The L^p Neumann Problem for the Heat Equation in Non-Cylindrical Domains

Steve Hofmann
Department of Mathematics
University of Missouri
Columbia, MO 65211
hofmann@math.missouri.edu

John L. Lewis
Department of Mathematics
University of Kentucky
Lexington, KY 40506-0027
john@ms.uky.edu

November 2, 2004

1 Introduction

In this note, we study the Neumann problem for the heat equation in non-cylindrical domains. The class of domains which we consider are those defined by the region above a time varying graph

$$\Omega \equiv \{(x_0, x, t) \in \mathbb{R} \times \mathbb{R}^{n-1} \times \mathbb{R} : x_0 > A(x,t)\}. \tag{1.1}$$

Here, A is Lipschitz in x, uniformly in t, i.e.

$$\sup_{(x,t)} |A(x,t) - A(x+h,t)| \leq \beta_0 |h|, \tag{1.2}$$

For some $\beta_0 < \infty$; furthermore A posesses a certain half-order smoothness in t, which we shall now describe.

Let $\|(x,t)\|$ denote the parabolic "norm" in $\mathbb{R}^n$; i.e., for $(x,t) \neq 0$, $\|(x,t)\|$ is defined to be the unique positive solution ρ of the equation

$$\frac{|x|^2}{\rho^2} + \frac{t^2}{\rho^4} = 1. \tag{1.3}$$

It is well known, and easy to verify, that

$$\|(x,t)\| \approx |x| + |t|^{1/2}$$

and that

$$\|(rx, r^2 t)\| = r\|(x,t)\|, \quad r > 0.$$

Furthermore, equipped with the metric induced by $\|\cdot\|$, parabolic $\mathbb{R}^n$ is a homogeneous space with homogeneous dimension $d = n+1$. Following Fabes and Riviere [12], we define a half-order time derivative, denoted $\mathbb{D}_n$, by

$$(\mathbb{D}_n a)^\wedge(\xi,\tau) \equiv \frac{i\tau}{\|\xi,\tau)\|}, \tag{1.4}$$

where ξ and τ denote the Fourier transform variables in space and time respectively. We shall assume that A satisfies the following half-order smoothness condition in t: $\mathbb{D}_n A \in$ (parablic) BMO, i.e. there exists $\beta_1 < \infty$ such that

$$\|\mathbb{D}_n A\|_* \le \beta_1. \tag{1.5}$$

(We shall discuss below the "naturalness" of this condition). Here, the parabolic BMO norm is defined, as usual, by

$$\|b\|_* \equiv \sup \frac{1}{|B|}\int_B |b(z) - b_B| dz,$$

where $z \equiv (x,t)$, and where the sup runs over all parabolic balls

$$B \equiv \{z \in \mathbb{R}^n : \|z - z_0\| < r\}.$$

Having defined the class of domains which we shall consider, we are now in a position to define a certain "surface measure" on the time-varying boundaries of such domains. For each fixed t, let Ω_t denote the cross-section

$$\Omega_t \equiv \{(x_0,x) \in \mathbb{R}\times\mathbb{R}^{n-1} : x_0 > A(x,t)\}.$$

By (1.2), Ω_t is a Lipschitz domain with Lipschitz constant no larger than β_0. We define $d\sigma_t$ to be the usual surface measure on the Lipschitz graph $\partial\Omega_t$; i.e. in graph co-ordinates,

$$d\sigma_t \equiv \sqrt{1 + |\nabla_x A(x,t)|^2} dx.$$

We then define "surface measure" as

$$d\sigma \equiv d\sigma_t dt.$$

In this paper we consider the Neumann problem

$$N_p \begin{cases} \Delta u - \frac{\partial u}{\partial t} = 0 & \text{in } \Omega \\ \frac{\partial u}{\partial N}\big|_{\partial\Omega} = f \in L^p(\partial\Omega, d\sigma) \\ N_*(\nabla u) \in L^p(\partial\Omega, d\sigma). \end{cases} \tag{1.6}$$

Here N_* denotes the parabolic non-tangential maximal operator

$$N_*(F)(A(x,t),x,t) \equiv \sup_\Gamma |F(y_0,y,s)|,$$

where $\Gamma \equiv \Gamma(A(x,t),x,t))$ is the parabolic cone

$$\Gamma \equiv \{(y_0,y,s) \in \mathbb{R}\times\mathbb{R}^{n-1}\times\mathbb{R} : \|(x-y,t-s)\| \le \mu(y_0 - A(x,t))\}.$$

For μ small enough, depending only on β_0 and β_1 in (1.2) and (1.5) respectively, we have that $\Gamma(A(x,t),x,t) \subseteq \Omega$, at every point $(A(x,t),x,t) \in \partial\Omega$. Indeed, by a result of Strichartz [29] on BMO Sobolev spaces (or rather a routine extension of that result to the parabolic case, see [16]), we see that (1.2) and (1.5) imply the $\mathrm{Lip}_{1,1/2}$ condition

$$\sup_{x,t} |A(x,t) - A(x+h, t+h^2)| \leq C(\beta_0 + \beta_1)|h|, \tag{1.7}$$

From which the non-tangential accessibility follows readily.

The main results of this paper are as follows:

Theorem 1.8. *Let $\beta_0 < \infty$, and $\beta_1 \leq \epsilon_0$. Suppose $\rho \in (1, 2+\delta)$, where $\delta = \delta(\beta_0)$ is a small positive number. Then the Neumann problem N_p is uniquely solvable, as long as ϵ_0 is small enough depending only on n, β_0 and p.*

Remark. The case $p = 2$ of this theorem has already been obtained in previous work of the authors [18], and furthermore it is shown there that the restriction on the size of β_1 cannot be avoided. The case $2 - \delta < \rho < 2 + \delta$ follows immediately from the result of [18], plus general functional analytic arguments. The point of the present paper, then, is to fill in the interval $1 < p < 2$.

Our next theorem states that Theorem 1.8 is "sharp".

Theorem 1.9. *The results above are "best possible", in the sense that, given $\epsilon > 0$, $\exists\ p_0 > 1$ and a domain $\Omega = \Omega(p_0)$ as above with $\beta_0 = 0$ and $\beta_1 \leq \epsilon$, and such that for all $p \in (1, p_0]$, N_p fails. In particular, Theorem 1.8 cannot hold in general, unless $\epsilon_0 \geq \beta_1$ is small enough, with $\epsilon_0(p) \to 0$ as $p \to 1$.*

Remark. We note that Theorem 1.9 implies that one cannot obtain, in general, the H^1 estimate $\|N_*(\nabla u)\|_{L^1} \leq C \left\|\frac{\partial u}{\partial N}\right\|_{H^1}$ for any given $\epsilon_0 > 0$. Indeed, if one could prove such an estimate, then interpolating with the L^2 results of [18], one would contradict Theorem 1.9.

To put these results into context, let us review a bit of recent history. The class of domains which we consider here was introduced (in slightly disguised form) by the second author and M. Murray [26], in their study of the L^p Dirichlet problem. This class of domains exhibits nearly minimal smoothness for the study of boundary value problems with L^p data. Indeed, an example of Kaufman and Wu [21], shows that if condition (1.5) is replaced by the slightly weaker condition that $A(x,t)$ is $\mathrm{Lip}_{1/2}$ in t, then caloric measure and "surface measure" $d\sigma$ may be mutually singular. Furthermore, in work of the first author [16], it is shown that the parabolic Calderón commutator $\left[\sqrt{\Delta - \frac{\partial}{\partial t}}, A\right]$ is bounded on L^2 if and only if A satisfies (1.2) and (1.5) with

$$\left\| \left[\sqrt{\Delta - \frac{\partial}{\partial t}}, A \right] \right\|_{L^2 \to L^2} \approx \beta_0 + \beta_1.$$

(The one dimensional analogue of this fact had previously been obtained by Murray [27]). Thus, our domains are the parabolic analogues of (elliptic) Lipschitz domains.

Prior to the work described in [26] and [18], only the case of cylindrical domains (i.e. $\beta_1 \equiv 0$) had been treated, by Fabes and Salsa [13] in the case of the L^p Dirichlet problem ($2-\delta < p$) and by R. Brown [2],[3] in the case of $N_p(1 < p < 2+\delta)$. We note that the range of p for which N_p holds in cylindrical domains is dual to that for which the Dirichlet problem is solvable (as had also been the case for harmonic functions in Lipschitz domains, see the papers of Dahlberg [7], Verchota [30] and Dahlberg and Kenig [10]). It is somewhat remarkable then, that in light of Theorem 1.9, the analogous duality statement fails utterly in non-cylindrical domains. Indeed, in [26] it is shown that the L^p Dirchlet problem holds, without restriction on β_0 and β_1, if p is sufficiently large. The corresponding dual statement for the Neumann problem would be that N_p should hold without restriction on β_0 and β_1, if p is close enough to 1, but in fact Theorem 1.9 says that, in general, we must have $\beta_1 \leq \epsilon_0(p) \to 0$ as $p \to 1$. It is interesting to note that the corresponding dual statement for the regularity problem *does* remain true in the non-cylindical case, as we showed in [19]. Thus, the optimal results for the Neumann and regularity problems are no longer parallel, in contrast to the situation in [10] and [3].

The paper is organized as follows. In the next section, we construct our counter-examples which establish Theorem 1.9. In Section 3, we make some preliminary reductions, and discuss the strategy of the Proof of Theorem 1.8. In Section 4, we give some basic estimates for solutions of certain divergence form parabolic equations. Sections 5-9 will contain various steps of the Proof of Theorem 1.8. Section 10 is an appendix containing the proofs of certain technical estimates.

Acknowledgements

We are grateful to Carlos Kenig for helpful and encouraging conversations regarding this (and related) work, and in particular for explaining to us his joint work with J. Pipher on perturbation theory for the elliptic Neumann problem.

2 Proof of Theorem 1.9

We shall prove the theorem in 2 dimensions; i.e. for Ω of the form $\Omega = \{(x_0, t) \in \mathbb{R}\times\mathbb{R} : x_0 > A(t)\}$. One can easily extend this result to higher dimensions by considering $\tilde{\Omega} \equiv \Omega \times \mathbb{R}^{n-1}$, where Ω is a two dimensional domain, i.e. $\tilde{\Omega} = \{(x_0, x, t) \in \mathbb{R} \times \mathbb{R}^{n-1} \times \mathbb{R} : x_0 > A(t)\}$. Furthermore, in our counter-example, $A(t)$ will have compact support, and we shall construct A so that $A(t) \leq 0$, for all t, (hence $\{(x_0, t) : x_0 > 0\} \subseteq \Omega$, and $\partial\Omega$ coincides with $\{(x_0, t) : x_0 = 0\}$ except on a compact set).

Given a $p \in (1, \infty)$, suppose that N_p holds for a domain of the type described in the previous paragraph. Since $\frac{\partial w}{\partial N} \equiv \frac{\partial w}{\partial x_0}$ in such a domain, we would have that given $f \in L^p(\partial\Omega)$, there exists a unique caloric w in Ω such that $\left.\frac{\partial w}{\partial x_0}\right|_{\partial\Omega} = f$ almost everywhere, and such that $N_*(\nabla w) \in L^p$. Furthermore,

$$\|N_*(\nabla w)\|_{L^p(\partial\Omega)} \leq C \left\| \frac{\partial w}{\partial x_0} \right\|_{L^p(\partial\Omega)}. \tag{2.1}$$

We claim that estimate (2.1) implies solvability of the Dirichlet problem

$$D_{\tilde{p}} \begin{cases} \Delta v - \frac{\partial v}{\partial t} = 0 & \text{in } \Omega \\ v|_{\partial\Omega} = g \in L^{\tilde{p}} \\ N_* v \in L^{\tilde{p}} \end{cases},$$

for $\tilde{p} = p$. Indeed, it is enough to establish the a priori estimate

$$\|N_* v\|_{L^p(\partial\Omega)} \leq C\|g\|_{L^p(\partial\Omega)}, \tag{2.2}$$

where v solves the classical Dirichlet problem $\Delta v - \frac{\partial v}{\partial t} = 0$ in Ω, $v|_{\partial\Omega} = g \in C_c(\partial\Omega)$ (here C_c denotes the space of continuous functions with compact support). But given $g \in C_c(\partial\Omega)$, there exists such a v which is unique among the class of v such that $\sup_t |v(x_o, t)| \to 0$ as $x_0 \to \infty$. Let $v^* = v|_{\Omega^*}$, where $\Omega^* = \{(x_0, t) : x_0 > 0)$. We have already noted that $\Omega^* \subseteq \Omega$, and that $\partial\Omega^*$ coincides with $\partial\Omega$ except on a compact set. Hence v^* is the unique solution of

$$D^* \begin{cases} \Delta v^* - \frac{\partial v^*}{\partial t} = 0 & \text{in } \Omega^* \\ v^*|_{\partial\Omega^*} = v|_{\partial\Omega^*} \in C_c(\partial\Omega^*) \\ \sup_t |v^*(x_0, t)| \to 0 & \text{as } x_0 \to \infty \end{cases}.$$

But by standard interior estimates, for $x_0 > 0$ we have for all $r > 1$,

$$\begin{aligned} |v(x_0, t)| \equiv |v^*(x_0, t)| &\leq c \left(x_0^{-3} \iint_{\|(x_0 - y_0, t-s)\| < \frac{x_0}{2}} |v^*(y_0, s)|^r dy ds \right)^{\frac{1}{r}} \\ &\leq C x_0^{-2/r} \sup_{y_0} \left(\int |v^*(y_0, s)|^r ds \right)^{\frac{1}{r}} \\ &\leq C x_0^{-2/r} \left(\int |v(0, s)|^r ds \right)^{\frac{1}{r}}, \end{aligned}$$

where in the last estimate we have used that the Dirichlet problem D_r is solvable in the half-plane Ω^* for all $r > 1$.

Hence, taking $r < 2$ in the previous estimate, we have that

$$w(x_0, t) \equiv - \int_{x_0}^{\infty} v(\alpha, t) d\alpha,$$

is well defined, and $\Delta w - \frac{\partial w}{\partial t} = 0$ in Ω. Furthermore, $\frac{\partial w}{\partial x_0} \equiv v$, $\left.\frac{\partial w}{\partial x_0}\right|_{\partial\Omega} = g$ and (2.2) follows immediately from (2.1). Thus, to produce a counter-example to N_p, it suffices to produce a counter-example to D_p; i.e. we shall establish the failure of (2.2) for domains of the sort described at the beginning of this section.

Let $\psi \in C_0^\infty(-2, 2)$, with $\psi \equiv 1$ on $[-1, 1]$. Define, for given $\delta \in (0, 1)$,

$$A(t) \equiv -\delta\psi(t)|t|^{1/2},$$

and let

$$\Omega \equiv \{(x,t) \in \mathbb{R} \times \mathbb{R} : x > A(t)\}$$

(for convenience, in this section only, we have dropped the subscript zero). Let $G(x,t) \equiv G(2,4,x,t)$ denote the adjoint Green's function with pole at $(2,4)$. Let $D^{1/2}$ denote the 1-dimensional half order derivative defined by

$$D^{1/2} f(t) \equiv (|\tau|^{1/2} \hat{f}(\tau))^{\vee}(t).$$

Lemma 2.3. *For A and Ω as above, we have*

$$\text{(i)} \quad \|D^{1/2} A\|_{BMO(\mathbb{R}^1)} \leq C\delta$$

$$\text{(ii)} \quad \int_{-1}^{1} \left(\frac{\partial G}{\partial x}(A(t),t)\right)^q dt = +\infty,$$

for some $q \in (1,\infty)$, depending on δ.

By our previous observations, Theorem 1.9 is an immediate consequence of this lemma, for the lemma implies that D_p fails, whenever $p \leq q'$, $\frac{1}{q'} + \frac{1}{q} = 1$.
Proof of Lemma 2.3. To prove (i), we recall the following characterization of Strichartz [29]:

$$\|D^{1/2} A\|_*^2 \approx \sup_I \frac{1}{|I|} \int_I \int_I \frac{|A(s) - A(t)|^2}{|s-t|^2} ds dt \equiv \Lambda, \tag{2.4}$$

where $\approx$ means that the two quantities have a ratio which is bounded above and below by two positive constants. To establish that the right hand side of (2.4) is no larger than $c\delta^2$, we consider two cases.

Case 1. $d \equiv \operatorname{dist}(I,0) > 5|I|$.
In this case, $|s| \approx |t| \approx d$, for all $s,t \in I$. Thus

$$|A(t) - A(s)| \leq C\delta \left(d^{-\frac{1}{2}} |s-t| + d^{\frac{1}{2}} |s-t| \left\{\chi_{[-2,2]}(s) + \chi_{[-2,2]}(t)\right\}\right).$$

Furthermore $d \leq 10$, otherwise $A(t) = 0 = A(s)$. The desired bound follows easily in this case.

Case 2. $d \equiv \operatorname{dist}(I,0) \leq 5|I|$.
In this case, we may replace I by $\tilde{I}$, where $\tilde{I}$ is centered at 0, and $|\tilde{I}| = C|I|$. Thus we may assume $WLOG$ that I is centered at 0. We consider two subcases:

Case 2a. $|s| > 2|t-s|$, or $|t| > 2|t-s|$.
By symmetry, it suffices to assume $|t| > 2|t-s|$. Then $|t| \approx |s|$, so

$$|A(s) - A(t)| \leq C\delta \left(|t|^{-\frac{1}{4}} |s|^{-\frac{1}{4}} |s-t| + \min\left(|s|^{\frac{1}{2}}, |t|^{\frac{1}{2}}\right) |s-t| \left\{\chi_{[-2,2]}(s) + \chi_{[-2,2]}(t)\right\}\right).$$

But a routine computation then shows that

$$\frac{1}{|I|} \int_I \int_I \chi\{|t| > 2|t-s|\} \frac{|A(s) - A(t)|^2}{|t-s|^2} ds dt \leq C\delta^2.$$

Case2(b). $|t| \le 2|t-s|$ and $|s| \le 2|t-s|$.

By symmetry, the corresponding part of Λ is then no larger than

$$C\delta^2 \cdot 2 \cdot \frac{1}{|I|} \int_I \int_{I \cap \{|t-s| > |t|/2\}} \frac{|t|}{|t-s|^2} ds dt \le C\delta^2.$$

Part (i) of Lemma 2.3 is now established.

We now turn to part (ii). As above, let

$$\Omega^* \equiv \{(x,t) : x > 0\}$$

and let $\Omega_r^* = \Omega^* \cap \{(x,t) : |x| < r, |t| < r^2\}$. We first show that, for $r \le 1$,

$$\begin{aligned} &\text{if } \frac{G(x,t)}{x} \ge k > 0 \text{ on } \Omega_r^*, \quad \text{then} \\ &\frac{G(x,t)}{x} \ge (1+\beta)k \text{ on } \Omega_{r/4}^* \text{ for some } \beta > 0. \end{aligned} \tag{2.5}$$

(Here β will depend on δ). To prove (2.5), we begin by setting

$$h(x,t) = G(x,t) - kx,$$

which is an adjoint solution in Ω^*. Furthermore $h \ge 0$ in Ω_r^*. Now, on $E_r \equiv \partial\Omega^* \cap \left\{\frac{r^2}{4} < |t| < r^2\right\}$, we have that

$$C(\delta) h \ge \max_{\partial\Omega_r^*} G. \tag{2.6}$$

Indeed, on $\partial\Omega^*$, $h = G$. Furthermore, for every point $(0,t) \in E_r$, we have that the parabolic $\operatorname{dist}((0,t), \partial\Omega) \approx C(\delta) r$. Hence, by Harnack's inequality, plus the boundary Harnack principle, plus the backward Harnack inequality (see Fabes, Garofalo and Salsa [14] for the latter), we obtain (2.6).

Next, let $\omega_r = \omega_r^{(x,t)}$ denote adjoint caloric measure for Ω_r^* with respect to the point (x,t), and set $u(x,t) \equiv w_r^{(x,t)}(E_r)$. Then by (2.6) and the maximum principle,

$$h \ge \frac{1}{C(\delta)} (\max_{\partial\Omega_r^*} G) u$$

in Ω_r^*. By the comparison principle [14],

$$\frac{u(x,t)}{x} \approx \frac{u\left(\frac{r}{4}, \frac{r^2}{16}\right)}{\frac{r}{4}} \approx \frac{c}{r},$$

for $(x,t) \in \overline{\Omega_{\frac{r}{4}}^*}$. Thus, on $\partial\Omega_{\frac{r}{4}}^*$,

$$h(x,t) \equiv G(x,t) - kx \ge \frac{1}{C(\delta)} \left(\max_{\partial\Omega_r^*} G\right) \frac{x}{r},$$

i.e.

$$\frac{G(x,t)}{x} > k(1+\beta), \quad \text{with } \beta = \frac{1}{C(\delta)}$$

since $\frac{1}{r}\max_{\partial\Omega_r^*} G \geq k$. This proves (2.5).

Next, we iterate (2.5), starting with $r=1$, to obtain for some $\gamma \in \left(0, \frac{1}{2}\right)$, depending on δ, and some $c = c(\delta)$,

$$G(x,0) \geq cx^{1-\gamma}, \quad \text{for } 0 < x \leq 1. \tag{2.7}$$

Indeed, we have that $0 \leq x \leq 1$ on $\partial\Omega_1^*$, whereas there exists $k(\delta) > 0$ such that $G(x,t) \geq k(\delta) \geq k(\delta)x$ on those portions of $\partial\Omega_1^*$ where $x > 0$. Hence, by the maximum principle, (2.5) holds with $r = 1$. Iterating, we obtain

$$\frac{G(x,0)}{x} \geq (1+\beta)^N k(\delta),$$

whenever $4^{-N-1} < x \leq 4^{-N}$, from which (2.7) follows readily.

We now show that (2.7) implies conclusion (ii) of Lemma 2.3. We define

$$S_r \equiv \left\{ (A(t),t) : \frac{r^2}{4} \leq |t| < r^2 \right\} \subseteq \partial\Omega,$$

and we let $\omega^{(x,t)}$ denote adjoint caloric measure for Ω with respect to (x,t). Then (see [14]),

$$\begin{aligned}
\left(\frac{G(r,0)}{r}\right)^q &\leq c\left(\frac{\omega^{(2,4)}(S_r)}{r^2}\right)^q \\
&\leq cr^{-2}\int_{\frac{r^2}{4}\leq|t|<r^2} \left|\frac{\partial G}{\partial x}(A(t),t)\right|^q dt.
\end{aligned}$$

We now set $r = 2^{-j}$, and then sum in j. The result is

$$C\int_{-\frac{1}{4}}^{1/4} \left|\frac{\partial G}{\partial x}(A(t),t)\right|^q dt \geq \sum_{j=1}^{\infty} 2^{-2j}\left(\frac{G(2^{-j},0)}{2^{-j}}\right)^q,$$

which in turn equals $+\infty$, for $q = q(\delta)$ large enough, by (2.7). This concludes the Proof of Lemma 2.3.

3 Proof of Theorem 1.8: Preliminaries.

In the sequel, we shall make the *a priori* assumption that $A \in C_0^\infty$. Of course, our estimates will be independent of our *a priori* assumption. We begin by mapping our domain Ω to a half space. We set

$$\rho(\lambda,x,t) \equiv (\lambda + P_{\gamma\lambda}A(x,t), x, t), \quad \lambda > 0, \tag{3.1}$$

where γ is a small parameter to be chosen, and P_λ is a "nice" parabolic approximate identity, i.e.

$$P_\lambda A \equiv \varphi_\lambda * A,$$

where $\varphi_\lambda(x,t) \equiv \lambda^{-d}\varphi\left(\frac{x}{\lambda}, \frac{t}{\lambda^2}\right)$, $\varphi \in C_0^\infty(B_1(0))$, with $\int \varphi = 1$. (Here we recall that $d = n+1$ is the homogeneous dimension of parabolic $\mathbb{R}^n$). Then for γ small enough, depending only on $\beta_0 + \beta_1$,

$$\rho : \mathbb{R}^{n+1}_+ = \{(\lambda, x, t) : \lambda > 0\} \to \Omega,$$

bijectively, with $\rho(\partial\mathbb{R}^{n+1}_+) = \partial\Omega$. Such mappings had appeared first in a paper of Dahlberg [8] (but this explicit construction is due to Kenig and Stein) in connection with generalized paraproducts above a Lipschitz graph. The mapping (3.1) has recently been useful in work of the present authors [18] on the case $p = 2$ of Theorem 1.8, as well as in work of Dahlberg, Kenig, Pipher and Verchota [11] on square function estimates for constant coefficient elliptic equations and systems in Lipschitz domains.

Given a caloric function u in Ω, we set $u_1 \equiv u \circ \rho$. Then u_1 is a weak solution in $\mathbb{R}^{n+1}_+$ of the equation

$$L_1 u_1 \equiv \operatorname{div}(a\nabla u_1) - \left(1 + \frac{\partial}{\partial\lambda}P_{\gamma\lambda}A\right)\frac{\partial}{\partial t}u_1 + \frac{\partial}{\partial t}P_{\gamma\lambda}A \cdot \frac{\partial}{\partial\lambda}u_1 = 0, \tag{3.2}$$

where a is the $n \times n$ matrix with coefficients

$$\begin{aligned} a_{00} &= \frac{1 + |\nabla_x P_{\gamma\lambda}A|^2}{1 + \frac{\partial}{\partial\lambda}P_{\gamma\lambda}A} \\ a_{jj} &= 1 + \frac{\partial}{\partial\lambda}P_{\gamma\lambda}A, && 1 \le j \le n-1 \\ a_{0j} &= a_{j0} = -\frac{\partial}{\partial x_j}P_{\gamma\lambda}A, && 1 \le j \le n-1 \\ a_{ij} &= 0, && \text{otherwise.} \end{aligned} \tag{3.3}$$

We note that the matrix a satisfies the ellipticity condition $\nu|\xi|^2 \le \langle a\xi, \xi\rangle \le \frac{1}{\nu}|\xi|^2$, with ν depending only on β_0. We also introduce the auxiliary equation

$$L_0 u_0 \equiv \operatorname{div}(a\nabla u_0) - \frac{\partial}{\partial t}u_0 = 0 \tag{3.4}$$

The strategy of our proof is as follows. To establish existence in Theorem 1.8, our goal is to prove the estimate

$$\|\nabla_{\tan} u\|_{L^p(\partial\Omega)} + \|[\mathbb{D}_n(u \circ \rho)] \circ \rho^{-1}\|_{L^p(\partial\Omega)} \le C(p, n, \beta_0)\left\|\frac{\partial u}{\partial N}\right\|_{L^p(\partial\Omega)}, \tag{3.5}$$

as long as $\beta_1 \le \epsilon_0$ is small enough. Here, $\nabla_{\tan}$ denotes the tangential gradient in the space variables. Since the opposite inequality to (3.5) has been established in [19] under much less restrictive hypotheses on β_1, one may then use a strategy introduced by Verchota [30] in the case of harmonic functions in a Lipschitz domain, to establish existence in Theorem 1.8 via the method of layer potentials. One obtains that

$$u = S\left(\left(\frac{1}{2}I + K^*\right)^{-1} g\right), \tag{3.6}$$

for some density $g \in L^p(\partial\Omega)$, where

$$Sf(X,t) \equiv \int_{\partial\Omega} W(X-Q,t-s)f(Q,s)d\sigma_s(Q)ds, \qquad (X,t) \in \Omega,$$

$$K^*f(P,t) \equiv p.v. \int_{\partial\Omega} \frac{\partial}{\partial N_{(P,t)}} W(P-Q,t-s)f(Q,s)d\sigma_s(Q)ds, \qquad (P,t) \in \partial\Omega,$$

and

$$W(X,t) \equiv (4\pi t)^{-\frac{n}{2}} \exp\left\{-\frac{|X|^2}{4t}\right\} \chi_{\{t>0\}},$$

$X = (x_0, x) \in \mathbb{R}^n$. Here $\frac{\partial}{\partial N_{(p,t)}}$ denotes differentiation in the direction of the outer unit normal at the point $P \in \partial\Omega_t$. Give the parabolic singular integral estimates of [26] and [17], it is by now a fairly standard matter to use the method of [30] to deduce (3.6) from (3.5) and its converse, and we omit the details. However, the reader may consult [18] in the case $p = 2$. The case $1 < p < 2$ is the same.

Thus, it is enough to prove (3.5). As mentioned above, (see the remark following Theorem 1.9) the $H^1 \to L^1$ analogue of (3.5) fails. Thus, a direct adaptation of the strategy of Dahlberg and Kenig [10], as used by Brown [3] in the case of parabolic cylinders, is impossible. However, when one views matters in terms of the pullback equation (3.2), it turns out that it is the singular drift term $\frac{\partial}{\partial t} P_{\gamma\lambda} A \cdot \frac{\partial}{\partial \lambda} u_1$ which causes the problems. Indeed, we will prove, for β_1 small enough, that for u_0 the analog of [10] and [3], and also of Kenig and Pipher [23] does hold:

$$\|\mathbb{D}_n u_0\|_{L^1(\mathbb{R}^n)} + \|\nabla_{\tan} u_0\|_{L^1(\mathbb{R}^n)} \le C(\beta_0, n) \left\|\frac{\partial u_0}{\partial \nu}\right\|_{H^1(\mathbb{R}^n)}, \tag{3.7}$$

where u_0 solves the modified equation (3.4), and where $\frac{du_0}{\partial \nu}$ denotes the co-normal derivative on $\partial\mathbb{R}^{n+1}_+$

$$\frac{\partial u_0}{\partial \nu} \equiv \langle -e_0, a\nabla_{\lambda,x} u_0 \rangle.$$

From (3.7), plus the L^2 estimate

$$\|\mathbb{D}_n u_0\|_{L^2(\mathbb{R}^n)} + \|\nabla_{\tan} u_0\|_{L^2(\mathbb{R}^n)} \le C(\beta_0, n) \left\|\frac{\partial u_0}{\partial \nu}\right\|_{L^2(\mathbb{R}^n)}, \tag{3.8}$$

which we shall also establish, we then obtain immediately by interpolation the bound

$$\|\mathbb{D}_n u_0\|_{L^p} + \|\nabla_{\tan} u_0\|_{L^p} \le C(\beta_0, n, p) \left\|\frac{\partial u_0}{\partial \nu}\right\|_{L^p(\mathbb{R}^n)}. \tag{3.9}$$

Thus the first main step in our proof is to establish (3.8) and (3.7). The proof of (3.7) will be an adaptation to the parabolic case of arguments in [23] for variable coefficient elliptic operators. The second main step of our proof is to pass from (3.9) to (3.5), or equivalently to the analogue of (3.5) for u_1, namely

$$\|\nabla_{\tan} u_1\|_{L^p(\mathbb{R}^n)} + \|\mathbb{D}_n u_1\|_{L^p(\mathbb{R}^n)} \leq C(\rho, n, \beta_0) \left\| \frac{\partial u_1}{\partial \nu} \right\|_{L^p(\mathbb{R}^n)} \tag{3.10}$$

if $\beta_1 \leq \epsilon_0(\rho)$ is small enough. The proof of this last step will be based on perturbation techniques for the Neumann Problem developed in [24] for elliptic equatitons. The method of [24] dealt with perturbations of the leading coefficients, but the same method adapts well to the case when one perturbs by certain sorts of lower order terms, as is the situation here.

The remainder of the Proof of Theorem 1.8 will be given as follows. In the next section, we give some basic interior estimates for solutions to a class of divergence form parabolic equations, which, in particular, includes L_0 (see (3.4)). In Section 5, we establish L^2 estimates for the Dirichlet problem for $L_0 u_0 = 0$ (this is an auxiliary step in the proof). Sections 6-10 are the heart of the matter. In Section 6, we state our "Main Lemma", which we then use to establish the L^2 estimate (3.8). We also show that the "Main Lemma" plus (3.9) imply (3.10) (we note that, as we have already observed, existence in Theorem 1.8 follows from (3.10)). We construct a variational solution to the Neumann problem in Section 7. In Section 8, we prove (3.7) (and thus also (3.9) by interpolation with (3.8)). The "Main Lemma" is proved in Section 9, modulo certain square function estimates. Finally, Section 10 is an appendix, in which we prove the square function estimates.

We remark that uniqueness in Theorem 1.8 can be established by the argument given in [18] to treat the case $p = 2$. We omit the details. Hence, in the sequel, we shall concentrate on the question of existence.

4 Estimates for solutions of certain divergence form parabolic equations

In this section we state some standard estimates for certain classes of parabolic operators defined in $\mathbb{R}^{n+1}_+$. Let $(X, t) \equiv (\lambda, x, t) \in \mathbb{R}^{n+1}_+ \equiv \{(\lambda, x, t) : \lambda > 0, x \in \mathbb{R}^{n-1}, t \in \mathbb{R}\}$, and set

$$Lu \equiv \operatorname{div}(a(X,t)\nabla u) - b\frac{\partial u}{\partial t} + \vec{B} \cdot \nabla u, \tag{4.1}$$

where the coefficient matrix a is real, symmetric. Furthermore we suppose that there exist $\nu, \mu > 0$ with

$$\begin{aligned} &(i) && \nu|\xi|^2 \leq \langle a\xi, \xi\rangle \leq \frac{1}{\nu}|\xi|^2 \\ &(ii) && \nu \leq b \leq \frac{1}{\nu} \\ &(iii) && |\vec{B}| \leq \frac{\sqrt{\mu}}{\lambda} \\ &(iv) && b(\cdot, t) \in C^1(\mathbb{R}^n_+),\ \text{for each } t,\ \text{with } |\nabla_{\lambda,x} b| \leq \frac{\sqrt{\mu}}{\lambda}. \end{aligned} \tag{4.2}$$

We then have the following Cacciopoli inequality.

Lemma 4.3. *Let L be as above. Suppose that u is a weak solution of the equation $Lu = 0$ in $B_{2r} \equiv B_{2r}(X^1, t^1) \subseteq \mathbb{R}^{n+1}_+$. Then*

$$\int_{B_r} |\nabla u|^2 \leq C(n, \nu)(1 + \mu)r^{-2} \int_{B_{2r}} |u|^2.$$

Sketch of Proof. The case $b \equiv 1$ is well known (see [1], for example); the linear dependence on μ, of the constant on the right hand side of the inequality, may be readily obtained by keeping track of the constants arising in the standard proof. We omit the details. The general case (i.e. $b \not\equiv 1$) is an immediate consequence of the case $b \equiv 1$. Indeed, consider the operator

$$\tilde{L}u \equiv \operatorname{div}\left(\frac{1}{b} a \nabla u\right) - \frac{\partial u}{\partial t} + \left(\frac{1}{b}\vec{B} - \left[\nabla\left(\frac{1}{b}\right)\right] a\right) \cdot \nabla u.$$

Clearly, $Lu = 0 \Rightarrow \tilde{L}u = 0$. Moreover, $\tilde{L}$ satisfies (4.2) with $\tilde{a} \equiv \frac{1}{b}a$, $\tilde{b} \equiv 1$, $\tilde{\vec{B}} \equiv \frac{1}{b}\vec{B} - \left(\nabla \frac{1}{b}\right) a$, and $\tilde{\nu} = C(\nu)$, $\tilde{\mu} \leq C(\nu)\mu$.

As a Corollary of Lemma 4.3, we shall obtain the following. In the sequel $\Delta_r(z)$ denotes the parabolic "surface" ball (i.e. a parabolic ball in $\mathbb{R}^n$ in the present case) with center z.

Lemma 4.4. *Let L be as in Lemma 4.3, and suppose that $Lu = 0$ in $\mathbb{R}^{n+1}_+$. Let $\Delta_r \equiv \Delta_r(x^1, t^1) \subseteq \mathbb{R}^n$. Then*

$$\int_0^r \int_{\Delta_r} |\nabla u|^2 dz \lambda^2 d\lambda \leq C(n, \nu)(1 + \mu) \int_0^{2r} \int_{\Delta_{2r}} |u|^2 dz d\lambda.$$

Sketch of Proof. Cover $\psi_r \equiv (0, r) \times \Delta_r$ by a non-overlapping grid of parabolic cubes Q_j such that the parabolic diameter d_j of Q_j is comparable to (parabolic) $\operatorname{dist}(Q_j, \partial\mathbb{R}^{n+1}_+)$. Then, for $(\lambda, z) \in Q_j$, we have $\lambda \approx d_j$. Applying Lemma 4.3 in each Q_j, we obtain the conclusion of the present Lemma. We leave the routine details to the reader.

Next, we note some additional estimates which hold in the case $b \equiv 1$, $\vec{B} \equiv 0$. Probably, these can be generalized, but the following results will suffice for our purposes. In addition to (4.2), we suppose that $a \in C^1(\mathbb{R}^{n+1}_+)$, with

$$\text{(4.5)} \qquad \begin{aligned} &(i) \quad |\nabla_{\lambda,x} a(\lambda, x, t)| \leq \frac{\sqrt{\mu}}{\lambda} \\ &(ii) \quad \left|\frac{\partial a}{\partial t}(\lambda, x, t)\right| \leq \frac{\sqrt{\mu}}{\lambda^2}. \end{aligned}$$

We then have the following:

Lemma 4.6. *Let L be as in* (4.1)*, with $b \equiv 1$, $\vec{B} \equiv 0$, and let the coefficient matrix a satisfy* (4.2)*(i) and* (4.5)*. Suppose that $Lu = 0$ in $B_{2r} \equiv B_{2r}(\lambda^1, x^1, t^1) \subseteq \mathbb{R}^{n+1}_+$. Then, for all $(\lambda, x, t) \in B_r$,*

$$|\nabla_{\lambda,x} u(\lambda, x, t)| \leq C(n, \nu)\frac{\sqrt{\mu}}{r} \sup_{B_{\frac{3r}{2}}} |u|$$

Lemma 4.7. *Let L be as in Lemma 4.6, and suppose that $Lu = 0$ in $\mathbb{R}^{n+1}_+$. Then for any constant $\tilde{C}$,*

$$|\nabla u(\lambda, x, t)| \leq C(n, \nu)\frac{\sqrt{\mu}}{\lambda} \sup_{\|(\rho,y,s)-(\lambda,x,t)\|<\frac{\lambda}{8}} |u(\rho, y, s) - \tilde{C}|.$$

Sketch of Proofs. The case $r = 1$ of Lemma 4.6 is a particular case of a result of Aronson [1]. The general case follows by dilation invariance. It is enough to prove Lemma 4.7 with $\tilde{C} = 0$, in which case we apply Lemma 4.6 in $B_r(\lambda, x, t)$, with $r = \frac{\lambda}{16}$.

As a corollary of the last lemma, we deduce

Lemma 4.8. *Let L, u be as in Lemma 4.7. Then*

$$\lambda^{d+1}|\nabla u(\lambda, x, t)|^2 \leq C(n, \nu)\mu \iint_{\|(\rho,y,s)-(\lambda,x,t)\|<\frac{\lambda}{4}} |\nabla u(\rho, y, s)|^2 d\rho dy ds.$$

Proof of Lemma 4.8. We apply, in order: (1) Lemma 4.7; (2) a well known estimate of Moser; (3) a Poincaré inequality for solutions of $Lu = 0$ due to R. Brown [3, (3.9)], to obtain that, for suitably chosen $\tilde{C}$,

$$\begin{aligned}
\lambda^{d+1}|\nabla u(\lambda, x, t)|^2 &\leq C(n, \nu)\frac{\mu}{\lambda^2}\left(\sup_{\|(\rho,y,s)-(\lambda,x,t)\|<\frac{\lambda}{8}} |u(\rho, y, s) - \tilde{C}|\right)^2 \\
&\leq C(n, \nu)\frac{\mu}{\lambda^2} \iint_{\|(\rho,y,s)-(\lambda,x,t)\|<\frac{\lambda}{4}} |u(\rho, y, s) - \tilde{C}|^2 d\rho dy ds \\
&\leq C(n, \nu)\mu \iint_{\|(\rho,y,s)-(\lambda,x,t)\|<\frac{\lambda}{4}} |\nabla u|^2 d\rho dy ds,
\end{aligned}$$

and the lemma is proved.

5 L^2 estimates for the Dirichlet problem for $L_0 u_0 = 0$.

Our goal in this section is to prove the following:

Lemma 5.1. *Let $q \geq 2$. The Dirichlet problem*

$$(D_q) \begin{cases} L_0 u_0 = 0 & \text{in } \tilde{\Omega} \equiv \mathbb{R}^{n+1}_+ \equiv \{(\lambda, x, t) \in \mathbb{R} \times \mathbb{R}^{n-1} \times \mathbb{R} : \lambda > 0\} \\ u_0|_{\partial\tilde{\Omega}} = f \in L^q \\ N_*(u_0) \in L^q \end{cases}$$

is uniquely solvable if β_1 and γ are small enough. Furthermore, we have the estimate

$$\|N_*(u_0)\|_{L^q(\mathbb{R}^n)} \leq C(n, \beta_0, q)\|f\|_{L^q(\mathbb{R}^n)}, \quad q \geq 2.$$

(Remark, here N_* is defined in $\mathbb{R}^{n+1}_+$ exactly as in Ω, with $A(x, t) \equiv 0$; see (1.6) *and the ensuing definition).*

Proof. The corresponding result for $L_1 u_1 = 0$ was obtained in [18]. To be more precise, L^2 solvability of the Dirichlet problem for the heat equation in Ω was established in [18], and this is equivalent to (D_2) for L_1 in $\tilde{\Omega}$. The case $q = 2$ immediately implies the case $q > 2$. The Proof of Lemma 5.1 involves perturbing off the result for L_1, via the same approach as had been used in [20] to treat the "small constant" case.

We recall that we are using the notational convention that $z \equiv (x,t) \in \mathbb{R}^n$, $v = (y,s) \in \mathbb{R}^n$, and that

$$\Delta_r(z) \equiv \{v \in \mathbb{R}^n : \|v - z\| < r\};$$

i.e. $\{0\} \times \Delta_r(z)$, which by abuse of notation we also denote as $\Delta_r(z)$, is the parabolic surface ball on $\partial\mathbb{R}^{n+1}_+$. Let $Z \equiv (\lambda, z) \equiv (\lambda, x, t) \in \mathbb{R}^{n+1}$, and set

$$K(Z,v) \equiv \frac{d\omega^Z}{dv},$$

where ω^Z denotes parabolic measure for L_0 in $\mathbb{R}^{n+1}_+$, with respect to the point Z. We recall that we have made the *a priori* assumption that the boundary of our original domain Ω was given by the graph of a C^∞ function A. Hence L_0 has C^∞ coefficients, so that ω^Z and surface (i.e. Lebesque) measure are mutually absolutely continuous. Thus $K(z,v)$ is well defined. Of course, our estimates will not depend on our *a priori* assumption of smoothness. We recall that $d = n+1$. We define

$$T \equiv \sup_{\substack{r>0 \\ z\in\mathbb{R}^n}} r^d \int_{\Delta_r(z)} (K(A_r(z),v))^2 dv, \tag{5.2}$$

where, if $z = (x,t)$,

$$A_r(z) \equiv (8r, x, t + 64r^2).$$

By our *a priori* assumptions, T is finite. Now fix r and z_0. Our goal is to prove that, for all $g \in C_0^\infty$, supported in $\Delta_r(z_0)$ with $\|g\|_2 = 1$,

$$\left| \int_{\Delta_r(z_0)} g(v) K(A_r(z_0), v) dv \right| \le C(n, \beta_0 + \beta_1) r^{-\frac{d}{2}} (1 + (\epsilon T)^{\frac{1}{2}}), \tag{5.3}$$

where $\epsilon = 0(\epsilon_0)$ as $\epsilon_0 \to 0$, and where $\beta_1 \le \epsilon_0 \le 1$. For ϵ_0 sufficiently small, depending only on n and β_0, (5.3) implies that $T \le C(n, \beta_0 + 1)$, and the latter estimate in turn implies the conclusion of Lemma 5.1, just as in the case of the heat equation (see [13]). We omit the details.

We now proceed to prove (5.3). Let θ_r be a smooth cut-off function which is 1 on $[0,r] \times \Delta_r(z_0)$, and which vanishes outside of $(-2r, 2r) \times \Delta_{2r}(z_0)$. Let u_1 satisfy $L_1 u_1 = 0$, $u_1|_{\partial\tilde{\Omega}} = g$, with

$$\|N_*(u_1)\|_2 \le C(\beta_0)\|g\|_2 \equiv C(\beta_0). \tag{5.4}$$

By [18], there exists such a u_1, if $\beta_1 \le \epsilon_0$ is small enough. Let $h \equiv u_1\theta_r$. Set $u_0 \equiv \int_{\mathbb{R}^n} g(v) K(\cdot, v) dv$, so that $L_0 u_0 = 0$, $u|_{\partial\tilde{\Omega}} = g$. We recall that G_0 denotes the Green's function for L_0. By the divergence theorem, and the fact that $\theta_r \equiv 1$ on $\operatorname{supp} g$, we have

$$u_0(A_r(z_0)) = \int_{\mathbb{R}^n} h(v) K(A_r(z_0), v) dv$$

$$
\begin{aligned}
&= -\iint_{\mathbb{R}^{n+1}_+} \operatorname{div}(ha\nabla G_0(A_r(z_0),\cdot)) \\
&= -\iint \langle \nabla h, a\nabla G_0\rangle - \iint h \operatorname{div}(a\nabla G_0) \\
&= -\iint \langle \nabla h, a\nabla G_0\rangle + \iint h\frac{\partial}{\partial t} G_0,
\end{aligned}
$$

since $h = u_1\theta_r$ is supported away from the pole at $A_r(z_0)$. Also, letting N denote the inner unit normal to $\partial\tilde{\Omega}$, we observe that

$$
\begin{aligned}
0 = -\int_{\mathbb{R}^n} \langle N, a\nabla h\rangle G_0 &= \iint \operatorname{div}(a\nabla h G_0) \\
&= \iint \operatorname{div}(a\nabla h)G_0 + \iint a\nabla h\cdot\nabla G_0.
\end{aligned}
$$

Since a is a self-adjoint matrix, we therefore obtain that

$$
\begin{aligned}
u_0(A_r(z_0)) &= \iint_{\mathbb{R}^{n+1}_+} h\frac{\partial}{\partial t}G_0 + \iint_{\mathbb{R}^{n+1}_+} \operatorname{div}(a\nabla h)G_0 \\
&= \iint (L_0 h)\cdot G_0 \\
&= \iint (L_0\theta_r)u_1 G_0 + \iint \theta_r(L_0 u_1)G_0 + 2\iint a\nabla\theta_r\cdot\nabla u_1 G_0 \\
&\equiv I + II + III.
\end{aligned}
\tag{5.5}
$$

Furthermore, since $L_1 u_1 = 0$,

$$
\begin{aligned}
II &= \iint_{\mathbb{R}^{n+1}_+} \theta_r(L_0 - L_1)u_1 G_0 \\
&= -\iint \frac{\partial}{\partial t}P_{\gamma\lambda}A\frac{\partial u_1}{\partial\lambda}G_0\theta_r + \iint \frac{\partial}{\partial\lambda}P_{\gamma\lambda}A\frac{\partial u_1}{\partial t}G_0\theta_r \\
&\equiv II_1 + II_2.
\end{aligned}
\tag{5.6}
$$

We first consider

$$
\begin{aligned}
I &\equiv \iint_{\mathbb{R}^{n+1}_+} \operatorname{div}(a\nabla\theta_r)u_1 G_0 - \iint_{\mathbb{R}^{n+1}_+} \frac{\partial\theta_r}{\partial t}u_1 G_0 \\
&\equiv I_1 + I_2.
\end{aligned}
$$

Let $T_{2r}(z_0) \equiv \operatorname{supp}\theta_r \cap \mathbb{R}^{n+1}_+$. Then

$$
\begin{aligned}
|I_2| &\le \frac{C}{r^2}\left(\iint_{T_{2r}(z_0)} |u_1|^2\right)^{\frac{1}{2}} \left(\iint_{T_{2r}(z_0)} |G_0|^2\right)^{\frac{1}{2}} \\
&\le Cr^{\frac{-(d+1)}{2}}\left(\int_0^{2r}\int_{\Delta_{2r}(z_0)} |N_*(u_1)|^2\right)^{\frac{1}{2}}
\end{aligned}
$$

(since $G_0 \le Cr^{1-d}$ in $T_{2r}(z_0)$)

$$\le C(\beta_0) r^{-\frac{d}{2}},$$

by (5.4), since $\|g\|_2 \equiv 1$. This establishes the desired bound (5.3) for I_2. To handle I_1, we integrate by parts to obtain

$$\begin{aligned} |I_1| &= \left| -\iint_{\mathbb{R}^{n+1}_+} a\nabla\theta_r \cdot \nabla u_1 G_0 - \iint_{\mathbb{R}^{n+1}_+} a\nabla\theta_r \cdot \nabla G_0 u_1 \right| \\ &\le \frac{c}{r}\left\{ \left(\iint_{T_{2r}(z_0)} |\nabla u_1|^2 \lambda d\lambda dz \right)^{\frac{1}{2}} \left(\iint_{T_{2r}(z_0)} \frac{G_0^2}{\lambda} d\lambda dz \right)^{\frac{1}{2}} \right. \\ &\quad \left. + \left(\iint_{T_{2r}(z_0)} |u_1|^2 \right)^{\frac{1}{2}} \left(\iint_{T_{2r}(z_0)} |\nabla G_0|^2 \right)^{\frac{1}{2}} \right\}. \end{aligned} \tag{5.7}$$

The second summand in this last expression satisfies the same estimate as I_2 above, by Caccipoli's inequality at the boundary, applied to ∇G_0. Furthermore since we have square function estimates for $|\nabla u_1|^2\lambda$, by [4] and [26] (or alternatively, by [18]), we have that the first summand is no larger than

$$\begin{aligned} &\frac{C(\beta_0+\beta_1)}{r} \|N_*(u_1)\|_2 \left(\iint_{T_{2r}(z_0)} \frac{G_0^2}{\lambda} d\lambda dz \right)^{\frac{1}{2}} \\ &\le C(\beta_0+\beta_1) r^{-1} \left(\iint_{T_{2r}(z_0)} \frac{G_0^2}{\lambda} d\lambda dz \right)^{\frac{1}{2}}, \end{aligned}$$

where we have used (5.4) in the second step. (We remark that the square function estimates of [4] can be applied above since u_1 is the pullback of a solution of the heat equation). Also, by a standard estimate which can be proved exactly as in the elliptic case (see [22, p. 10] for the elliptic case), we have, for all $(\lambda, z) \in T_{2r}(z_0)$,

$$\begin{aligned} \frac{G_0(A_r(z_0),\lambda,z)}{\lambda} &\le C(\beta_0+\beta_1,n)\lambda^{-d} \int_{\Delta_\lambda(z)} K(A_r(z_0),\cdot) \\ &\approx C(\beta_0+\beta_1,n) P_\lambda(K(A_r(z_0),\cdot))(z), \end{aligned} \tag{5.8}$$

where P_λ denotes a nice parabolic approximate identity, which is of course bounded on $L^1(\mathbb{R}^n)$, uniformly in λ. Hence,

$$\begin{aligned} r^{-1} \left(\iint_{T_{2r}(z_0)} \frac{G_0^2}{\lambda} d\lambda dz \right)^{\frac{1}{2}} &\le C r^{\frac{-(d+1)}{2}} \left(\int_0^{2r} \int_{\mathbb{R}^n} K \right)^{\frac{1}{2}} \\ &= C r^{\frac{-d}{2}}. \end{aligned}$$

Altogether then, we obtain (5.3) for I, and also for III, since III is the same as the first term in (5.7). It remains to treat II. That is, we consider the splitting of II into $II_1 + II_2$ as in (5.6). We have

$$|II_1| \le \left(\iint_{T_{2r}(z)} \left(\frac{G_0}{\lambda}\right)^2 \left| \frac{\partial}{\partial t} P_{\gamma\lambda} A \right|^2 \lambda d\lambda dz \right)^{\frac{1}{2}} \left(\iint_{\mathbb{R}^{n+1}_+} |\nabla u_1|^2 \lambda d\lambda dz \right)^{\frac{1}{2}}.$$

Now, as above, by (5.4) and the square function estimates of [4] and the caloric measure estimates of [26] (but see also [18]), the second factor is no larger than $C(\beta_0+\beta_1)\|g\|_2 = C(\beta_0+\beta_1)$. Moreover, by (5.8), and the definition (5.2), the first factor is no larger than

$$C(\beta_0+\beta_1,n)r^{\frac{-d}{2}}T^{\frac{1}{2}}\left(\sup_{\substack{r>0\\ z\in\mathbb{R}^n}}\int_0^r r^{-d}\int_{\Delta_r(z)}\left|\frac{\partial}{\partial t}P_{\gamma\lambda}A\right|^2 dv\lambda d\lambda\right)^{\frac{1}{2}}.$$

Since the Carleson measure in the last term has norm no larger than

$$C\gamma^{-2}\|\mathbb{D}_n A\|_*^2 \le C\gamma^{-2}\beta_1^2$$

we therefore obtain the desired bound (5.3) for II_1, as long as β_1 is taken suitably small compared to γ.

Finally, we observe that

$$\begin{aligned}|II_2| &\le C\left(\iint_{T_{2r}(z_0)}\left(\frac{G_0}{\lambda}\right)^2\left|\frac{\partial}{\partial\lambda}P_{\gamma\lambda}A\right|^2\frac{d\lambda}{\lambda}dz\right)^{\frac{1}{2}}\left(\iint_{\mathbb{R}^{n+1}_+}\left|\frac{\partial u_1}{\partial t}\right|^2\lambda^3 d\lambda dz\right)^{\frac{1}{2}}\\ &\le C(\beta_0+\beta_1,n)r^{-\frac{d}{2}}T^{\frac{1}{2}}\gamma,\end{aligned}$$

by the same reasoning as above, since

$$\sup_{\substack{r>0\\ z\in\mathbb{R}^n}}\int_0^r r^{-d}\int_{\Delta_r(z)}\left|\frac{\partial}{\partial\lambda}P_{\gamma\lambda}A\right|^2\frac{d\lambda}{\lambda}dz \le C(\beta_0+\beta_1)\gamma^2.$$

Choosing γ small, we obtain (5.3). This concludes the proof of existence in Lemma 5.1.

The proof of uniqueness is standard (it is not hard to adapt the proof given in [22, p.30] in the elliptic case), and we omit the details.

6 The "Main Lemma" implies (3.8); "Main Lemma" plus (3.9) imply (3.10).

We recall that it suffices (ultimately) to prove (3.10). Thus, the present section reduces matters to proving the "Main Lemma" and (3.7) in the sequel (since (3.7) and (3.8) imply (3.9) by interpolation).

We begin by introducing some notation. Following [15], [23],[24] set

$$\begin{aligned}F(X,t) &\equiv \iint_{\mathbb{R}^{n+1}_+} G_0(X,t,Y,s)\frac{\partial}{\partial t}P_{\gamma\lambda}A\frac{\partial}{\partial\lambda}u_1 dY ds\\ &\quad - \iint_{\mathbb{R}^{n+1}_+} G_0(X,t,Y,s)\frac{\partial}{\partial\lambda}P_{\gamma\lambda}A\frac{\partial u_1}{\partial t}dY ds\\ &\equiv F_1(X,t)-F_2(X,t).\end{aligned} \tag{6.1}$$

We note that

$$|u_1-u_0| \equiv |F|, \quad \text{if } \ u_1|_{\partial\mathbb{R}^{n+1}_+} = u_0|_{\partial\mathbb{R}^{n+1}_+}. \tag{6.2}$$

For $z \in \mathbb{R}^n$, and for f defined in $\mathbb{R}^{n+1}_+$, set

$$N_* f(z) \equiv \sup_{(\lambda,v):\|v-z\|<\lambda} (f(\lambda, v)). \tag{6.3}$$

We recall also that $\|\nabla_x A\|_\infty \le \beta_0$, $\|\mathbb{D}_n A\|_* \le \beta_1$, $\frac{\partial}{\partial\nu} f \equiv \langle -e_0, a\nabla f\rangle$ and that we are taking $\beta_1 \le \epsilon_0$. We now state our "Main Lemma":

Lemma 6.4. *Suppose $\epsilon_0 \le \gamma^{M_0}$, for M_0 sufficiently large (M_0 depends on n and p). Then*

$$\left\|\frac{\partial F}{\partial \nu}\right\|_p \le C(p, \beta_0)\gamma \|N_*(\nabla u_1)\|_p, \quad 1 < p \le 2.$$

Let us defer the Proof of Lemma 6.4, and establish first estimate (3.8). Our strategy follows that of [23, 24]. We take $u_0|_{\partial\mathbb{R}^{n+1}_+} \equiv u_1|_{\partial\mathbb{R}^{n+1}_+}$, with $L_0 u_0 = 0$, $L_1 u_1 = 0$. Then

$$\begin{aligned}\|\mathbb{D}_n u_0\|_{L^2(\mathbb{R}^n)} + \|\nabla_{\tan} u_0\|_{L^2(\mathbb{R}^n)} &\equiv \|\mathbb{D}_n u_1\|_2 + \|\nabla_{\tan} u_1\|_2 \\ &\le C(\beta_0, n)\left\|\frac{\partial u_1}{\partial \nu}\right\|_2\end{aligned} \tag{6.5}$$

(the latter estimate is simply the result of [18] for the Neumann problem, pulled back to the half-space), and in turn,

$$\begin{aligned}\left\|\frac{\partial u_1}{\partial \nu}\right\|_2 &\le \left\|\frac{\partial u_0}{\partial \nu}\right\|_2 + \left\|\frac{\partial F}{\partial \nu}\right\|_2 \\ &\le \left\|\frac{\partial u_0}{\partial \nu}\right\|_2 + C(\beta_0)\gamma\|N_*(\nabla u_1)\|_2\end{aligned}$$

(by Lemma 6.4)

$$\le \left\|\frac{\partial u_0}{\partial \nu}\right\|_2 + C(\beta_0)\gamma(\|\mathbb{D}_n u_1\|_2 + \|\nabla_{\tan} u_1\|_2),$$

(by the result of [18] for the regularity problem). Choosing γ small, and hiding the small term on the left side of (6.5), we obtain (3.8).

Next, we show that the Main Lemma plus (3.9) imply (3.10). Again we choose $u_0|_{\partial\mathbb{R}^{n+1}_+} \equiv u_1|_{\partial\mathbb{R}^{n+1}_+}$. Then the left side of (3.10) equals

$$\begin{aligned}\|\nabla_{\tan} u_1\|_{L^p(\mathbb{R}^n)} + \|\mathbb{D}_n u_1\|_{L^p(\mathbb{R}^n)} &\equiv \|\nabla_{\tan} u_0\|_p + \|\mathbb{D}_n u_0\|_p \\ &\le C(\beta_0, n, p)\left\|\frac{\partial u_0}{\partial \nu}\right\|_p\end{aligned}$$

(by (3.9))

$$\le C(\beta_0, n, p)\left(\left\|\frac{\partial u_1}{\partial \nu}\right\|_p + \left\|\frac{\partial F}{\partial \nu}\right\|_p\right)$$

$$\leq C(\beta_0, n, p) \left(\left\| \frac{\partial u_1}{\partial \nu} \right\|_p + \gamma \| N_*(\nabla u_1) \|_p \right)$$

(by Lemma 6.4)

$$\leq C(\beta_0, n, p) \left(\left\| \frac{\partial u_1}{\partial \nu} \right\|_p + \gamma [\| \mathbb{D}_n u_1 \|_p + \| \nabla_{\tan} u \|_p] \right),$$

where the last inequality follows from the result of [19] for the regularity problem for $L^p, 1 < p < 2$. Choosing γ small, depending now on β, n and p, and hiding the small term on the left, we obtain (3.10).

7 Variational Solution to the Neumann Problem

In this section we construct a weak solution to the Neumann problem (see (7.7) below) for L_0. To this end fix q, $\frac{1}{2} = \frac{1}{q} - \frac{1}{2d}$, $d = n+1$. By the parabolic version of the Hardy-Littlewood-Sobolev Theorem (see [28, p. 119] for the elliptic version; the proof in the parabolic case is essentially the same), we have that

$$\|I_{\frac{1}{2}} f\|_{L^2(\mathbb{R}^n)} \leq C \|f\|_{L^q(\mathbb{R}^n)}, \tag{7.1}$$

where the parabolic fractional integral $I_{\frac{1}{2}}$ and fractional derivative $\mathbb{D}^\alpha$ are defined, respectively, by

$$\begin{aligned} I_\alpha f &\equiv (\|\zeta\|^{-\alpha} \hat{f}(\zeta))^\vee \\ \mathbb{D}^\alpha f &\equiv (\|\zeta\|^{\alpha} \hat{f}(\zeta))^\vee. \end{aligned} \tag{7.2}$$

(here we recall that $\| \cdot \|$ denotes the parabolic "norm", and that $\wedge, \vee$ denote, respectively, the Fourier and inverse Fourier Transforms). We remark that in the case $\alpha = 1$, we shall simply write $\mathbb{D}$, rather than $\mathbb{D}^1$.

Thus, (7.1) implies that $L^q \subseteq L^2_{-\frac{1}{2}} \equiv (L^2_{\frac{1}{2}})^*$, where, $L^2_{\frac{1}{2}} \equiv I_{\frac{1}{2}}(L^2)$ is the usual homogeneous parabolic Sobolev space having $\frac{1}{2}$ of a spatial derivative and $\frac{1}{4}$ of a time derivative in L^2. We shall construct our variational solution in a certain Hilbert space. Let H denote the closure of $C_0^\infty(\mathbb{R}^{n+1})$ under the norm

$$\|\nabla \varphi\|_{L^2(\mathbb{R}^{n+1}_+)} + \|D_t^{\frac{1}{2}} \varphi\|_{L^2(\mathbb{R}^{n+1}_+)} \equiv \|\varphi\|_H.$$

Here, $D_t^\alpha \varphi \equiv (|\tau|^\alpha \hat{\varphi}(\tau))^\vee$. We have the following,

Lemma 7.3. *Suppose $f \in L^q(\mathbb{R}^n)$, $\frac{1}{2} = \frac{1}{q} - \frac{1}{2d}$, and for $\varphi \in H$, define a linear functional on $\varphi|_{\partial \mathbb{R}^{n+1}_+}$ by*

$$\Lambda_f(\varphi) \equiv \int_{\mathbb{R}^n} f \varphi.$$

Then $\varphi \to \Lambda_f(\varphi)$ defines a bounded linear functional on H, with

$$|\Lambda_f(\varphi)| \leq C\|f\|_q\|\varphi\|_H.$$

Proof. As we have observed, $f \in L^q \Rightarrow f = \mathbb{D}^{\frac{1}{2}}g$, where $\|g\|_2 \leq C\|f\|_q$. Thus, by self-adjointness of $\mathbb{D}^\alpha$,

$$\left|\int f\varphi\right| \equiv \left|\int g\mathbb{D}^{\frac{1}{2}}\varphi\right| \leq C\|f\|_q\|\mathbb{D}^{\frac{1}{2}}\varphi\|_2.$$

Hence, it is enough to show that

$$\int_{\mathbb{R}^n} |\mathbb{D}^{\frac{1}{2}}\varphi|^2 dz \leq C\|\varphi\|_H^2. \tag{7.4}$$

But it is enough to prove this for $\varphi \in C_0^\infty(\mathbb{R}^{n+1})$, in which case the left hand side of (7.4) equals

$$\begin{aligned} -\int_0^\infty \int_{\mathbb{R}^n} \frac{\partial}{\partial\lambda}(\mathbb{D}^{\frac{1}{2}}\varphi)^2 dz d\lambda &= -2\int_0^\infty \int_{\mathbb{R}^n} \mathbb{D}\varphi\frac{\partial\varphi}{\partial\lambda} dz d\lambda \\ &\leq C\|\varphi\|_H^2, \end{aligned}$$

as desired. This proves the Lemma.

Next, we define certain bilinear functionals on H as follows. Fix $\epsilon > 0$ (to be chosen), and let H_t denote the Hilbert transform in t, i.e.

$$H_t f \equiv (i \operatorname{sgn} \tau \hat{f}(\tau))^\vee$$

For $u, \varphi \in H$, define

$$\begin{aligned} B_\epsilon(u,\varphi) &\equiv -\int_0^\infty \int_{\mathbb{R}^n} D_t^{\frac{1}{2}} u H_t D_t^{\frac{1}{2}} (I + \epsilon H_t)\varphi dz d\lambda \\ &\quad + \int_0^\infty \int_{\mathbb{R}^n} a\nabla u \cdot \nabla(I + \epsilon H_t)\varphi dz d\lambda \\ &\equiv -\iint D_t^{\frac{1}{2}} u H D_t^{\frac{1}{2}} \varphi_\epsilon - \iint a\nabla u \cdot \nabla\varphi_\epsilon \\ &\equiv B(u, \varphi_\epsilon), \end{aligned} \tag{7.5}$$

where $\varphi_\epsilon \equiv (I + \epsilon H_t)\varphi$, $\quad I =$ identity. Note that for ϵ small enough $I + \epsilon H$ is invertible on H and on $L^p(\mathbb{R}^n)$, $q \leq p \leq q'$. Also,

$$|B_\epsilon(u,\varphi)| \leq C\|u\|_H\|\varphi\|_H,$$

with C independent of $\epsilon \in [0,1]$.

We observe that, since $\int fH_t f dt = 0$, and since $H_t^2 = -I$, we have

$$\begin{aligned} B_\epsilon(u,u) &= \epsilon\iint |D_t^{\frac{1}{2}} u|^2 + \iint a\nabla u \cdot \nabla u \\ &\quad + \epsilon\iint a\nabla u \cdot H_t\nabla u \\ &\approx C_\epsilon\|u\|_H^2, \end{aligned} \tag{7.6}$$

if ϵ is chosen small enough. Furthermore,

$$\begin{aligned}\Lambda_f^\epsilon(\varphi) \equiv \Lambda_f(\varphi_\epsilon) &\equiv \int_{\mathbb{R}^n} f\varphi_\epsilon \\ &\equiv \int_{\mathbb{R}^n} f(I+\epsilon H_t)\varphi\end{aligned}$$

defines a bounded linear functional on H, with $|\Lambda_f^\epsilon(\varphi)| \leq C\|f\|_q\|\varphi\|_H$, with C independent of $\epsilon \in [0,1]$. Hence, by the Lax-Milgram Lemma, there exists $u \in H$ such that

$$B(u,\varphi_\epsilon) \equiv B_\epsilon(u,\varphi) \equiv \Lambda_f^\epsilon(\varphi) \equiv \Lambda_f(\varphi_\epsilon)$$

for all $\varphi \in H$. But we have chosen ϵ small enough that $I+\epsilon H_t$ is invertible on H, hence $B(u,\psi) \equiv \Lambda_f(\psi)$, for all $\psi \in H$; i.e., this u satisfies

$$\int_{\mathbb{R}^n} f\psi = -\iint_{\mathbb{R}^{n+1}_+} D_t^{\frac{1}{2}} u H_t D_t^{\frac{1}{2}}\psi + \iint_{\mathbb{R}^{n+1}_+} a\nabla u \cdot \nabla\psi. \tag{7.7}$$

Bearing in mind that $D_t H_t \equiv \frac{\partial}{\partial t}$, we take (7.7) to be our definition of a weak solution to the Neumann problem.

We now prove several estimates for our weak solutions which will be useful in the sequel.

Lemma 7.8. *Suppose $f \in L^q(\mathbb{R}^n)$, and u satisfies (7.7). Then*

$$\|u\|_H \leq C_\epsilon\|f\|_q.$$

Proof. By (7.6)

$$\begin{aligned}\|u\|_H^2 \leq C_\epsilon B_\epsilon(u,u) &\equiv C_\epsilon B(u,u_\epsilon) \\ &= C_\epsilon \Lambda_f(u_\epsilon) \\ &\equiv C_\epsilon \int_{\mathbb{R}^n} f u_\epsilon \\ &\leq C_\epsilon\|f\|_q\|u_\epsilon\|_H\end{aligned}$$

(the latter inequality follows from Lemma 7.3). Since $\|u_\epsilon\|_H \leq C\|u\|_H$, the Lemma now follows.

Lemma 7.9. *Let $u \in H$, and let $\frac{1}{p} = \frac{1}{2} - \frac{1}{d}$. Then*

$$\begin{aligned}&(i) \quad \int_0^\infty \left(\int_{\mathbb{R}^n} |u|^p\right)^{\frac{2}{p}} d\lambda \leq C\|u\|_H^2 \\ &(ii) \quad \lim_{R\to\infty} \frac{1}{R^2}\int_0^\infty \int_{R\leq\|z\|\leq 2R} |u|^2 dz d\lambda = 0 \\ &(iii) \quad \forall z_0 \in \mathbb{R}^n, \quad \frac{1}{R^2}\int_0^\infty \int_{\|z-z_0\|<R} \|u\|^2 dz d\lambda \leq C\|u\|_H^2.\end{aligned}$$

Proof. By the parabolic Hardy-Littlewood-Sobolev Theorem,

$$\left(\int_{\mathbb{R}^n} |u|^p\right)^{\frac{2}{p}} \leq C\int_{\mathbb{R}^n} |\mathbb{D}u|^2$$
$$\approx \int_{\mathbb{R}^n} |D_t^{\frac{1}{2}} u|^2 + \int_{\mathbb{R}^n} |\nabla_x u|^2,$$

and (i) now follows. To prove (ii), observe that, since $\frac{2}{\rho} + \frac{2}{d} = 1$, we have

$$\frac{1}{R^2}\int_{R\leq\|z\|\leq 2R} |u|^2 \leq C\frac{1}{R^2}\left(\int_{R\leq\|z\|\leq 2R} |u|^p\right)^{\frac{2}{p}} (R^d)^{\frac{2}{d}}$$
$$= \left(\int_{R\leq\|z\|\leq 2R} |u|^p\right)^{\frac{2}{p}}.$$

Integrating in λ, and using (i) and dominated convergence, we obtain (ii). The proof of (iii) is similar, since

$$\frac{1}{R^2}\int_{\|z-z_0\|<R} |u|^2 \leq C\left(\int_{\mathbb{R}^n} |u|^p\right)^{\frac{2}{p}},$$

and we again may integrate in λ and invoke (i).

8 Proof of (3.7).

It remains to prove (3.7), and Lemma 6.4. Taking the latter for granted, we carry out the former task in this section.

Our approach in this section follows analogous arguments of [23] for elliptic equations, and given the methods of that paper, this part of our proof will be fairly routine. The complications which arise in passing from the elliptic to the parabolic case are merely technical. The essentials of the proof remain the same as in [23]. We begin by defining a Neumann function for L_0. As above, let $G_0(X,t,Y,s)$ denote the Green's function for L_0. By Lemma 5.1 (L^2 solvability of the Dirichlet problem for L_0) we have that

$$K(X,t,v) \equiv -\frac{\partial}{\partial\nu} G_0(X,t,\cdot)(v)$$

belongs to L^2_{loc} in v (and even in $L^2(\mathbb{R}^n)$). Moreover, by our *a priori* assumption of smoothness, $K(X,t,\cdot) \in C^\infty$, although of course, none of our estimates will depend upon this smoothness. Let $D_t^{\frac{1}{2}}$ denote the one-dimensional $\frac{1}{2}$ order time derivative

$$D_t^{\frac{1}{2}} f \equiv (|\tau|^{\frac{1}{2}} \hat{f}(\tau))^\vee, \tag{8.1}$$

and as in Section 7 let H denote the closure of $C_0^\infty(\mathbb{R}^{n+1})$ under the norm

$$\|\nabla\varphi\|_{L^2(\mathbb{R}^{n+1}_+)} + \|D_t^{\frac{1}{2}}\varphi\|_{L^2(\mathbb{R}^{n+1}_+)} \equiv \|\varphi\|_H.$$

Then, given $f \in L^q(\mathbb{R}^n)$, $\frac{1}{2} = \frac{1}{q} - \frac{1}{2d}$, $d = n+1$, we say that $u \in H$ is a weak solution to the Neumann problem for L_0 with data f if, for all $\psi \in H$, we have

$$\int_{\mathbb{R}^n} f\psi = \iint_{\mathbb{R}^{n+1}_+} H_t D_t^{\frac{1}{2}} u D_t^{\frac{1}{2}} \psi + \iint_{\mathbb{R}^{n+1}_+} a\nabla u \cdot \nabla \psi, \tag{8.2}$$

where $H_t f \equiv (i \operatorname{sgn} \tau \hat{f}(\tau))^\vee$ denotes the Hilbert transform in t, so that

$$D_t^{\frac{1}{2}} H_t D_t^{\frac{1}{2}} \equiv \frac{\partial}{\partial t}.$$

In Section 7, we saw that, given $f \in L^q(\mathbb{R}^n)$, there exists such a u, and that it is unique in H. Furthermore, the arguments of Section 7 may be applied with a minus sign in front of the middle integral in (8.2), so that we also have weak solutions to the Neumann problem for L_0^*, the adjoint of L_0.

With our weak solutions in hand, we now proceed to construct a Neumann function. By the results of Section 5, we have that, for each fixed (X,t) in $\mathbb{R}^{n+1}_+$, the density of parabolic measure for L_0 satisfies

$$K(X,t,\cdot) \equiv -\frac{\partial G}{\partial \nu}(X,t,\cdot) \in L^p(\mathbb{R}^n) \quad 1 \le p < 2+\epsilon.$$

In particular, we may take $p = q$, $\frac{1}{2} = \frac{1}{q} - \frac{1}{2d}$. Thus there exists a variational solution for the Neumann problem

$$\begin{cases} L_0^* V(X,t,\cdot,\cdot) = 0 & \text{in } \mathbb{R}^{n+1}_+ \\ \frac{\partial V}{\partial \nu}(X,t,\cdot) = \frac{\partial G}{\partial \nu}(X,t,\cdot) & \text{on } \partial\mathbb{R}^{n+1}_+ \end{cases}.$$

We now define the Neumann function

$$N(X,t,Y,s) \equiv G(X,t,Y,s) - V(X,t,Y,s),$$

and observe that $\frac{\partial N}{\partial \nu}(X,t,\cdot) \equiv 0$ on $\partial\mathbb{R}^{n+1}_+$.

Estimate (3.7) will be an easy consequence of the following two results.

Lemma 8.3. *Let $(X,t) \equiv (\lambda,x,t)$, $\quad (Y,s) \equiv (\rho,y,s)$. The Neumann function satisfies*

$$|N(X,t,Y,s)| \le C(\beta_0+\beta_1,n)\|(X,t)-(Y,s)\|^{1-d},$$

(where $d \equiv n+1$ is the homogeneous dimension of $\mathbb{R}^n = \partial\mathbb{R}^{n+1}_+$) whenever

$$\rho \le \frac{1}{2}\|(x,t)-(y,s)\|$$

and

$$\lambda \le 2\|(x,t)-(y,s)\|.$$

Remark. Actually, the lemma holds under more general circumstances then stated, but this is all we'll need.

Lemma 8.4 (Localization Lemma). *Let Δ_r be any parabolic surface ball, and set $\psi_r \equiv (0,r)\times\Delta_r$. Suppose $L_0u_0 = 0$ in $\mathbb{R}^{n+1}_+$. Then*

$$\begin{aligned}\int_{\Delta_r} |\mathbb{D}_n u_0|^2 + \int_{\Delta_r} |\nabla_{\tan} u_0|^2 \\ &\leq C(\beta_0, n)\left[\int_{\Delta_{4r}} \left|\frac{\partial u_0}{\partial \nu}\right|^2 + \frac{1}{r}\iint_{\psi_{4r}} |\nabla u_0|^2 \right.\\ &\left.+ \frac{1}{r^2}\int_{\Delta_{4r}} |u_0|^2 + \frac{1}{r^3}\iint_{\psi_{4r}} |u_0|^2\right].\end{aligned}$$

For the moment, let us take Lemma 8.3 and Lemma 8.4 for granted, and deduce (3.7). It is enough to show that

$$\|\mathbb{D}_n u_0\|_1 + \|\nabla_{\tan} u_0\|_1 \leq C(\beta_0 + \beta_1, n), \tag{8.5}$$

whenever $\frac{\partial u_0}{\partial \nu} \equiv a$ is a smooth H^1 atom (since sums of the form $\Sigma\lambda_k a_k$, with a_k a smooth atom for each k, are dense in H^1). We suppose now that $\operatorname{supp} a \subseteq \Delta_r \equiv \Delta_r(z^0)$, with $\|a\|_\infty \leq cr^{-d}$. The technique is the usual one (see [10],[3],[23]). We have

$$\begin{aligned}\int_{\Delta_{64r}} |\mathbb{D}_n u_0| + \int_{\Delta_{64r}} |\nabla_{\tan} u_0| &\leq Cr^{\frac{d}{2}}(\|\mathbb{D}_n u_0\|_{L^2(\mathbb{R}^n)} + \|\nabla_{\tan} u_0\|_{L^2(\mathbb{R}^n)}) \\ &\leq C(\beta_0, n) r^{\frac{d}{2}} \|a\|_2,\end{aligned}$$

by (3.8). We write $\mathbb{R}^n\backslash\Delta_{64r} \equiv \cup_{k=6}^\infty R_k$, where $R_k \equiv \Delta_{2^{k+1}r}\backslash\Delta_{2^k r}$. Furthermore, each $R_k \subseteq \cup_{j=1}^N \Delta_{2^{k-4}r}(z^j)$, for some collection $\{z^j\} \subseteq \overline{R_k}$, where N depends only on dimension. Now,

$$\begin{aligned}\int_{R_k} (|\mathbb{D}_n u_0| + |\nabla_{\tan} u_0|) &\leq \sum_{j=1}^N \int_{\Delta_{2^{k-4}r}(z^j)} (|\mathbb{D}_n u_0| + |\nabla_{\tan} u_0|) \\ &\leq C\sum_{j=1}^N (2^k r)^{\frac{d}{2}} \left(\int_{\Delta_{2^{k-4}r}(z^j)} (|\mathbb{D}_n u_0|^2 + |\nabla_{\tan} u_0|^2)\right)^{\frac{1}{2}} \\ &\leq C\sum_{j=1}^N (2^k r)^{\frac{d}{2}} \left\{\left(\frac{1}{2^k r}\iint_{\psi_{2^{k-2}r}(z^j)} \left[|\nabla u_0|^2 + \left(\frac{1}{2^k r}|u_0|\right)^2\right]\right)^{\frac{1}{2}}\right. \\ &\left.+ \left(\frac{1}{(2^k r)^2}\int_{\Delta_{2^{k-2}r}(z^j)} |u_0|^2\right)^{\frac{1}{2}}\right\},\end{aligned}$$

by Lemma 8.4, since $\frac{\partial u_0}{\partial \nu}$ vanishes on $\mathbb{R}^n\backslash\Delta_r(z^0)$. Furthermore, as in [3], we can extend L_0 and u_0 by even reflection across the portion of the boundary where $\frac{\partial u_0}{\partial \nu}$ vanishes. Thus, we have a Cacciopoli inequality at the boundary which implies that

$$\frac{1}{2^k r}\iint_{\psi_{2^{k-2}r}(z^j)} |\nabla u_0|^2 \leq C(\beta_0 + \beta_1, n)\frac{1}{(2^k r)^3}\iint_{\psi_{2^{k-1}r}(z^j)} |u_0|^2.$$

Hence, to prove (8.5) (and therefore (3.7)), it suffices to show that, on $\overline{\psi_{2^{k-1}r}(z^j)}$,

$$|u_0| \le C(\beta_0 + \beta_1, n)(2^k r)^{-d+1} 2^{-k\alpha}, \tag{8.6}$$

for some $\alpha > 0$, for in that case we can add up a geometric series in k. But

$$u_0(X,t) = \int_{\mathbb{R}^n} N(X,t,z) a(z) dz, \tag{8.7}$$

where N is the Neumann function.

Following [3] we observe that, by definition $\frac{\partial N}{\partial \nu}(X,t,\cdot) = 0$, so by even reflection $N(X,t,\cdot,\cdot)$ is an adjoint solution of the extended operator $\tilde{L}_0$, in $\psi_{2^{k-3}r}(z^0)$, since (X,t) lies in $\overline{\psi_{2^{k-4}r}(z^j)} \subseteq (\psi_{2^{k-1}r}(z^0))^c$. Thus, by Nash's Theorem and Lemma 8.3, for some $\alpha > 0$,

$$\begin{aligned} |N(X,t,z) - N(X,t,z^0)| &\le C(\beta_0 + \beta_1, n) \left(\frac{\|z - z^0\|}{2^k r} \right)^{\alpha} (2^k r)^{1-d} \\ &\le C(\beta_0 + \beta_1, n) 2^{-k\alpha} (2^k r)^{1-d}, \end{aligned} \tag{8.8}$$

whenever $z \in \Delta_r(z^0)$. Since $\int a(z)dz = 0$, we have that $u_0(X,t) = \int [N(X,t,z) - N(X,t,z_0)] a(z) dz$, so by (8.8) and the fact that $\|a\|_1 \le 1$, we obtain (8.6), as desired. This proves (3.7), modulo Lemmas 8.3 and 8.4. We give the proof of the latter now, and defer the proof of Lemma 8.3 until the end of this section.

To prove Lemma 8.4, we first require a characterization of the symmetry properties of $N(X,t,Y,s)$. Recall that $N(X,t,Y,s) \equiv G(X,t,Y,s) - V(X,t,Y,s)$, where $V(X,t,\cdot,\cdot)$ solves the Neumann problem for L_0^* with data $\frac{\partial G}{\partial \nu}(X,t,\cdot,\cdot)$.

Lemma 8.9. *Suppose that $\tilde{V}(X,t,Y,s)$ solves, for each fixed (Y,s), the Neumann problem for L_0 with data $-\tilde{K}(\cdot,Y,s) \equiv \left.\frac{\partial G(\cdot,\cdot,Y,s)}{\partial \nu}\right|_{\partial \mathbb{R}^{n+1}_+}$. Then $\tilde{V}(X,t,Y,s) = V(X,t,Y,s)$.*

Proof of Lemma 8.9. We continue to follow [23]. By definition of $\tilde{V}, V$, and the definition of our variational solutions (and adjoint solutions), we have

$$\begin{aligned} &- \left\{ \int_{\mathbb{R}^n} K(X,t,\cdot)\tilde{V}(\cdot,Y,s) - \int_{\mathbb{R}^n} \tilde{K}(\cdot,Y,s) V(X,t,\cdot) \right\} \\ &= - \iint_{\mathbb{R}^{n+1}_+} H_t D_t^{\frac{1}{2}} V D_t^{\frac{1}{2}} \tilde{V} + \iint_{\mathbb{R}^{n+1}_+} a \nabla V \cdot \nabla \tilde{V} \\ &\quad - \left[\iint_{\mathbb{R}^{n+1}_+} H_t D_t^{\frac{1}{2}} \tilde{V} D_t^{\frac{1}{2}} V + \iint_{\mathbb{R}^{n+1}_+} a \nabla V \cdot \nabla \tilde{V} \right] \\ &= 0, \end{aligned} \tag{8.10}$$

by the anti-symmetry of H_t. Thus, it is enough to show that the first two integrals in (8.10) equal $\tilde{V}(X,t,Y,s)$, and $V(X,t,Y,s)$, respectively. The proof is the same in each case, so let us consider the former. Fix (Y,s) and let $\tilde{u}(X,t) \equiv \tilde{u}_{Y,s}(X,t) \equiv \int_{\mathbb{R}^n} K(X,t,\cdot)\tilde{V}(\cdot,Y,s)$. Of course $\tilde V$, and $\tilde V = \tilde V(\cdot, Y, s)$ are both solutions, with $\tilde u|_{\partial\mathbb{R}^{n+1}_+} \equiv \tilde V\Big|_{\partial\mathbb{R}^{n+1}_+}$, so by the maximum principle it suffices to show that, for some $\epsilon > 0$,

$$\sup_{\substack{R\le\|(X,t)\|\le 2R \\ (X,t)\in\mathbb{R}^{n+1}_+}} |\tilde u(X,t) - \tilde V(X,t)| = 0(\mathbb{R}^{-\epsilon}) \tag{8.11}$$

as $R \to \infty$. Let $w \equiv \tilde u - \tilde V$. By Moser's interior estimates,

$$\begin{aligned}
|w(\lambda, X, t)|^2 &\le C\lambda^{-d-1} \int_{B_{\frac{\lambda}{2}}(\lambda,x,t)} |w|^2 \\
&\le C\lambda^{-d-1} \int_0^{\frac{3\lambda}{2}} \int_{\|(x,t)\|\le\lambda} |w|^2 \\
&\le C\left(\lambda^{-d+1}\lambda^{-2}\int_0^{\frac{3\lambda}{2}}\int_{\|(x,t)\|\le\lambda} |\tilde V|^2 + \lambda^{-d-1}\int_0^{\frac{3\lambda}{2}}\int_{\|(x,t)\|<\lambda} |N_*\tilde u|^2\right).
\end{aligned}$$

The first of these is $0(\lambda^{-d+1})$ by Lemma 7.9 (iii). The second term is no larger than

$$\begin{aligned}
\lambda^{-d}\int_{\|(x,t)\|<\lambda} |N_*\tilde u|^2 &\le \left(\lambda^{-d}\int_{\|(x,t)\|<\lambda} |N_*\tilde u|^p\right)^{\frac{2}{p}} \\
&\le C\left(\lambda^{-d}\int_{\mathbb{R}^n} |\tilde u|^p\right)^{\frac{2}{p}}
\end{aligned}$$

where $\frac{1}{p} = \frac{1}{2} - \frac{1}{d}$, and where we have used the fact that Lemma 5.1 also implies L^p estimates for the Dirichlet problem when $p > 2$. But by the parabolic Hardy-Littlewood-Sobolev Theorem, the last expression is dominated by

$$C\lambda^{-\frac{2d}{p}} \int_{\mathbb{R}^n} |\mathbb{D}\tilde u|^2,$$

and we claim that this last expression is $0(\lambda^{-\frac{2d}{p}})$. Indeed $\mathbb{D}\tilde u|_{\partial\mathbb{R}^{n+1}_+} \equiv \mathbb{D}\tilde V\Big|_{\partial\mathbb{R}^{n+1}_+}$. Also,

$$\int_{\mathbb{R}^n} |\mathbb{D}\tilde V|^2 \le c\int_{\mathbb{R}^n}\left|\frac{\partial\tilde V}{\partial\nu}\right|^2 \equiv C\int_{\mathbb{R}^n}\left|\frac{\partial G(\cdot, Y, s)}{\partial\nu}\right|^2 < \infty,$$

where the first inequality is (3.8) (we recall that we have already established (3.8), modulo the "Main Lemma") and the last inequality is essentially Lemma 5.1. This establishes (8.11) in the case that $\|(x,t)\| < \lambda$.

To treat the case that $\|(x,t)\| > \lambda$, we note that since $w \equiv \tilde u - \tilde V \equiv 0$ on $\mathbb{R}^{n+1}_+$, by Moser's estimates we have

$$\sup_{T_R} |w|^2 \le R^{-d-1}\int_{T_R^*} w^2,$$

where $T_R \equiv (0,R)\times\{(x,t) : R \le \|x,t\| \le 2R\}$, and $T_R^* \equiv (0,2R)\times\big\{(x,t) : \frac{R}{2} \le \|(x,t)\| \le 4R\big\}$. But $\int_{T_R^*} w^2 \le \int_0^{2R} \int_{\|(x,t)\|\le 4R} |w|^2$, and we may repeat our previous argument with R in place of λ to obtain $R^{-d-1}\int_{T_R^*} w^2 = O(R^{-\epsilon})$. The proof of Lemma 8.9 is now complete.

With Lemma 8.9 in hand, we now return to the proof of Lemma 8.4 (the "localization lemma"). Again, the strategy here is that of [23]. Let $\theta \equiv \theta_{3r} \in C_0^\infty$, $\theta \equiv 1$ on $\overline{\psi}_{3r}$, $\theta \equiv 0$ outside of $(-4r,4r)\times\Delta_{4r}$. By our *a priori* assumptions of smoothness, we may use Green's Theorem to write

$$\begin{aligned}
\theta u_0(X,t) &\equiv -\iint_{\mathbb{R}^{n+1}_+} L_0^* N(X,t,\cdot,\cdot)\theta u_0 \\
&= -\int_{\partial\mathbb{R}^{n+1}_+} \partial_\nu N(X,t,\cdot)\theta u_0 \\
&\quad + \int_{\partial\mathbb{R}^{n+1}_+} N(X,t,\cdot)\frac{\partial\theta}{\partial\nu}u_0 + \int_{\partial\mathbb{R}^{n+1}_+} N(X,t,\cdot)\theta\frac{\partial u_0}{\partial\nu} \\
&\quad - \iint_{\mathbb{R}^{n+1}_+} N(X,t,\cdot)L_0(\theta u_0) \\
&\equiv I + II + III + IV.
\end{aligned}$$

The term I is O, since $\frac{\partial N}{\partial\nu}\big|_{\partial\mathbb{R}^{n+1}_+} \equiv 0$ by construction. The term III solves the Neumann problem with data $\theta\frac{\partial u_0}{\partial\nu}$, so the desired bound for this term follows immediately from (3.8). Since $L_0 u_0 = 0$,

$$\begin{aligned}
IV &= -\iint N(X,t,\cdot)(L_0\theta)u_0 - 2\iint N(X,t,\cdot)a\nabla\theta\cdot\nabla u_0 \\
&\equiv IV_1 + IV_2.
\end{aligned}$$

It is therefore enough to show that

$$\tag{8.12} \begin{aligned}
\left|\int_{\partial\mathbb{R}^{n+1}_+} DgIV_2\right| &\le C(\beta_0,n)\left(\frac{1}{r}\iint_{\psi_{4r}} |\nabla u_0|^2\right)^{\frac12} \\
\left|\int_{\partial\mathbb{R}^{n+1}_+} Dg(IV_1 + II)\right| &\le C(\beta_0,n)\left[\left(\frac{1}{r}\iint_{\psi_{4r}} |\nabla u_0|^2\right)^{\frac12} + \left(\frac{1}{r^3}\iint_{\psi_{4r}} |u_0|^2\right)^{\frac12}\right]
\end{aligned}$$

where D denotes either $\mathbb{D}_n$ or $\frac{\partial}{\partial x_j}$, $1 \le j \le n-1$, $g \in C_0^\infty$, $\operatorname{supp} g \subseteq \Delta_r$, and $\|g\|_2 = 1$.

The first term in (8.12) equals two times the absolute value of

$$\tag{8.13} \begin{aligned}
&\int_{\partial\mathbb{R}^{n+1}_+} Dg(z)\iint_{\mathbb{R}^{n+1}_+} N(z,Y,s)a(Y,s)\nabla\theta(Y,s)\cdot\nabla u_0(Y,s)dY\,ds\,dz \\
&\quad = \iint_{\mathbb{R}^{n+1}_+} a\nabla\theta\cdot\nabla u_0\tilde{N}(Dg)dY\,ds,
\end{aligned}$$

where, by slight abuse of notation, we shall use $\tilde{N}$ to denote

$$\tilde{N}(X,t,Y,s) \equiv N(Y,s,X,t),$$

and also the operator

$$f \to \int \tilde{N}(Y,s,z)f(z)dz$$
$$\equiv \int N(z,Y,s)f(z)dz.$$

We now claim that

$$\|\tilde{N}(Dg)\|_{L^2(\partial\mathbb{R}^{n+1}_+)} \le C,$$

where C depends only on ellipticity (and hence only on β_0 (see (1.2)), and on dimension.
Proof of the Claim. Let $\eta \in L^2(\partial\mathbb{R}^{n+1}_+)$, $\|\eta\|_2 = 1$, $\eta \in C_0^\infty$. Then

$$\begin{aligned}\int_{\mathbb{R}^n} \eta(v)\tilde{N}(Dg)(v)dv &= \int \left(\int N(z,v)\eta(v)dv\right) Dgdz \\ &= -\int D\left(\int N(z,v)\eta(v)dv\right) g(z)dz \\ &\equiv -\int Dw_0 g,\end{aligned}$$

where $L_0 w_0 = 0$, $\frac{\partial w_0}{\partial \nu} = \eta$. By (3.8), the last expression is dominated in absolute value by $C(\beta_0,n)\|\eta\|_2\|g\|_2$, and the claim follows. Moreover since $L_0^*\tilde{N}(Dg) = 0$, Lemma 5.1 implies that

$$\|N_*(\tilde{N}(Dg))\|_2 \le C(\beta_0,n), \tag{8.14}$$

where we recall that N_* denotes the non-tangential maximal function. Consequently, since $|\nabla\theta| \le \frac{c}{r}$, the right hand side of (8.13) is dominated in absolute value by

$$\begin{aligned}&C(\beta_0,n)\left(\frac{1}{r}\iint_{\psi_{4r}} |\nabla u_0|^2\right)^{\frac{1}{2}} \left(\frac{1}{r}\int_0^{4r}\int_{\mathbb{R}^n} |N_*(\tilde{N}(Dg))|^2\right)^{\frac{1}{2}} \\ &\quad \le C(\beta_0,n)\left(\frac{1}{r}\iint_{\psi_{4r}} |\nabla u_0|^2\right)^{\frac{1}{2}},\end{aligned}$$

as desired.

We now turn to the second term in (8.12). Notice that II equals

$$\begin{aligned}\iint \operatorname{div}(a\nabla\theta u_0 N(X,t,\cdot)) &= \iint (\operatorname{div} a\nabla\theta)u_0 N(X,t,\cdot) \\ &\quad + \iint a\nabla\theta\cdot\nabla N(X,t,\cdot)u_0 \\ &\quad + \iint a\nabla\theta\cdot\nabla u_0 N(X,t,\cdot) \\ &\equiv II_1 + II_2 + II_3.\end{aligned}$$

But $II_3 = \frac{-1}{2} IV_2$, which we have already treated. Also $IV_1 + II_1 =$

$$\iint N(X,t,\cdot)\frac{\partial\theta}{\partial t}u_0,$$

which can also be handled like IV_2, since $\left|\frac{\partial\theta}{\partial t}\right| \leq \frac{c}{r^2}$, except that now one obtains the bound $C\left(\frac{1}{r^3}\iint_{\psi_{4r}}|u_0|^2\right)^{\frac{1}{2}}$. Thus, it remains only to consider

$$\begin{aligned}\left|\int_{\mathbb{R}^n} DgII_2\right| &= \left|\iint u_0 a\nabla\theta\cdot\nabla\tilde{N}(Dg)\right| \\ &\leq C(\beta_0,n)\frac{1}{r}\left(\iint_{\psi_{4r}}|u_0|^2\right)^{\frac{1}{2}}\left(\iint_{\operatorname{supp}\theta\cap\mathbb{R}^{n+1}_+}|\nabla\tilde{N}(Dg)|^2\right)^{\frac{1}{2}}.\end{aligned} \tag{8.15}$$

But $\operatorname{supp}\theta\cap\mathbb{R}^{n+1}_+ \subseteq \psi_{4r}\backslash\psi_{3r}$, and $\operatorname{supp} g \subseteq \Delta_r$. Thus, if $D = \frac{\partial}{\partial x_j}$, then by Lemma 8.9, $\tilde{N}\left(\frac{\partial g}{\partial x_j}\right)$ is the solution of the adjoint Neumann problem with data $\frac{\partial g}{\partial x_j}$, which is supported in Δ_r. We can therefore make an even reflection across the boundary outside of Δ_{2r}, to obtain a Cacciopoli inequality at the boundary which implies that

$$\begin{aligned}\left(\iint_{\psi_{4r}\backslash\psi_{3r}}\left|\nabla\tilde{N}\left(\frac{\partial g}{\partial x_j}\right)\right|^2\right)^{\frac{1}{2}} &\leq \frac{C}{r^{\frac{1}{2}}}\left(\frac{1}{r}\int_0^{8r}\int_{\Delta_{5r}}\left|\tilde{N}\left(\frac{\partial g}{\partial x_j}\right)\right|^2\right)^{\frac{1}{2}} \\ &\leq C(\beta_0,n)r^{-\frac{1}{2}},\end{aligned}$$

where we have used (8.14) to obtain the last inequality. Hence (8.15) is no larger than $C(\beta_0,n)\left(\frac{1}{r^3}\iint_{\psi_{4r}}|u_0|^2\right)^{\frac{1}{2}}$ as desired.

If $D \equiv \mathbb{D}_n$, then we need to modify our treatment of (8.15). In this case, let us write

$$\begin{aligned}\mathbb{D}_n g(z) &\equiv p.v.\int_{\|z-v\|\leq r/4}\mathbb{D}_n(z-v)g(v)dv + \int_{\|z-v\|>r/4}\mathbb{D}_n(z-v)g(v)dv, \\ &\equiv D_1g(z) + D_2g(z).\end{aligned}$$

where, by [12], the kernel $\mathbb{D}_n(z)$ satisfies

$$\begin{gathered}\mathbb{D}_n(\lambda x,\lambda^2 t) = \lambda^{-d-1}\mathbb{D}_n(x,t) \\ \mathbb{D}_n(z) \in C^\infty(\mathbb{R}^n\backslash\{0\}).\end{gathered}$$

In particular $|\mathbb{D}_n(z)| \leq C\|z\|^{-d-1}$, so that

$$|D_2 g| \leq \frac{c}{r}Mg. \tag{8.16}$$

where M denotes the parabolic Hardy-Littlewood maximal operator. Now, $\operatorname{supp} D_1 g \subseteq \Delta_{\frac{5r}{4}}$, so that we may apply Cacciopoli's inequality as above, to obtain that

$$\left|\int_{\mathbb{R}^n} D_1 gII_2\right| \leq C(\beta_0,n)\left(\frac{1}{r^3}\iint_{\psi_{4r}}|u_0|^2\right)^{\frac{1}{2}}\left(\frac{1}{r}\int_0^{8r}\int_{\Delta_{5r}}|\tilde{N}(D_1g)+C_0|^2\right)^{\frac{1}{2}},$$

for any constant C_0. It then suffices to show that, for an appropriate choice of C_0, we have

$$\frac{1}{r}\int_0^{8r}\int_{\Delta_{5r}} |\tilde{N}(D_1 g) + C_0|^2 \le C(\beta_0, n). \tag{8.17}$$

We write $\tilde{N}(D_1 g) + C_0 = \tilde{N}(Dg) - (\tilde{N}(D_2 g) - C_0)$, and note that by (8.14), we obtain immediately that (8.17) holds for the contribution of $\tilde{N}(Dg)$. Thus, by R. Brown's Poincaré inequality for solutions of $L_0 u_0 = 0$ [3, (3.9)], it is enough to show that

$$r\int_0^{8r}\int_{\Delta_{5r}} |\nabla \tilde{N}(D_2 g)|^2 \le C(\beta_0, n). \tag{8.18}$$

Moreover, we note that (8.18) will imply the conclusion of Lemma 8.4, because it also implies the desired bound for $\int_{\mathbb{R}^n} D_2 g II_2$ (see (8.15)). Let us denote by u the function $\tilde{N}(D_2 g)$, which is a solution of $L_0 u = 0$. Indeed, this u is the variational solution to the Neumann problem with data $D_2 g$, constructed in Section 7. Thus, by Lemma 7.8 and (8.16), we have that

$$\begin{aligned}\left(r\int_0^{8r}\int_{\Delta_{5r}} |\nabla u|^2\right)^{\frac{1}{2}} &\le C r^{\frac{1}{2}} \|u\|_H \\ &\le C r^{\frac{1}{2}} \|D_2 g\|_q \\ &\le C r^{-\frac{1}{2}} \|g\|_q \\ &\le C\|g\|_2 = C,\end{aligned}$$

as desired, since $\frac{1}{2} = \frac{1}{q} - \frac{1}{2d}$. This concludes the Proof of Lemma 8.4.

To conclude this section, it remains to give the Proof of Lemma 8.3. At this point we shall deviate from the approach of [23]. The solvability of the Dirichlet problem (Lemma 5.1) permits a significant simplification in our case. We set

$$(X,t) \equiv (\lambda, x, t), \quad (Y,s) \equiv (\rho, y, s),$$

and recall that, by hypothesis, $\rho \le \frac{1}{2}\|(x,t) - (y,s)\|$, $\lambda \le 2\|(x,t) - (y,s)\|$. Thus,

$$\|(X,t) - (Y,s)\| \approx \|(x,t) - (y,s)\| \equiv r,$$

and $(Y,s) \equiv (\rho, y, s) \in \Psi_{\frac{3r}{4}}(y,s) \equiv \Psi_{\frac{3r}{4}}$. Since $\frac{\partial N}{\partial \nu}$ vanishes on $\partial\mathbb{R}^{n+1}_+$, we may extend N by even reflection across the boundary. Thus, by the estimates of Moser,

(8.19)

$$\begin{aligned}|N(X,t,Y,s)| &\le C\left(r^{-d-1}\iint_{\Psi_{\frac{7r}{8}}(y,s)} |N(X,t,\tilde{Y},\tilde{s})|^2 d\tilde{Y} d\tilde{s}\right)^{\frac{1}{2}} \\ &\le C\left(r^{-d-1}\iint_{\Psi_{\frac{7r}{8}}(y,s)} |V_{X,t}|^2\right)^{\frac{1}{2}} + C\left(r^{-d-1}\iint_{\Psi_{\frac{7r}{8}}(y,s)} |G(X,t,\tilde{Y},\tilde{s})|^2 d\tilde{Y} d\tilde{s}|^2\right)^{\frac{1}{2}}.\end{aligned}$$

Now, by Aronson's estimates for G, [1], the second term is dominated by $C(\beta_0,n)r^{1-d}$, as desired. As regards the first term, we are using the notation that $V_{X,t}$ is the variational solution to the Neumann problem for L_0^*, with data $\frac{\partial G(X,t,\cdot)}{\partial\nu}$. Thus, by Lemma 7.9 (iii),

$$r^{-d-1}\iint_{\Psi_{\frac{7r}{8}}(y,s)} |V_{X,t}|^2 \le Cr^{1-d}\|V_{X,t}\|_H^2.$$

But by Lemma 7.8 and Lemma 5.1,

$$\|V_{X,t}\|_H \le C\left\|\frac{\partial G}{\partial\nu}\right\|_q \le C(\beta_0,n),$$

and the conclusion of Lemma 8.3 follows.

9 Proof of the "Main Lemma"

We begin by stating some preliminary facts and estimates. Recall that $u_1 \equiv u\circ\rho$, where $\rho(\lambda,x,t)$ is defined in (3.1), and u solves the heat equation in Ω. Thus

$$\begin{aligned}
\frac{\partial u_1}{\partial t} &\equiv \frac{\partial u}{\partial t}\circ\rho + \left(\frac{\partial u}{\partial x_0}\circ\rho\right)\frac{\partial}{\partial t}P_{\gamma\lambda}A \\
\frac{\partial u_1}{\partial\lambda} &\equiv \left(\frac{\partial u}{\partial x_0}\circ\rho\right)\left(1+\frac{\partial}{\partial\lambda}P_{\gamma\lambda}A\right) \\
\frac{\partial^2 u_1}{\partial\lambda^2} &\equiv \left(\frac{\partial^2 u}{\partial x_0^2}\circ\rho\right)\left(1+\frac{\partial}{\partial\lambda}P_{\gamma\lambda}A\right)^2 + \left(\frac{\partial u}{\partial x_0}\circ\rho\right)\left(\frac{\partial^2 P_{\gamma\lambda}A}{\partial\lambda^2}\right) \\
\frac{\partial u_1}{\partial x_j} &\equiv \frac{\partial u}{\partial x_j}\circ\rho + \left(\frac{\partial u}{\partial x_0}\circ\rho\right)\frac{\partial P_{\gamma\lambda}A}{\partial x_j} \\
\frac{\partial^2 u_1}{\partial\lambda\partial x_j} &\equiv \left(\frac{\partial^2 u}{\partial x_0\partial x_j}\circ\rho\right)\left(1+\frac{\partial P_{\gamma\lambda}A}{\partial\lambda}\right) + \left(\frac{\partial^2 u}{\partial x_0^2}\circ\rho\right)\left(\frac{\partial P_{\gamma\lambda}A}{\partial\lambda}\right)\left(1+\frac{\partial}{\partial\lambda}P_{\gamma\lambda}A\right) \\
&\quad + \left(\frac{\partial u}{\partial x_0}\circ\rho\right)\left(\frac{\partial^2 P_{\gamma\lambda}A}{\partial\lambda\partial x_j}\right) \\
\frac{\partial^2 u_1}{\partial x_j\partial x_k} &= \frac{\partial^2 u}{\partial x_j\partial x_k}\circ\rho + \left(\frac{\partial^2 u}{\partial x_j\partial x_0}\circ\rho\right)\left(\frac{\partial}{\partial x_k}P_{\gamma\lambda}A\right) + \left(\frac{\partial^2 u}{\partial x_0\partial x_k}\circ\rho\right)\left(\frac{\partial}{\partial x_j}P_{\gamma\lambda}A\right) \\
&\quad + \left(\frac{\partial^2 u}{\partial x_0^2}\circ\rho\right)\left(\frac{\partial P_{\gamma\lambda}A}{\partial x_j}\right)\left(\frac{\partial P_{\gamma\lambda}A}{\partial x_k}\right) + \left(\frac{\partial u}{\partial x_0}\circ\rho\right)\left(\frac{\partial^2 P_{\gamma\lambda}A}{\partial x_j\partial x_k}\right).
\end{aligned} \tag{9.1}$$

Moreover, the following Carleson measure and pointwise estimates hold. The unweighted case is [18, Lemma 2.8]. The extension to the weighted case is routine (see the appendix to [17]). The weighted Carleson measure estimates will be useful in the proof of certain L^p inequalities, $p\neq 2$.

Lemma 9.2. *For every parabolic A_∞ weight w, there exists a constant C depending only on n, β_0, and the A_∞ constant of w such that for all parabolic balls Δ_r, $0 < r < \infty$, and for $\beta_1 \equiv \|\mathbb{D}_n A\|_* \le \epsilon_0 \le \gamma^{M_0}$, we have*

$$(i) \qquad \frac{1}{w(\Delta_r)} \int_0^r \int_{\Delta_r} \left| \frac{\partial P_{\gamma\lambda} A}{\partial t} \right|^2 \lambda w(z) dz d\lambda \le C\gamma^{2(M_0-1)}$$

$$(ii) \qquad \frac{1}{w(\Delta_r)} \int_0^r \int_{\Delta_r} \left(\left| \frac{\partial P_{\gamma\lambda} A}{\partial \lambda} \right|^2 + \left| \lambda \frac{\partial^2 P_{\gamma\lambda} A}{\partial \lambda^2} \right|^2 \right) w(z) dz \frac{d\lambda}{\lambda} \le C\gamma^2$$

$$(iii) \qquad \frac{1}{w(\Delta_r)} \int_0^r \int_{\Delta_r} \left| \frac{\partial^2 P_{\gamma\lambda} A}{\partial \lambda \partial x_j} \right|^2 w(z) dz \lambda d\lambda \le C$$

$$(iv) \qquad \frac{1}{w(\Delta_r)} \int_0^r \int_{\Delta_r} \left| \frac{\partial^2 P_{\gamma\lambda} A}{\partial x_j \partial x_k} \right|^2 w(z) dz \lambda d\lambda \le C\gamma^{-2}$$

$$(v) \qquad \frac{1}{w(\Delta_r)} \int_0^r \int_{\Delta_r} \left(\left| \frac{\partial P_{\gamma\lambda} b}{\partial \lambda} \right|^2 + \left| \frac{1}{\lambda} \mathbb{D}^{-1} \frac{\partial P_{\gamma\lambda} b}{\partial \lambda} \right|^2 + \left| \mathbb{D}^{-1} \frac{\partial^2 P_{\gamma\lambda} b}{\partial \lambda^2} \right|^2 \right) w(z) dz \lambda d\lambda \le C\|b\|_*^2,$$

for all $b \in$ (parabolic) BMO. Moreover,

$$(vi) \qquad \left| \frac{\partial}{\partial \lambda} P_{\gamma\lambda} A \right| \le C(\beta_0, n)\gamma$$

$$(vii) \qquad \left| \frac{\partial^2 P_{\gamma\lambda} A}{\partial \lambda^2} \right| \le C(\beta_0, n)\frac{\gamma}{\lambda}$$

$$(viii) \qquad \left| \frac{\partial P_{\gamma\lambda} A}{\partial x_j} \right| \le \beta_0$$

$$(ix) \qquad \left| \frac{\partial^2 P_{\gamma\lambda} A}{\partial x_j \partial \lambda} \right| \le C(\beta_0, n)\lambda^{-1}$$

$$(x) \qquad \left| \frac{\partial^2 P_{\gamma\lambda} A}{\partial x_j \partial x_k} \right| \le C(\beta_0, n)(\gamma\lambda)^{-1}$$

$$(xi) \qquad \left| \mathbb{D}^{-1} \frac{\partial P_{\gamma\lambda} b}{\partial \lambda} \right| \le C(\beta_0, n)\|b\|_*$$

for $b \in$ parabolic BMO.

Next, we note the following elementary fact:

Lemma 9.3. *Suppose that for $w \in A_\infty$, we have the weighted Carleson measure estimate*

$$\frac{1}{w(\Delta_r)} \int_0^r \int_{\Delta_r} h(\lambda, z) w(z) dz d\lambda \le C(A_\infty).$$

Then, for all $1 < p < \infty$, we have, for all measurable F and $h \ge 0$ in $\mathbb{R}^{n+1}_+$,

$$(9.4) \qquad \left\| \left(\int_0^\infty |F(\lambda, z)|^2 h(\lambda, z) d\lambda \right)^{\frac{1}{2}} \right\|_{L^p(\mathbb{R}^n)} \le C_q \|N_*(F)\|_{L^p(\mathbb{R}^n)}.$$

Proof. We consider first the case $p \geq 2$ (the case $p = 2$ is of course well known). The left side of the claimed inequality is bounded by

$$\text{(9.5)} \qquad \sup \left(\int_0^\infty \int_{\mathbb{R}^n} |F(\lambda, z)|^2 h(\lambda, z) \varphi(z) dz d\lambda \right)^{\frac{1}{2}},$$

where the sup runs over all $\varphi \in L^{\left(\frac{p}{2}\right)'}$, $\|\varphi\|_{L^{(\frac{p}{2})'}} = 1$, $\varphi \geq 0$. But $\varphi \leq w \equiv (M(\varphi^{1+\delta})^{\frac{1}{1+\delta}} \in A_1$ (by a result of Coifman and Rochberg [6]), where we choose δ small enough, depending only on p so that $\|w\|_{L^{(\frac{p}{2})'}} \leq C$. Moreover, the A_1 constant of w depends only on p. We then replace φ by w in (9.5), and apply the weighted Carleson measure estimate and then Hölder's inequality to obtain that (9.4) is no larger than

$$C_p \left(\int_{\mathbb{R}^n} (N_*(F))^2 w(z) dz \right)^{\frac{1}{2}} \leq C_p \|N_*(F)\|_p,$$

as desired.

Next, we treat the case $1 < p < 2$. Let $\Lambda(F)(z) \equiv \left(\int_0^\infty |F(\lambda, z)|^2 h(\lambda, z) d\lambda \right)^{\frac{1}{2}}$. Then the left hand side of (9.4) equals

$$\text{(9.6)} \qquad \begin{aligned} \|\Lambda(F)\|_p &\leq \|\Lambda(F) + N_*(F)\|_p \\ &\leq \|M(\Lambda(F) + N_*(F))\|_p \\ &\equiv \left(\int_{\mathbb{R}^n} (M(\Lambda(F) + N_*(F)))^2(z) w(z) dz \right)^{\frac{1}{p}}, \end{aligned}$$

where M denotes the maximal function, and

$$w \equiv (M(\Lambda(F) + N_*(F)))^{p-2}.$$

Since $0 < 2 - p < 1$, we have that w is an A_2 weight (indeed, it is the reciprocal of an A_1 weight), and, in fact, its A_2 constant may be taken to depend only on p. Thus, we may apply Muckenhoupt's theorem to remove the maximal function, and then use the weighted Carleson measure estimate to replace $\Lambda(F)$ by $N_*(F)$, so that we may dominate (9.6) by

$$\begin{aligned} &C_q \left(\int_{\mathbb{R}^n} (N_*(F)(z))^2 w(z) dz \right)^{\frac{1}{p}} \\ &\leq C_q \|N_*(F)\|_p, \end{aligned}$$

as desired, since $N_*(F)^{2-p} w \leq 1$. This concludes the Proof of Lemma 9.3.

Next, combining (9.1), Lemma 9.2, Lemma 9.3, and the square function estimates of R. Brown [4] for caloric functions in $\text{Lip}_{1,\frac{1}{2}}$ domains (applied to the caloric functions $\frac{\partial u}{\partial x_j}$, $0 \leq j \leq n-1$), and using also the fact that by [26], the Euclidean measure in $\mathbb{R}^n$ and the "pullback" of caloric measure are A_∞ equivalent, we obtain

Lemma 9.7. *Let $1 < p < \infty$. Then*

$$\left\| \left(\int_0^\infty \left\{ \left| \frac{\partial u_1}{\partial t} \right|^2 + \left| \frac{\partial}{\partial \lambda} \nabla u_1 \right|^2 \right\} \lambda d\lambda \right)^{\frac{1}{2}} \right\|_p \leq C(\beta_0, n, p) \| N_*(\nabla u_1) \|_p.$$

$$\left\| \left(\int_0^\infty \left| \frac{\partial^2 u_1}{\partial x_j \partial x_k} \right|^2 \lambda d\lambda \right)^{\frac{1}{2}} \right\|_p \leq C(\beta_0, n, p)(1 + \gamma^{-1}) \| N_*(\nabla u_1) \|_p.$$

Moreover, in Section 10 we shall prove the following square function estimates for solutions of $L_0 u_0 = 0$, and $L_0^* u_0^* = 0$, where L_0^* denotes the adjoint of L_0.

Lemma 9.8. *Suppose that $L_0 u_0 = 0$, $L_0^* u_0^* = 0$ in $\mathbb{R}_+^{n+1}$. Let $q \geq 2$. Then there exists a number $N(q)$ such that*

$$\left\| \left(\int_0^\infty |\nabla u_0|^2 \lambda d\lambda \right)^{\frac{1}{2}} \right\|_q \leq C(\beta_0, n, q) \gamma^{-N(q)} \| N_*(u_0) \|_q.$$

In addition, the same estimate holds for u_0^* in place of u_0.

We defer the proof of Lemma 9.8 until the next section. Taking this lemma for granted, we shall now give the Proof of Lemma 6.4. The strategy of our proof will follow that of [24].

To begin the proof, let $1 < p \leq 2$, and let $\frac{1}{p} + \frac{1}{q} = 1$. Let $g \in C_0^\infty$, $\|g\|_{L^q(\mathbb{R}^n)} = 1$. Following [24], we consider

$$\begin{aligned} \int_{\mathbb{R}^n} g \frac{\partial F}{\partial \nu} &= \iint_{\mathbb{R}_+^{n+1}} \frac{\partial P_{\gamma\lambda} A}{\partial t} \frac{\partial u_1}{\partial \lambda} \int_{\mathbb{R}^n} \tilde{K}^{Y,s}(z) g(z) dz dY ds \\ &\quad - \iint_{\mathbb{R}_+^{n+1}} \frac{\partial P_{\gamma\lambda} A}{\partial \lambda} \frac{\partial u_1}{\partial t} \int_{\mathbb{R}^n} \tilde{K}^{Y,s}(z) g(z) dz dY ds \\ &\equiv I + II, \end{aligned} \tag{9.9}$$

where $\tilde{K}^{Y,s} \equiv \frac{\partial G}{\partial \nu}(\cdot, Y, s)$, and $(Y, s) \equiv (\lambda, y, s) \equiv (\lambda, v)$. Set $u_0^* \equiv \int_{\mathbb{R}^n} \tilde{K}^{(\cdot)} g$, so that $L_0^* u_0^* = 0$, $u_0^*|_{\mathbb{R}^n} \equiv g$. Then $|II|$ is no larger than

$$\begin{aligned} &\left\| \left(\int_0^\infty \left| \frac{\partial u_1}{\partial t} \right|^2 \lambda d\lambda \right)^{\frac{1}{2}} \right\|_p \left\| \left(\int \left| \frac{\partial P_{\gamma\lambda} A}{\partial \lambda} \right|^2 |u_0^*|^2 \frac{d\lambda}{\lambda} \right)^{\frac{1}{2}} \right\|_q \\ &\leq C(\beta_0, n, p) \| N_*(\nabla u_1) \|_p \| N_*(u_0^*) \|_q \gamma, \end{aligned}$$

by Lemmas 9.7, 9.2(ii), and 9.3. But $q \geq 2$, so by Lemma 5.1 we may majorize $\|N_*(u_0^*)\|_q$ by $C(\beta_0, n, q)\|g\|_q = C(\beta_0, n)$. This establishes the desired bound for II. To handle I, we integrate by parts in λ to obtain that

$$\begin{aligned} I &\equiv - \iint_{\mathbb{R}_+^{n+1}} \frac{\partial^2 u_1}{\partial_\lambda^2} \frac{\partial P_{\gamma\lambda} A}{\partial t} u_0^* \lambda d\lambda dv \\ &\quad - \iint_{\mathbb{R}_+^{n+1}} \frac{\partial u_1}{\partial \lambda} \frac{\partial P_{\gamma\lambda} A}{\partial t} \frac{\partial u_0^*}{\partial \lambda} \lambda d\lambda dv - \iint_{\mathbb{R}_+^{n+1}} \frac{\partial u_1}{\partial \lambda} \frac{\partial^2 P_{\gamma\lambda} A}{\partial \lambda dt} u_0^* \lambda d\lambda dv \\ &\equiv I_1 + I_2 + I_3 \end{aligned}$$

(we have used here that the "boundary terms" as $\lambda \to +\infty$ are zero). We can handle I_1 exactly like II, using now part (i) of Lemma 9.2. A similar argument, involving Lemmas 9.8, 9.3 and 9.2(i) (as well as Lemma 5.1) yields the bound

$$|I_2| \le C(\beta_0, n, p)\gamma^{M_0-N(q)-1}\|N_*(\nabla u)\|_p,$$

which again suffices, if we take M_0 large enough, depending on q.

It therefore remains to consider I_3. We recall that

$$\frac{\partial}{\partial t} = \mathbb{D}\circ\mathbb{D}_n,$$

where $\mathbb{D}g \equiv (\|\xi\|\hat{g}(\xi))^{\vee}$, and $\mathbb{D}_n \equiv \mathbb{D}^{-1}\circ\frac{\partial}{\partial t}$. Furthermore,

$$\mathbb{D} \equiv -\sum_{j=1}^{n-1}\frac{\partial}{\partial x_j}R_j - \mathbb{D}_n R_n,$$

where the parabolic Riesz transforms are given by

$$\begin{aligned} R_j &\equiv \mathbb{D}^{-1}\circ\frac{\partial}{\partial x_j}, \quad 1\le j\le n-1 \\ R_n &\equiv \mathbb{D}^{-2}\circ\frac{\partial}{\partial t}. \end{aligned}$$

We recall that each R_j, $1 \le j \le n$, defines a bounded operator on $L^p(\mathbb{R}^n)$, and on parabolic BMO. We can now write

$$I_3 = \sum_{j=1}^{n}\iint \frac{\partial u_1}{\partial\lambda}\mathbb{D}_j\frac{\partial P_{\gamma\lambda}R_j\mathbb{D}_nA}{\partial\lambda}u_0^*\lambda d\lambda dv,$$

where $\mathbb{D}_j \equiv \frac{\partial}{\partial x_j}$, $1 \le j \le n-1$, and $\mathbb{D}_n$ is defined as above. Consider first the case $1 \le j \le n-1$. In this case we integrate by parts to move $\mathbb{D}_j$ onto either $\frac{\partial u_1}{\partial\lambda}$ or onto u_0^*. If the former occurs, we apply Lemmas 9.7, 9.3, 9.2 (v) with $b \equiv R_j\mathbb{D}_nA$, and 5.1 to obtain the bound $C(\beta_0, n, p)\|N_*(\nabla u_1)\|_p\gamma^{M_0}\|N_*u_0^*\|_q \le$

$$C(\beta_0, n, p)\|N_*(\nabla u_1)\|_p\gamma^{M_0},$$

as desired. If $\mathbb{D}_j$ hits u_0^*, we apply Lemmas 9.8, 9.3, 9.2(v) and 5.1 to obtain essentially the same bound, except with $\gamma^{M_0-N(q)}$, which still suffices, with M_0 large enough depending on q. The case $j = n$ is a bit more complicated. In this case, recalling that $\mathbb{D}_n \equiv \frac{\partial}{\partial t}\circ\mathbb{D}^{-1}$ we need to consider

$$\begin{aligned} &\iint \frac{\partial u_1}{\partial\lambda}\frac{\partial}{\partial t}\mathbb{D}^{-1}\frac{\partial P_{\gamma\lambda}R_n\mathbb{D}_nA}{\partial\lambda}u_0^*\lambda d\lambda dv \\ &= -\iint\frac{\partial^2 u_1}{\partial t\partial\lambda}\mathbb{D}^{-1}\frac{\partial P_{\gamma\lambda}R_n\mathbb{D}_nA}{\partial\lambda}u_0^*\lambda d\lambda dv \\ &\quad -\iint\frac{\partial u_1}{\partial\lambda}\mathbb{D}^{-1}\frac{\partial P_{\gamma\lambda}R_n\mathbb{D}_nA}{\partial\lambda}\frac{\partial u_0^*}{\partial t}\lambda d\lambda dv \\ &\equiv I_{31} + I_{32}. \end{aligned} \tag{9.10}$$

Integrating by parts in λ, we obtain that

$$\begin{aligned}
I_{31} = &\iint \frac{\partial u_1}{\partial t}\mathbb{D}^{-1}\frac{\partial^2 P_{\gamma\lambda}R_n\mathbb{D}_n A}{\partial\lambda^2}u_0^*\lambda d\lambda dv \\
&+ \iint \frac{\partial u_1}{\partial t}\mathbb{D}^{-1}\frac{\partial P_{\gamma\lambda}R_n\mathbb{D}_n A}{\partial\lambda}\frac{\partial u_0^*}{\partial\lambda}\lambda d\lambda dv \\
&+ \iint \frac{\partial u_1}{\partial t}\mathbb{D}^{-1}\frac{\partial P_{\gamma\lambda}R_n\mathbb{D}_n A}{\partial\lambda}u_0^* d\lambda dv \\
&\equiv I'_{31} + I''_{31} + I'''_{31}.
\end{aligned}$$

The desired bounds for I'_{31} and I'''_{31} follow immediately from Lemmas 9.2(v), 9.3, 9.7, and 5.1. Moreover, we can handle I''_{31} by Lemmas 9.2(xi), 9.7, 9.8 and 5.1.

To handle the term I_{32} in (9.10), we use that $\frac{\partial u_0^*}{\partial t} = -\operatorname{div}(a\nabla u_0^*)$, and then integrate by parts in x, λ, so that (recalling that $\nabla \equiv \nabla_{x,\lambda}$) we have,

$$\begin{aligned}
I_{32} = &-\iint \nabla\frac{\partial u_1}{\partial\lambda}\cdot a\nabla u_0^*\mathbb{D}^{-1}\frac{\partial P_{\gamma\lambda}R_n\mathbb{D}_n A}{\partial\lambda}\lambda d\lambda dv \\
&- \iint \frac{\partial u_1}{\partial\lambda}a\nabla u_0^*\cdot\nabla\mathbb{D}^{-1}\frac{\partial P_{\gamma\lambda}R_n\mathbb{D}_n A}{\partial\lambda}\lambda d\lambda dv \\
&- \iint \frac{\partial u_1}{\partial\lambda}\langle e_0, a\nabla u_0^*\rangle\frac{\mathbb{D}^{-1}\partial P_{\gamma\lambda}R_n\mathbb{D}_n A}{\partial\lambda}d\lambda dv \\
&\equiv I'_{32} + I''_{32} + I'''_{32}.
\end{aligned}$$

By Lemmas 9.7, 9.8, and 9.2(xi), and 5.1, we have that

$$|I'_{32}| \le C(n,\beta_0,p)\gamma^{M_0-N(q)}\|N_*(\nabla u_1)\|_p,$$

and similar estimates hold for I''_{32}, I'''_{32} by Lemmas 9.8, 9.3, 9.2(v), 5.1 and the fact that $\frac{\partial}{\partial x_j}\mathbb{D}^{-1}$ is a Riesz transform. Modulo the Proof of Lemma 9.8, this concludes the Proof of Lemma 6.4 and thus also that of Theorem 1.8.

10 Appendix: Proof of square function estimates (Lemma 9.8)

For reasons which shall become apparent, it will be convenient to consider a more general class of operators, of which L_0 is a special case. We consider the class of operators of the type

$$Lu \equiv \operatorname{div}(a(X,t)\nabla u) - b\frac{\partial u}{\partial t} + \vec{B}\cdot\nabla u, \tag{10.1}$$

$(X,t) \equiv (\lambda, x, t) \in \mathbb{R}^{n+1}_+$, where $a, b \in C^1(\mathbb{R}^{n+1}_+)$, the coefficient matrix a is real symmetric, and

$$\text{(10.2)} \qquad (i) \quad \nu|\xi|^2 \leq \langle a\xi, \xi\rangle \leq \frac{1}{\nu}|\xi|^2$$

$$(ii) \qquad \nu \leq b \leq \frac{1}{\nu}$$

$$\text{(10.3)} \qquad (i)\ |\nabla_{\lambda,x} a| + |\nabla_{\lambda,x} b| + |\vec{B}| \leq \sqrt{\mu}/\lambda$$

$$(ii) \quad \left|\frac{\partial a}{\partial t}\right| + \left|\frac{\partial b}{\partial t}\right| \leq \sqrt{\mu}/\lambda^2$$

$$\text{(10.4)} \qquad \sup_{\Delta_r} \int_0^r r^{-d} \int_{\Delta_r} (|\nabla_{\lambda,x} a|^2 + |\nabla_{\lambda,x} b|^2 + |\vec{B}|^2) dx dt \lambda d\lambda \leq \mu$$

$$\text{(10.5)} \qquad \sup_{\Delta_r} \int_0^r r^{-d} \int_{\Delta_r} \left(\left|\frac{\partial a}{\partial t}\right|^2 + \left|\frac{\partial b}{\partial t}\right|^2 \right) d\lambda dt \lambda^3 d\lambda \leq \mu.$$

We observe that L_0 belongs to this class, with $b \equiv 1$, $\vec{B} \equiv 0$, $\nu = \nu(\beta_0)$, and $\mu \approx \gamma^{-2}$. Consequently, the case $q = 2$ of Lemma 9.8 is an immediate corollary of the following result; which may be of some independent interest:

Lemma 10.6. *Let L be as in* (10.1)-(10.5)*, and suppose that $Lu = 0$ in $\mathbb{R}^{n+1}_+ \equiv \{(\lambda, x, t) \in \mathbb{R} \times \mathbb{R}^{n-1} \times \mathbb{R} : \lambda > 0\}$. Then*

$$(i) \quad \iint_{\mathbb{R}^{n+1}_+} |\nabla u|^2 \lambda d\lambda dz \leq C(n, \nu)(1 + \mu^2) \int_{\mathbb{R}^n} |N_* u|^2$$

$$(ii) \quad \iint_{\mathbb{R}^{n+1}_+} \left|\frac{\partial u}{\partial t}\right|^2 \lambda^3 d\lambda dz \leq C(n, \nu)(1 + \mu) \left[\int_{\mathbb{R}^n} (N_* u)^2 + \iint_{\mathbb{R}^{n+1}_+} |\nabla u|^2 \lambda d\lambda dz \right].$$

Remark. Operators where coefficients satisfy (10.2)-(10.5) have been studied previously in [20], but with an additional technical condition on the transverse (i.e. λ) derivatives of the coefficients. This technical hypothesis is actually satisfied by our pullback operators, but it is somewhat complicated and less natural in appearance, so we prefer to avoid it here. We also note that the elliptic analogue of Lemma 10.6, without the extra technical condition of [20], has been obtained recently by Kenig and Pipher [25], and it was their work which prompted us to attempt to dispense with this extra hypothesis, in our present formulation of Lemma 10.6.

Proof of Lemma 10.6. We claim that without loss of generality, we may assume that $a_{00} \equiv 1$. Indeed, set $\tilde{a} \equiv \frac{1}{a_{00}} a$, where $a \equiv (a_{ij})_{0 \leq i,j \leq n-1}$. Thus $\tilde{a}_{00} \equiv 1$. Set

$$\tilde{L}u \equiv \operatorname{div} \tilde{a} \nabla u - \frac{b}{a_{00}} \frac{\partial u}{\partial t} + \frac{\vec{B}}{a_{00}} \cdot \nabla u - \nabla \left(\frac{1}{a_{00}}\right) a \cdot \nabla u.$$

We note that $Lu = 0 \Rightarrow \tilde{L}u = 0$. Then $\tilde{L}$ also satisfies (10.1)-(10.5), with constants $\tilde{\nu}$ depending only on ν, and $\tilde{\mu} \leq C(\nu)\mu$. Thus, we may assume that $a_{00} \equiv 1$.

We shall begin by proving estimate (ii) of Lemma 10.6, which is an auxiliary step in the proof of the main estimate (i). Actually, to make the argument rigorous, we shall prove a local version of estimate (ii) (and estimate (i)). We shall need these local versions in the sequel, in any case, to prove the case $q > 2$ in Lemma 9.8. By passing to a limit, we then obtain the conclusion of Lemma 10.6 from the local estimates. To be more precise, fix $\Delta_r = \Delta_r(z^1) \subseteq \mathbb{R}^n$. Set $\psi_r \equiv (0,r) \times \Delta_r$. Let $\theta_r \in C_0^\infty(\mathbb{R}^{n+1})$, $\theta_r \equiv 1$ on $(-r,r) \times \Delta_r$, $\operatorname{supp} \theta_r \subseteq \left(\frac{-3r}{2}, \frac{3r}{2}\right) \times \Delta_{\frac{3r}{2}}$, with $|\nabla \theta_r| \leq \frac{c}{r}$, $\left|\frac{\partial \theta_r}{\partial t}\right| \leq c/r^2$. Define

$$N_*^r(F) \equiv N_*(F\chi_{\{\lambda<r\}}).$$

We have the following

Lemma 10.7. *Let θ_r, N_*^r be defined as above, and let u be as in Lemma 10.6. Then*

$$(i) \quad \iint_{\mathbb{R}_+^{n+1}} |\nabla u|^2 \theta_r^2 \lambda d\lambda dz \leq C(n,\nu)(1+\mu^2) \int_{\Delta_{5r}} (N_*^{5r}(u))^2$$

$$(ii) \quad \iint_{\mathbb{R}_+^{n+1}} \left|\frac{\partial u}{\partial t}\right|^2 \theta_r^2 \lambda^3 d\lambda dz \leq C(n,\nu)(1+\mu) \left[\int_{\Delta_{5r}} (N_*^{5r} u)^2 + \iint |\nabla u|^2 \theta_r^2 \lambda d\lambda dz\right].$$

We note that Lemma 10.6 follows immediately from Lemma 10.7, by letting $r \to \infty$. Thus, it is enough to prove Lemma 10.7. We begin by proving (ii). By (10.2)(ii), we have that

$$\begin{aligned}
\iint_{\mathbb{R}_+^{n+1}} \left|\frac{\partial u}{\partial t}\right|^2 \lambda^3 \theta_r^2 d\lambda dz &\leq C(\nu) \iint \frac{\partial u}{\partial t} b \frac{\partial u}{\partial t} \theta_r^2 \lambda^3 d\lambda dz \\
&= C(\nu) \iint \frac{\partial u}{\partial t} \operatorname{div}(a \nabla u) \theta_r^2 \lambda^3 d\lambda dz \\
&\quad + C(\nu) \iint \frac{\partial u}{\partial t} \vec{B} \cdot \nabla u \theta_r^2 \lambda^3 d\lambda dz
\end{aligned} \tag{10.8}$$

The second of these summands is dominated by

$$C(\nu)\left(\epsilon \iint \left|\frac{\partial u}{\partial t}\right|^2 \lambda^3 \theta_r^2 d\lambda dz + \frac{\mu}{\epsilon} \iint |\nabla u|^2 \theta_r^2 \lambda d\lambda dz\right),$$

where ϵ is at our disposal, and where we have used (10.3)(i) to control the size of $|\vec{B}|^2$. Choosing ϵ small enough, depending only ν, we may hide the small term on the left side of (10.8), to obtain the upper bound $C(\nu)\mu \iint |\nabla u|^2 \theta_r^2 \lambda d\lambda dz$, as desired. It remains to consider the first summand on the right hand side of (10.8). Integrating by parts, we obtain that this term equals $C(\nu)$ times

$$\begin{aligned}
&- \iint \nabla \frac{\partial u}{\partial t} \cdot a \nabla u \theta_r^2 dz \lambda^3 d\lambda - 3 \iint \frac{\partial u}{\partial t} \langle e_0, a \nabla u \rangle \lambda^2 \theta_r^2 d\lambda dz \\
&\quad - \iint \frac{\partial u}{\partial t} \nabla \left(\theta_r^2\right) \cdot a \nabla u \lambda^3 dz d\lambda \\
&\equiv A + B + C.
\end{aligned} \tag{10.9}$$

Integrating by parts in t, we have that (since a is self-adjoint)

$$2A = \iint \nabla u \cdot \frac{\partial a}{\partial t} \nabla u \theta_r^2 \lambda^3 dz d\lambda + \iint \nabla u \cdot a \nabla u \frac{\partial}{\partial t} \left(\theta_r^2\right) \lambda^3 dz d\lambda.$$

By (10.3)(ii), we therefore have that

$$\begin{aligned} |A| &\le \sqrt{\mu} \iint |\nabla u|^2 \theta_r^2 \lambda d\lambda dz + \frac{c}{r^2} \int_0^{\frac{3r}{2}} \int_{\Delta_{\frac{3r}{2}}} |\nabla u|^2 \lambda^3 d\lambda dz. \\ &\le \sqrt{\mu} \iint |\nabla u|^2 \theta_r^2 \lambda d\lambda dz + \frac{c(n,\nu)}{r} (1+\mu) \int_0^{3r} \int_{\Delta_{3r}} |u|^2 d\lambda dz \end{aligned}$$

(by Lemma 4.4), and the desired bound for A now follows readily.

To treat the remaining terms in (10.9)

$$|B| + |C| \le C(\nu) \left(\epsilon \iint \left| \frac{\partial u}{\partial t} \right|^2 \theta_r^2 \lambda^3 d\lambda dz + \frac{1}{\epsilon} \iint |\nabla u|^2 \lambda \theta_r^2 d\lambda dz + \frac{1}{\epsilon} \cdot \frac{1}{r^2} \int_0^{\frac{3r}{2}} \int_{\Delta_{\frac{3r}{2}}} |\nabla u|^2 \lambda^3 d\lambda dz \right).$$

Choosing ϵ small enough, depending only on ν, we may hide the small term on the left hand side of Lemma 10.7(ii). The other two summands satisfy the desired bound (again we use Lemma 4.4 to handle the very last term). We have therefore established part (ii) of Lemma 10.7.

Let us now turn to part (i). By the ellipticity hypothesis (10.2)(i), we have that

$$\begin{aligned} \iint |\nabla u|^2 \theta_r^2 \lambda d\lambda dz &\le C(\nu) \iint \nabla u \cdot a \nabla u \theta_r^2 \lambda d\lambda dz \\ &= -C(\nu) \left(\iint u \operatorname{div}(a\nabla u) \theta_r^2 dz \lambda d\lambda + \iint u \langle e_0, a \nabla u \rangle \theta_r^2 d\lambda dz \right. \\ &\quad \left. + \iint u \nabla(\theta_r^2) \cdot a \nabla u \lambda d\lambda dz \right) \\ &\equiv I + II + III \end{aligned} \tag{10.10}$$

(here, $e_0 \equiv (1, 0, \dots 0)$ denotes the unit vector in the $x_0 \equiv \lambda$ direction). We observe that

$$\begin{aligned} |III| &\le C(\nu) \left(\frac{1}{r} \int_0^{\frac{3r}{2}} \int_{\Delta_{\frac{3r}{2}}} |u|^2 d\lambda dz + \frac{1}{r} \int_0^{\frac{3r}{2}} \int_{\Delta_{\frac{3r}{2}}} |\nabla u|^2 \lambda^2 d\lambda dz \right) \\ &\le C(\nu) \int_{\Delta_{3r}} (N_*^{3r} u)^2 dz \end{aligned}$$

(where we have used Lemma 4.4 to treat the second integral), as desired. Since $Lu = 0$, we have that

$$
\begin{aligned}
I &= -C(\nu) \iint ub \frac{\partial u}{\partial t} \theta_r^2 dz \lambda d\lambda \\
&\quad + C(\nu) \iint u\vec{B} \cdot \nabla u \theta_r^2 dz \lambda d\lambda \\
&\equiv I' + I''.
\end{aligned} \tag{10.11}
$$

To handle these two terms, we shall require the following localized version of a well known lemma of Carleson.

Lemma 10.12. *Let dm be a measure on $\mathbb{R}^{n+1}_+$ which satisfies*

$$\sup_{\Delta_r} \int_0^r \lambda^{-d} \int_{\Delta_r} dm(\lambda, z) \leq \mu.$$

Then for all measurable F defined on $\mathbb{R}^{n+1}_+$, and for all $r > 0$ and all parabolic balls $\Delta_r \subseteq \mathbb{R}^n$, we have

$$\int_0^r \int_{\Delta_r} |F(\lambda, z)|^2 dm(\lambda, z) \leq C(n)\mu \int_{\Delta_{3r}} (N_*^{3r}(F))^2 dz.$$

Proof. Apply the usual Carleson Lemma to $F\chi$, where χ denotes the characteristic function of $(0, r) \times \Delta_r$. We omit the routine details.

With Lemma 10.12 in hand, we return to the terms I' and I'' in (10.11). By Lemma 10.12, and (10.4), we have that

$$
\begin{aligned}
|I''| \leq C(\nu) \Big\{ \epsilon & \iint |\nabla u|^2 \theta_r^2 \lambda d\lambda dz \\
&+ C(n) \frac{\mu}{\epsilon} \int_{\Delta_{5r}} |N_*^{5r}(u)|^2 dz \Big\}.
\end{aligned}
$$

Choosing ϵ small enough depending only on ν, we may hide the small term on the left hand side of Lemma 10.7(i). The other term yields the desired bound.

To handle I', we integrate by parts in λ to obtain

$$
\begin{aligned}
I' &\equiv \frac{-C(\nu)}{2} \iint b \frac{\partial}{\partial t}(u^2) \theta_r^2 dz \lambda d\lambda \\
&\equiv \frac{C(\nu)}{4} \Big\{ \iint \frac{\partial b}{\partial \lambda} \frac{\partial (u^2)}{\partial t} \theta_r^2 dz \lambda^2 d\lambda + \\
&\quad \iint b \frac{\partial}{\partial t}(u^2) \frac{\partial}{\partial \lambda}(\theta_r^2) dz \lambda^2 d\lambda + \iint b \frac{\partial^2}{\partial t \partial \lambda}(u^2) \theta_r^2 dz \lambda^2 d\lambda \Big\} \\
&\equiv I_1' + I_2' + I_3'.
\end{aligned}
$$

By (10.4), Lemma 10.12, Lemma 10.7(ii), and the fact that $\frac{1}{2}\frac{\partial}{\partial t}(u^2) = u\frac{\partial u}{\partial t}$, we obtain that

$$
\begin{aligned}
|I_1'| &\leq C(\nu) \left\{ \epsilon \iint \left| \frac{\partial u}{\partial t} \right|^2 \lambda^3 \theta_r^2 d\lambda dz + C(n) \frac{\mu}{\epsilon} \int_{\Delta_{5r}} (N_*^{5r}(u))^2 \right\} \\
&\leq C(n, \nu) \left\{ \epsilon(1 + \mu) \iint |\nabla u|^2 \theta_r^2 \lambda d\lambda dz + \frac{(1 + \mu)}{\epsilon} \int_{\Delta_{5r}} (N_*^{5r}(u))^2 \right\}.
\end{aligned}
$$

We now choose $\epsilon \equiv \epsilon'(1+\mu)^{-1}$, where ϵ' is chosen small enough, depending only on n, ν, that we may hide the small term on the left side of Lemma 10.7(i). This contributes a term on the right hand side of (i) on the order of

$$C(n,\nu)(1+\mu^2)\int_{\Delta_{5r}}(N_*^{5r}(u))^2,$$

as desired. Next,

$$\begin{aligned}|I_2'| &\le C(\nu)\left\{\epsilon\iint\left|\frac{\partial u}{\partial t}\right|^2\lambda^3\theta_r^2 d\lambda dz + \frac{1}{\epsilon}\frac{1}{r^2}\int_0^{\frac{3r}{2}}\int_{\Delta_{\frac{3r}{2}}}|u|^2\lambda dz d\lambda\right\}\\ &\le C(\nu)\left\{\epsilon(1+\mu)\iint|\nabla u|^2\theta_r^2\lambda d\lambda dz + \frac{1}{\epsilon}\int_{\Delta_{2r}}(N_*^{2r}(u))^2 dz\right\},\end{aligned}$$

where again we have invoked Lemma 10.7(ii). Choosing $\epsilon = \epsilon'(1+\mu)^{-1}$ as above, we may again hide the small term, and the desired bound for I_2' now follows.

To handle I_3', we integrate by parts in t, to obtain that

$$\begin{aligned}|I_3'| &= \left|\frac{-C(\nu)}{2}\left\{\iint\frac{\partial b}{\partial t}u\frac{\partial u}{\partial\lambda}\theta_r^2 dz\lambda^2 d\lambda + \iint bu\frac{\partial u}{\partial\lambda}\frac{\partial}{\partial t}(\theta_r^2)dz\lambda^2 d\lambda\right\}\right|\\ &\le C(\nu)\left\{\epsilon\iint|\nabla u|^2\theta_r^2\lambda dz d\lambda + \frac{1}{\epsilon}\left[\mu\int_{\Delta_{5r}}(N_*^{5r}(u))^2 + \frac{1}{r^4}\int_0^{\frac{3r}{2}}\int_{\Delta_{\frac{3r}{2}}}|u|^2 dz\lambda^3 d\lambda\right]\right\},\end{aligned}$$

where in the last inequality we have invoked (10.5) and Lemma 10.12. Since the second summand inside the square brackets is clearly dominated by the first, we may again choose ϵ small, depending only on ν, and hide the small term on the left side of Lemma 10.7(i). This yields the desired bound for I.

To finish the Proof of Lemma 10.7, it remains only to consider term II in (10.10). Observe that, using the notational convention $x_0 \equiv \lambda$, we have that

$$II = -C(\nu)\sum_{j=0}^{n-1}\iint ua_{0j}\frac{\partial u}{\partial x_j}\theta_r^2 d\lambda dz. \tag{10.13}$$

Since $a_{00} \equiv 1$ (recall that there was no loss of generality in assuming this, and that is why it is convenient to consider the class of operators defined in (10.1)-(10.5)), we have that the $j=0$ term equals

$$\frac{-C(\nu)}{2}\iint\frac{\partial}{\partial\lambda}(u^2)\theta_r^2 d\lambda dz =$$
$$\frac{C(\nu)}{2}\left\{\int_{\mathbb{R}^n}u^2\theta_r^2 dz + \iint u^2\frac{\partial}{\partial\lambda}(\theta_r)^2 d\lambda dz\right\}.$$

But each of these summands is easily seen to be majorized by $C(\nu)\int_{\Delta_{5r}}(N_*^{5r}(u))^2 dz$, as desired.

On the other hand, if $j > 0$ in (10.13), we integrate by parts in λ to obtain

$$\frac{C(\nu)}{2}\left\{\iint \frac{\partial}{\partial x_j}(u^2)\frac{\partial a_{0j}}{\partial \lambda}\theta_r^2 \lambda d\lambda dz + \right.$$
$$\left.\iint \frac{\partial^2}{\partial x_j \partial \lambda}(u^2) a_{0j}\theta_r^2 \lambda d\lambda dz + \iint \frac{\partial}{\partial x_j}(u^2) a_{0j}\frac{\partial \theta_r^2}{\partial \lambda}\lambda d\lambda dz\right\}.$$

The first and last of these summands can be handled exactly like term I'' in (10.11), and term III in (10.10), respectively. Moreover, the middle summand can also be handled in the same fashion, after first integrating by parts to more $\frac{\partial}{\partial x_j}$ onto either a_{0_j} or θ_r^2. This concludes the Proof of Lemma 10.7, and hence also that of Lemma 10.6, as well as the case $q = 2$ of Lemma 9.8.

We now turn to the proof of the case $q > 2$ of Lemma 9.8. For $u(\cdot, t) \in L_1^2(\mathbb{R}_+^n)$, for each t, define

$$g(u) \equiv \left(\int_0^\infty |\nabla u|^2 \lambda d\lambda\right)^{\frac{1}{2}},$$
$$S(u)(z) \equiv \left(\iint_{\|z-v\|<\lambda} |\nabla u|^2 \lambda^{1-d} d\lambda dv\right)^{\frac{1}{2}},$$

where $\nabla \equiv \nabla_{x,\lambda}$. Of course, our goal is to control the L^q norm of $g(u)$ when u is a solution of $L_0 u = 0$. The following fact is well known (see, e.g. [29, p. 91], where it is established for g_λ^* — the same argument applies to S).

Lemma 10.14. *Let g, S be defined as above. Then for $q \geq 2$, we have that*

$$\|S(u)\|_{L^q(\mathbb{R}^n)} \leq C(n,q)\|g(u)\|_{L^q(\mathbb{R}^n)}.$$

Remark. This lemma is an immediate consequence of the fact that $\int_{\mathbb{R}^n}(S(u))^2 w(z)dz \leq C(n)\int_{\mathbb{R}^n}(g(u))^2 Mw$, which inequality in turn follows easily from the definition of $S(u)$. We omit the details.

We shall now obtain the case $q > 2$ of Lemma 9.8 as a consequence of the following

Lemma 10.15. *Let L be as in* (10.1)-(10.5), *with $b \equiv 1$ and $\vec{B} \equiv 0$. Suppose that $Lu = 0$ in $\mathbb{R}_+^{n+1}$. Then, for all $q \geq 2$, there exists a number $M(q)$ such that*

$$\|g(u)\|_{L^q(\mathbb{R}^n)} \leq C(n,\nu)(1+\mu^{M(q)})\|N_* u\|_{L^q(\mathbb{R}^n)}.$$

Remark. Lemma 9.8 is an immediate consequence of Lemma 10.15; indeed, L_0 satisfies the hypothesis of Lemma 10.15, with $\mu \approx \gamma^{-2}$, and ν depending only on β_0.

Proof of Lemma 10.15. We may assume that $q > 2$: the case $q = 2$ is a special case of Lemma 10.6. Consequently, by Lemma 10.14, it is enough to establish the following "good-λ" inequality, for all small $\eta > 0$:

$$\begin{aligned} |\{g(u) > (1 + C(n,\nu)\sqrt{\mu})\alpha, (M(N_* u)^2)^{\frac{1}{2}} \leq \eta\alpha\}| \\ \leq C(n,\nu)\eta^2(1+\mu^2)|\{S(u) > \alpha\}| \end{aligned} \tag{10.16}$$

Indeed, we then choose $\eta \equiv \epsilon(n,\nu,q)(1+\mu)^{-\Lambda(q)}$, with $\epsilon(n,\nu,q)$ sufficiently small, $\Lambda(q)$ sufficiently large, and the conclusion of Lemma 10.15 then follows by a standard argument. The proof of (10.16) with a blend of well known arguments, see, e.g., [5], [9], and [4].

We now prove (10.16). As usual, we perform a Whitney decomposition of the set $E_\alpha \equiv \{S(u) > \alpha\}$, and write, for α fixed,

$$E_\alpha = \cup I_j^\alpha \equiv \cup I_j,$$

where the parabolic cubes I_j are disjoint, $|I_j| \equiv r_j^d$, and (parabolic) $\operatorname{dist}(I_j, E_\alpha^c) \approx r_j$. To establish (10.16), it is enough to show that for each I_j,

$$\begin{aligned} &|\{z \in I_j : g(u) > (1 + C(n,\nu)\sqrt{\mu})\alpha, (M(N_* u)^2)^{\frac{1}{2}} \le \eta\alpha\}| \\ &\qquad \le c(n,\nu)\eta^2(1+\mu^2)|I_j|. \end{aligned} \tag{10.17}$$

Fix I_j. We may assume that there exists $z^1 \in I_j$ such that $(M(N_* u)^2)^{\frac{1}{2}}(z^1) \le \eta\alpha$, otherwise (10.17) is vacuously true. Since I_j is now fixed, we simplify notation by writing $I = I_j$, $r \equiv r_j$, and we split $g(u) \equiv g_1(u) + g_2(u)$, where

$$g_1(u) \equiv \left(\int_0^{Nr} |\nabla u|^2 \lambda d\lambda\right)^{\frac{1}{2}}, g_2(u) \equiv \left(\int_{Nr}^{\infty} |\nabla u|^2 \lambda d\lambda\right)^{\frac{1}{2}},$$

and N is to be chosen. We now claim that, for $z \in I$,

$$g_2(u)(z) \le C(n,\nu)\sqrt{\mu}\alpha. \tag{10.18}$$

Indeed, by the Whitney instruction, there exists $z^* \in E_\alpha^c$, $\operatorname{dist}(z^*, I) \le C_n r$. Moreover, by Lemma 4.8, we have that

$$\begin{aligned} |g_2(u)(z)|^2 &\equiv \int_{Nr}^{\infty} |\nabla u(\lambda, z)^2 \lambda d\lambda \\ &\le C(n,\nu)\mu \int_{Nr}^{\infty} \lambda^{-d-1} \int\int_{\|(\rho,v)-(\lambda,z)\|<\frac{\lambda}{4}} |\nabla u(\rho, v)|^2 d\rho dv \lambda d\lambda. \end{aligned}$$

But, in particular $|\rho - \lambda| < \frac{\lambda}{4} \Rightarrow \frac{3\lambda}{4} < \rho < \frac{5\lambda}{4}$. Hence, the last expression is no larger than

$$\begin{aligned} &C(n,\nu)\mu \iint_{\substack{\|v-z\|<\frac{\rho}{3} \\ \rho > \frac{3Nr}{4}}} |\nabla u(\rho,v)|^2 \rho^{-d} \int_{|\rho-\lambda|<\frac{\rho}{3}} d\lambda d\rho dv \\ &= C(n,\nu)\mu \int_{\frac{3Nr}{4}}^{\infty} \int_{\|z-v\|<\frac{\rho}{3}} |\nabla u(\rho,v)|^2 \rho^{1-d} d\rho dv \\ &\le C(n,\nu)\mu \int_{\frac{3Nr}{4}}^{\infty} \int_{\|v-z^*\|<\frac{\rho}{3}+c_n r} |\nabla u(\rho,v)|^2 \rho^{1-d} d\rho dv \end{aligned}$$

(since, for all $z \in I$, $\|z - z^*\| \le C_n r$)

$$\le C(n,\nu)\mu(S(u)(z^*))^2 \le C(n,\nu)\mu\alpha^2,$$

if we chose N large enough depending on dimension.

This establishes the claim (10.18). Consequently, the left hand side of (10.17) is bounded by

$$\begin{aligned}&|\{z \in I : g_1(u) > \alpha, (M(N_*u)^2)^{\frac{1}{2}} \leq \eta\alpha\}| \\ &\quad \leq \frac{1}{\alpha^2}\int_I (g_1(u))^2 \equiv \frac{1}{\alpha^2}\int_I\int_0^{N_r} |\nabla u|^2 \lambda d\lambda dz \\ &\quad \leq \frac{C(n,\nu)(1+\mu^2)}{\alpha^2}\int_{\Delta_{C_n r}(z^1)} (N_*(u))^2 \\ &\leq C(n,\nu)(1+\mu^2)\eta^2|I|,\end{aligned}$$

where in the next to last inequality we have used Lemma 10.7(i), and the fact that $z^1 \in I$, and in the last inequality we have used that $M(N_*u)^2 \leq \eta^2\alpha^2$. We have therefore proved (10.17), and thus also (10.16). Consequently, Lemma 9.8 is true, and the Proof of Theorem 1.8 is now complete.

References

[1] D. Aronson, Non-negative solutions of linear parabolic equation, Ann. Scuola Norm. Sup. Pisa 22 (1968), 607-694.

[2] R. Brown, The method of layer potentials for the heat equation in Lipschitz cylinders, Amer. J. Math 111 (1989), 359-379.

[3] R. Brown, The initial-Neumann problem for the heat equation in Lipschitz cylinders, Trans. Amer. Math. Soc. 320 (1990), 1-52.

[4] R. Brown, Area integral estimates for caloric functions, Trans. Amer. Math. Soc. 315 (1989), 565-589.

[5] R. Coifman and C. Fefferman, Weighted norm inequalities for maximal functions and singular integrals, Studia Math. 51 (1974), 241-250.

[6] R. Coifman and R. Rochberg, Another characterization of BMO, Proc. Amer. Math. Soc. 79 (1980), 249-254.

[7] B. Dahlberg, On estimates for harmonic measure, Arch. Rat. Mech. Anal. 65 (1977), 272-288.

[8] B. Dahlberg, Poisson semigroups and singular integrals, Proc. Amer. Math. Soc. 97 (1986), 41-48.

[9] B. Dahlberg, D. Jerison and C. Kenig, Area integral estimates for elliptic differential operators with nonsmooth coefficients, Arkiv Mat. 22 (1984), 97-108.

[10] B. Dahlberg and C. Kenig, Hardy spaces and the Neumann problem in L^p for Laplace's equation in Lipschitz domains, Annals of Math. 125 (1987), 437-466.

[11] B. Dahlberg, C. Kenig, J. Piper and G. Verchota, Area integral estimates and maximum principles for solutions of higher order elliptic equations and systems on Lipschitz domains, Ann. Inst. Fourier 47 (1997), no 5, 1425 - 1461.

[12] E.B. Fabes and N. Riviere, Symbolic calculus of kernels with mixed homogeneity, in **Singular Integrals**, A.P. Calderón, ed., Proc. Symp. Pure Math., Vol. 10, Amer. Math. Soc., Providence, 1967, 106-127.

[13] E.B. Fabes and S. Salsa, Estimates of caloric measure and the initial-Dirichlet problem for the heat equation in Lipschitz cylinders, Trans. Amer. Math. Soc. 279 (1983), 635-650.

[14] E.B. Fabes, N. Garofalo and S. Salsa, Comparison theorems for temperatures in non-cylindrical domains, Atti. Acad. Naz. Lincei, Rend. Cl. Sci. Fis., Mat. Nat. 8-77 (1984), 1-12.

[15] R. Fefferman, C. Kenig and J. Pipher, The theory of weights and the Dirichlet problem for elliptic equations, Annals of Math. 134 (1991), 65-124.

[16] S. Hofmann, A characterization of commutators of parabolic singular integrals, in **Fourier Analysis** and Partial Differential Equations, J. Garcá-Cuerva, E. Hernández, F. Soria, and J.-L. torrea, eds., Studies in Advanced Mathematics, CRC Press, Boca Raton, 1995, 195-210.

[17] S. Hofmann, Parabolic singular integrals of Calderón type, rough operators, and caloric layer potentials, Duke Math. J. 90 (1997), 209-259.

[18] S. Hofmann and J.L. Lewis, L^2 Solvability and representation by caloric layer potentials in time-varying domains, Annals of Math. 144 (1996), 349-420.

[19] S. Hofmann and J.L. Lewis, The L^p regularity problem for the heat equation in non-cylindrical domains, Illinois J. Math. 43 (1999), no. 4, 752-769.

[20] S. Hofmann and J.L. Lewis, The Dirichlet problem for parabolic operators with singular drift terms, Mem. Amer. Math. Soc. no. 719, 151 (2001), 1-113.

[21] R. Kaufman and J.-M. Wu, Parabolic measure on domains of class $\mathrm{Lip}_{\frac{1}{2}}$, Comp. Math. 65 (1988), 201-207.

[22] C. Kenig, **Harmonic Analysis Techniques for Second Order Elliptic Boundary Value Problems**, CBMS Regional Conference Series in Mathematics, Number 83, Amer. Math. Soc., Providence, 1994.

[23] C. Kenig and J. Pipher, The Neumann problem for elliptic equations with non-smooth coefficients, Invent. Math. 113 (1993), 447-509.

[24] C. Kenig and J. Pipher, The Neumann problem for elliptic equations with non-smooth coefficients. II A celebration of John F. Nash, Jr. Duke Math. J, 81 (1995), no. 1, 227-250.

[25] C. Kenig and J. Pipher, The Dirichlet problem for elliptic equations with drift terms, Pub. Mat. 45 (2001), no. 1, 199 - 217.

[26] J.L. Lewis and M.A.M. Murray, The method of layer potentials for the heat equation in time-varying domains, Memoirs of the Amer. Math. Soc. 545 (1995), 1-157.

[27] M.A.M. Murray, Commutators with fractional differentiations and BMO Sobolev spaces, Indiana Univ. Math. J. 34 (1985), 205-215.

[28] E.M. Stein, Singular Integrals and Differentiability Properties of Functions, Princeton Univ. Press, 1970.

[29] R. Strichartz, Bounded mean oscillation and Sobolev spaces, Indiana Univ. Math. J. 29 (1980), 539-558.

[30] G. Verchota, Layer potentials and regularity for Laplace's equation in Lipschitz domains, J. of Functional Anal. 59 (1984), 572-611.