International Journal of Pure and Applied Mathematics

Volume 24 No. 1 2005, 59-106

WEIGHTED NORM INEQUALITIES FOR PARABOLIC GRADIENTS ON NON-SMOOTH DOMAINS

Caroline Sweezy[1 §], J. Michael Wilson[2]

[1]Department of Mathematical Sciences, MSC 3MB
College of Arts and Sciences
New Mexico State University
P.O. Box 30001, Las Cruces, NM 88003-8001, USA
e-mail: csweezy@nmsu.edu

[2]Department of Mathematics
University of Vermont
Burlington, VT 05405, USA
e-mail: wilson@cem.uvm.edu

Abstract: For a bounded domain Ω in $\mathbb{R}^{d+1}$ sufficient conditions on measures $d\mu(x,t)$ and $v(y',s)d\omega(y',s)$ are established to prove the inequality

$$\left(\int_\Omega |\nabla u(x,t)|^q \, d\mu(x,t)\right)^{1/q} \leq C \left(\int_{\partial_p\Omega} \left|f(y',s)\right|^p v(y',s) d\omega(y',s)\right)^{1/p}$$

for $1 < p \leq q < \infty$ and $q \geq 2$, where $(\partial/\partial t - L)u = 0$ in Ω and $u\mid_{\partial_p\Omega} = f$ on the part of $\partial\Omega$ that is called the parabolic boundary of Ω and designated by $\partial_p\Omega$. Ω is a non-cylindrical domain in $\mathbb{R}^{d+1} = \{(x,t) : x = (x', x_d), x' \in \mathbb{R}^{d-1}, x_d \in \mathbb{R}, t \in \mathbb{R}\}$. The lateral boundary of Ω can be as rough as the parabolic analogue of a Lipschitz domain, i.e. it can be a $\mathrm{Lip}(1,1/2)$ domain. $\partial/\partial t - L$ is a strictly parabolic divergence form operator with time dependent coefficients that are bounded and measurable, and $\omega(y',s)$ is parabolic measure on the parabolic boundary, $\partial_p\Omega$, of the domain Ω. The method of proof follows Wheeden and Wilson in using the dual operator and a Littlewood-Paley type inequality. The Littlewood-Paley inequality is proved for functions on $\partial_p\Omega$, that

Received: August 10, 2005

§Correspondence author

can be expressed as $\sum \lambda_I \varphi_{(I)}$, where $\{\varphi_{(I)}\}$ is a family of functions which have good decay properties, and are indexed by parabolic cubes I.

AMS Subject Classification: 35K20, 42B25
Key Words: parabolic equations, boundary value problems, weighted norm inequalities, Littlewood-Paley theory, non-smooth domains

1. Introduction

In [14], Wheeden and Wilson considered this general question: Suppose we solve the Dirichlet problem on the upper half space $\mathbb{R}^{d+1}_+$ with "reasonable" boundary data (e.g. bounded and measurable with compact support). To what extent is the smoothness of the solution controlled by the size of the boundary values? There are many ways to make this rather vague question quantitative. Wheeden and Wilson approached it by looking for conditions on measures μ (defined on $\mathbb{R}^{d+1}_+$) and weights ν (defined on $\mathbb{R}^d$) which ensured that the inequality

$$\left(\int_{\mathbb{R}^{d+1}_+} |\nabla u|^q \, d\mu\right)^{1/q} \leq C \left(\int_{\mathbb{R}^d} |f|^p \, v dx\right)^{1/p} \tag{1}$$

will hold for all f in the test class, for numbers p and q which are assumed to lie strictly between 1 and ∞. As usual we are using u to denote f's Poisson integral, with ∇u being the full gradient:

$$\nabla u =< \partial u/\partial x_0, \partial u/\partial x_1, \partial u/\partial x_2, ..., \partial u/\partial x_d >, \quad \text{where } x_0 \text{ is } y.$$

It is clear that (1) should hold if, roughly speaking, μ does not put too much mass near sets where ν gets small; or, phrased another way, if μ does not put too much mass near places where some negative power of ν is large. The sufficient conditions obtained in [14] were exact quantitative statements of this intuitive truth.

Wheeden and Wilson attacked (1) by, first considering one derivative at a time, and then rephrasing their simplified version of (1) in a dual form. This dual formulation is fairly straightforward. Any of the partial derivatives $\partial u/\partial x_i$ is given by a convolution:

$$\frac{\partial u}{\partial x_i}(t, y) = f * (y^{-1}\psi_y(t)),$$

where ψ is obtained by taking a derivative (or in the case of x_0 several derivatives) of the Poisson kernel, and the subscript y denotes the usual L^1 dilation. The dual of the operator that takes f into $\partial u/\partial x_i$ is an operator T, defined on bounded compactly supported functions $g : \mathbb{R}^{d+1}_+ \to \mathbb{R}$, and given by the formula

$$Tg(x) = \int_{\mathbb{R}^{d+1}_+} g(t,y)y^{-1}\psi_y(t-x)d\mu(t,y).$$

Functional analysis implies that (1) will hold for all bounded compactly supported f if

$$\left(\int_{\mathbb{R}^d} |Tg(x)|^{p'} (\nu(x))^{1-p'} dx\right)^{1/p'} \leq \left(\int_{\mathbb{R}^{d+1}_+} |g(t,x)|^{q'} d\mu(t,x)\right)^{1/q'} \tag{2}$$

holds for all bounded compactly supported g, where we are using p' and q' to denote the dual indices of p and q.

Inequality (2) is really an object from Littlewood-Paley theory. This will become clearer if we rewrite the integral that defines Tg. Recall that a dyadic cube Q is a Cartesian product of the form

$$Q = [\frac{j_1}{2^k}, \frac{j_1+1}{2^k}) \times [\frac{j_2}{2^k}, \frac{j_2+1}{2^k}) \times ... \times [\frac{j_d}{2^k}, \frac{j_d+1}{2^k}),$$

where the numbers k and j_i are integers. We say such a cube has side length – denoted by $l(Q)$ – equal to 2^{-k}. The collection of all dyadic cubes we denote by $\mathcal{D}$. To each dyadic cube we associate a subset $T(Q) \stackrel{\circ}{=} Q \times [l(Q)/2, l(Q))$ of $\mathbb{R}^{d+1}_+$. Since the family $\{T(Q)\}_{Q\in\mathcal{D}}$ clearly tiles $\mathbb{R}^{d+1}_+$, we can rewrite Tg's integral as:

$$\begin{aligned} Tg(x) &= \sum_{Q\in\mathcal{D}} \int_{T(Q)} g(t,y)y^{-1}\psi_y(t-x)d\mu(t,y) \\ &= \sum_{Q\in\mathcal{D}} b_{(Q)}(x) = \sum_{Q\in\mathcal{D}} \lambda_Q \phi_{(Q)}(x), \end{aligned} \tag{3}$$

where the λ_Q's are non-negative numbers (depending on g and μ) and the $\phi_{(Q)}$'s are continuously differentiable functions satisfying certain decay and smoothness conditions that depend on Q, as well as the cancellation condition:

$$\int \phi_{(Q)}(x)dx = 0. \tag{4}$$

The decay and smoothness conditions we have in mind are:

$$\left|\phi_{(Q)}(x)\right| \leq |Q|^{-1/2} \left(1 + |x - x_Q| / l(Q)\right)^{-M} , \tag{5}$$

$$\left|\nabla\phi_{(Q)}(x)\right| \leq l(Q)^{-1} |Q|^{-1/2} \left(1 + |x - x_Q| / l(Q)\right)^{-M-1} , \tag{6}$$

where M is a fixed number larger than d, x_Q is Q's center, and $|Q|$ denotes Q's Lebesgue measure.

Written as it is in (3), Tg is seen to be a (finite!) linear sum of smoothly varying functions with good decay and cancellation. We are trying to control its size, in some weighted space $L^{p'}(\mathbb{R}^d, \nu^{1-p'} dx)$, in terms of the sizes of its coefficients λ_Q; and that is the business of weighted Littlewood-Paley theory.

In [14], the authors were able to reduce the analysis of sums of the form (3) to the case in which the $\phi_{(Q)}$'s have compact supports (bounded dilates of the Q's). This approach made liberal use of the extra decay (to order $-M-1$) in the derivatives of the $\phi_{(Q)}$ and the translation invariance of $\mathbb{R}^d$.

In this paper we generalize the study of inequalities such as (1) to solutions of parabolic equations on domains with minimally smooth boundaries. That is, suppose we have a domain $\Omega \subseteq \mathbb{R}^{d+1}$ and a strictly parabolic operator $\partial/\partial t - L$ defined on Ω. We look at solutions to the boundary value problem:

$$\begin{aligned} &(\partial/\partial t - L)u = 0 \ \text{ in } \ \Omega\,, \\ &u = f \ \text{ on } \ \partial_p\Omega, \end{aligned}$$

where $\partial_p\Omega$ is the parabolic boundary of Ω, and f is assumed to belong to a reasonable class of test functions defined on the parabolic boundary. The exact nature of the domains Ω under consideration, as well as the precise definitions of "strictly parabolic" and "parabolic boundary" will be given in the next section; they are not important right now. However, a good first approximation to our situation is given by taking $\Omega = \{(x,t) : |x| < 1, 0 < t < T\}$, $L = \Delta =$ the usual Laplacian operator, and $\partial_p\Omega = \{(x,0) : |x| \leq 1\} \cup \{(x,t) : |x| = 1, 0 \leq t < T\}$. The biggest changes we make are that Δ is replaced by L, a more general kind of elliptic operator with time dependent coefficients, and Ω is no longer a cylinder domain, but resembles a cylinder in having "top" and "bottom" parts of its boundary that are contained in single hyperplanes.

We prove the sufficiency of conditions on measures μ, defined on Ω, and certain weights $d\rho$, defined on $\partial_p\Omega$, which ensure that

$$\left(\int_\Omega |\nabla u|^q \, d\mu\right)^{1/q} \leq \left(\int_{\partial_p\Omega} |f|^p \, d\rho\right)^{1/p} , \tag{7}$$

will hold for all f in our test class and for appropriate p and q lying between 1 and ∞. This is the subject of the paper.

We retain the basic duality approach of [14]. By suitably localizing, we can even reduce our problem to the study of linear sums such as that in (3), where the $\phi_{(Q)}$'s are continuous functions, with controlled decay and smoothness, and satisfying a cancellation condition similar to (4), in which the integral is taken with respect to an appropriate measure ω defined on a part of the boundary of our domain. The measure ω is one of the (infinitely many) "harmonic measures" of the parabolic operator. The non-isotropic geometry of the operator makes it more natural to index the $\phi_{(Q)}$'s, not over dyadic cubes, but over certain non-isotropic analogues of dyadic cubes. The exact definition of these sets will be given in the next section. Right now we note the important property they share with standard dyadic cubes, namely that, for any two "non-isotropic dyadic cubes" Q and Q', either Q and Q' are disjoint, or one is a subset of the other. The measure ω is known to satisfy a corresponding doubling property: there is a constant C so that, for all non-isotropic cubes Q, $\omega(2Q) \leq C\omega(Q)$, where $2Q$ denotes a non-isotropic analogue of Q's concentric double.

In this more general setting, we lose translation invariance. Smoothness also becomes more problematic. In particular, it is no longer useful to talk about $\nabla\phi_{(Q)}$. Instead we refer to $\phi_{(Q)}$'s modulus of Holder continuity (of some fixed order $\alpha > 0$).

The decay condition is rephrased as follows. For every dyadic cube Q, there is a function $\phi_{(Q)}$ such that, if $x \in Q$, then

$$\left|\phi_{(Q)}(x)\right| \leq \frac{1}{\sqrt{\omega(Q)}};$$

and, for $j \geq 1$, if $x \in 2^j Q \setminus 2^{j-1}Q$, then

$$\left|\phi_{(Q)}(x)\right| \leq 2^{-j\beta}\frac{\sqrt{\omega(Q)}}{\omega(2^j Q)},$$

where $\beta > 0$ is fixed and $2^j Q$ denotes Q's 2^j-fold dilate. We can express this more compactly by defining $R_0(Q) = Q$ and $R_j(Q) = 2^j Q \setminus 2^{j-1}Q$ for $j \geq 1$, and writing

$$\left|\phi_{(Q)}(x)\right| \leq \sqrt{\omega(Q)}\sum_{j=0}^{\infty}\frac{2^{-j\beta}}{\omega(2^j Q)}\chi_{R_j(Q)}(x). \tag{8}$$

For the decay condition, we ask that, for some fixed $\alpha > 0$, and all x and y in $\mathbb{R}^d$,

$$\left|\phi_{(Q)}(x) - \phi_{(Q)}(y)\right| \\ \leq \left(\frac{d_L(x,y)}{l(Q)}\right)^{\alpha} \sqrt{\omega(Q)} \sum_{j=0}^{\infty} \frac{2^{-j\beta}}{\omega(2^j Q)} (\chi_{R_j(Q)}(x) + \chi_{R_j(Q)}(y)), \quad (9)$$

where d_L denotes the metric associated with our operator. If the reader puts d_L equal to the usual Euclidean metric, ω equal to Lebesgue measure, and carefully rewrites (8), he will get (5), up to a multiplication by a positive constant. But if he sets $\alpha = 1$ and similarly rewrites (9), he will not get (6), even if he reinterprets (6) in terms of moduli of continuity. This is because (9) has decay no better than that in (8). This unfortunate fact is an essential part of the more general setting, and it puts a roadblock to the approach taken in [14].

What is needed is a way to obtain Littlewood-Paley estimates for linear sums of functions that satisfy (8), (9) and a cancellation condition (which, note, has yet to be specified), *without* reducing to the case of locally supported $\phi_{(Q)}$'s.

Such a method exists. It is based on a stopping-time argument, first described in [15]. It is worthwhile to show how this argument will play out in our present setting.

We suppose we have a function $f = \sum_{Q \in \mathcal{D}} \lambda_Q \phi_{(Q)}$, a finite linear sum of functions satisfying (8), (9) and the following "cancellation condition": For every finite linear combination $\sum_Q \gamma_Q \phi_{(Q)}$,

$$\int | \sum_Q \gamma_Q \phi_{(Q)} |^2 \, d\omega \leq \sum_Q |\gamma_Q|^2 . \quad (10)$$

The stopping time argument yields the following: For all $0 < p < \infty$,

$$\int |f|^p \, d\omega \leq C \int (g^*(f))^p d\omega, \quad (11)$$

for

$$g^*(f)(x) = \left(\sum_Q |\lambda_Q|^2 \left(\sum_{j=0}^{\infty} \frac{2^{-j(2\beta-\tau)}}{\omega(2^j Q)} \chi_{R_j(Q)}(x) \right) \right)^{1/2},$$

where τ is some small positive number strictly smaller than β, and the constant C depends on p, ω, τ, d, α and β. Moreover, (11) holds for measures ν that are sufficiently "regular" with respect to ω. This type of regularity, called the A_∞ property, is defined below.

Inequalities similar to (11) were a key part of the method of [14], since it is fairly straightforward to use (11) to prove an inequality of the nature of (2). They play a similar role here.

Notice that (11) is really a Littlewood-Paley inequality: the function $g^*(f)$ is a "discretized" version of the $g^*_\lambda-$function from the classical theory.

Our proof of (11) will proceed by means of good-λ inequalities, and thus will imply the analogous inequality

$$\int |f|^p \, d\nu \le C \int (g^*(f))^p d\nu, \tag{12}$$

for all measures ν which are A_∞ with respect to ω. We write this in symbols as "$\nu \in A_\infty(\omega)$"; it simply means that there are positive constants a and b so that, for all cubes Q and measurable sets $E \subset Q$,

$$\frac{\nu(E)}{\nu(Q)} \le a \left(\frac{\omega(E)}{\omega(Q)} \right)^b .$$

For almost all of this paper, these cubes Q will be non-isotropic, but we will have occasion (see below) to mention the A_∞ property with respect to ordinary cubes as well. We will try to make it clear from the context what kind of cubes we are dealing with. The family of dyadic cubes, non-isotropic or otherwise, will be denoted by $\mathcal{D}$.

We want to say a few words about the stopping-time argument we will use to prove (12). It is designed to surmount a particular problem, namely, finding a good way to partially sum $f = \sum_{Q \in \mathcal{D}} \lambda_Q \phi_{(Q)}$. The difficulty is that the $\phi_{(Q)}$'s do not have local supports, which means that each $\phi_{(Q)}$ that is "centered" around a distant cube can seriously affect the behavior of f on a given cube Q_0. A kind of partial sum that has turned out to work pretty well is the following. For any cube Q in $\mathcal{D}$, let $S(Q) = \{Q' \in \mathcal{D} : Q' \not\subseteq Q\}$. The partial sum of f, conditioned for the cube Q, is defined to be

$$F(Q) \doteq \sum_{Q' \in S(Q)} \lambda_{Q'} \phi_{Q'}(x_Q).$$

Notice that we are evaluating the sum at the center of the cube Q. We similarly define a "partial sum" for $g^*(f)$:

$$G(Q) \doteq \left(\sum_{Q' \in S(Q)} \left|\lambda_{Q'}\right|^2 \left(\sum_{j=0}^{\infty} \frac{2^{-j(2\beta-\tau)}}{\omega(2^j Q')} \chi_{R_j(Q')}(x_Q) \right) \right)^{1/2} .$$

We define maximal functions from these objects:

$$F^*(x) \mathrel{\dot=} \sup_{Q:\, x\in Q} |F(Q)|\ , \quad G^*(x) \mathrel{\dot=} \sup_{Q:\, x\in Q} G(Q).$$

After some juggling – standard manipulations – the inequality (12) is seen to be a consequence of the following lemma.

Main Lemma. *Let* $\{\phi_{(Q)}\}_Q$ *be a family satisfying* (8), (9) *and* (10) *with respect to a doubling measure* ω. *For every* $\epsilon > 0$ *there is a* $\gamma > 0$ *such that the following is true: If* Q_0 *is any non-isotropic dyadic cube and* $f = \sum_Q \lambda_Q \phi_{(Q)}$ *is a fintie sum such that the only non-zero* λ_Q*'s correspond to subcubes of* Q_0, *then*

$$\omega(\{x \in Q_0 : F^*(x) > 1,\, G^*(x) \le \gamma\}) \le \epsilon\omega(Q_0).$$

Most of the paper is taken up with the proof of our present version of Main Lemma.

It is now time to make our "parabolic" setting precise. We will be dealing with solutions to the Dirichlet problem on a bounded domain $\Omega_T \subseteq \mathbb{R}^{d+1}$. For most of our results we will be working on a domain that is the analogue of a Lipschitz domain for the heat equation when the side boundary of the domain is allowed to vary with time. We will be using the parabolic metric; for (y, s) and (x, t) points in $\mathbb{R}^{d+1}$, the distance between them is

$$d_p(x, t; y, s) \mathrel{\dot=} |x - y| + \sqrt{|t - s|}.$$

To define the boundary of the domain we first define a local cylinder,

$$\psi_R(x_0, t_0) \mathrel{\dot=} \{(x, t) : d_p(x, t; x_0, t_0) < R/2\}.$$

The boundary of the domain, $\partial\Omega = T\Omega \cup B\Omega \cup S\Omega$, consists of three parts, a top, $T\Omega = \{(x, t) \in \partial\Omega :$ there exists some $\delta > 0,$ so that $\psi_\delta(x, t + \frac{\delta^2}{4}) \cap \Omega_T = \emptyset\}$, a bottom $B\Omega = \{(x, t) \in \partial\Omega :$ there exists some $r > 0,$ so that $\psi_r(x, t + \frac{1}{4}r^2) \cap \mathbb{R}^{d+1}\backslash\Omega_T = \emptyset\}$ and the side $S\Omega = \partial\Omega\backslash\{T\Omega \cup B\Omega\}$. We will assume that the top and bottom lie in single hyperplanes: $T\Omega \subset \{t = T\}$ and $B\Omega \subset \{t = 0\}$. The parabolic boundary of Ω is written $\partial_p\Omega \mathrel{\dot=} S\Omega \cup B\Omega$. $S\Omega$ will also have the property that for some fixed $r_0 > 0$, if $(x, t) \in S\Omega$ then the set $\psi_{r_0}(x, t) \cap S\Omega$ can be described as the graph of a function $\varphi(y', s)$, where (y', s) lies in a neighborhood of the origin, $N(0)$ in $\mathbb{R}^d$, after an appropriate rotation of axes. The map $\varphi(y', s)$ has the property that $|\varphi(y', s) - \varphi(x', t)| \le N\left(|x' - y'| + \sqrt{|t - s|}\right)$. In other words, $\psi_{r_0}(x, t) \cap S\Omega = \{(y', y_d, s) : y_d =$

$\varphi(y', s)$ for $(y', s) \in N(0)\}$, and we also assume that $\Omega_T \cap \psi_{r_0}(x, t) = \{(y', y_d, s) : y_d > \varphi(y', s)$ for $(y', s) \in N(0)\}$. Note: By a mild abuse of notation we will write (x', t) and (y', s) for points on $S\Omega$, even though $y_d \neq 0$ here. We will also use the notation (Q_n, s_n), $n \in \mathbb{N} \cup \{0\}$, to denote boundary points.

The next order of business is to define the operators we will be employing: $\partial/\partial t - L = \partial/\partial t - \sum_{i,j=1}^{d} \frac{\partial}{\partial x_i}(a_{i,j}(x, t)\frac{\partial}{\partial x_j})$ is the kind of operator we are considering. It is a strictly parabolic, divergence form operator whose coefficients are bounded and measurable. In other words it satisfies the condition $(1/\theta)\,|\xi|^2 \leq \sum_{i,j} \xi_i a_{i,j}(x, t)\xi_j \leq \theta\,|\xi|^2$ for a fixed constant $\theta \geq 1$, and for all (x, t) in Ω_T. A solution to the Dirichlet problem on Ω_T will be a function $u(x, t)$ that satisfies

$$\begin{aligned} &(\partial/\partial t - L)u(x, t) = 0 \text{ in } \Omega_T\,, \\ &u(x', t)\,|_{S\Omega} = f(x', t) \;\text{ and }\; u(x, 0) = 0\,. \end{aligned}$$

The assumption of the initial data being zero is for convenience. Our solution $u(x, t)$ has the property that it can always be written as an integral of the boundary function multiplied by the kernel function of $\partial/\partial t - L$ on Ω_T. This representation will be discussed further in Section 2.

Our "parabolic dyadic cubes" in $S\Omega$ can be defined as the images of ordinary non-isotropic cubes in $\mathbb{R}^d$ under the map φ that defines the lateral boundary of Ω_T. A non-isotropic dyadic cube in $\mathbb{R}^d$ has the form $[\frac{k_1}{2^n}, \frac{k_1+1}{2^n}) \times [\frac{k_2}{2^n}, \frac{k_2+1}{2^n}) \times \ldots \times [\frac{k_{d-1}}{2^n}, \frac{k_{d-1}+1}{2^n}) \times [\frac{k_t}{2^{2n}}, \frac{k_t+1}{2^{2n}})$, k_i and $n \in \mathbb{Z}$; the only difference from ordinary Euclidean cubes comes in the time dimension which is always the square of the space dimension. These cubes have the property that at each level (i.e. each value of n) they tile the plane, as a set they are nested or disjoint and their dimension in the parabolic metric is 2^n. On the parts of $S\Omega$, where there are overlapping images of dyadic cubes from $\mathbb{R}^d$, one can adjust the boundary maps to obtain a set of cubes in $\mathbb{R}^d$ that are disjoint. Or one can make a choice of parts of cubes whose dimensions in all directions are $\geq \frac{1}{2}$ the dimension of the original cube, and/or take regions that are the union of several parts of intersecting cubes, where the dimension of the region created by the union does not exceed a fixed multiple of the dimension of the original cube. To simplify the expressions in the proofs of the main results, our indexing of "dyadic parabolic" boundary cubes will refer to the collection of non-isotropic cubes in $\mathbb{R}^d$ rather than to the actual cubes on $\partial_p\Omega_T$. We will also be using "Carleson-type" regions inside Ω_T which are associated to each dyadic cube. These regions can be taken to be the images under the map φ of the Carleson boxes in $\mathbb{R}^{d+1}_+$associated to each non-isotropic cube in $\mathbb{R}^d$; they will be discussed further below. Both the parabolic "cubes" on $\partial_p\Omega_T$ and the associated Carleson-type regions in Ω_T will retain the property that their dimensions on a single level are comparable, and we can retain the property of the regions being nested or disjoint. In fact to prove Theorem 2 for solutions to general parabolic equations such as $(\partial/\partial t - L)u = 0$, where $L = \sum \frac{\partial}{\partial x_i}(a_{ij}(x,t)\frac{\partial}{\partial x_j})$, we can take these top halves of Carleson type regions be actual $(d+1)$-dimensional parabolic rectangles or cylinders, or, at least, a finite union of such rectangles (with at most a fixed number of them in a single top half of a Carleson-type region) of dimension comparable to that of the top half of the Carleson type region. It is easy to see that one can impose this restriction and still have a collection of parabolic rectangles or cylinders (possibly with bounded overlap) that covers a region that lies within a certain distance of the parabolic boundary of Ω_T. Dilations of these regions up to a fixed maximal dilation will have bounded overlap; the regions form a cover of the interior of Ω_T and provide the usual collection of regions whose dimension shrinks proportionally to their distance from the boundary of Ω_T. Similar constructions have been used by several authors; see [1], [10], [11] and [13] for details. The center-doubling property of our parabolic measure will allow us to carry out the geometric estimations on the parabolic cubes needed to obtain our results.

Choosing a fixed point (X_0, T) inside the top of Ω_T, the parabolic measure associated to the operator $\partial/\partial t - L$ on $\partial_p\Omega_T$ will be designated by $\omega = \omega^{(X_0,T)}$. As is well known ω is a Borel measure defined on sets in $\partial_p\Omega_T$. For any fixed Borel set $E \subseteq \partial_p\Omega_T$, $\omega^{(x,t)}(E)$ is the unique solution to the Dirichlet problem $(\partial/\partial t - L)(\omega^{(x,t)}(E)) = 0$ in Ω_T, $\omega^{(x,t)}(E) = \chi_E(x', t)$ for $(x,t) = (x', t) \in S\Omega$, where $\chi_E(x', t)$ is the characteristic function of the set E.

As with (1), it is clear that (7) will hold if μ does not put too much mass near places, where $d\rho$ is too small – or where some negative power of $d\rho$ (appropriately interpreted) gets too big. Since the only $d\rho$'s we consider will have the form $d\rho = \nu d\omega$, where ν is a non-negative function, a "negative power of $d\rho$" will be a weight of the form $\nu^{-\delta} d\omega$.

To see what sort of sufficient condition we should aim for in the parabolic context, let us take a look at a *simple case* of one of the sufficient conditions for (1) obtained in [14]. For this simple example, we assume that $\nu^{1-p'}dx$ is A_∞ with respect to Lebesgue measure. Given this assumption, we can state the following: If $1 < p \leq q < \infty$ and $q \geq 2$, then there exist a positive constant

c and a positive exponent R so that if

$$\mu(T(Q))^{1/q}\left(\int \frac{\nu^{1-p'}(x)}{(1+|x-x_Q|/l(Q))^R}dx\right)^{1/p'} \leq cl(Q)^{d+1} \qquad (13)$$

holds for all dyadic cubes Q, then (1) holds for all reasonable f, where the constants c and R depend on p, q, d and the A_∞ parameters (the a and b) of $\nu^{1-p'}$. Inequality (13) is a precise way of saying that, if μ puts a lot of mass on $T(Q)$, then $\nu^{1-p'}$ cannot get too big near Q.

Our sufficient conditions for (7) also have this form. After localizing (i.e., cutting up $\partial_p\Omega_T$), we can discuss the condition as if we were working with non-isotropic cubes on $\mathbb{R}^d$. As described above we can define regions $T(Q) \subset \Omega_T$, analogous to the top halves of Carleson regions, the $T(Q)$'s we defined in $\mathbb{R}^{d+1}_+$. We assume that our weight $d\rho$ has the form $\nu d\omega$, and we *require* that $d\sigma \doteq \nu^{1-p'}d\omega \in A_\infty(\omega)$. This last assumption is strong, but perhaps not so strong as it looks: because of ω's doubling property, if $\nu^{1-p'} \in L^r(\omega)$ for some $r > 1$, then $\nu^{1-p'}$ has an $A_\infty(\omega)$ pointwise majorant with an L^r norm bounded by a constant times that of $\nu^{1-p'}$.

Given all this, the conditions we obtain are the following: For certain p's and q's (specified later), there are positive constants c' and τ_0, depending on p, q, L, τ, Ω, and the A_∞ parameters (a and b) of σ, such that if

$$\mu(T(Q))^{1/q}\left(\int_{\partial_p\Omega}\left(\omega(Q)\sum_{j=0}^{\infty}\frac{2^{-j(2\beta-\tau)}}{\omega(2^jQ)}\chi_{R_j(Q)}\right)^{p'/2}d\sigma\right)^{1/p'} \leq c'l(Q)\omega(Q) \qquad (14)$$

holds for all non-isotropic cubes Q, then (7) holds for all f in our test class.

The reader will see the relation between (13) and (14) if he rewrites the integral in (14) as one taken over $\mathbb{R}^d$, and replaces ω with Lebesgue measure.

The rest of the paper is organized as follows: In Section 2. we mention some background results on parabolic kernels and parabolic measure that will be crucial to our arguments, and we give a few more definitions that will be needed for the rest of the paper. We then state the two main theorems of the paper. Theorem 1 is the Littlewood-Paley inequality and Theorem 2 gives the weighted norm inequality for the gradient of a parabolic solution on a nonsmooth domain. In Section 3, assuming that Theorem 1 is valid, we prove Theorem 2. A discussion of special cases in which better results can be obtained, such as when $L = \Delta$, and for Ω_T a cylinder domain or a box domain, appears here. Section 4 of the paper contains the proof that the particular functions $\phi_{(Q)}$, to which we apply Theorem 1 to obtain the result of Theorem 2 and which are defined in terms of $\partial/\partial t - L$'s kernel function, satisfy the three conditions (8), (9) and (10). Finally in Section 5, the last section of the paper, Theorem 1 is proved. The proof follows the method established in [15] and [13] with non-trivial technical changes that are necessary for dealing with the parabolic setting. We need seven elementary lemmas to establish local relations for the "partial sum" functions, $F(Q;x',t)$ and $G(Q;x',t)$ in order to prove the crucial good-λ inequality, which is given in the Parabolic Main Lemma and its corollary.

2. Section

Our parabolic solutions are assumed to have the form

$$u(x,t) = \int_{S\Omega} K(x,t;y',s) f(y',s) d\omega(y',s),$$

where ω is the parabolic measure described above and $K(x,t;y',s)$ is the kernel function for the operator $\partial/\partial t - L$ on Ω_T. That is, $K(x,t;y',s)$ is a solution to $(\partial/\partial t - L)K = 0$ in Ω_T; $K(x,t;y',s) > 0$ in Ω_T, and $K(x',t;y',s) = 0$ when $(x',t) \in S\Omega \backslash \{(y',s)\}$. $K(x,t;y',s)$ is normalized to equal 1 at (X_0,T). The existence and uniqueness of such a kernel function on a Lip(1,1/2) domain was proved by Kaj Nystrom in [10]. He proved certain basic estimates for the kernel function such as geometric decay along with the proof that the parabolic measure satisfies a center-doubling condition on $S\Omega$. The results in this paper depend heavily on these geometric properties of parabolic measure and parabolic kernels. In addition to geometric decay, the fact that the kernel function is Holder continuous in the adjoint variables at the parabolic boundary of the domain is crucial to being able to obtain the Littlewood-Paley inequality. Holder continuity for harmonic functions that vanish continuously at the boundary of a rough domain was proved by Jerison and Kenig [7] (as were geometric decay estimates for a harmonic kernel) on non-tangentially accessible domains. Their approach was used to prove similar estimates for parabolic functions by Fabes and Safonov on Lipschitz cylinders [3] and by Nystrom on time-varying domains [10] (see also [11] in this connection). We discuss Holder continuity for parabolic functions that vanish at the boundary of a rough domain in greater detail below; the proof of Holder continuity at the boundary for such functions on a Lip(1,1/2) domain is given in Section 4. for the sake of completeness.

Remember we will be using the parabolic metric $d_p(x,t;y,s) = |x-y| + |t-s|^{1/2}$. Local estimates are often obtained on parabolic cylinders which have the form $\Psi_r(x_0,t_0) = \{(y,s) : d_p(x_0,t_0;y,s) < r\}$ and/or parabolic cubes. When the point $(x_0',t_0) = (Q_0,s_0)$ is taken to lie on the lateral boundary of a domain, then a boundary disk will be $\Delta_r(Q_0,s_0) \doteq \{(y',s) : d_p(y',s;Q_0,s_0) < r, (y',s) \in \partial_p\Omega\} = \Psi_r(Q_0,s_0) \cap \partial_p\Omega$. Notice that in most situations estimates on the dyadic parabolic cubes can be replaced by estimates on these boundary disks and vice versa. This is also true in interior regions for cubes of the form $Q_R(x_0,t_0) \doteq \{(y,s) : |y_i - x_{0,i}| < R, i = 1,2,...,d \text{ and } \sqrt{|t-s|} < R\}$. We will have occasion to refer to parabolic annular regions in the boundary; these are designated as $R_j(I) = \Delta_{2^j r} \backslash \Delta_{2^{j-1}r}$, when $I = \Delta_r$. Space cubes will be denoted by $Q_R(x_0,t_0) \doteq \{(y,s) : |y_i - x_{0,i}| < R, i = 1,2,...,d \text{ and } t = t_0\}$; cross-sections of local cylinders will be written as $B_R(x_0,t_0) \doteq \{(y,t_0) : |y - x_0| < R\}$.

For a Lip(1,1/2) domain as described above, the points $A_r^+(Q_0,s_0) = (Q_0 + c(N)r, s_0 + c'(N)r^2)$ and $A_r^-(Q_0,s_0) = (Q_0 + c(N)r, s_0 - c'(N)r^2)$ for $r \le r_0$ are "Harnack-type" points in Ω; the distance of each of these points to $S\Omega$ compares with r. r_0 is a fixed constant that depends on Ω and on the size of the local Lipschitz graphs that describe the lateral boundary of the domain. In fact as described above r_0 is chosen so that for any fixed point $(Q_0,s_0) \in S\Omega$, the region in $S\Omega$ that lies within a distance of r_0 of (Q_0,s_0) can be described as a single Lipschitz graph.

For Theorem 1 we will be dealing with finite linear combinations of functions taken from a particular family of functions $\{\varphi_{(I)}\}_{I\in\mathcal{D}}$ which are indexed by parabolic "dyadic cubes", $\mathcal{D}$, on the lateral part of the parabolic boundary of the given domain, Ω. We will need the definition of the $\varphi_{(I)}$ in the proof of Theorem 2. It is, in its simplest form,

$$\varphi_{(I)}(y',s) = \sqrt{\omega(I)}\left(K(x^1_{(I)},t^0_{(I)};y',s) - K(x^2_{(I)},t^0_{(I)};y',s)\right).$$

The points $(x^1_{(I)},t^0_{(I)})$ and $(x^2_{(I)},t^0_{(I)})$ are fixed and lie inside $T(I)$ or a fixed dilation of $T(I)$, the top half of the Carleson-type region for the dyadic parabolic boundary cube I.

The family of functions $\{\varphi_{(I)}\}_{I\in\mathcal{D}}$ satisfies three conditions (8), (9) and (10) given above. We note that these conditions are the parabolic analogue of conditions given in [13] and [15] for similar families of functions.

Theorem 1. *Let $\{\varphi_{(I)}\}_{I\in\mathcal{D}}$ be a family of functions defined on $\partial_p^+\Omega = S\Omega$, the lateral part of $\partial_p\Omega$, so that the $\varphi_{(I)}$ satisfy conditions* (8), (9) *and* (10). *Let $\sigma \in A^\infty(\omega)$. For $0 < p < \infty$ there is a constant $C = C(\alpha,\beta,\lambda,d,\tau,N,r_0,p)$ such that, for any finite sum $f(x',t) = \sum_{I\in\mathcal{F}} \lambda_I \phi_{(I)}$*

(x', t), *then*

$$\int_{S\Omega} \left|f(x',t)\right|^p \sigma(x',t)d\omega(x',t) \leq C \int_{S\Omega} \left|g^*(f)(x',t)\right|^p \sigma(x',t)d\omega(x',t).$$

Recall that

$$g^*(f)(x',t) = \left(\sum |\lambda_I|^2 \cdot \sum_{j=0}^{\infty} (2^{-j(2\beta-\tau)}/\omega(2^j I)) \cdot \chi_{R_j(I)}(x',t)\right)^{1/2}$$

and $\sigma \in A^\infty(\omega)$ means that the weight σ lies in the Muckenhaupt class A-infinity with respect to the measure ω. We note that this is an A-infinity condition with respect to the collection of parabolic cubes in $S\Omega$, i.e. the cubes that are projections of parabolic cubes from $\mathbb{R}^d$.

Using Theorem 1 and, following the general method in [14] of using a dual operator inequality, we can prove the result stated in Theorem 2. Proceeding as in [13] we divide Ω_T into two connected regions, $\Omega_{T,\delta}$ which is the part of Ω_T that is near the parabolic boundary and $\Omega_T \backslash \Omega_{T,\delta}$. Specifically $\Omega_{T,\delta} = \{(y,s) : (y,s) \in \Omega_T \text{ and } d_p((y,s); S\Omega) < \delta\}$.

We state the condition in Theorem 2 for Carleson-type regions in both $\Omega_{T,\delta}$ and $\Omega_T \backslash \Omega_{T,\delta}$. The reason for doing this will become clear in the proof of Theorem 2.

Theorem 2. *For Ω_T a* Lip*(1,1/2) domain in $\mathbb{R}^{d+1}$, let μ be a measure defined on Ω_T, $\nu d\omega$ a measure defined on $S\Omega$ such that for $\sigma = \nu^{(1-p')}$, $\sigma d\omega \in A^\infty(\partial_p \Omega_T, d\omega)$ and*

$$\mu(T(I))^{1/q} \times \left(\int_{S\Omega} \left\{\sum_{j=0}^{\infty} \{2^{-j(2\beta-\tau)}/\omega(2^j I)\} \cdot \chi_{R_j(I)}(x',t)\right\}^{p'/2} \sigma(x',t)d\omega(x',t)\right)^{1/p'} \leq \omega(I)^{1/2} l(I)$$

for all parabolic cubes $I \subset S\Omega$. $T(I)$ is the top half of the Carleson region associated with I. Then there is a constant θ so that for $1 < p \leq q$ and $2 \leq q < (2+\theta)$, and there is a constant C_2, independent of $u(x,t)$ and $f(x',t)$, so that

$$\left(\int_{\Omega_{T,\delta}} |\nabla u(x,t)|^q \, d\mu(x,t)\right)^{1/q}$$

$$\leq C_2 \left(\int_{S\Omega} \left| f(x', t) \right|^p v(x', t) d\omega(x', t) \right)^{1/p}, \quad (15)$$

whenever $u(x,t)$ is a solution to the Dirichlet problem on Ω_T; that is, $(\partial/\partial t - L)u(x,t) = 0$ in Ω_T, $u(x',t) = f(x',t)$ for $(x',t) \in S\Omega$. We assume that $f(x,0) = 0$.

3. Section

In [13] Sweezy and Wilson extended inequality (1) to harmonic and elliptic functions on a bounded Lipschitz domain. They found a condition in the spirit of Wheeden and Wilson's condition for $\mu(x)$ and $v(x')d\omega(x')$ that guarantees (1) is valid for harmonic $u(x)$ with dx replaced by $d\omega$, ω being the harmonic measure on $\partial\Omega$, for the same range of p and q. Similar results are valid when $u(x)$ is elliptic, although the range for q must be limited to $[2, 2+\epsilon)$. The lack of smoothness in the domain necessitates technical changes in the proof from the situation for harmonic functions on a half space dealt with in [14] and briefly described above, although the overall outline, namely the use of a dual operator and a Littlewood-Paley inequality, remain the same. The Littlewood-Paley inequality on a Lipschitz domain is proved using methods more along the lines of [15].

In the work described above, Wheeden and Wilson also established sufficient conditions on the measures for (1) to hold when $u(x,t)$ is caloric in $\mathbb{R}^{d+1}_+$. In [12] the question was examined for caloric functions in the right half space and on a box domain. As is well known the situation in the right half space for solutions to the heat equation is significantly different from the heat equation in the upper half space, largely because the boundary of the domain has time as a variable. A key factor in obtaining their results for the right half space was that the heat kernel has excellent decay and smoothness properties. In this paper the authors start to examine the situation for caloric and for strictly parabolic functions on domains whose boundaries are not assumed to be smooth and in fact may vary with time. The lack of smoothness in the kernel function on such domains led the authors to use methods more closely allied with the argument for harmonic functions on Lipschitz domains, [13], than it is with the half space results in [12] for the heat equation. It turns out that, even for solutions to the heat equation on a standard right circular cylinder domain, the heat kernel does not have enough smoothness to use the method of [12] in handling the requisite Littlewood-Paley inequality. The differences in what can be proved for solutions to the heat equation and what can be proved for solutions to general parabolic equations are discussed in the proof of Theorem 2 below.

We assume that $u(x,t)$ is extended to a domain that contains Ω_T, but also contains a region below $t=0$ on which $u_{EXT}(x,t)=0$. Uniqueness guarantees that this extended function u_{EXT} is the solution to $(\partial/\partial t - L)u_{EXT}(x,t)=0$ in Ω_{EXT}, $u_{EXT}(x',t)=f_{EXT}(x',t)$ for $(x',t)\in\partial_p\Omega_{EXT}$ where $f_{EXT}(x',t)=f(x',t)$ for $(x',t)\in S\Omega$ and $f_{EXT}(x',t)=0$ for $(x',t)\in\partial_p\Omega_{EXT}$, $t\le 0$. The reason for defining this extension is to justify treating estimates for $u(x,t)$ on cubes touching the bottom boundary of Ω_T as interior estimates for $u_{EXT}(x,t)$.

To prove (15) for Ω_T and $u(x,t)$ as in Theorem 2, we have the following estimate for the interior region $\Omega_T\backslash\Omega_{T,\delta}$, in which $\Omega_T\backslash\Omega_{T,\delta}$ is essentially treated as a single parabolic box. We have the Cacciopoli type inequality

$$\frac{1}{|Q_{r_0/2}|}\int_{Q_{r_0/2}}|\nabla u(y,s)|^2\,dyds\le C\frac{1}{r_0^2\,|Q_{r_0}|}\int_{Q_{r_0}}|u(y,s)|^2\,dyds$$

or, we will frequently be using the version of this inequality in Lemma 2.1 of [5], which is

$$\int_{Q_{r_0/2}}|\nabla u(y,s)|^2\,dyds\le C\frac{1}{r_0^2}\int_{Q_{r_0}}|u(y,s)-\widetilde{u}_{r_0}(t)|^2\,dyds.$$

Here we take

$$\widetilde{u}_{r_0}(t)=\frac{1}{\chi^2(Q_{r_0}(x_{(I)},t))}\int\int_{Q_{r_0}(x_{(I)},t)}u(y,t)\chi^2(y)dy$$

as in [5]. ($\chi^2(y)$ is a smooth bump function supported on $Q_{r_0}(x_I,t)$.)

For the case of caloric $u(x,t)$ we have the pointwise estimate $|\nabla u(x,t)|^2\le\frac{1}{|Q_{r_0}|}\int_{Q_{r_0}}|\nabla u(y,s)|^2\,dyds$ if (x,t) lies inside $Q_{r_0/2}$. Writing

$$u(y,s)=\int_{\partial_p\Omega}f(x',\tau)d\omega^{(y,s)}(x',\tau),$$

using a standard estimate on the kernel $K(y,s;Q,\tau)$ when $d_p((y,s);\partial_p\Omega_T)\ge C(r_0)$ and Holder's inequality now gives

$$\left(\int_{\Omega_T\backslash\Omega_{T,\delta}}|\nabla u(y,s)|^q\,d\mu(y,s)\right)^{1/q}$$
$$\lesssim\left(\int_{\partial_p\Omega_T}\left|f(x',\tau)\right|^p\nu(x',\tau)d\omega(x',\tau)\right)^{1/p}$$

whenever

$$\mu(\Omega_T \backslash \Omega_{T,\delta})^{1/q} \left(\int_{\partial_p \Omega_T} \sigma(x', \tau) d\omega(x', \tau) \right)^{1/p'} \leq C(\Omega_T, r_0, d)$$

(see [12] for details). But this is merely the condition stated in Theorem 2 for parabolic cubes and their associated top halves of Carleson boxes (for general parabolic $u(x,t)$ we also need to use reverse Holder for parabolic $u(x,t)$ and the $B^{\epsilon_0'}$ condition for μ (see below) to get the same result).

When the operator $\partial/\partial t - L$ has the property that any solution $u(x,t)$ has a gradient that satisfies a sub-mean value inequality, for example when $L = \Delta$ (first proved by Hattemer in [6]), then Theorem 2 is valid for all p and q such that $1 < p \leq 2 \leq q < \infty$ and for $2 < p \leq q < \infty$. However, for operators whose solutions do not satisfy such a condition, for example when the coefficients of L are only assumed to be bounded and measurable, q must be restricted to a much smaller range of indices. In this case we must have $2 \leq q < (1+\epsilon_0)2$ for some small $\epsilon_0 > 0$. In addition, to obtain the gradient norm inequality of the theorem as it appears above, one must put some additional restriction on the measure μ, for example, that $\mu \in B^{\epsilon_0'}(dxdt)$, where $\frac{1}{1+\epsilon_0} + \frac{1}{\epsilon_0'} = 1$ (that is, μ is a parabolic Muckenhaupt class A^p measure for some index p not related to the index p appearing in Theorem 2). The reason for these added conditions is that the integral $\int |\nabla u(x,t)|^q \, d\mu$ has to be dominated by the discrete expression $\sum_{I \in \mathcal{D}} \{ \left| u(x^1_{(I)}, t_{(I)}) - u(x^2_{(I)}, t_{(I)}) \right| / l(I) \}^q \cdot \mu(T(I))$ in order to proceed with the dual operator proof. The same expression can be used in the case of the heat equation. In this latter case we know that $\nabla u(x,t)$ has the sub-mean value property, so

$$\begin{aligned}
\int_{\Omega_T} |\nabla u(x,t)|^q \, d\mu(x,t) &\lesssim \sum_{I \in \mathcal{PD}} \int_{T(I)} |\nabla u(x,t)|^q \, d\mu(x,t) \\
&\lesssim \sum_{I \in \mathcal{D}} \int_{T(I)} \left(\frac{1}{|T(I)|} \int_{2T(I)} |\nabla u(x,t)|^2 \, dxdt \right)^{q/2} d\mu \\
&= \sum_{I \in \mathcal{D}} \left(\frac{1}{|T(I)|} \int_{2T(I)} |\nabla u(x,t)|^2 \, dxdt \right)^{q/2} \mu(T(I)) \\
&\lesssim \sum_{I \in \mathcal{D}} \{ \left| u(x^1_{(I)}, t_{(I)}) - u(x^2_{(I)}, t_{(I)}) \right| / l(I) \}^q \cdot \mu(T(I)).
\end{aligned}$$

The second inequality is from using the sub-mean value property of ∇u; the last inequality follows from Caccioppoli's inequality and from choosing $(x^1_{(I)}, t_{(I)})$

and $(x^2_{(I)}, t_{(I)})$ to maximize $\left|u(x^1_{(I)}, t_{(I)}) - u(x^2_{(I)}, t_{(I)})\right|$ over an appropriate dilate of $T(I)$ (see Lemma 2.1 in [5]).

From this point one can proceed as in Wheeden and Wilson, [14], to use a dual operator argument to estimate

$$\sum_{I\in\mathcal{PD}} \{\left|u(x^1_{(I)}, t_{(I)}) - u(x^2_{(I)}, t_{(I)})\right| / l(I)\}^q \cdot \mu(T(I)).$$

To obtain a similar majorization of $\int_{\Omega_T} |\nabla u(x,t)|^q \, d\mu(x,t)$ by a discrete sum when $u(x,t)$ is a solution to a general parabolic equation, the mean value inequality is no longer valid. So some way around this must be found. One way to accomplish this is to use Holder's inequality, reverse Holder for ∇u (Theorem 2.1 in [5]), and to assume that $\mu \in B^{\epsilon_0'}(dxdt)$. Instead of the inequality $|\nabla u(x,t)| \lesssim (\frac{1}{|2T(I)|}\int_{2T(I)} |\nabla u|^2 \, dyds)^{1/2}$ for $(x,t) \in T(I)$, one can obtain

$$\begin{aligned}
&\left(\int_{T(I)} |\nabla u|^q \, d\mu\right) \\
&\quad\lesssim \left(\int_{T(I)} |\nabla u|^{q(1+\epsilon_0)} \, dxdt\right)^{1/(1+\epsilon_0)} \cdot \left(\int_{T(I)} (d\mu/(dxdt))^{\epsilon_0'} dxdt\right)^{1/\epsilon_0'} \\
&\quad\lesssim |T(I)| \cdot (\mu(T(I))/|T(I)|) \cdot \left(\frac{1}{|2T(I)|}\int_{2T(I)} |\nabla u|^2 \, dxdt\right)^{q/2} \lesssim \\
&\left(\sup_{t_{(I)}-16l(I)^2<t<t_{(I)}} \left(\frac{1}{|B_{4l(I)}(x_{(I)},t)|}\int_{B_{4l(I)}(x_{(I)},t)} (\left|u(x,t) - \widetilde{u}_{4l(I)}(t)\right| / l(I))^2\right)^{q/2} \\
&\qquad\qquad\times \mu(T(I)).
\end{aligned}$$

The second inequality follows from reverse Holder for parabolic solutions (see Theorem 2.1 in [5]) and from the hypothesis that $\mu \in B^{\epsilon_0'}(dxdt)$, the third is from Lemma 2.1 in [5]. This last expression can be dominated by $\sum_{I\in\mathcal{D}} \{\left|u(x^1_{(I)}, t^0_{(I)}) - u(x^2_{(I)}, t^0_{(I)})\right| / l(I)\}^q \cdot \mu(T(I))$ with the right choice of the points $(x^i_{(I)}, t^0_{(I)})$, $i = 1, 2$.

We proceed with the estimates for the discrete expression

$$\begin{aligned}
&\sum_{I\in\mathcal{D}} \{\left|u(x^1_{(I)}, t^0_{(I)}) - u(x^2_{(I)}, t^0_{(I)})\right| / l(I)\}^q \cdot \mu(T(I)) \\
&\qquad\qquad = \left\|\left\{(u(x^1_{(I)}, t^0_{(I)}) - u(x^2_{(I)}, t^0_{(I)}))/l(I)\right\}_{I\in\mathcal{D}}\right\|^q_{l^q(\mu)}.
\end{aligned}$$

By switching the points $(x^1_{(I)}, t_{(I)})$ and $(x^2_{(I)}, t_{(I)})$ if necessary, we want to find

$$\sup_{\|\{g(I)\}\|_{l^{q'}(\mu)}=1} \sum_{I\in\mathcal{G}} \left\{(u(x^1_{(I)}, t^0_{(I)}) - u(x^2_{(I)}, t^0_{(I)}))/l(I)\right\} g(I)\cdot\mu(T(I))$$

$$= \sup_{\|\{g(I)\}\|_{l^{q'}(\mu)}=1} \sum_{I\in\mathcal{G}} \frac{g(I)\mu(T(I))}{l(I)}$$

$$\int_{\partial_p\Omega} (K(x^1_{(I)}, t^0_{(I)}; y', s) - K(x^2_{(I)}, t^0_{(I)}; y', s)) f(y', s) d\omega(y', s)$$

$$= \sup_{\|\{g(I)\}\|_{l^{q'}(\mu)}=1} \int_{\partial_p\Omega} f(y', s) Tg(y', s) d\omega(y', s)\,,$$

where

$$Tg(y', s)$$

$$= \sum_{I\in\mathcal{G}} \frac{g(I)\mu(T(I))}{l(I)\sqrt{\omega(I)}} \cdot \left(\sqrt{\omega(I)}(K(x^1_{(I)}, t^0_{(I)}; y', s) - K(x^2_{(I)}, t^0_{(I)}; y', s))\right)$$

$$= \sum_{I\in\mathcal{G}} \lambda_I \varphi_{(I)}(y', s).$$

Here $\lambda_I = \frac{g(I)\mu(T(I))}{l(I)\sqrt{\omega(I)}}$ and recall that

$$\varphi_{(I)}(y', s) = \sqrt{\omega(I)}(K(x^1_{(I)}, t^0_{(I)}; y', s) - K(x^2_{(I)}, t^0_{(I)}; y', s))\,,$$

where $K(x, t; y', s)$ is the kernel function for the operator $\partial/\partial t - L$ on the domain Ω_T. The family $\{g(I)\}_{I\in\mathcal{G}}$ can be taken to be finite so $Tg(y', s) = \sum_{I\in\mathcal{D}} \lambda_I \varphi_{(I)}(y', s)$ is a finite sum. At this point it is convenient to use Theorem 1 on $\|Tg\|_{L^{p'}(\sigma d\omega)}$. To do this conditions (8), (9) and (10) must be verified for the above choice of the functions $\varphi_{(I)}$. Once this is done, then by Holder's inequality and Theorem 1,

$$\int_{\partial_p\Omega} f(y', s) Tg(y', s) d\omega(y', s) \le \|f\|_{L^p(vd\omega)} \cdot \|Tg\|_{L^{p'}(\sigma d\omega)}$$

$$\lesssim \|f\|_{L^p(vd\omega)} \cdot \|g^*(Tg)\|_{L^{p'}(\sigma d\omega)}\,.$$

Then, in the case $p = q = 2$, the condition on μ given in Theorem 2 is easily seen to imply that $\|g^*(Tg)\|_{L^{p'}(\sigma d\omega)} \le C\,\|\{g(I)\}\|_{l^{q'}(d\mu)}$. We write

$$\|g^*(Tg)\|^2_{L^2(\sigma d\omega)} = \int_{S\Omega}\left(\sum_{I\in\mathcal{G}}|\lambda_I|^2\left[\sum_{j=0}^{\infty}\frac{2^{-(2\beta-\tau)j}}{\omega(2^jI)}\chi_{R_j(I)}(x',t)\right]\right)\sigma d\omega$$
$$\le C\sum_{I\in\mathcal{G}}\left(\frac{g(I)\mu(T(I))}{l(I)\sqrt{\omega(I)}}\right)^2\int_{S\Omega}\left[\sum_{j=0}^{\infty}\frac{2^{-(2\beta-\tau)j}}{\omega(2^jI)}\chi_{R_j(I)}(x',t)\right]\sigma d\omega.$$

To have the last integral dominated by $C\,\|\{g(I)\}\|^2_{l(d\mu)} = C\sum_{I\in\mathcal{G}}(g(I))^2$ $\mu(T(I))$, one can achieve this by comparing each term and requiring that, for each $I\in\mathcal{G}$,

$$\left(\frac{g(I)\mu(T(I))}{l(I)\sqrt{\omega(I)}}\right)^2\int_{S\Omega}\left[\sum_{j=0}^{\infty}\frac{2^{-2\tau j}}{\omega(2^jI)}\chi_{R_j(I)}(x',t)\right]\sigma d\omega \le (g(I))^2\,\mu(T(I)).$$

Simplifying, multiplying through by $l(I)^2\omega(I)/\mu(T(I))$ and taking square roots gives the condition

$$\mu(T(I))^{1/2}\left(\int_{S\Omega}\left[\sum_{j=0}^{\infty}\frac{2^{-(2\beta-\tau)j}}{\omega(2^jI)}\chi_{R_j(I)}(x',t)\right]\sigma d\omega\right)^{1/2}\le l(I)\sqrt{\omega(I)}.$$

This is exactly the condition of Theorem 2.

For the case $2\le p\le q$ the reasoning is similar to the case $p=q=2$, except that we must use the fact that $p'/2$ and q'/p' are both ≤ 1. Remember that $\sum_{I\in\mathcal{G}}$ is a finite sum. We can dominate

$$\|g^*(Tg)\|^{p'}_{L^{p'}(\sigma d\omega)} = \int_{S\Omega}\left(\sum_{I\in\mathcal{G}}|\lambda_I|^2\left[\sum_{j=0}^{\infty}\frac{2^{-(2\beta-\tau)j}}{\omega(2^jI)}\chi_{R_j(I)}(x',t)\right]\right)^{p'/2}\sigma d\omega$$

by

$$\int_{S\Omega}\sum_{I\in\mathcal{G}}|\lambda_I|^{p'}\left[\sum_{j=0}^{\infty}\frac{2^{-(2\beta-\tau)j}}{\omega(2^jI)}\chi_{R_j(I)}(x',t)\right]^{p'/2}\sigma d\omega$$
$$\le\sum_{I\in\mathcal{G}}\left|\frac{g(I)\mu(T(I))}{l(I)\sqrt{\omega(I)}}\right|^{p'}\int_{S\Omega}\left[\sum_{j=0}^{\infty}\frac{2^{-(2\beta-\tau)j}}{\omega(2^jI)}\chi_{R_j(I)}(x',t)\right]^{p'/2}\sigma d\omega\le$$

$$\left(\sum_{I\in\mathcal{G}}\left|\frac{g(I)\mu(T(I))}{l(I)\sqrt{\omega(I)}}\right|^{q'}\left(\int_{S\Omega}\left[\sum_{j=0}^{\infty}\frac{2^{-(2\beta-\tau)j}}{\omega(2^jI)}\chi_{R_j(I)}(x',t)\right]^{p'/2}\sigma d\omega\right)^{q'/p'}\right)^{p'/q'}.$$

Now, taking the $1/p'$ of the last expression, and comparing term-by-term with $\left(\sum_{I\in\mathcal{G}}(g(I))^{q'}\mu(T(I))\right)^{1/q'}$, simplifying, and taking $1/q'$ roots gives the sufficient condition

$$\mu(T(I))^{1/q}\left(\int_{S\Omega}\left[\sum_{j=0}^{\infty}\frac{2^{-(2\beta-\tau)j}}{\omega(2^jI)}\chi_{R_j(I)}(x',t)\right]^{p'/2}\sigma d\omega\right)^{1/p'}\leq l(I)\sqrt{\omega(I)}.$$

The case where $1<p\leq 2\leq q<\infty$ is more involved. The argument is similar to [13] for the elliptic case; one needs only to replace the maximal operator appearing there by the parabolic maximal operator, and verify that key results hold with parabolic cubes replacing Euclidean cubes. To begin with, when $p\leq 2$, the dual exponent $p'\geq 2$ so $p'/2\geq 1$. If s is the dual exponent to $p'/2$, then there is a non-negative function $h\in L^s(\sigma d\omega)$ so that $\int_{\partial_p\Omega}h^s\sigma d\omega\leq 1$, and $\int_{\partial_p\Omega}|Tg|^{p'}\sigma d\omega=\left(\int_{\partial_p}|Tg|^2 h\sigma d\omega\right)^{p'/2}$. Recall that we assumed $\sigma d\omega\in A_\infty(d\omega)$, where this is parabolic A_∞. Unfortunately, we do not know that $h\sigma d\omega\in A_\infty(d\omega)$. However, we can dominate $h\sigma d\omega$ by a weight $\zeta d\omega$ that is in $A_\infty(d\omega)$. Then we can apply Theorem 1 to $\left(\int_{\partial_p}|Tg|^2\zeta d\omega\right)^{p'/2}$. After some manipulation we will be able to see that the condition given above emerges as the natural sufficient condition for $\left(\int_{\partial_p\Omega}|Tg|^{p'}\sigma d\omega\right)^{1/p'}\leq\left\|\{g_{(I)}\}_{I\in\mathcal{G}}\right\|_{l^{q'}(\mu,\Omega)}$. To define $\zeta d\omega$, we must define a parabolic maximal function for any measure $d\nu$ on $\partial_p\Omega$ that satisfies a center-doubling condition. Taking $1<r<s$, we let

$$M_{r,\nu}(h)(x',t)=\left(\sup_{Q,(x',t)\in Q}\frac{1}{\nu(Q)}\int_Q\left|h(y',s)\right|^r d\nu(y',s)\right)^{1/r},$$

where the supremum is taken over all parabolic boundary cubes Q that contain (x',t). In the Appendix we show that $M_{r,\nu}(h)(x',t)d\nu(x',t)\in A_\infty(d\nu)$. Then, since parabolic A_∞ is a transitive condition, this gives $M_{r,\nu}(h)(x',t)$ $d\nu(x',t)\in A_\infty(d\omega)$ if $d\nu\in A_\infty(d\omega)$. In our situation we will be taking $d\nu=\sigma d\omega$. We also have that $h(x',t)\leq M_{r,\nu}(h)(x',t)$ ν a.e.

We have the weak $L^1(d\nu)$ property for the parabolic maximal function $M_{1,\nu}(h)(x',t)$, and this can be used to show that the strong $L^p(d\nu)$ property holds for $M_{r,\nu}(h)(x',t)$ when $p > r$. In other words, for $1 < r < s < \infty$,

$$\left(\int_{\partial_p\Omega} \left|M_{r,\nu}(h)(x',t)\right|^s d\nu(x',t)\right)^{1/s} \leq C\left(\int_{\partial_p\Omega} \left|h(x',t)\right|^s d\nu(x',t)\right)^{1/s}, \tag{16}$$

where C depends on r, s, d, and on ν's doubling constant.

For $1 < r < s < \infty$, if h and s are chosen as above, we have

$$\int_{\partial_p\Omega} |Tg|^{p'} \sigma d\omega = \left(\int_{\partial_p\Omega} |Tg|^2 h\sigma d\omega\right)^{p'/2} \leq \left(\int_{\partial_p} |Tg|^2 M_{\nu,r}(h)\sigma d\omega\right)^{p'/2}.$$

We can use Theorem 1 to bound the last integral. It is less than or equal to a constant times

$$\left(\int_{\partial_p\Omega}\left(\sum_{Q\in\mathcal{G}} |\lambda_Q|^2 \left[\sum_{j=0}^{\infty} \frac{2^{-j(2\beta-\tau)}}{\omega(2^jQ)}\chi_{R_j}(x',t)\right]\right) M_{\nu,r}(h)\sigma d\omega\right)^{p'/2}$$

$$= \left(\sum_{Q\in\mathcal{G}} |\lambda_Q|^2 \int_{\partial_p\Omega}\left[\sum_{j=0}^{\infty} \frac{2^{-j(2\beta-\tau)}}{\omega(2^jQ)}\chi_{R_j}(x',t)\right] M_{\nu,r}(h)\sigma d\omega\right)^{p'/2}.$$

Using Holder's inequality and (16) we can dominate the integral in the last expression by

$$\left(\int_{\partial_p\Omega}\left[\sum_{j=0}^{\infty} \frac{2^{-j(2\beta-\tau)}}{\omega(2^jQ)}\chi_{R_j}(x',t)\right]^{p'/2} \sigma d\omega\right)^{2/p'} \cdot \left(\int_{\partial_p\Omega} (M_{\nu,r}(h))^s \sigma d\omega\right)^{1/s}$$

$$\leq C\left(\int_{\partial_p\Omega}\left[\sum_{j=0}^{\infty} \frac{2^{-j(2\beta-\tau)}}{\omega(2^jQ)}\chi_{R_j}(x',t)\right]^{p'/2} \sigma d\omega\right)^{2/p'},$$

where we have also used the fact that $\int_{\partial_p\Omega} h^s\sigma d\omega \leq 1$. So far we have

$$\left(\int_{\partial_p\Omega} |Tg|^{p'} \sigma d\omega\right)^{1/p'}$$

$$\leq C\left(\sum_{Q\in\mathcal{G}}|\lambda_Q|^2\left(\int_{\partial_p\Omega}\left[\sum_{j=0}^{\infty}\frac{2^{-j(2\beta-\tau)}}{\omega(2^jQ)}\chi_{R_j}(x',t)\right]^{p'/2}\sigma d\omega\right)^{2/p'}\right)^{1/2}.$$

Using the fact that $q'\leq 2$, so $q'/2\leq 1$, we have

$$\left(\int_{\partial_p\Omega}|Tg|^{p'}\,\sigma d\omega\right)^{1/p'}$$
$$\leq C\left(\sum_{Q\in\mathcal{G}}|\lambda_Q|^{q'}\left(\int_{\partial_p\Omega}\left[\sum_{j=0}^{\infty}\frac{2^{-j(2\beta-\tau)}}{\omega(2^jQ)}\chi_{R_j}(x',t)\right]^{p'/2}\sigma d\omega\right)^{q'/p'}\right)^{1/q'}.$$

We want the quantity on the right hand side of the inequality sign to be bounded by $\left(\sum_{Q\in\mathcal{G}}|g(Q)|^{q'}\,\mu(T(Q))\right)^{1/q'}$. If we substitute back for $\lambda_Q=C\frac{g(Q)\mu(T(Q))}{l(Q)\sqrt{\omega(Q)}}$ and compare corresponding terms in both sums, we get the condition

$$\left(\frac{g(Q)\mu(T(Q))}{l(Q)\sqrt{\omega(Q)}}\right)^{q'}\left(\int_{\partial_p\Omega}\left[\sum_{j=0}^{\infty}\frac{2^{-j(2\beta-\tau)}}{\omega(2^jQ)}\chi_{R_j}(x',t)\right]^{p'/2}\sigma d\omega\right)^{q'/p'}$$
$$\leq c\,|g(Q)|^{q'}\,\mu(T(Q)).$$

Simplifying this inequality, moving the $\mu(T(Q))$ to the left-hand side of the inequality and $l(Q)^{q'}$ to the right hand side and taking q' roots gives

$$\mu(T(Q))^{1/q}\left(\int_{\partial_p\Omega}\left[\sum_{j=0}^{\infty}\frac{2^{-j(2\beta-\tau)}}{\omega(2^jQ)}\chi_{R_j}(x',t)\right]^{p'/2}\sigma d\omega\right)^{1/p'}\leq cl(Q)\sqrt{\omega(Q)}.$$

Once again this is exactly the condition stated in Theorem 2.

4. Section

In this section we verify that (8), (9) and (10) are valid for

$$\varphi_{(I)}(y',s)=\sqrt{\omega(I)}\left(K(x^1_{(I)},t^0_{(I)};y',s)-K(x^2_{(I)},t^0_{(I)};y',s)\right).$$

We first restate the conditions for reference:

$$\left|\phi_{(I)}(x',t)\right| \lesssim (\omega(I))^{1/2}(2^{-j\beta}/\omega(2^j I)) \text{ for } (x',t) \in R_j(I)\,. \tag{8}$$

For (x',t) and (y',s) both on

$$S\Omega \cap R_k(I),\ \left|\phi_{(I)}(x',t) - \phi_{(I)}(y',s)\right| \lesssim \left(\frac{d_p(x',t;y',s)}{l(I)}\right)^{\alpha} \times \omega(I)^{\frac{1}{2}} \cdot \sum_{j=0}^{\infty}((2^{-j\beta}/\omega(2^j I)) \cdot (\chi_{R_j(I)}(x',t) + \chi_{R_j(I)}(y',s))\,. \tag{9}$$

For any

$$\{\lambda_I\}_{I\in\mathcal{D}}, \int_{S\Omega}\left|\sum \lambda_I \phi_{(I)}(x',t)\right|^2 d\omega(x',t) \lesssim \sum |\lambda_I|^2\,. \tag{10}$$

Remark. (9) is only needed in the arguments below when (x',t) and (y',s) both lie in a single cube J, and $l(J) \le l(I)$.

We start by proving (10), almost orthogonality. It holds from the mean value property for $u(x,t)$ and $\nabla u(x,t)$ when $u(x,t)$ is caloric [6]. When L is smooth we still have both the calculus mean value property for $u(x,t)$ and a sub-mean value property for non-negative subsolutions, proved by Jurgen Moser. When L's coefficients are bounded and measurable we can still obtain the same kind of estimate as we use for smooth solutions since we can approximate the rough coefficient solutions by solutions to $(\partial/\partial t - L_m)u_m = 0$, where the coefficients of L_m are C^∞ mollifications of those of L. When u_m is a solution so is ∇u_m, and $|\nabla u_m|^2$ is an obviously non-negative subsolution. By Theorem 3 in [9], $|\nabla u_m|^2 \le C\left(\int_{T(I)} |\nabla u_m|^{2p}\right)^{1/p}$ for $p > 1$. Now when p is close enough to 1, we can use reverse Holder to obtain $\left(\int_{T(I)} |\nabla u_m|^{2p}\right)^{1/p} \le C'\left(\int_{T(I)} |\nabla u_m|^2\right)^{1/2}$. This means we have

$$\left|\left((u_m(x^1_{(I)}, t^0_{(I)}) - u_m(x^2_{(I)}, t^0_{(I)}))/l(I)\right)\right|^2 \le \left|\nabla u_m(x^*, t_{(I)})\right|^2 \le C\left(\frac{1}{|T(I)|}\int_{4T(I)} |\nabla u_m(x,t)|^2\, dxdt\right)$$

for smooth solutions. It is well known that the smooth solutions u_m converge to u pointwise and in the local Sobolev space $H^2(\eta T(I))$ for $\eta T(I) \subsetneq \Omega_T$. So we have the inequality

$$\left|\left((u(x^1_{(I)},t^0_{(I)})-u(x^2_{(I)},t^0_{(I)}))/l(I)\right)\right|^2 \leq C\left(\frac{1}{|T(I)|}\int_{4T(I)}|\nabla u(x,t)|^2\,dxdt\right)$$

for rough coefficient solutions. Then we can use Green's Theorem, as is shown below. We now prove (10) for the parabolic $\phi_{(I)}$'s.

Take $f(y',s)=\sum_{I\in\mathcal{F}}\lambda_I\varphi_{(I)}(y',s)$ and write

$$\int_{\partial_p\Omega}\left|f(y',s)\right|^2 d\omega(y',s)=\int_{\partial_p\Omega}f(y',s)\cdot\sum_{I\in\mathcal{F}}\lambda_I\varphi_{(I)}(y',s)d\omega(y',s)=$$

$$\sum_{I\in\mathcal{F}}\lambda_I\int_{\partial_p\Omega}f(y',s)\cdot\sqrt{\omega(I)}\cdot\left(K(x^1_{(I)},t^0_{(I)};y',s)-K(x^2_{(I)},t^0_{(I)};y',s)\right)d\omega(y',s)$$

$$=\sum_{I\in\mathcal{F}}\lambda_I\sqrt{\omega(I)}\cdot(u(x^1_{(I)},t^0_{(I)})-u(x^2_{(I)},t^0_{(I)}))$$

$$\leq\left(\sum_{I\in\mathcal{F}}\lambda_I^2\right)^{1/2}\cdot\left(\sum_{I\in\mathcal{F}}\omega(I)l(I)^2\left|\left((u(x^1_{(I)},t^0_{(I)})-u(x^2_{(I)},t^0_{(I)}))/l(I)\right)\right|^2\right)^{\frac{1}{2}}.$$

Now by the argument given above we have

$$\left|\left((u(x^1_{(I)},t^0_{(I)})-u(x^2_{(I)},t^0_{(I)}))/l(I)\right)\right|^2\leq C\left(\frac{1}{|T(I)|}\int_{4T(I)}|\nabla u(x,t)|^2\,dxdt\right)$$

and by a well-known estimate for parabolic measure

$$\omega(I)\simeq G(X_0,T;A_{l(I)}(x'_{(I)},t_{(I)})\cdot l(I)^d$$

($A_{l(I)}(x'_{(I)},t_{(I)})$ was defined in Section 2), so by Cauchy-Schwarz and the above estimates we have

$$\left(\sum_{I\in\mathcal{F}}\omega(I)l(I)^2\left|\left((u(x^1_{(I)},t^0_{(I)})-u(x^2_{(I)},t^0_{(I)}))/l(I)\right)\right|^2\right)^{1/2}$$

$$\leq C\left(\sum_{I\in\mathcal{F}}|T(I)|\cdot\frac{1}{|T(I)|}\cdot\int_{3T(I)}|\nabla u(x,t)|^2\,G(X_0,T;x,t)dxdt\right)^{1/2}$$

$$\leq C'\left(\int_{\Omega}|\nabla u(x,t)|^2\,G(X_0,T;x,t)dxdt\right)^{1/2}.$$

Green's Theorem gives the estimate

$$\int_{\Omega} |\nabla u(x,t)|^2 G(X_0,T;x,t)dxdt \leq \int_{\partial_p\Omega} \left|f(y',s)\right|^2 d\omega(y',s),$$

and, finally, dividing by $\left(\int_{\partial_p\Omega} |f(y',s)|^2 d\omega(y',s)\right)^{1/2}$ gives the estimate in (10), $\left(\int_{\partial_p\Omega} |f(y',s)|^2 d\omega(y',s)\right)^{1/2} \lesssim \left(\sum_{I\in\mathcal{F}} \lambda_I^2\right)^{1/2}$.

Green's Theorem is valid for general parabolic operators of the type described for Theorem 2 on a Lip(1,1/2) domain by standard arguments.

Next, (8) holds because the kernel $K(x,t;y',s)$ has geometric decay for $(y',s) \in S\Omega$. This estimate is proved in Theorem 4.2 in Kaj Nystrom's paper [10] for strictly parabolic divergence form operators on Lip(1,1/2) domains. The constants C_j in the estimate

$$\sup_{(Q,s)\in R_j(Q_0,s_0)} [K(A_r(Q_0,s_0);y',s)] \leq C_j/\omega^{(X_0,T_0)}(\Delta_{2^j r}(Q_0,s_0))$$

are easily seen to be exponentially decreasing in j (see the proof of Theorem 2.12 in [2]).

Finally, this geometric decay, combined with Holder continuity in (y',s), is enough to prove that (9) holds for the discrete functions $\phi_{(I)}$ as defined above.

Holder continuity at the lateral parabolic boundary of Ω_T was proved for positive parabolic solutions by Fabes and Safanov in [3] when Ω_T is a cylinder domain of the form $D\times(0,T)$, and D is a Lipschitz domain in $\mathbb{R}^d$. The geometric estimates that can be used to prove Holder continuity at the boundary of a Lip(1,1/2) domain for ratios of positive solutions are proved in Nystrom's paper [10]. The key step in this proof is ratio Harnack for solutions vanishing locally at $S\Omega$ (see Lemma 2.10 in [10]). Given this result the argument of Fabes and Safonov in [3], originally due to J. Moser [9] for parabolic functions in the interior of a domain, gives a uniform estimate in the adjoint variable for ratios of Green's functions on the interior of a local boundary cylinder, $\Psi_{r/8}(y',s)\cap\Omega_T$. This is enough to prove Holder continuity for the kernel function $K(x,t;y',s)$ in the variables (y',s) whenever (y',s) lies in $\partial_p\Omega_T \cap \{(y\hat{}',s\hat{}) \in \partial_p\Omega_T : 0 < r_0^2 \leq s\hat{} < t - r_0^2\}$. (See Kenig [8] for the same result for strictly elliptic functions.) However, here (9) must hold at all points $(y',s) \in \partial_p\Omega_T$, including points where s approaches t or 0. So we extend the range of allowable values for s as shown below.

Specifically, to prove Holder continuity for $K(x,t;y',s)$ at the lateral boundary of Ω_T, one needs to prove the following theorem.

Theorem 3. *For $(\partial/\partial t - L)u = 0$ and $(\partial/\partial t - L)v = 0$ in Ω_T, where $u, v > 0$ in $\Psi_{2r}(Q_0, s_0) \cap \Omega_T$, and u and v both vanish continuously on $\Delta_{2r}(Q_0, s_0) = \Psi_{2r}(Q_0, s_0) \cap S\Omega$, then there is a constant C,* where $C = C(\lambda, d, N, r_0, T)$ *so that*

$$\left| \frac{u(x,t)}{v(x,t)} - \frac{u(y,s)}{v(y,s)} \right| \le C \frac{u(A_{r\_}(Q_0, s_0))}{v(\hat{A_r}(Q_0, s_0))} \left(\frac{d_p(x,t;y,s)}{r} \right)^{\alpha}$$

whenever (x,t) and $(y,s) \in \Psi_{r/8}(Q_0, s_0) \cap \Omega_T$ (recall that $\hat{A_r}(Q_0, s_0) = (Q_0 + c(N)r, s_0 + c(N^2)r^2)$ and $A_{r\_}(Q_0, s_0) = (Q_0 + c(N)r, s_0 - c(N^2)r^2)$ for $r < r_0$, we also take $A_r(Q_0, s_0) = (Q_0 + c(N)r, s_0))$.

Proof. Theorem 3 follows from the results in [10], Lemmas 2.8, 2.9, 2.10 and Theorem 4.2, along with Moser's method of proving Holder continuity given that Harnack's inequality for positive solutions is known. Jerison and Kenig [7] first adapted this proof to the case of positive harmonic functions. Since Lemma 2.10 in [10] is ratio Harnack at the boundary for parabolic solutions, positive on the interior of a domain, that vanish continuously at the boundary, this method works to give the result of Theorem 3 for a Lip(1,1/2) domain. As mentioned above, Theorem 3 is proved in [3] for Lipschitz cylinders.

Given the adjoint version of Theorem 3, one can proceed very much as in [8], Chapter 1.3, with $u(y_i, \hat{s_i}) = G(x,t;y_i,\hat{s_i})$ and $v(y_i, \hat{s_i}) = G(X_0, T; y_i, \hat{s_i})$. One has

$$\left| \frac{G(x,t;y_1,\hat{s_1})}{G(X_0,T;y_1,\hat{s_1})} - \frac{G(x,t;y_2,\hat{s_2})}{G(X_0,T;y_2,\hat{s_2})} \right| \lesssim \frac{G(x,t;A_{r\_}(Q_0,s_0))}{G(X_0,T;\hat{A_r}(Q_0,s_0))} \cdot \left(\frac{d_p(y_1,\hat{s_1};y_2,\hat{s_2})}{r} \right)^{\alpha}$$

for $(y_i, \hat{s_i}) \in \Psi_{r/8}(Q_0, s_0) \cap \Omega_T$, $i = 1, 2$, with (x,t) such that $d_p(x,t;Q_0,s_0) \gtrsim r$, $t \ge s_0 + c(N^2)r^2$.

Since

$$\lim_{(y_i,\hat{s_i}) \to (Q_i,s_i)} \frac{G(x,t;y_i,\hat{s_i})}{G(X_0,T;y_i,\hat{s_i})} = K(x,t;Q_i,s_i)$$

using uniqueness of the kernel function and taking limits one obtains

$$|K(x,t;Q_1,s_1) - K(x,t;Q_2,s_2)| \le C \left(\frac{d_p(Q_1,s_1;Q_2,s_2)}{r} \right)^{\alpha} \cdot \frac{G(x,t;A_{r\_}(Q_0,s_0))}{G(X_0,T;\hat{A_r}(Q_0,s_0))}.$$

Now $\frac{G(x,t;A_{r\_}(Q_0,s_0))}{G(X_0,T;\hat{A_r}(Q_0,s_0))} \simeq K(x,t;Q_0,s_0)$ by the following Lemma (take $\delta << r$).

Lemma. *If* $t > s + C(N^2)r^2$, *then*

$$\omega^{(x,t)}(\Delta_r(y',s)) \simeq \omega^{(x,t)}(\Delta_\delta(y',s))/\omega^{A_r(y',s)}(\Delta_\delta(y',s)).$$

The lemma follows from standard comparisons of the Green's function with parabolic measure, (see Lemmas 2.8 and 2.9 in [10]), and ratio Harnack at the boundary applied to the Green's function. By definition

$$K(x,t;y',s) = \lim_{\delta\downarrow 0}\left(\omega^{(x,t)}(\Delta_\delta(y',s))/\omega^{(X_0,T_0)}(\Delta_\delta(y',s))\right),$$

so

$$\frac{G(x,t;A_{r\_}(Q,s))}{G(X_0,T;A_{\hat{r}}(Q,s))} \simeq \frac{\omega^{(x,t)}(\Delta_r(y',s))}{\omega^{(X_0,T_0)}(\Delta_r(y',s))} \simeq \frac{\omega^{(x,t)}(\Delta_\delta(y',s))}{\omega^{(X_0,T_0)}(\Delta_\delta(y',s))}$$
$$\to K(x,t;y',s) \quad \text{as} \quad \delta\downarrow 0.$$

Reverse Harnack for the Green's function means

$$G(x,t;A_{r\_}(Q,s)) \simeq G(x,t;A_{\hat{r}}(Q,s)):$$

this fact has been used along with the geometric estimates $G(x,\tau;A_r(y',s)) \simeq \omega^{(x,\tau)}(\Delta_r(y',s))/r^d$ for the first comparison. The lemma gives the second comparison.

Finally, geometric decay for $K(x,t;y',s)$ gives the estimate in (9) for

$$\phi_{(I)}(y',s) = \sqrt{\omega(I)}\left(K(x_I^1,\hat{t_I};y',s) - K(x_I^2,\hat{t_I};y',s)\right),$$

whenever $(x_I^j,\hat{t_I})$ are a fixed distance above (y',s), and (y',s) is a fixed distance above the bottom of the domain Ω_T. To obtain the estimate in (9) for points (y',s) with $s\to 0$, as mentioned above, one can use an extension of the domain, $\Omega_{EXT} = \Omega_{[-1,T]}$. This device is used in [10] in the proof of Theorem 3.2. For points (y',s) when $s\uparrow t$, ratio Holder continuity (Theorem 3) is still valid when the numerator function $u(x,t)$ vanishes on the upper boundary of $\Psi_{r/8}(Q_0,s_0)$; this happens because one can use $u(x,t)+\eta b(x,t)$ as the solution instead of $u(x,t)$ with $b(x,t)$ a solution that vanishes on $\Delta_{2r}(Q_0,s_0)$ and $b(x,t)>0$ on the upper half of $\Psi_{2r}(Q_0,s_0)$, and then let $\eta\to 0$. To prove geometric decay for $K(x,t;y',s)$ with s near t, one can use Harnack's inequality to replace $K(x,t;y',s)$ with $K(\hat{x},\hat{t};y',s)$ at a point $(\hat{x},\hat{t})$ lying above the point (x,t), and then use the geometric decay of $K(\hat{x},\hat{t};y',s)$.

For the particular functions

$$\phi_{(I)}(y',s) = \sqrt{\omega(I)}\left(K(x_I^1,\hat{t_I};y',s) - K(x_I^2,\hat{t_I};y',s)\right),$$

of course, $\phi_{(I)}(y', s) = 0$ whenever $s > \hat{t_I}$. However, (8), (9) and (10) are also valid for the points where the $\phi_{(I)}$ are zero. The argument used to prove Theorem 1 is technically valid for any family of functions that satisfy (8), (9) and (10), if we define the G and g^* functions to be symmetric in the time variable (i.e. they have not been truncated to fit the lack of symmetry in time that holds for the kernel function). We can proceed with the proof of Theorem 1 obtaining upper bounds for the $\phi_{(I)}(y', s)$'s as if they were non-zero when $s \geq \hat{t_{(I)}}$. The reason for using g^*, G, and G^* functions that are not truncated in time to fit the support of the $\phi_{(I)}$'s is to simplify the statement and proofs of Lemmas 2 through 7.

Another limitation imposed by our choice of the $\phi_{(I)}(y', s)$ is that the cubes I are all bounded in size, $l(I) \leq r_0$, where r_0 is determined by the domain Ω_T. So in fact the sums over j for the regions $R_j(I)$ or $R_j(J)$ are not infinite, but merely go to some fixed j_0, so that $2^{j_0} l(I) \leq r_0$. Also the second condition (9), is only valid when (x', t) and (y', s) are within r_0 of each other; this does not create any problem proving Theorem 1 since the second condition is only used in regions of dimension $\lesssim r_0$.

The proof of Theorem 1 will appear in the next section. Right now we discuss the special situation for the heat equation on a smooth cylinder domain. When the kernel $K(x, t; y', s)$ is the kernel for the heat equation on Ω, and the domain Ω is sufficiently regular, one can prove the estimates in (A) and (B) when $(x, t) \in I$:

$$K(x, t; y', s) \lesssim l(I)^{-(d+1)} \left(1 + \frac{d_p(x, t; y', s)}{l(I)}\right)^{-(d+2)}, \qquad \text{(A)}$$

$$\begin{aligned} &|K(x, t; Q_1, s_1) - K(x, t; Q_2, s_2)| \\ &\quad \lesssim l(I)^{-(d+1)} \left(1 + \frac{d_p(x, t; Q_0, s_0)}{l(I)}\right)^{-(d+2)} \cdot \left(\frac{d_p(Q_1, s_1; Q_2, s_2)}{l(I)}\right)^{\alpha}, \qquad \text{(B)} \end{aligned}$$

if $(Q_i, s_i) \in R_j(I)$. $R_j(I) = \left\{(y, s) : 2^{j-1} l(I) \leq d_p(y, s; x'_I, t_I) \leq 2^j l(I)\right\}$ and (Q_0, s_0) is the center of the sub-region in $R_j(I)$ that contains both (Q_1, s_1) and (Q_2, s_2).

(A) and (B) are valid because for the heat kernel K, on a smooth cylinder domain, $K(x, t; y', s) = \lim_{r \downarrow 0} \left(\frac{G(x,t;A_r(y',s))}{r}\right)$ (see [4]). The heat kernel on the right half space also has the same identification as the limit of the Green's function divided by the distance to the lateral boundary of the domain; consequently (A) and (B) hold on any domain that has the property that $\partial_p^+ \Omega_T$

is made up of vertical lines that are tangent to a half space along their entire length. This follows from the half space result and the fact that the maximum principle applied to the Green's functions of the respective domains gives the estimate $K_\Omega(x,t;y',s) \lesssim K_{RHS}(x,t;y',s)$. The upper bounds on the heat kernel for the right half space, given in [16], provide the rest of the proof for (A) and (B). Holder continuity for $K_\Omega(x,t;y',s)$ in the variables (y',s) can be derived directly from boundary Holder continuity for the Green's function. One result of this is that the constants in (A) and (B) do not depend on T, so the domain Ω can be semi-infinite in time. However, one does have to assume that any solution $u(x,t) \longrightarrow 0$ as $t \to -\infty$ to verify the almost orthogonality condition in (10).

Using $\sqrt{|I|} = l(I)^{\frac{d+1}{2}}$ instead of $\sqrt{\omega(I)}$, the functions $\phi_{(I)}(y',s)$ can be defined as $l(I)^{\frac{d+1}{2}}\left(K(x_I^1,\hat{t_I};y',s) - K(x_I^2,\hat{t_I};y',s)\right)$. Then estimates (A) and (B) provide the necessary bounds on the $\phi_{(I)}$ to prove that

$$\left(\int_{\partial_p\Omega} \left|f(x',t)\right|^p \sigma(x',t)d\omega(x',t)\right)^{1/p} \leq C\left(\int_{\partial_p\Omega} \left|g^*(f)(x',t)\right|^p \sigma(x',t)d\omega(x',t)\right)^{1/p}, \quad \text{(LP)}$$

for any finite sum of the form $f(x',t) = \sum \lambda_I \phi_{(I)}(x',t)$, and for $g^*(f)(y',s) = \left(\sum_{I\in\mathcal{F}} \left(\lambda_I^2/|I|\right) \cdot [1 + (d_p(y',s;x_I',t_I)/l(I)]^{-(2M-d-1-\epsilon)}\right)^{1/2}$. As above we assume that $M > d+1$.

Again the Littlewood-Paley type inequality (LP) can be used to prove the result of Theorem 2 with $d\omega(x',t)$ replaced by surface measure, $dm(x',t) = dx'dt =$ Lebesgue measure on $\partial_p\Omega$. (LP) is valid from the proof in [12].

The proof of Theorem 1 in its original formulation that is given below, however, follows the same general outline in [15] with technical changes necessary for dealing with parabolic functions on rough domains.

Remark. For the heat operator $\partial/\partial t - \Delta$ on a Lipschitz cylinder, i.e. on a domain of the form $D \times (0,T]$, with D being a Lipschitz domain in $\mathbb{R}^d$, Theorem 1 and Theorem 2 must be proved using caloric measure, not surface measure.

5. Section

To prove Theorem 1, we want to prove the Parabolic Main Lemma. So we must first establish versions of Lemmas 1 - 7 in [15]. The main difference here from [15] is that our parabolic cubes are no longer Euclidean cubes in $\mathbb{R}^d$. We express this fact by writing $l(I) \simeq 2^n l(J)$ when I and J are the images under the boundary map φ of actual dyadic, non-isotropic cubes in $\mathbb{R}^d$, whose dimensions in $\mathbb{R}^d$ have the exact relation $l(\varphi^{-1}(I)) = 2^n l(\varphi^{-1}(J))$, and so on. One can think of the sums as being taken over cubes in $\mathbb{R}^d$, since the essential properties of being nested and disjoint have been preserved, although the exact correspondence of length and distance is only comparative on $S\Omega$.

Let $\mathcal{S}(I) = \bigcup_{J \subsetneqq I,\, J \in \mathcal{D}} J$. The functions $F(I,(x',t))$, $F(I)$, $F^*(x',t)$ and $G(I;x'.t)$, $G(I)$ were defined in the introduction to this paper for $(x',t) = (x'_{(I)}, t_{(I)})$. For any point $(x',t) \in I$, we take $F(I,(x',t)) \doteq \sum_{J \in \mathcal{S}(I)} \lambda_J \phi_{(J)}(x',t)$, and so on. For (x',t) lying outside of I, the functions $F(I,(x',t))$ and $G(I,(x',t))$ are not defined. We note that these "partial sum" functions have almost exactly the same definitions here that they have in [15]. Notice that we take $F(I) = F(I,(x'_{(I)}, t_{(I)})) = \sum_{J\in\mathcal{S}(I)} \lambda_J \phi_{(J)}(x'_{(I)}, t_{(I)})$, where $(x'_{(I)}, t_{(I)})$ is in the geometric center of the parabolic cube I. As defined previously, $\mathcal{D}$ is the collection of all parabolic dyadic cubes J on the lateral boundary of the domain Ω_T. The constants that appear below depend on $d, \lambda, N, \epsilon, \alpha, T, r_0$.

Lemma 1. *$f(x',t) \le F^*(x',t)$, ω-a.e. when $(x',t) \in S\Omega$.*

Proof. This follows directly from the definition of $F^*(x',t) = \sup_{I \ni (x',t)} F(I)$.

Recall that the regions $2^j J = \Delta_{2^j l(J)}(x'_J, t_J)$ and $R_j(J) = 2^j J \setminus 2^{j-1} J = \{(y',s) : d_p(y',s;J) \simeq 2^j l(J)$.

Lemma 2. *$G^*(x',t) \le C g^*(x',t)$, for ω-a.e. $(x',t) \in S\Omega$.*

Proof.

$$
\begin{aligned}
G^*(x',t) &= \sup_{I \ni (x',t)} G(I) \\
&= \sup_{I \ni (x',t)} \left(\sum_{J \in \mathcal{S}(I)} |\lambda_J|^2 \cdot \sum_{j=0}^{\infty} (2^{-j(2\beta-\tau)} \chi_{R_j(J)}(x'_I, t_I)/\omega(\Delta_{2^j l(J)}(x'_J, t_J)) \right)^{1/2} \\
&\lesssim \sup_{I \ni (x',t)} \left(\sum_{J \in \mathcal{S}(I)} |\lambda_J|^2 \cdot \sum_{j=0}^{\infty} (2^{-j(2\beta-\tau)} \chi_{R_j(J)}(x',t)/\omega(\Delta_{2^j l(J)}(x'_J, t_J)) \right)^{1/2}
\end{aligned}
$$

because if $(x',t) \in I$ and $J \in \mathcal{S}(I)$, then either $I \subsetneq J$, in which case both (x',t) and (x'_I, t_I) lie in $R_j(J)$ (which is J in this case since $j=0$), so either

$$\sum_{j=0}^{\infty}(2^{-j(2\beta-\tau)}\chi_{R_j(J)}(x'_I,t_I)/\omega(\Delta_{2^j l(J)}(x'_J,t_J))$$
$$\lesssim \sum_{j=0}^{\infty}(2^{-j(2\beta-\tau)}\chi_{R_j(J)}(x',t)/\omega(\Delta_{2^j l(J)}(x'_J,t_J)),$$

or $J \cap I$ is empty, in which case $d_p(x',t;J) \lesssim d_p(x'_I,t_I;J)$. This last inequality implies that

$$\sum_{j=0}^{\infty}(2^{-j(2\beta-\tau)}\chi_{R_j(J)}(x'_I,t_I)/\omega(\Delta_{2^j l(J)}(x'_J,t_J))$$
$$\leq C\sum_{j=0}^{\infty}(2^{-j(2\beta-\tau)}\chi_{R_j(J)}(x',t)/\omega(\Delta_{2^j l(J)}(x'_J,t_J)).$$

Lemma 3. *If for $0<\eta<1$, $(x',t) \in (1-\eta)I$, then $G(I) = G(I;x'_I,t_I) \simeq G(I;x',t)$.*

Proof. $G(I;x',t) = \left(\sum_{J\in\mathcal{S}(I)} |\lambda_J|^2 \sum_{j=0}^{\infty}(2^{-j(2\beta-\tau)}\chi_{R_j(J)}(x',t)/\omega(2^j J))\right)^{1/2}$. If $I \subsetneq J$, then for $(x',t) \in (1-\eta)I$, (x'_I,t_I) also lies in $(1-\eta)I$, so both (x',t) and (x'_I,t_I) lie inside $2^j J$ for $j=0$. This gives that

$$\sum_{j=0}^{\infty}(2^{-j(2\beta-\tau)}\chi_{R_j(J)}(y',s)/\omega(2^j J)) = 1/\omega(J)$$

for $(y',s) = (x',t)$ or (x'_I,t_I). If $J \cap I$ is empty, then either $l(J) \gtrsim l(I)$, or $l(J) < l(I)$. In the first case, $I \subseteq R_j(J) \cup R_{j+1}(J)$ for some $j \geq 1$, and this means that both (x',t) and (x'_I,t_I) lie inside $R_j(J) \cup R_{j+1}(J)$ for the same index j. If $l(J) < l(I)$, then for J located far from I, say distance $(J,I) \gtrsim 2^k l(I)$, both (x',t) and (x'_I,t_I) will lie inside some $R_{j_0}(J)$ or in $R_{j_0 \pm l}(J)$ for $l \leq l_0$, where l_0 is fixed. This happens because if $2^m l(J) \simeq l(I)$, then the distance from (x',t) to J is $\delta l(I) + d_p(I,J)$ and $d_p((x'_I,t_I);J) \simeq \frac{1}{2}l(I) + d_p(I,J)$. Also $\eta/2 \leq |\delta| \leq C'(d)(1-\eta/2)$, so $d_p((x',t);J) = d_p((x'_I,t_I);J) \pm Cl(I) = d_p(I,J) \pm C''l(I)$. Now $d_p(I,J) \simeq 2^k l(I)$ and $l(I) \simeq 2^m l(J)$, consequently $d_p(I,J) \pm C''l(I) \simeq (2^k l(I) \pm C2^m l(J))$; this quantity is $\lesssim (2^k + C(d))2^m l(J)$ and is also $\gtrsim (2^k - C(d))2^m l(J)$, because C'' is a dimensional constant. This means that we must choose k large enough so that $2^k - C(d) > 0$. That is o.k. Now both (x', t) and (x'_I, t_I) will lie inside $2^{m+k+l}J$ and outside $2^{m+k-l}J$, for some $l \le l_0$, where $2^{l_0} \simeq C(d)$, i.e. both points lie in neighboring J annuli. The doubling condition on $\omega^{(X_0,T)}$ implies that $\sum_{j=0}^{\infty}(2^{-j(2\beta-\tau)}\chi_{R_j(J)}(x', t)/\omega(2^j J)) \simeq \sum_{j=0}^{\infty}(2^{-j(2\beta-\tau)}\chi_{R_j(J)}(x'_I, t_I)/\omega(2^j J))$ for this situation.

If $d_p(I, J) \lesssim l(I)$, then obviously $d_p((x', t); J) \simeq d_p((x'_I, t_I); J) \simeq 2^m l(J)$, so $\sum_{j=0}^{\infty}(2^{-j(2\beta-\tau)}\chi_{R_j(J)}(x', t)/\omega(2^j J)) \simeq \sum_{j=0}^{\infty}(2^{-j(2\beta-\tau)}\chi_{R_j(J)}(x'_I, t_I)/\omega(2^j J))$ for this situation also.

Lemma 4. *If $(x', t) \in (1-\eta)I$, then there is a constant C, where $C = C(d, \lambda, N, T, r_0, \alpha, \eta, \beta, \tau)$ such that $|F(I; (x', t)) - F(I)| \le CG(I)$.*

Proof.

$$\begin{aligned}
\left|F(I; (x', t)) - F(I)\right| &= \left|\sum_{J\in\mathcal{S}(I)} \lambda_J(\phi_J(x', t) - \phi_J(x'_I, t_I))\right| \\
&\le \left|\sum_{J\in\mathcal{S}(I), l(J)>l(I)} \lambda_J(\phi_J(x', t) - \phi_J(x'_I, t_I))\right| \\
&\quad + \left|\sum_{J\in\mathcal{S}(I), l(J)\lesssim l(I)} \lambda_J(\phi_J(x', t) - \phi_J(x'_I, t_I))\right| = I + II.
\end{aligned}$$

I will be controlled by using Holder continuity; II will be estimated using the brute force estimate

$$\begin{aligned}
&\left|\sum_{J\in\mathcal{S}(I), l(J)\lesssim l(I)} \lambda_J(\phi_J(x', t) - \phi_J(x'_I, t_I))\right| \\
&\qquad \le \sum_{l(J)\lesssim l(I), I\cap J=\emptyset} \left|\lambda_J\phi_J(x', t)\right| + \sum_{l(J)\lesssim l(I), I\cap J=\emptyset} \left|\lambda_J\phi_J(x'_I, t_I)\right|
\end{aligned}$$

and geometric decay for the ϕ_I's ((8)).

First we control I as follows:

$$I \le \sum_{J\in\mathcal{S}(I), l(J)>l(I)} |\lambda_J| \sqrt{\omega(J)} \left(\frac{d_p((x', t); (x'_I, t_I))}{l(J)}\right)^{\alpha}$$

$$\times \sum_{j=0}^{\infty} \left(2^{-j\beta}\chi_{R_j(J)}(x^*,t^*)\right))/\omega(2^jJ)$$

$$\lesssim \left(\sum_{J\in\mathcal{S}(I),l(J)>l(I)} |\lambda_J|^2 \sum_{j=0}^{\infty} \left(2^{-j(2\beta-\tau)}\chi_{R_j(J)}(x^*,t^*)\right)/\omega(2^jJ) \right)^{1/2}$$

$$\times \left(\sum_{J\in\mathcal{S}(I),l(J)>l(I)} (l(I)/l(J))^{\alpha}\omega(J) \sum_{j=0}^{\infty} \left(2^{-j\tau}\chi_{R_j(J)}(x^*,t^*)\right)/\omega(2^jJ) \right)^{1/2}.$$

Here (x^*,t^*) can be either (x',t) or (x'_I,t_I), depending on which point is closer to J. The first term in the last expression is $\lesssim G(I)$ by Lemma 3. The second term can be written as two different sums, $\sum\limits_{J\in\mathcal{S}(I),J\supsetneqq I} + \sum\limits_{J\in\mathcal{S}(I),J\cap I=\emptyset}$. For the first sum

$$\sum_{J\in\mathcal{S}(I),J\supsetneqq I} (l(I)/l(J))^{\alpha}\omega(J) \sum_{j=0}^{\infty} \left(2^{-j\tau}\chi_{R_j(J)}(x^*,t^*)\right)/\omega(2^jJ)$$

$$= \sum_{k=0}^{\infty} \left(\sum_{J\in\mathcal{S}(I),l(J)\simeq 2^k l(I)} 2^{-\alpha k}\omega(J) \sum_{j=0}^{\infty} \left(2^{-j\tau}\chi_{R_j(J)}(x^*,t^*)\right)/\omega(2^jJ) \right)$$

$$= \sum_{k=0}^{\infty} 2^{-\alpha k} \sum_{l(J)\simeq 2^k l(I)} \omega(J) \sum_{j=0}^{\infty} \left(2^{-j\tau}\chi_{R_j(J)}(x^*,t^*)\right)/\omega(2^jJ).$$

Now $(x^*,t^*)\in(1-\eta)I$ and J has the property that $l(J)\simeq 2^k l(I)$, $I\subset J$, so $(x^*,t^*)\in 2^0J=J$ and $\chi_{R_0(J)}(x^*,t^*)=1$, but $\chi_{R_j(J)}(x^*,t^*)=0$ if $j>0$. So the first sum is bounded by $C\sum\limits_{k=0}^{\infty} 2^{-\alpha k}\left(\frac{\omega(J)}{\omega(J)}\right) = C(\alpha,\beta,N,\tau,T,\lambda,d)$.

The second sum is also bounded by a constant since

$$\sum_{J\in\mathcal{S}(I),l(J)>l(I),J\cap I=\emptyset} (l(I)/l(J))^{\alpha}\,\omega(J) \sum_{j=0}^{\infty} \left(2^{-j\beta}\chi_{R_j(J)}(x^*,t^*)\right)/\omega(2^jJ)$$

$$= \sum_{k=0}^{\infty} 2^{-\alpha k} \sum_{l(J)\simeq 2^k l(I)} \omega(J) \sum_{j=0}^{\infty} \left(2^{-j\beta}\chi_{R_j(J)}(x^*,t^*)\right))/\omega(2^jJ)$$

as before.

Writing the second sum over $l(J)\simeq 2^k l(I)$ as sums over J such that $2^m l(I)\simeq d_p(x'_I,t_I;x'_J,t_J)$, for $m\geq k$, gives

$$\sum_{k=0}^{\infty} 2^{-\alpha k} \sum_{m=k}^{\infty} \sum_{\substack{d_p(x'_I,t_I;J)\simeq 2^m l(I) \\ l(J)\simeq 2^k l(I)}} \omega(J) \sum_{j=0}^{\infty} \left(2^{-j\beta}\chi_{R_j(J)}(x^*,t^*)\right)/\omega(2^j J).$$

To estimate this sum we note that if $l(J) = 2^k l(I)$ and if $d_p(x'_I, t_I; x'_J, t_J) \simeq 2^m l(I)$, then $(x'_I, t_I) \in R_{m-k\pm l}(J)$ for $l \leq l_0$ fixed, because the distance from (x'_I, t_I) to $(x'_J, t_J) \simeq 2^{m-k} l(J)$. Also $(x^*, t^*) \in (1-\eta)I$ so $\chi_{R_j(J)}(x^*, t^*) = 1$ when $j \sim (m-k)$. As a result the last sum is bounded above by a constant times

$$\sum_{k=0}^{\infty} 2^{-\alpha k} \sum_{m=k;}^{\infty} \sum_{d_p(x'_I,t_I;x'_J,t_J)\simeq 2^m l(I), l(J)\simeq 2^k l(I)} \omega(J) \left(\frac{2^{-(m-k)\beta}}{\omega(2^{(m-k)}J)}\right).$$

Using the center doubling property of ω (see [10] and [3]) the last sum is bounded by

$$C \sum_{m=1}^{\infty} 2^{-m\beta} \sum_{k=0}^{m} 2^{-k(\alpha-\beta)} \leq C'(\alpha, \beta).$$

This happens because for m and k fixed, $m \geq k$,

$$\begin{aligned} &\sum_{d_p(x'_I,t_I;x'_J,t_J)\simeq 2^m l(I), l(J)\simeq 2^k l(I)} \omega(J) \cdot \left(\frac{2^{-(m-k)\beta}}{\omega(2^{(m-k)}J)}\right) \\ &\qquad = \frac{C}{\omega(2^m I)} \cdot \sum_J \omega(J) 2^{-(m-k)\beta} \\ &\qquad \leq C 2^{-(m-k)\beta} \omega(2^m + 2^k)I)/\omega(2^m I) \\ &\qquad\qquad \leq C 2^{-(m-k)\beta} \omega(2^{m+1} I)/\omega(2^m I) \leq C 2^{-(m-k)\beta}. \end{aligned}$$

Altogether

$$\left| \sum_{J\in\mathcal{S}(I), l(J)>l(I)} \lambda_I(\phi_I(x',t) - \phi_I(x'_I, t_I)) \right| \leq CG(I).$$

For the second term, II, it is enough to estimate

$$\sum_{J\in\mathcal{S}(I), l(J)\lesssim l(I)} \lambda_J(\phi_J(x',t)), \text{ for } (x',t) \in (1-\eta)I,$$

since $(x_I', t_I) \in (1-\eta)I$. Then using (8) and Cauchy Schwartz twice gives

$$\left| \sum_{J\in\mathcal{S}(I), l(J)\lesssim l(I), J\cap I=\emptyset} \lambda_J(\phi_J(x',t)) \right|$$
$$\lesssim \sum_{J\in\mathcal{S}(I), l(J)\lesssim l(I), J\cap I=\emptyset} |\lambda_J| \sqrt{\omega(J)} \left(\sum_{j=0}^{\infty} \left(2^{-j\beta} \chi_{R_j(J)}(x',t)/\omega(2^j J) \right) \right)$$
$$\lesssim \left(\sum_{l(J)\lesssim l(I), J\cap I=\emptyset} |\lambda_J|^2 \left(\sum_{j=0}^{\infty} \left(2^{-j(2\beta-\tau)} \chi_{R_j(J)}(x',t)/\omega(2^j J) \right) \right) \right)^{1/2}$$
$$\times \left(\sum_{l(J)\lesssim l(I), J\cap I=\emptyset} \omega(J) \left(\sum_{j=0}^{\infty} \left(2^{-j\tau} \chi_{R_j(J)}(x',t)/\omega(2^j J) \right) \right) \right)^{1/2}.$$

The first factor in the last expression is $\le CG(I)$ by Lemma 3. The square of the second term equals

$$\sum_{k=0}^{\infty} \sum_{l(J)\simeq 2^{-k}l(I), J\cap I=\emptyset} \omega(J) \left(\sum_{j=0}^{\infty} \left(2^{-j\tau} \chi_{R_j(J)}(x',t)/\omega(2^j J) \right) \right)$$
$$= \sum_{k=0}^{\infty} \sum_{l(J)\simeq 2^{-k}l(I), J\cap I=\emptyset} \sum_{m=-m_0}^{\infty} \omega(J) \left(\sum_{j=0}^{\infty} \left(2^{-j\tau} \chi_{R_j(J)}(x',t)/\omega(2^j J) \right) \right).$$

The index m is determined by $d_p(x_J', t_J; (1-\eta)I) \simeq 2^m l(I)$; $m_0 = m_0(\eta)$ is fixed. The last expression equals

$$C \sum_{k=0}^{\infty} \sum_{m=-m_0}^{\infty} \sum_{l(J)\simeq 2^{-k}l(I)} \omega(J) \cdot \left(\frac{2^{-(m+k)\tau}}{\omega(2^{(m+k)}J)} \right)$$

because $(x',t) \in (1-\eta)I$, $l(J) = 2^{-k}l(I)$, and $d_p(J; (1-\eta)I) \simeq 2^m l(I)$ implies that $\chi_{R_j(J)}(x',t) \neq 0$ only when $j \simeq m+k$.

Now $d_p(x',t;J) \simeq 2^m l(I)$, and this implies there is an m_1 such that m_1 is independent of m and k, and $2^m I \subset 2^{m+k+m_1} J$, and $2^{m+k} J \subset 2^{m+m_1} I$. As noted above ω is a measure that satisfies a center doubling condition, so $\omega(2^{m+k}J) \simeq \omega(2^m I)$. This means that we can estimate the last sum by

$$C\sum_{k=0}^{\infty}\sum_{m=-m_0}^{\infty}2^{-(k+m)\tau}/\omega(2^mI)\sum_{l(J)\simeq 2^{-k}l(I),d_p(x'_I,t_I;x'_J,t_J)\simeq 2^ml(I)}\omega(J)$$
$$\lesssim\sum_{k=0}^{\infty}\sum_{m=-m_0}^{\infty}\left(2^{-(k+m)\tau}/\omega(2^mI)\right)\cdot\omega(2^{m+m_1}I)$$
$$\lesssim\sum_{k=0}^{\infty}2^{-k\tau}\sum_{m=-m_0}^{\infty}2^{-m\tau}\le C(\eta,\lambda,N,T,r_0,\alpha,\tau,d).$$

The next three lemmas deal with nearest neighbor estimates. We define $I^{\hat{}}\in N(I)$ to mean that $I^{\hat{}}\subsetneqq I$ and that $l(I^{\hat{}})\simeq\frac{1}{2}l(I)$.

Lemma 5. $G(I)\le C(\lambda,N,T,r_0,\alpha,\beta,\tau,d)G(I^{\hat{}})$.

Proof. Lemma 3 says that $G(I;(x_{I^{\hat{}}},t_{I^{\hat{}}}))\simeq G(I)$ because $(x_{I^{\hat{}}},t_{I^{\hat{}}})\in(1-\eta)I$ for $\eta<\frac{1}{4}$, for example. But $I^{\hat{}}\subsetneqq I$, so $\mathcal{S}(I)\subset\mathcal{S}(I^{\hat{}})$ and so $G(I;(x_{I^{\hat{}}},t_{I^{\hat{}}}))\lesssim G(I^{\hat{}})$. □

Lemma 6. *For all* $(x',t)\in I$, $G^*(x',t)\ge CG(I^{\hat{}})$.

Proof. If $(x',t)\in I$, then when $(x',t)\in(1-\eta)I^{\hat{}}$, the result is obvious by Lemma 3 and the definition of G^*. If (x',t) is close to the boundary of $I^{\hat{}}$, choose $J\subsetneqq I^{\hat{}}$ so that $(x',t)\in J$. Then $\mathcal{S}(I^{\hat{}})\subset\mathcal{S}(J)$ and $G(J;(x',t))$ have all the terms that $G(I^{\hat{}};(x'_{I^{\hat{}}},t_{I^{\hat{}}}))$ does, plus possibly more. Also for $Q\in\mathcal{S}(I^{\hat{}})\cap\mathcal{S}(J)$, either $d_p(Q;(x',t))\simeq d_p(Q;(x'_{I^{\hat{}}},t_{I^{\hat{}}}))$ or $d_p(Q;(x',t))\ll d_p(Q;(x'_{I^{\hat{}}},t_{I^{\hat{}}}))$. Either way we have

$$\sum_{j=0}^{\infty}(2^{-j(2\beta-\tau)}\chi_{R_j(Q)}(x',t))/\omega(2^jQ)$$
$$\ge C\sum_{j=0}^{\infty}(2^{-j(2\beta-\tau)}\chi_{R_j(Q)}(x'_{I^{\hat{}}},t_{I^{\hat{}}}))/\omega(2^jQ).$$

So $G^*(x',t)\ge CG(I^{\hat{}})$ in this case also.

In the remaining case, that $(x',t)\in I-I^{\hat{}}$, suppose that $(x',t)\in L\subset I-I^{\hat{}}$. The cubes in $\mathcal{S}(I^{\hat{}})-\mathcal{S}(L)$ are those cubes that lie inside L. We intend to bound $G(I^{\hat{}})$ by $G(L;(x'_L,t_L))$ plus the sum of terms over $Q\subset L$. $G(I^{\hat{}})\le CG(L;(x'_{I^{\hat{}}},t_{I^{\hat{}}}))+\sum_{Q\subset L}\omega(Q)\sum_{j=0}^{\infty}(2^{-j(2\beta-\tau)}\chi_{R_j(Q)}(x'_{I^{\hat{}}},t_{I^{\hat{}}}))/\omega(2^jQ)$, so we need to show that $(x'_{I^{\hat{}}},t_{I^{\hat{}}})$ can be replaced by (x'_L,t_L) in the first expression. Now if $J\in\mathcal{S}(I^{\hat{}})$, then

$$d_p((x'_J,t_J);(x'_L,t_L))\le d_p((x'_J,t_J);(x'_{I^{\hat{}}},t_{I^{\hat{}}}))+d_p((x'_{I^{\hat{}}},t_{I^{\hat{}}});(x'_L,t_L))$$

$$\leq d_p((x'_J, t_J); (x'_{I^\wedge}, t_{I^\wedge})) + Cl(I^\wedge) \leq C' d_p((x'_J, t_J); (x'_{I^\wedge}, t_{I^\wedge})),$$

where $C' = C'(N, d)$, N being the Lipschitz constant of the domain. Then $(x'_{I^\wedge}, t_{I^\wedge}) \in R_j(J)$, and one has that $d_p((x'_J, t_J); (x'_{I^\wedge}, t_{I^\wedge})) \simeq 2^j l(J)$ and

$d_p((x'_J, t_J); (x'_L, t_L)) \leq C 2^j l(J)$, so $\chi_{R_{j+m_0}(J)}(x'_L, t_L) = 1$. Using the doubling property for ω and the fact that C is fixed implies m_0 must be fixed, gives

$$\sum_{j=0}^{\infty} (2^{-j(2\beta-\tau)} \chi_{R_j(J)}(x'_{I^\wedge}, t_{I^\wedge}))/\omega(2^j J)$$

$$\leq C \sum_{j=0}^{\infty} (2^{-(j+m_0)(2\beta-\tau)} \chi_{R_j(J)}(x'_L, t_L))/\omega(2^{j+m_0} J)$$

So

$$G(I^\wedge) \leq C \left(G(L; (x'_L, t_L)) + \sum_{Q \subset L} \omega(Q) \sum_{j=0}^{\infty} (2^{-j(2\beta-\tau)} \chi_{R_j(Q)}(x'_{I^\wedge}, t_{I^\wedge}))/\omega(2^j Q) \right).$$

Letting $l(L) \to 0$ implies that $G(I^\wedge) \leq CG^*(x', t)$ because $G(L; (x'_L, t_L)) \leq G^*(x'_L, t_L)$, and letting $l(L) \to 0$ forces $(x'_L, t_L) \to (x', t)$ and $\sum_{Q \subset L} \to 0$ as well. □

Lemma 7. $\left|F(I) - F(I^\wedge)\right| \leq CG(I^\wedge)$.

Proof. Since $(x'_{I^\wedge}, t_{I^\wedge}) \in (1-\eta)I$, $\left|F(I; (x'_I, t_I)) - F(I; (x'_{I^\wedge}, t_{I^\wedge}))\right| \leq CG(I)$ by Lemma 4; Lemma 5 gives that $G(I) \leq CG(I^\wedge)$. So

$$\left|F(I) - F(I^\wedge)\right| \leq \left|F(I; (x'_{I^\wedge}, t_{I^\wedge})) - F(I^\wedge; (x'_{I^\wedge}, t_{I^\wedge}))\right| + CG(I^\wedge).$$

Now

$$\left|F(I; (x'_{I^\wedge}, t_{I^\wedge})) - F(I^\wedge; (x'_{I^\wedge}, t_{I^\wedge}))\right|$$

$$\leq \sum_{J \in \mathcal{S}(I^\wedge) \setminus \mathcal{S}(I)} |\lambda_J| \sqrt{\omega(J)} \left(\sum_{j=0}^{\infty} (2^{-j\beta} \chi_{R_j(J)}(x'_{I^\wedge}, t_{I^\wedge}))/\omega(2^j J) \right)$$

$$\leq G(I^\wedge) \cdot \left(\sum_{J \subset I \setminus I^\wedge} \omega(J) \sum_{j=0}^{\infty} (2^{-j\tau} \chi_{R_j(J)}(x'_{I^\wedge}, t_{I^\wedge}))/\omega(2^j J) \right)^{1/2} \leq CG(I^\wedge).$$

The second term in the last product was bounded by a constant exactly the same way $\left(\sum_{l(J)\lesssim l(I),J\cap I=\emptyset}\omega(J)\left(\sum_{j=0}^{\infty}\left(2^{-j\tau}\chi_{R_j(J)}(x',t)/\omega(2^jJ)\right)\right)\right)$ was estimated in the proof of Lemma 4. $\square$

We now state a parabolic version of Main Lemma and its proof (see [13]).

Parabolic Main Lemma. *For any* $\beta > 0$, *there is a* $\gamma = \gamma(d, \Omega_T, \beta, \alpha, \tau)$ *so that on a fixed cube* I_0, *if* $f(x',t) = \sum_{I\in\mathfrak{F}}\lambda_I\phi_I(x',t)$ *is any finite sum with* $\lambda_I = 0$ *whenever* $I \nsubseteq I_0$, *with the* ϕ_I *satisfying* (8), (9), (10) *then*

$$\omega(\{(x',t)\in I_0 : F^*(x',t) > 1, G^*(x',t)\le\gamma\}) \le \beta\omega(I_0).$$

Proof of Parabolic Main Lemma. (see [13], [15]): Let $E = \{(x',t) \in I_0 : F^*(x',t) > 1, G^*(x',t) \le \gamma\}$. To prove that for any given value of $\beta > 0$, there is a γ so that $\omega(E) \le \beta\omega(I_0)$, we begin by creating a family of maximal, disjoint parabolic cubes. First, mimicking the proof in [15], we choose maximal parabolic subcubes $\{I_k\}$ in I_0 so that $G(I^*) > A\gamma$, for some "next generation down" subcube $I^* \subsetneqq I_k \subseteq I_0$. $A > 0$ is a fixed constant, chosen so that $\cup I_k \subset \{(x',t) : G^*(x',t) > \gamma\}$; we will see momentarily that it is possible to choose A so this happens. A also depends on the constant in Lemma 6. Notice that $G(I_k) \le A\gamma$ and $G^*(x',t) \le A\gamma$ if $(x',t) \notin \cup I_k$.

Next we choose another family of parabolic dyadic cubes $\{J_l\}$. Each $J_l \subset I_0$ and is chosen to be the maximal such cube so that $F(J_l) > 1$ and so that J_l is not contained in any I_k. This means that a cube J_l is either disjoint from $\cup I_k$ or that it "straddles" one or more of the I_k. Notice that $E \subset \cup J_l$. This is true because if $(x',t) \in E$, but $(x',t) \notin \cup J_l$, then there is some $Q \ni (x',t)$ so that $F(Q) > 1$ and $Q \subset I_{k_0}$for some k_0. But then by Lemma 6 $G^*(x',t) \ge G(I_{k_0}) \ge CG(I^*) \ge CA\gamma$, so if $AC > 1$, this implies (x',t) cannot lie in E.

Now the $\{I_k\}$ and the $\{J_l\}$ are put into one family and the maximal cubes are chosen to form a new family $\{P_j\}$; and we have $G(P_j) \le A\gamma$. To bound $\omega(E)$, since $E \subset \cup_{F(P_j)>1}P_j$ and since ω satisfies a center doubling condition, we need only bound the sum $\sum_{F(P_j)>1}\omega(c(P_j))$, where $c(P_j) = \{(y',s) \in P_j : d_p(x'_{P_j}, t_{P_j}; y', s) < 0.1 l(P_j)\}$.

Proceeding as in [13] and [15] we define two collections of parabolic dyadic cubes, $\mathfrak{C}_1 = \{Q \subset I_0 : Q \nsubseteq P_j$ for any $j\}$ and $\mathfrak{C}_2 = \{Q \subset I_0 : Q \subseteqq P_j$ for some $j\}$. Now defining $f_i(x',t) = \sum_{I, I\in\mathfrak{F}\cap\mathfrak{C}_i}\lambda_I\phi_I(x',t)$ gives that $f(x',t) = f_1(x',t) + f_2(x',t)$.

Let $F(Q;x',t)$, $F(Q)$, $F^*(x',t)$, $F_i(Q;x',t)$, $F_i(Q)$, $F_i^*(x',t)$ and $G(Q;x',t)$, $G(Q)$, $G^*(x',t)$, $G_i(Q;x',t)$, $G_i(Q)$, $G_i^*(x',t)$ be the localized functions, corresponding to f and f_i, as defined in the introduction, with x replaced by (x',t). We have the simple inequalities $F^*(x',t) \leq F_1^*(x',t) + F_2^*(x',t)$ and $G_i(P;x',t) \leq G(P;x',t) \leq G_1(P;x',t) + G_2(P;x',t)$. In addition note that if $(x',t) \in P_j$ for one of the maximal cubes, then $F_1(P_j;x',t) = f_1(x',t)$ whenever $(x',t) \in P_j$ by definition of $\mathfrak{C}_1$. These facts will be used in the following argument.

We now proceed to replace $\sum_{F(P_j)>1} \omega(c(P_j))$ by

$$\sum_j \omega(\{(x',t) \in c(P_j) : F_1(P_j;x',t) > 0.25\}) + \sum_j \omega(\{(x',t) \in c(P_j) : F_2(P_j;x',t) > 0.25\}).$$

We can do this since Lemma 4 shows that for $(x',t) \in c(P_j)$, if $i = 1,2$, then $|F_i(P_j;x',t) - F_i(P_j)| \leq CG(P_j) \leq CA\gamma$. For γ sufficiently small the fact that $F_i(P_j) > \frac{1}{2}$ implies $F_i(P_j;x',t) > \frac{1}{4}$. Together with the fact that $1 < F(P_j) \leq F_1(P_j) + F_2(P_j)$ only when at least one of the terms $F_i(P_j) > \frac{1}{2}$, this implies that

$$\sum_{F(P_j)>1} \omega(c(P_j)) \leq \sum_j \omega(\{(x',t) \in c(P_j) : F_1(P_j;x',t) > .25\}) + \sum_j \omega(\{(x',t) \in c(P_j) : F_2(P_j;x',t) > .25\}).$$

First, Chebyshev's inequality implies that $\omega(\{(x',t) \in H : F_i(P_j;x',t) > \frac{1}{4}\}) \leq 16 \int_H |F_i(P_j;x',t)|^2 \, d\omega(x',t)$. From here on the two sums are estimated by different methods. For the first sum we take $H = c(P_j)$; for the second sum we will have to cut out a (fortunately) small set of points in $c(P_j)$. Proceeding to estimate the first sum we now use the fact noted above, that if $(x',t) \in c(P_j)$, then $F_1(x',t) = f_1(x',t)$. So $16 \int_{c(P_j)} |F_1(P_j;x',t)|^2 \, d\omega(x',t) = 16 \int_{c(P_j)} |f_1(P_j;x',t)|^2 \, d\omega(x',t)$. It is at this point that we use condition (10), the almost orthogonality condition. It gives $\sum_{P_j} \int_{c(P_j)} |f_1(P_j;x',t)|^2 \, d\omega(x',t) \leq$

$\sum\limits_{I, I\in\mathfrak{F}\cap\mathfrak{C}_1} \lambda_I^2$. Now

$$\sum_{I, I\in\mathfrak{F}\cap\mathfrak{C}_1} \lambda_I^2 = \sum_{I, I\in\mathfrak{F}\cap\mathfrak{C}_1} \lambda_I^2 \cdot \frac{\omega(I)}{\omega(I)} = \int_{I_0} (\sum_{\substack{I, (x',t)\in I \\ I\in\mathfrak{F}\cap\mathfrak{C}_1}} \frac{\lambda_I^2}{\omega(I)}) d\omega(x', t).$$

We claim that

$$\int_{I_0} (\sum_{\substack{I, (x',t)\in I \\ I\in\mathfrak{F}\cap\mathfrak{C}_1}} \frac{\lambda_I^2}{\omega(I)}) d\omega(x', t) \le CA^2\gamma^2\omega(I_0).$$

To prove the claim, notice that for $(x', t) \in P_j$,

$$\sum_{I, (x',t)\in I\in\mathfrak{F}\cap\mathfrak{C}_1} \frac{\lambda_I^2}{\omega(I)} \le \sum_{I, (x',t)\in I\in\mathfrak{F}\cap\mathfrak{C}_1} \lambda_I^2 (\sum_{k=0}^{\infty} \frac{2^{-k(2\beta-\tau)}}{\omega(2^k I)} \chi_{R_k(I)}(x', t))$$
$$= \sum_{I, I\in\mathfrak{F}\cap\mathfrak{C}_1} \lambda_I^2 (\sum_{k=0}^{\infty} \frac{2^{-k(2\beta-\tau)}}{\omega(2^k I)} \chi_{R_k(I)}(x'_{P_j}, t_{P_j})) = G(P_j)^2 \le CA^2\gamma^2.$$

If $(x', t) \notin \cup P_j$, then $\sum\limits_{I, (x',t)\in I\in\mathfrak{F}\cap\mathfrak{C}_1} \frac{\lambda_I^2}{\omega(I)} \le G^*(x', t)^2 \le (A\gamma)^2$. So the claim follows.

We have shown that $\sum\limits_j \omega(\{(x', t) \in c(P_j) : F_1(P_j; x', t) > \frac{1}{4}\}) \le \frac{\beta}{2}\omega(I_0)$, if γ is small enough.

To estimate $\sum\limits_j \omega(\{(x', t) \in c(P_j) : F_2(P_j; x', t) > \frac{1}{4}\})$, we first need to cut out the part of I_0 that gives bad edge effects. Using the doubling property of ω, for every $\frac{\beta}{4} > 0$, there is a $\kappa > 1$ so that for all cubes $Q \subset I_0$, $\omega(\kappa Q \backslash Q) \le \frac{\beta}{4}\omega(Q)$. Applying this to the $P_j's$ means that $\sum\limits_j \omega(\kappa P_j \backslash P_j) \le \frac{\beta}{4}\omega(I_0)$. This leaves the quantity $\sum\limits_j \omega(\{(x', t) \in c(P_j)\backslash D : F_2(P_j; x', t) > \frac{1}{4}\})$, where $D = \cup(\kappa P_l \backslash P_l)$, to be estimated. Or, as shown above, we want to bound $\sum\limits_j 16 \int_{c(P_j)\backslash D} |F_2(P_j; x', t)|^2 d\omega(x', t)$. We claim that for $(x', t) \in c(P_j)\backslash D$, $|F_2(P_j; x', t)| \le CG(P_j) \cdot H_j(x', t)$, where

$$(H_j(x', t))^2 \le C \sum_l \omega(P_l) \sum_{k=0}^{\infty} \frac{2^{-j\tau}}{\omega(2^k P_l)} \chi_{R_k(P_l)}(x', t).$$

Given the claim,

$$\sum_j \int_{c(P_j)\setminus D} \left|F_2(P_j;x',t)\right|^2 d\omega(x',t)$$

$$\leq C\gamma^2 \sum_l \omega(P_l) \int \sum_{k=0}^{\infty} \frac{2^{-k\tau}}{\omega(2^k P_l)} \chi_{R_k(P_l)}(x',t) d\omega(x',t)$$

$$\leq C\gamma^2 \sum_l \omega(P_l) = C\gamma^2 \omega(I_0).$$

For γ sufficiently small, this means that

$$\sum_j \omega(\{(x',t) \in c(P_j)\setminus D : F_2(P_j;x',t) > 0.25\}) \leq 0.25\omega(I_0)\,.$$

Combining this estimate with the estimate for $\omega(D)$, shows that

$$\sum_j \omega(\{(x',t) \in c(P_j) : F_2(P_j;x',t) > 0.25\}$$

$$\leq 0.25\beta\omega(I_0) + 0.25\beta\omega(I_0) = 0.5\beta\omega(I_0).$$

To prove the claim we write

$$\left|F_2(P_j;x',t)\right| = \left| \sum_{\substack{I,I\in\mathfrak{C}_2 \\ I\subset P_l, l\neq j}} \lambda_I \phi_I(x',t) \right| = \left| \sum_{l,l\neq j} \sum_{\substack{I,I\in\mathfrak{C}_2 \\ I\subset P_l}} \lambda_I \phi_I(x',t) \right|$$

$$\leq \sum_{l,l\neq j} \sum_{\substack{I,I\in\mathfrak{C}_2 \\ I\subset P_l}} |\lambda_I| \sqrt{\omega(I)} \sum_{k=0}^{\infty} \frac{2^{-j\beta}}{\omega(2^k I)} \chi_{R_k(I)}(x',t)$$

$$\leq \left(\sum_{l,l\neq j} \sum_{\substack{I,I\in\mathfrak{C}_2 \\ I\subset P_l}} |\lambda_I|^2 \left(\sum_{k=0}^{\infty} \frac{2^{-k(2\beta-\tau)}}{\omega(2^k I)} \chi_{R_k(I)}(x',t) \right) \right)^{1/2}$$

$$\times \left(\sum_{l,l\neq j} \sum_{\substack{I,I\in\mathfrak{C}_2 \\ I\subset P_l}} \omega(I) \sum_{k=0}^{\infty} \frac{2^{-k\tau}}{\omega(2^k I)} \chi_{R_k(I)}(x',t) \right)^{1/2} = G_2(P_j;x',t) \cdot H_j(x',t)\,.$$

Now $(x',t) \in c(P_j)$ implies that $G_2(P_j;x',t) \leq CG_2(P_j) \leq CA\gamma$, by Lemma 3 and the estimate for $G(P_j)$. To estimate $H_j(x',t)$, fix $l \neq j$ and look at $\sum_{I, I \subset P_l} \omega(I) \sum_{k=0}^{\infty} \frac{2^{-k\tau}}{\omega(2^k I)} \chi_{R_k(I)}(x',t)$ for $(x',t) \in c(P_j) \backslash D$. This sum can be written as $\sum_{n=0}^{\infty} \sum_{I: l(I) \simeq 2^{-n} l(P_l)} \omega(I) \sum_{k=0}^{\infty} \frac{2^{-k\tau}}{\omega(2^k I)} \chi_{R_k(I)}(x',t)$. For n fixed, we can show that

$$\sum_{I: l(I) \simeq 2^{-n} l(P_l)} \omega(I) \sum_{k=0}^{\infty} \frac{2^{-k\tau}}{\omega(2^k I)} \chi_{R_k(I)}(x',t)$$
$$\leq C 2^{-n\epsilon} \omega(P_l) \sum_{i=0}^{\infty} \frac{2^{-i\tau}}{\omega(2^i P_l)} \chi_{R_i(P_l)}(x',t).$$

To see this last inequality, notice that for $(x',t) \in c(P_j) \backslash D$ and $I \subset P_l$, $l \neq j$, then $d_p(x',t;x'_I,t_I) \simeq d_p(x'_{P_j},t_{P_j};x'_{P_l},t_{P_l})$. This means that $(x',t) \in R_k(I)$ if and only if $(x',t) \in R_i(P_l)$ for some index i, where $|i+n-k| \leq C'$ for a fixed constant C'. The doubling property of ω now gives that $\omega(2^k I) \simeq \omega(2^i P_l)$ and so when $(x',t) \in c(P_j) \backslash D$, we have

$$\frac{2^{-k\tau}}{\omega(2^k I)} \chi_{R_k(I)}(x',t) \leq C 2^{-n\tau} \sum_{i, |i+n-k| \leq C'} \frac{2^{-i\tau}}{\omega(2^i P_l)} \chi_{R_i(P_l)}(x',t).$$

This implies that

$$\sum_{\substack{I, I \subset P_l \\ l(I) \simeq 2^{-n} l(P_l)}} \omega(I) \sum_{k=0}^{\infty} \frac{2^{-k\tau}}{\omega(2^k I)} \chi_{R_k(I)}(x',t)$$
$$\leq C \sum_{k=0}^{\infty} \sum_{i, |i+n-k| \leq C'} \frac{2^{-n\tau} 2^{-i\tau}}{\omega(2^i P_l)} \chi_{R_i(P_l)}(x',t) \sum_{\substack{I, I \subset P_l \\ l(I) \simeq 2^{-n} l(P_l)}} \omega(I)$$
$$\leq C'' 2^{-n\tau} \omega(P_l) \sum_{i=0}^{\infty} \frac{2^{-i\tau}}{\omega(2^i P_l)} \chi_{R_i(P_l)}(x',t).$$

Summing over n and l gives the estimate

$$H_j(x',t)^2 \leq C \sum_l \omega(P_l) \sum_{i=0}^{\infty} \frac{2^{-i\tau}}{\omega(2^i P_l)} \chi_{R_i(P_l)}(x',t)$$

as needed to finish the proof of the claim. □

To prove Theorem 1 we need the following general version of the good lambda inequality.

Corollary. (The General Good-λ Inequality) *Let ν be an A^∞ weight with respect to ω. For any $\beta > 0$, there is a $\gamma > 0$, so that for any $\lambda > 0$ and $f(x',t) = \sum \lambda_I \phi_I(x',t)$ any finite sum with the ϕ_I satisfying* (8), (9) and (10), *one has*

$$\nu(\{(x',t) \in S\Omega : F^*(x',t) > 2\lambda, G^*(x',t) \le \gamma\lambda\}) \\ \le \beta\nu\left(\{(x',t) \in S\Omega : F^*(x',t) > \lambda\}\right).$$

Sketch of the proof of the Corollary to Parabolic Main Lemma. Following the argument in [13] and [15], we can reduce to the case of proving $\omega(\{(x'.t) \in I_j : F^*(x',t) > 2\lambda, G^*(x',t) \le \gamma\lambda\}) \le \beta\omega(I_j)$, where the I_j are maximal cubes such that $|F(I_j)| > \lambda$. Then using Lemma 4 and Lemma 7 to obtain that, $|F(I_j) - F(I_j,(x',t))| \le CG(I_j)$ when $(x',t) \in (1-\eta)I_j$, and that $\left|F(I_j) - F(\hat{I_j})\right| \le CG(I_j)$, when $I_j \in N(\ \hat{I_j})$; and combining these estimates with the fact that $\left|F(\hat{I_j})\right| \le \lambda$ by maximality of the I_j, we can guarantee that $|F(I_j,(x',t))| \le 1.2\lambda$ when $(x',t) \in (1-\eta)I_j$, by taking γ to be sufficiently small.

For η chosen so that $\omega(I_j \backslash (1-\eta)I_j) \le (\beta/3)\omega(I_j)$, we have that $\omega(\{(x'.t) \in I_j : F^*(x',t) > 2\lambda, G^*(x',t) \le \gamma\lambda\}) \le \omega((x',t) \in (1-\eta)I_j : H^*(x',t) \ge 0.8\lambda, G^*(x',t) \le \gamma\lambda) + (\beta/3)\omega(I_j)$, where $H^*(x',t) = \sup_{J \ni (x',t)} H(J) = \sup_{J \ni (x',t)} H(J;(x'_J,t$ $H(I;(x',t)) = \sum_{J \not\subseteq I, J \subseteq I_j} \lambda_J \phi_J(x',t)$, in other words $H(I;(x',t))$ has the same relation to $h(x',t) = \sum_{J \in \mathcal{F}, J \subseteq I_j} \lambda_J \phi_J(x',t)$ that $F(I;(x',t))$ has to $f(x',t) = \sum_{J \in \mathcal{F}} \lambda_J \phi_J(x',t)$. But estimating $\omega((x',t) \in (1-\eta)I_j : H^*(x',t) \ge 0.8\lambda, G^*(x',t) \le \gamma\lambda)$ is exactly what is done in Parabolic Main Lemma.

Given Main Lemma and the Corollary, Theorem 1 can be proved assuming that $F^* \in L^p(\partial_p\Omega_T, d\omega)$ whenever $g^* \in L^p(\partial_p\Omega_T, d\omega)$. The Corollary with Lemma 1 and Lemma 2 now give the result. □

References

[1] R. Brown, Area integral estimates for caloric functions, *Transactions of the AMS*, **315**, No. 2 (1989), 565-589.

[2] E. Fabes, N. Garofalo, S. Salsa, A backward Harnack inequality and Fatou Theorems for non-negative solutions of parabolic equations, *Illinois J. of Math.*, **30** (1986), 536-565.

[3] E. Fabes, M. Safanov, Behavior near the boundary of positive solutions of second order parabolic equations, *Journal of Fourier Analysis and Applications*, **3** (1997), 871-882.

[4] E. Fabes, S. Salsa, Estimates of caloric measure and the initial Dirichlet problem for the heat equation in Lipschitz cylinders, *Transactions of the AMS*, **179**, No. 2 (1983), 625-650.

[5] M. Giaquinta, M. Struwe, On the partial regularity of weak solutions of non-linear parabolic systems, *Math. Zeitschrift*, **179** (1982), 437-451.

[6] J. Hattemer, Boundary behavior of temperatures I, *Studia Math.*, **25** (1964), 111-155.

[7] D. Jerison, C. Kenig, Boundary behavior of harmonic functions in nontangentially accessible domains, *Advances in Math.*, **146** (1982), 80-147.

[8] C. Kenig, *Harmonic Analysis Techniques for Second Order Elliptic Boundary Value Problems*, CBMS series, **83**, AMS (1994).

[9] J. Moser, A Harnack inequality for parabolic differential equations, *Comm. in Pure and Applied Math.*, **17** (1964), 101-134.

[10] K. Nystrom, The Dirichlet problem for second order parabolic operators, *Indiana U. Math. J.*, **46**, No. 1 (1997), 183-245.

[11] C. Sweezy, B^q for parabolic measures, *Studia Math.*, **131**, No. 2 (1998), 115-135.

[12] C. Sweezy, M. Wilson, Weighted inequalities for caloric functions on classical domains, *WSEAS Transactions on Math.*, **3**, No. 3 (2004), 578-583.

[13] C. Sweezy, M. Wilson, Weighted inequalities for gradients on nonsmooth domains, *Preprint* (2004).

[14] R. Wheeden, M. Wilson, Weighted norm estimates for gradients of half-space extensions, *Indiana U. Math. J.*, **44** (1995), 917-969.

[15] M. Wilson, Global orthogonality implies local almost-orthogonality, *Revista Math. Ib.*, **16**, No. 1 (2000), 29-48.

Appendix

To simplify notation we work on $\mathbb{R}^d = \mathbb{R}^{d-1}\times \mathbb{R}^1$. Suppose that ν is a doubling measure on $\mathbb{R}^d = \mathbb{R}^{d-1}\times \mathbb{R}^1$and h is a non-negative function that is locally integrable with respect to ν. We define, for $(x',t) \in \mathbb{R}^d$,

$$M_\nu(h)(x',t) \doteq \sup_{Q,(x',t)\in Q} \frac{1}{\nu(Q)} \int_Q h(y',s) d\nu(y',s) ,$$

where Q is a non-isotropic cube in $\mathbb{R}^d$.

Theorem 1A. *Let* $0 < \beta < 1$. *Then either* $M_\nu(h) = \infty$, ν*-a.e., or* $M_\nu(h)d\nu \in A_\infty(d\nu)$.

Proof. First we claim that, if $M_\nu(h) = \infty$ on a set of positive ν measure, then $M_\nu(h) = \infty$ ν-a.e. To prove the claim we use the dyadic maximal function, $M_{\nu,d}(h)(x',t) \doteq \sup_{Q,(x',t)\in Q} \frac{1}{\nu(Q)} \int_Q h(y',s)d\nu(y',s)$, where $Q \in \mathcal{D}$, i.e. Q is a dyadic, non-isotropic cube. Every non-isotropic cube in $\mathbb{R}^d$ can be covered by $3^{d-1} \times 5$ (or 3^{d+1}) dyadic non-isotropic cubes Q', each of whose side lengths, $l(Q')$, is $\le l(Q)$. One can use this fact and the center-doubling property of ν, to show that for any $\lambda > 0$, there are positive constants c_1 and c_2 such that

$$\nu\left(\{(x',t) : M_\nu(h)(x',t) > \lambda\}\right) \le c_1 \nu\left(\{(x',t) : M_{\nu,d}(h)(x',t) > c_2\lambda\}\right) .$$

We also have that $M_{\nu,d}(h) \le M_\nu(h)$ pointwise.

Suppose $M_\nu(h) = \infty$ on a set of positive ν measure. Then $M_{\nu,d}(h) = \infty$ on a set of positive ν measure, also; and there is a dyadic cube Q_0, containing a subset E such that $\nu(E) > 0$ and $M_{\nu,d}(h) = \infty$ on E. Let $\mathcal{F} \doteq \{Q \in \mathcal{D}$: $Q \cap Q_0 \ne \emptyset\}$. Let R be a large, positive number. Then either there is a maximal cube in $\mathcal{F}$ so that

$$\frac{1}{\nu(Q)} \int_Q h d\nu > R , \tag{A.1}$$

or there is no such maximal cube (the fact that the set E exists guarantees there is at least one Q exists so that $(A.1)$ holds). Consequently if there is no such maximal cube, then $(A.1)$ must hold for a nested sequence of dyadic cubes with infinitely many of them containing Q_0. These cubes might not cover $\mathbb{R}^d$, but their concentric triples will. Also ν is a center-doubling measure, so if $\widetilde{Q}$ $= 3Q$ is the threefold dilate of Q that is concentric with Q, then there is some fixed $c > 0$ such that

$$\frac{1}{\nu(\widetilde{Q})} \int_{\widetilde{Q}} h d\nu > cR , \tag{A.2}$$

and $M_\nu(h) > cR$ on all of $\mathbb{R}^d$. So, if, for every $R > 0$, no maximal cube exists of the type described above, then $M_\nu(h) = \infty$ everywhere and we are done.

We are left to deal with the case when, for some $R > 0$, $(A.1)$ holds for some maximal $Q \in \mathcal{F}$. Notice this implies that $(A.1)$ holds for all $R >$ some fixed R'. We will denote the maximal Q by $Q(R)$, when $R > R'$ and $(A.1)$ holds for Q.

$$\text{If } R' < R_1 < R_2, \text{ then } Q(R_2) \subset Q(R_1)\,. \tag{$\bigstar$}$$

Now either $Q_0 \subset Q(R)$, or $Q(R) \subset Q_0$. If the former occurs for all $R > R'$, then, because of $(\bigstar)$, there must be a cube $Q \in \mathcal{F}$ so that $Q_0 \subset Q$, and $\frac{1}{\nu(Q)}\int_Q h d\nu = \infty$; however, this contradicts the hypothesis that h is locally integrable.

Therefore we may assume that for some large $R > 0$ (and so for all $R > R''$ for some fixed $R'' > 0$), every $Q \in \mathcal{F}$ that satisfies $(A.1)$ lies inside Q_0. Let $R > R''$ be very large and suppose that $\{Q_k\}_{k\geq 1}$ are the maximal dyadic subcubes of Q_0 so that $\frac{1}{\nu(Q_k)}\int_{Q_k} h d\nu > R$. Notice that $E \subset \cup_{k\geq 1} Q_k$; this fact gives that $0 < \nu(E) < \sum\limits_{k\geq 1} \nu(Q_k)$. Then for all $R > R''$, since the Q_k are disjoint, $\int_{Q_0} h d\nu \geq \sum\limits_{k\geq 1} \int_{Q_k} h d\nu > \sum\limits_{k\geq 1} R\nu(Q_k) \geq R\nu(E)$. This implies that $\int_{Q_0} h d\nu = \infty$, which is a contradiction as above.

In conclusion we can say that if $M_\nu h$ is not identically infinite, then $M_\nu h$ is finite ν-a.e. Henceforth we will assume that $M_\nu h < \infty$ ν-a.e.

Let Q be a non-isotropic cube, and let $\widetilde{Q}$ be Q's concentric triple. Following standard procedure we write $h = h\chi_{\widetilde{Q}} + h\chi_{\mathbb{R}^d\setminus\widetilde{Q}} = h_1 + h_2$. Since ν is a center-doubling measure, we can use standard arguments to show that

$$M_\nu(h_2)(x', t) \leq C_1 M_\nu(h_2)(y', s) \leq C_2 M_\nu(h_2)(x', t)\,,$$

where (x', t) and (y', s) are any two points in Q, with C_1 and C_2 being fixed constants that depend only on d and ν's doubling constant. Therefore if $E \subset Q$ and $\nu(E) \leq \delta\nu(Q)$,

$$\int_E (M_\nu(h_2))^\beta d\nu \leq C\delta \int_Q (M_\nu(h_2))^\beta d\nu \leq C\delta \int_Q (M_\nu(h))^\beta d\nu.$$

To handle h_1, we recall that the parabolic maximal function M_ν is weak $L^1(d\nu)$. Consequently, for any $0 < r < 1$, we have

$$\frac{1}{\nu(Q)}\int_Q (M_\nu(h_1))^r d\nu \leq C\left(\frac{1}{\nu(\widetilde{Q})}\int_{\widetilde{Q}} h_1 d\nu\right)^r.$$

Given $0 < \beta < 1$, choose $p > 1$ such that $r \mathrel{\mathring{=}} p\beta < 1$, and let p' denote p's dual index. For $E \subset Q$

$$\frac{1}{\nu(Q)} \int_E (M_\nu(h_1))^\beta d\nu \leq \left(\frac{1}{\nu(Q)} \int_Q (M_\nu(h_1))^{\beta p} d\nu \right)^{1/p} \cdot (\nu(E)/\nu(Q))^{1/p'}.$$

The second factor on the right is small if $\nu(E)/\nu(Q)$ is small – and this is what we want! The first factor is bounded by a constant times

$$\left(\frac{1}{\nu(\widetilde{Q})} \int_{\widetilde{Q}} h_1 d\nu \right)^\beta \leq \frac{1}{\nu(Q)} \int_Q (M_\nu(h))^\beta d\nu.$$

This gives the following estimate

$$\int_E (M_\nu(h_1))^\beta d\nu \leq C \left(\int_Q (M_\nu(h))^\beta d\nu \right) \cdot (\nu(E)/\nu(Q))^{1/p'}.$$

Combining this estimate with the estimate for h_2 yields the result.

Corollary 2A. *If $1 < r < s < \infty$, ν is a center-doubling measure, and h, a non-negative function, belongs to $L^s(\nu)$, and if we set*

$$M_{r,\nu}(h)(x', t) \mathrel{\mathring{=}} \left(\sup_{Q:(x',t)\in Q} \frac{1}{\nu(Q)} \int_Q h^r d\nu \right)^{1/r},$$

then $M_{r,\nu}(h) d\nu \in A_\infty(d\nu)$, with A_∞ constants depending only on r, s, d, and ν's doubling constant.

Proof. Apply Theorem 1A, with h replaced by h^r and β replaced by $1/r$. □