - \rho_Q|. 
\]

Step 2.a: Estimating \( \sup_{Q \text{ dyadic}} \|\rho\|_{L^\infty(Q)} \frac{1}{|Q|} \int_Q |\psi - \psi_Q| \).

In all the following cases we bound \( \|\rho\|_{L^\infty(Q)} \leq 1 \).

Step 2.a.i: Case \( l(Q) \leq r \).

As before, by the reflection and tiling in time, we may assume \( Q \subset \{|t| \leq r^2\} \) and so \( \nabla \phi = \nabla \phi \) on \( Q \). Hence

\[
\frac{1}{|Q|} \int_Q |\psi - \psi_Q| = \frac{1}{|Q|} \int_Q |\nabla \phi - (\nabla \phi)_{Q}| = \frac{1}{|Q|} \int_Q |\nabla \phi - (\nabla \phi)_{Q}| \lesssim \eta. 
\]

Step 2.a.ii: Case \( r < l(Q) \leq 2R \).

Applying claim 8.3 gives

\[
\frac{1}{|Q|} \int_Q |\psi - \psi_Q| \leq |\psi_Q| \lesssim \eta^{1-\varepsilon}. 
\]

Step 2.a.iii: Case \( 2R < l(Q) \).

From step 1.c it follows that

\[
\frac{1}{|Q|} \int_Q |\psi - \psi_Q| \leq |\psi_Q| \lesssim \eta^{1-\varepsilon}. 
\]

Step 2.b: Estimating \( \sup_{Q \text{ dyadic}} \frac{\psi_Q}{|Q|} \int_Q |\rho - \rho_Q| \).

We have the following three cases to consider.

Step 2.b.i: Case \( Q \subset Q_{2R} \), \( l(Q) \leq r \) and \( Q \subset \{|t| \leq r^2\} \).

Because the cube \( Q \) might not be touching the boundary we can’t follow step 1.a and bound \( \frac{1}{|Q|} \int_Q |\rho - \rho_Q| \) by \( \|\rho\|_{L^\infty(Q)} \), which here is likely be 1. However, we can use the mean value theorem and get a better bound. By the intermediate value theorem there exists \( (z, \tau) \in Q \) such that \( \rho(z) = \rho_Q \) and using that \( \rho \) is independent of time and \( |\nabla \rho| \lesssim 1/R \) we have

\[
|\rho(x) - \rho_Q| = |\rho(x) - \rho(z)| \leq |\nabla \rho||l(Q)| \lesssim \frac{l(Q)}{R} \leq \frac{l(Q)}{r}. 
\]

Then applying Lemma 8.2 gives

\[
\frac{|\psi_Q|}{|Q|} \int_Q |\rho - \rho_Q| \lesssim \frac{l(Q)}{r} \frac{1}{|Q|} \int_Q |\nabla \phi - \nabla \phi_{Q}|| \leq \frac{l(Q)}{r} \frac{1}{|Q|} \int_Q |\nabla \phi - \nabla \phi_{Q}|| \lesssim \frac{l(Q)}{r} \log \left( 2 + \frac{r}{l(Q)} \right) \eta \lesssim \eta. 
\]

Step 2.b.ii: Case \( Q \subset Q_{2R} \) and \( r < l(Q) \leq 2R \).

This case is a straightforward application of claim 8.3

\[
\frac{|\psi_Q|}{|Q|} \int_Q |\rho - \rho_Q| \leq |\psi_Q| \lesssim \eta^{1-\varepsilon}. 
\]

Step 2.b.iii: Case \( Q_{2R} \subset Q \) so \( l(Q) > 2R \).
This follows similarly to step 1.c; let \( N \) be defined as there and
\[
\frac{|\psi_Q|}{|Q|} \int_Q |\rho - \rho_Q| \leq \frac{1}{|Q|} \left| \int_Q \nabla \phi - \nabla \phi_{Q2n} \right| \leq \frac{2^N}{2^{(n+1)}} \|\nabla \phi\|_{*,Q2n} \leq \eta.
\]

Therefore by Lemma 8.1, \( \|\rho \psi\|_* \leq \eta^{1-\varepsilon} + \eta \log(1 + \ell) \).

It remains to tackle the harder piece \( \nabla \rho \theta = \nabla \rho(\tilde{\phi} - x \cdot \nabla \tilde{\phi}_{Q_r}) \). Recall that \( \text{supp}(\nabla \rho) = \{ r \leq |x| \leq 2R \} \).

Step 3: (8.14) holds; that is: \( \text{supp}(Q_1, Q_2) \| (\nabla \rho \theta)_{Q_{1}} - (\nabla \rho \theta)_{Q_{2}} \| \leq C(\eta) \) where \( Q_1, Q_2 \) are dyadic with a touching edge and \( l(Q_1) = l(Q_2) \).

There are two different cases to consider:

1. \( Q_1 \cap \text{supp}(\nabla \rho) \neq \emptyset \) and \( Q_2 \cap \text{supp}(\nabla \rho) \neq \emptyset \)
2. \( Q_1 \cap \text{supp}(\nabla \rho) \neq \emptyset \) and \( Q_2 \cap \text{supp}(\nabla \rho) = \emptyset \)

Again case (1) is controlled by \( \|\nabla \rho\|_* \), dyadic, \( Q_{2n} \) by Lemma 8.2 so we only have to deal with case (2) and bound \( \sup \text{ sup } Q_1 \text{ dyadic } \| (\nabla \rho \theta)_{Q_{1}} \| \).

Step 3.a: Case \( Q_1 \subset Q_{2R} \) and \( l(Q_1) \leq \frac{R}{2\eta} \).

In this case \( Q_1 \) touches the boundary of the support of \( \nabla \rho \) so we have the estimate \( \|\nabla \rho\|_{L^\infty(Q_1)} \leq \frac{l(Q_1)}{R^2} \). Also \( \phi(0,0) = 0 \) and \( \phi \in \text{Lip}(1,1/2) \) so \( \|\phi(x,t)\|_{L^\infty(Q_1)} \leq \|\phi(x,t)\|_{L^\infty(Q_{2n})} \leq \ell_R \). Finally \( \|x \cdot \nabla \tilde{\phi}_{Q_r}\|_{L^\infty(Q_{2n})} \leq \ell_R \). Therefore
\[
\| (\nabla \rho \theta)_{Q_{1}} \| \leq \|\nabla \rho\|_{L^\infty(Q_1)} |\theta|_{Q_{1}} \leq \frac{l(Q_1)}{R^2} \frac{1}{|Q_1|} \int_{Q_1} |\tilde{\phi}(x,t) - x \cdot \nabla \tilde{\phi}_{Q_r}| \, dx \, dt \leq \frac{l(Q_1)}{R^2} \ell_R \leq \eta.
\]

Step 3.b: Case \( Q_1 \subset Q_{2R} \) and \( \frac{R \eta}{\ell} \leq l(Q_1) \leq 2R \).

By the fundamental theorem of calculus, we may write
\[
\tilde{\phi}(x,t) - \tilde{\phi} \left( r \frac{x}{|x|}, t \right) = x \cdot \int_{r/|x|}^1 \nabla \tilde{\phi}(\lambda x, t) \, d\lambda.
\]

Therefore
\[
\| (\nabla \rho \theta)_{Q_{1}} \| \leq \|\nabla \rho\|_{L^\infty(Q_1)} \| \theta \|_{Q_{1}} \leq \|\nabla \rho\|_{L^\infty(Q_1)} \| \phi \|_{L^\infty(Q_{2n})} \| \phi(0,0) \|_{L^\infty(Q_{2n})} \leq \eta \frac{R \ell}{R} + \eta \frac{R \ell}{R}.
\]

Since \( \tilde{\phi} \) defined by (8.10) is tiled and reflected in time on cubes of scale \( r \), and \( (rx/|x|, 0) \in Q_r \), we control the first term above by
\[
\frac{1}{R} \left\| \tilde{\phi} \left( r \frac{x}{|x|}, t \right) - 0 \right\|_{L^\infty(Q_1)} \leq \frac{1}{R} \|\phi - \phi(0,0)\|_{L^\infty(Q_{2n})} \leq \ell \frac{R}{R} \leq \ell \eta.
\]
Recall that \( r \sim \eta R, \frac{R^2}{r^2} \lesssim l(Q_1) \leq 2R \) and \( r \leq |x|_\infty \leq 2R \) so \( \eta/2 \leq \lambda \leq 1 \). We apply Fubini to the second term
\[
\frac{1}{|Q_1|} \int_{Q_1} \left( \int_{r/|x|}^1 |\nabla \tilde{\phi}(\lambda x, t) - \nabla \tilde{\phi}_{Q_1}| \, d\lambda \right) \, dx \, dt
= \frac{1}{|Q_1|} \int_{Q_1} |\nabla \tilde{\phi}(\lambda x, t) - \nabla \tilde{\phi}_{Q_1}| \, dx \, dt \, d\lambda.
\]

Let \( \tilde{Q}_1 \) be the set formed by \( Q_1 \) under the transformation \((x, t) \mapsto (\lambda x, t)\). We may further cover \( \tilde{Q}_1 \) by \( \sim \lambda^{-2} \) translations of \( \lambda Q_1 \) with \( |\lambda Q_1|/|\tilde{Q}_1| \lesssim \lambda^2 \). Therefore a similar proof to claim 8.3, using Lemma 8.2, gives
\[
\frac{1}{|Q_1|} \int_{\tilde{Q}_1} |\nabla \tilde{\phi}(\lambda x, t) - \nabla \tilde{\phi}_{Q_1}| \, dx \, dt = \frac{1}{|Q_1|} \int_{\tilde{Q}_1} |\nabla \phi - \nabla \phi_{Q_1}| \lesssim \lambda^{-2} \frac{\lambda^2}{|Q_1|} \int_{Q_1} |\nabla \phi - \nabla \phi_{Q_1}| \lesssim \eta \log \left( 2 + \frac{r}{s l(Q_1)} \right) \lesssim \eta \log \left( 1 + \frac{\ell}{\eta^2} \right)
\lesssim \varepsilon \eta^{1-\varepsilon} + \eta \log(1 + \ell)
\]
and hence after harmlessly integrating in \( \lambda \) we can control the second term by
\[
\int_{\eta/2}^{1} \eta \log \left( 1 + \frac{\ell}{\eta^2} \right) \, d\lambda \lesssim \varepsilon \eta^{1-\varepsilon} + \eta \log(1 + \ell).
\]

**Step 3.c:** Case \( l(Q_1) \geq 2R \).
As before in step 1.c, \( |(\nabla \rho \theta)_{Q_1}| \leq |(\nabla \rho \theta)_{Q_{2R}}| \), which can be further controlled by cubes that tile \( \text{supp}(\nabla \rho) \). Therefore, this case is bounded as in step 3.b.

**Step 4:** (8.13) holds; that is: \( \|\nabla \rho \theta\|_{*,\text{dyadic}} \lesssim c(\eta) \)

Here we have 3 cases to consider:

1. \( Q \subset Q_{2R} \)
2. \( Q \subset \mathbb{R}^n \setminus \text{supp}(\nabla \rho) \)
3. \( Q_{2R} \subset Q \)

Case (2) is obvious. Case (3) reduces down to case (1) by step 1.c, the reflection and tiling of \( \phi \), and the \( \text{supp}(\nabla \rho) \).

Case (1): Using Lemma 8.4 this reduces down to showing that
\[
(a) \frac{\theta_{|Q|}}{|Q|} \int_Q |\nabla \rho - (\nabla \rho)_{Q}| \lesssim c(\eta)
\]
\[
(b) \frac{1}{|Q|} \int_Q |\nabla \rho(\theta - \theta_{Q})| \lesssim c(\eta)
\]
for \( Q \) dyadic and \( Q \subset Q_{2R} \).

**Step 4.a:** (a) holds for \( Q \) dyadic and \( Q \subset Q_{2R} \).

**Step 4.a.i:** Case \( Q \subset Q_{2R} \) and \( l(Q) \lesssim R^2 \).

By the naive bounds in step 3.a \( |\theta_{|Q|}| \lesssim \ell R \). If we use the mean value theorem for \( \nabla \rho \) similar to Step 2.2.b.i then
\[
\frac{1}{|Q|} \int_Q |\nabla \rho - (\nabla \rho)_{Q}| \lesssim |\nabla^2 \rho| |l(Q)| \lesssim \frac{l(Q)}{R^2}.
\]
Therefore
\[
\frac{|\theta_{|Q|}|}{|Q|} \int_Q |\nabla \rho - (\nabla \rho)_{Q}| \lesssim \ell R \frac{l(Q)}{R^2} \lesssim \eta.
\]

**Step 4.a.ii:** Case \( Q \subset Q_{2R} \) and \( \frac{R^2}{r^2} \lesssim l(Q) \leq 2R \).
Here we apply the same technique as step 3.b
\[
\frac{|\theta_{|Q|}|}{|Q|} \int_Q |\nabla \rho - (\nabla \rho)_{Q}| \leq |\theta_{|Q|}| |\nabla \rho| \lesssim \varepsilon \eta^{1-\varepsilon} + \eta \log(1 + \ell).
\]
Step 4.b: (b) holds for \( Q \) dyadic and \( Q \subset Q_{2R} \).

\[
\frac{1}{|Q|} \int_Q |\nabla \rho(\theta - \theta_Q)| \lesssim \frac{1}{R} \frac{1}{|Q|} \int_Q |\theta - \theta_Q|.
\]

We split this into the now usual cases.

Step 4.b.i: Case \( l(Q) \lesssim \frac{Rn}{\ell} \).

By the intermediate and mean value theorems \( |\hat{\phi} - \hat{\phi}_Q| \lesssim \ell(Q) \ell \) and \( |x-x_Q| \lesssim l(Q) \) so

\[
\frac{1}{|Q|} \frac{1}{R} \frac{1}{|Q|} \int_Q |\theta - \theta_Q| = \frac{1}{|Q|} \frac{1}{R} \frac{1}{|Q|} \int_Q |\hat{\phi} - \hat{\phi}_Q - x \cdot \nabla \hat{\phi}_Q + (x \cdot \nabla \hat{\phi}_Q)_{|Q}| \\
\lesssim \frac{1}{R} l(Q)\ell \lesssim \eta.
\]

Step 4.b.ii: Case \( \frac{Rn}{\ell} \lesssim l(Q) < 2R \).

\[
\frac{1}{|Q|} \frac{1}{R} \frac{1}{|Q|} \int_Q |\theta - \theta_Q| \lesssim \frac{1}{R} |\theta|_Q
\]

then applying the result from step 3.b gives

\[
\frac{1}{|Q|} \int_Q |\nabla \rho(\theta - \theta_Q)| \lesssim \eta^{1-\varepsilon} + \eta \log(1 + \ell).
\]

Therefore by Lemma 8.1 we have shown \( \nabla \Phi \in \text{BMO}(\mathbb{R}^n) \) and the bound (8.21) holds.

To finish proving Theorem 2.8 we need to establish property (iv).

Lemma 8.6. Let \( \Phi : \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R} \) be defined in (8.11) with

\[
\sup_{Q_s = J_s \times I_s, \, Q_s \subset Q_{\text{hy}c, x \leq r}} \frac{1}{|Q_s|} \int_Q \int_{I_s} \frac{|\phi(x, t) - \phi(x, \tau)|^2}{|t - \tau|^2} \, d\tau \, dt \, dx \leq \eta^2
\]

then \( \Phi \) satisfies

\[
\sup_{Q_s = J_s \times I_s} \frac{1}{|Q_s|} \int_Q \int_{I_s} \frac{|\Phi(x, t) - \Phi(x, \tau)|^2}{|t - \tau|^2} \, d\tau \, dt \, dx \lesssim \eta^2.
\]

Proof. Trivially since \( \Phi \) is defined globally

\[
\sup_{Q_s = J_s \times I_s} \frac{1}{|Q_s|} \int_Q \int_{I_s} \frac{|\Phi(x, t) - \Phi(x, \tau)|^2}{|t - \tau|^2} \, d\tau \, dt \, dx \\
\leq \sup_{Q_s = J_s \times I_s} \frac{1}{|Q_s|} \int_Q \int_{I_s} \frac{|\hat{\phi}(x, t) - \hat{\phi}(x, \tau)|^2}{|t - \tau|^2} \, d\tau \, dt \, dx,
\]

where we interpret the value of \( \hat{\phi} \) where it is undefined as 0, i.e. \( \hat{\phi}(x, t) = 0 \) when \( (x, t) \notin \text{supp}(\hat{\phi}) \). It remains to establish

\[
\sup_{I_s \subset I} \frac{1}{|I_s|} \int_I \int_{I_s} \frac{|\phi(x, t) - \hat{\phi}(x, \tau)|^2}{|t - \tau|^2} \, d\tau \, dt \, dx \\
\lesssim \sup_{I_s \subset I} \frac{1}{|I_s|} \int_I \int_{I_s} \frac{|\phi(x, t) - \phi(x, \tau)|^2}{|t - \tau|^2} \, d\tau \, dt \, dx
\]

pointwise in \( x \), where \( Q_r = J_r \times I_r \) and is used to define \( \Phi \) in (8.11). To simplify our notation, we drop the dependence on the spatial variables in \( \hat{\phi} \) and \( \phi \). We also
set $A := I_s$. Recall from (8.10) that

$$
\tilde{\phi}(t) = \begin{cases} 
\phi(t) & t \in [-r^2, r^2], \\
\phi(2r^2 - t) & t \in [r^2, 3r^2] + 4kr^2,
\end{cases}
$$

for $k \in \mathbb{Z}$. Let $I_k = [-r^2, r^2] + 4kr^2$ and $J_k = [r^2, 3r^2] + 4kr^2$ be intervals in time for $k \in \mathbb{Z}$. We partition $A$ into disjoint pieces $A = \bigcup_i I_i \cup \bigcup_j J_j \cup A_1 \cup A_2$, where $A_1$ and $A_2$ are pieces that don’t contain either $I_i$ or $J_j$.

If $A = A_1 \cup A_2$ we may as well assume (by translation and reflection) that $A_1 = [a, r^2]$, $A_2 = [r^2, b]$. Let $\tau', b'$ and $A_2'$ be the images of $\tau, b$ and $A_2$ respectively under the map $\tau \mapsto 2r^2 - \tau$. Without loss of generality we only consider the case $|A_1| > |A_2|$. Since $|t - \tau| = |t - r^2| + |r' - r^2| \geq |t - \tau'|$ we have for $t \in A_1$, $\tau \in A_2$

$$
\int_{A_1} \int_{A_2} \frac{|\tilde{\phi}(t) - \tilde{\phi}(\tau)|^2}{|t - \tau|^2} \, d\tau \, dt = \int_{a}^{r^2} \int_{a'}^{r^2} \frac{|\phi(t) - \phi(\tau')|^2}{|t - (2r^2 - \tau')|^2} \, d\tau' \, dt \\
\leq \int_{a}^{r^2} \int_{a'}^{r^2} \frac{|\phi(t) - \phi(\tau')|^2}{|t - \tau'|^2} \, d\tau' \, dt \leq \int_{A_1} \int_{A_1} \frac{|\phi(t) - \phi(\tau')|^2}{|t - \tau'|^2} \, d\tau' \, dt.
$$

Therefore

$$
\frac{1}{|A|} \int_{A} \int_{A} \frac{|\tilde{\phi}(t) - \tilde{\phi}(\tau)|^2}{|t - \tau|^2} \, d\tau \, dt = \frac{1}{|A|} \left( \int_{A_1} \int_{A_1} + 2 \int_{A_1} \int_{A_2} + \int_{A_2} \int_{A_2} \right) \frac{|\phi(t) - \phi(\tau')|^2}{|t - \tau'|^2} \, d\tau' \, dt \lesssim \eta^2.
$$

In the general case when $A = \bigcup_{i \in \mathbb{Z}} I_i \cup \bigcup_{j \in \mathcal{J}} J_j \cup A_1 \cup A_2$ we write the double integral over $A$ in terms of integrals

$$
\sum_{i, k \in \mathbb{Z}} \int_{I_i} \int_{I_k} \frac{|\tilde{\phi}(t) - \tilde{\phi}(\tau)|^2}{|t - \tau|^2} \, d\tau \, dt, \quad \sum_{i \in \mathbb{Z}, j \in \mathcal{J}} \int_{I_i} \int_{J_j} \frac{|\tilde{\phi}(t) - \tilde{\phi}(\tau)|^2}{|t - \tau|^2} \, d\tau \, dt
$$

and integrals that involve sets $A_1$ or $A_2$ or both (those are handled similar to the earlier calculation).

Dealing with the first case, if $i \neq k$, $t \in I_i$ and $\tau \in I_k$ then $|t - \tau| \sim r^2|i - k|$; if $i = k$ then $|t - \tau| \sim |t' - \tau'|$. Therefore

$$
\sum_{i, k \in \mathbb{Z}} \int_{I_i} \int_{I_k} \frac{|\tilde{\phi}(t) - \tilde{\phi}(\tau')|^2}{|t - \tau'|^2} \, d\tau \, dt \\
\lesssim \sum_{i \in \mathbb{Z}} \int_{I_0} \int_{I_0} \frac{|\phi(t) - \phi(\tau)|^2}{|t - \tau|^2} \, d\tau \, dt + \sum_{i, k \in \mathbb{Z}} \frac{1}{r^4|i - k|^2} \int_{I_0} \int_{I_0} \frac{|\phi(t) - \phi(\tau)|^2}{|t - \tau|^2} \, d\tau \, dt
$$

$$
\leq \sum_{i \in \mathbb{Z}} \int_{I_0} \int_{I_0} \frac{|\phi(t) - \phi(\tau)|^2}{|t - \tau|^2} \, d\tau \, dt + \sum_{i, k \in \mathbb{Z}} \frac{1}{|i - k|^2} \int_{I_0} \int_{I_0} \frac{|\phi(t) - \phi(\tau)|^2}{|t - \tau|^2} \, d\tau \, dt
$$

$$
\lesssim |Z| \int_{I_0} \int_{I_0} \frac{|\phi(t) - \phi(\tau)|^2}{|t - \tau|^2} \, d\tau \, dt.
$$
In the second case
\[
\sum_{i \in I, j \in J} \int_{I_i} \int_{J_j} \frac{\phi(t) - \phi(\tau)}{|t - \tau|^2} \, d\tau \, dt
\]
\[
\lesssim \sum_{i \in I, j \in J, |i - j| \leq 1} \int_{I_0} \int_{J_0} \frac{\phi(t) - \phi(\tau)}{|t - \tau|^2} + \sum_{i \in I, j \in J, |i - j| \geq 2} \frac{1}{r^4(|i - j| - 1)^2} \int_{I_0} \int_{J_0} \frac{\phi(t) - \phi(\tau)}{|t - \tau|^2} \, d\tau \, dt.
\]
\[
\lesssim (|I| + |J|) \int_{I_0} \int_{J_0} \frac{\phi(t) - \phi(\tau)}{|t - \tau|^2} \, d\tau \, dt.
\]
Since $|A| \sim (|I| + |J|)|I_0|$ and $I_0$ is one of the time intervals considered in the supremum of (8.24)
\[
\frac{1}{|A|} \int_A \int_A \frac{\phi(t) - \phi(\tau)}{|t - \tau|^2} \, d\tau \, dt \sim \frac{1}{|I_0|} \int_{I_0} \int_{I_0} \frac{\phi(t) - \phi(\tau)}{|t - \tau|^2} \, d\tau \, dt \lesssim \eta^2.
\]

\section*{REFERENCES}


School of Mathematics, The University of Edinburgh and Maxwell Institute of Mathematical Sciences, UK
Email address: m.dindos@ed.ac.uk

School of Mathematics, The University of Edinburgh and Maxwell Institute of Mathematical Sciences, UK
Email address: luke321321@gmail.com

Department of Mathematics, Yonsei University, Republic of Korea
Email address: sukjung_hwang@yonsei.ac.kr