

## ON AREA INTEGRAL ESTIMATES FOR SOLUTIONS TO PARABOLIC SYSTEMS IN TIME-VARYING AND NON-SMOOTH CYLINDERS

KAJ NYSTRÖM

ABSTRACT. In this paper we prove results relating the (parabolic) non-tangential maximum operator and appropriate square functions in  $L^p$  for solutions to general second order, symmetric and strongly elliptic parabolic systems with real valued and constant coefficients in the setting of a class of time-varying, non-smooth infinite cylinders  $\Omega$ . In particular we prove a global as well as a local and scale invariant equivalence between the parabolic non-tangential maximal operator and appropriate square functions for solutions of our system. The novelty of our approach is that it is not based on singular integrals, the prevailing tool in the analysis of systems in non-smooth domains. Instead the methods explored have recently proved useful in the analysis of elliptic measure associated to non-symmetric operators through the work of Kenig-Koch-Pipher-Toro and in the analysis of caloric measure without the use of layer potentials.

### 1. INTRODUCTION AND STATEMENT OF MAIN RESULTS

In this paper we prove results relating the non-tangential maximum operator and appropriate square functions in  $L^p$  for solutions to general second order, symmetric and strongly elliptic parabolic systems with real valued and constant coefficients in time-varying non-smooth domains of the form

$$\Omega = \{(x_0, x, t) \in \mathbf{R} \times \mathbf{R}^{n-1} \times \mathbf{R} : x_0 > A(x, t)\}$$

where  $n \geq 2$  and where the function  $A(x, t) : \mathbf{R}^n \rightarrow \mathbf{R}$  is compactly supported. The function  $A(x, t)$  is assumed to be  $\text{Lip}(1, 1/2)$  but we also impose an additional regularity condition. To formulate the additional regularity assumption we need to introduce some more notation. Let  $z = (x, t) \in \mathbf{R}^{n-1} \times \mathbf{R}$  and let  $\|z\|$  be the unique positive solution  $\rho$  of the equation

$$\frac{t^2}{\rho^4} + \sum_{i=0}^{n-1} \frac{x_i^2}{\rho^2} = 1.$$

Note that  $\|(\delta x, \delta^2 t)\| = \delta\|(x, t)\|$  and we will call  $\|z\|$  the parabolic norm of  $z$ . By definition parabolic  $BMO$  is the space of locally integrable functions modulo

---

Received by the editors March 7, 2006.

2000 *Mathematics Subject Classification*. Primary 35B65, 35C15, 35K20, 35K40, 35K50.

*Key words and phrases*. Second order parabolic systems, time-varying cylinder, Carleson measure, maximal function, non-tangential maximal function, square function.

constants satisfying

$$\|b\|_* := \sup_B \frac{1}{|B|} \int_B |b(z) - m_B b| dz < \infty.$$

Here  $z = (x, t)$ ,  $B$  denotes the parabolic ball  $B = B_r(z_0) = \{z \in \mathbf{R}^n = \|z - z_0\| < r\}$  and  $m_B b$  denotes the average of the function  $b$  over the ball  $B$ . Let  $\widehat{f}(\xi, \tau)$  be the Fourier transform of a function  $f$  defined on  $\mathbf{R}^n$ , and let  $\xi, \tau$  denote the phase variables. For a function  $g \in C_0^\infty(\mathbf{R}^n)$  we introduce the fractional differentiation operator  $D_{1/2}$  by

$$(\widehat{D_{1/2}g})(\tau) := |\tau|^{1/2} \widehat{g}(\tau).$$

If  $h \in C_0^\infty(\mathbf{R}^n)$ , then by  $D_{1/2}^t h : \mathbf{R}^n \rightarrow \mathbf{R}$  we will mean  $D_{1/2} h(x, \cdot)$  defined a.e. for each  $x \in \mathbf{R}^{n-1}$ .

**Definition 1.** A Time-Varying domain with defining function  $A$  and parameters  $a_1$  and  $a_2$ , denoted  $TV(A, a_1, a_2)$ , is a domain  $\Omega \subset \mathbf{R}^{n+1}$  of the form

$$\Omega = \{(x_0, x, t) \in \mathbf{R} \times \mathbf{R}^{n-1} \times \mathbf{R} : x_0 > A(x, t)\}$$

where  $n \geq 2$  and where the function  $A(x, t) : \mathbf{R}^n \rightarrow \mathbf{R}$  is compactly supported and satisfies

- (i)  $|A(x, t) - A(y, t)| \leq a_1 |x - y|, x, y \in \mathbf{R}^{n-1}, t \in \mathbf{R},$
- (ii)  $D_{1/2}^t A \in BMO(\mathbf{R}^n), \|D_{1/2}^t A\|_* \leq a_2.$

One can prove that if  $\Omega$  is  $TV(A, a_1, a_2)$  for some  $a_1, a_2$ , then

$$|A(x, t) - A(y, s)| \leq \delta_0 (|x - y| + |t - s|^{1/2}), \quad x, y \in \mathbf{R}^n, \quad t, s \in \mathbf{R},$$

for some  $\delta_0 = \delta_0(a_1, a_2)$ . The geometric conditions stated in Definition 1 can also be expressed in a different but equivalent form, and in order to formulate this we define a parabolic half-order time derivative by

$$\widehat{\mathbf{D}_n A}(\xi, \tau) := \frac{\tau}{\|(\xi, \tau)\|} \widehat{A}(\xi, \tau).$$

We let  $\|\cdot\|_\infty$  be the supremum norm and define  $\|A\|_{\text{comm}} = \|\nabla_x A\|_\infty + \|\mathbf{D}_n A\|_*$  where  $\nabla_x = \left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_{n-1}}\right), \|\nabla_x A\|_\infty := \sup_t \|\nabla_x A(\cdot, t)\|_\infty.$  In [HL] it is proven that the condition  $\|A\|_{\text{comm}} < \infty$  is equivalent to the statement that  $\Omega$  is  $TV(A, a_1, a_2)$  for some  $a_1, a_2$ .

We define a surface measure on  $\partial\Omega$ ,  $d\sigma$ , as  $d\sigma_t dt$ , where  $d\sigma_t$  is the naturally defined surface measure on the Lipschitz graph  $\partial\Omega_t$ . Here  $\Omega_t = \{(x_0, x, t) \in \mathbf{R} \times \mathbf{R}^{n-1} \times \{t\}; x_0 > A(x, t)\}$  and the unit outer normal to  $\Omega_t$  is denoted by  $N_t = (N_t^0, \dots, N_t^{n-1})$ .  $L^p(\partial\Omega)$  denotes the  $L^p$ -space defined w.r.t. the measure  $d\sigma$ .

Through the works in [LM], [HL] it has become clear that from the perspective of parabolic singular integrals and caloric measure the parabolic analogue of the notion of Lipschitz domains, explored in elliptic partial differential equations, is graph domains satisfying the conditions stated in Definition 1 (or equivalently as

$\|A\|_{\text{comm}} < \infty$ ). In particular if  $\Omega$  is  $TV(A, a_1, a_2)$  for some  $a_1, a_2$  it was proven in [LM] that the parabolic Poisson kernel associated to the heat operator is in a certain  $L^p$  reverse Hölder class for some  $p > 1$ .

The systems we consider in this paper are general second order symmetric parabolic systems with real valued and constant coefficients having the following form for relevant  $n, m$  and constant coefficient matrices  $A^{rs}$ :

$$(1) \quad \frac{\partial u^r}{\partial t} = \frac{\partial}{\partial x_i} A_{ij}^{rs} \frac{\partial u^s}{\partial x_j}, \quad 0 \leq i, j \leq n - 1, \quad 1 \leq r, s \leq m,$$

$$(2) \quad A_{ij}^{rs} = A_{ji}^{sr}.$$

Concerning ellipticity we will assume that

$$(3) \quad A_{ij}^{rs} \eta_i^r \eta_j^s \geq c \sum_{(l,q)} |\eta_l^q|^2 \quad \text{for all } \eta_i, \eta_j \in \mathbf{R}^m.$$

This condition implies that the matrix  $\{A_{ij}^{rs}\}$  is positive definite. As many systems can be represented in an infinite number of ways our results will apply to any system having at least one representation in terms of a positive definite matrix  $\{A_{ij}^{rs}\}$ . The condition in (3) is stronger than the ellipticity condition encoded in the commonly used Legendre-Hadamard ellipticity condition:

$$(4) \quad A_{ij}^{rs} \nu_i \nu_j \eta^r \eta^s \geq c |\nu|^2 |\eta|^2 \quad \text{for all } \nu \in \mathbf{R}^n, \quad \eta \in \mathbf{R}^m.$$

We are now ready to formulate our results. If  $h$  is a function defined on  $\Omega$  the non-tangential maximal function  $\tilde{N}_*(h) = \tilde{N}_*^a(h) : \partial\Omega \rightarrow \mathbf{R}$  as well as the area integral or square function  $S(h) = S^{\tilde{a}}(h) : \partial\Omega \rightarrow \mathbf{R}$  are defined in the bulk of the paper. Here  $a$  as well as  $\tilde{a}$  determines the aperture of the underlying cones.

**Theorem 1.** *Let  $\Omega$  be a time-varying domain described as  $TV(A, a_1, a_2)$  for some  $a_1, a_2$  and assume that  $\{A_{ij}^{rs}\}$  are real constants satisfying (2) and (3). Assume that  $\vec{u} = (u^1, \dots, u^m)$  satisfies the system*

$$\frac{\partial u^r}{\partial t} = \frac{\partial}{\partial x_i} A_{ij}^{rs} \frac{\partial u^s}{\partial x_j}, \quad 0 \leq i, j \leq n - 1, \quad 1 \leq r, s \leq m, \quad \text{in } \Omega.$$

Then there exist constants  $C_1$  and  $C_2$  independent of  $\vec{u}$  such that

$$\sum_s \int_{\partial\Omega} |S(u^s)(X, t)|^p d\sigma(X, t) \leq C_1 \sum_k \int_{\partial\Omega} |\tilde{N}_*(u^k)(X, t)|^p d\sigma(X, t)$$

for  $p \in (0, \infty)$  and such that

$$\sum_k \int_{\partial\Omega} |\tilde{N}_*(u^k)(X, t)|^p d\sigma(X, t) \leq C_2 \sum_s \int_{\partial\Omega} |S(u^s)(X, t)|^p d\sigma(X, t)$$

for  $p \in (2, \infty)$  whenever the terms on the right hand sides are finite.

We note that we are currently not able to prove the second inequality for all  $p$  in the range  $(0, \infty)$  without additional assumptions on the constants  $a_1, a_2$  defining our time-varying domain. Before we comment more on this we note that the proof of Theorem 1 is based on a localized and scale invariant equivalence between truncated

(parabolic) non-tangential maximal operators and truncated square functions. To formulate this result, which is an interesting result in itself, we let  $\Delta = \Delta_r$  be a cube on  $\mathbf{R}^{n-1} \times \mathbf{R}$  of dimensions  $r \times \dots \times r \times r^2$  and let

$$\tilde{\Delta}_r = \{(A(x, t), x, t) : (x, t) \in \Delta_r\}$$

be the associated surface cube on the boundary of our domain  $\Omega$ . We furthermore define an associated Carleson box above  $\tilde{\Delta}_r$ , denoted by  $\tilde{T}(\Delta_r)$ , as

$$\tilde{T}(\Delta_r) = \{(x_0, x, t) : (x, t) \in \Delta_r, A(x, t) \leq x_0 \leq A(x, t) + r\}.$$

If  $(x_r, t_r)$  is the center of  $\Delta$  we define  $P_{\tilde{\Delta}} = (A(x_r, t_r) + \lambda_0 r, x_r, t_r)$  where  $0 < \lambda_0 < 1$  is a small fixed number. The truncated versions of the operators  $\tilde{N}_*(\cdot)$  and  $S(\cdot)$ , denoted  $\tilde{N}_*^{a,r}(\cdot)$  and  $S^{a,r}(\cdot)$ , are defined in the bulk of the paper. We prove the following

**Theorem 2.** *Let  $\Omega$  be a time-varying domain described as  $TV(A, a_1, a_2)$  for some  $a_1, a_2$  and let  $\vec{u}$  be as in Theorem 1. Then there exists a constant  $C_1 = C_1(n, a_1, a_2)$  such that*

$$\sum_k \int_{\tilde{\Delta}_r} |S^{a,r}(u^k)(X, t)|^2 d\sigma(X, t) \leq C_1 \sum_s \int_{2\tilde{\Delta}_r} |\tilde{N}^{a,2r}(u^s)(X, t)|^2 d\sigma(X, t).$$

Furthermore, for every  $\rho > 0$  there exist constants  $C_2 = C_2(n, p, a_1, a_2, \rho)$  and  $\tilde{C}_2 = \tilde{C}_2(n, p, a_1, a_2)$  such that if  $\vec{u}(P_{\tilde{\Delta}_r}) = \vec{0}$ ,  $p > 2$ , then

$$\begin{aligned} \sum_k \int_{\tilde{\Delta}_r} |\tilde{N}_*^{a,r}(u^k)(X, t)|^p d\sigma(X, t) &\leq C_2 \sum_s \int_{2\tilde{\Delta}_r} |S^{4a,2r}(u^s)(X, t)|^p d\sigma(X, t) \\ &+ \tilde{C}_2 \rho \sum_s \int_{2\tilde{\Delta}_r} |\tilde{N}_*^{a,2r}(u^s)(X, t)|^p d\sigma(X, t). \end{aligned}$$

Note that the constants of Theorem 1 and Theorem 2 also depend on the coefficients  $\{A_{ij}^r\}$  as well as the aperture of the cones used in the definition of the non-tangential maximal operator and the square function.

In the case of elliptic equations and systems, of second order and higher, satisfying an ellipticity condition of Legendre-Hadamard type, Theorem 1 and Theorem 2 (with the appropriate stronger version of the second inequality) were proven in [DKPV] for the full range of  $p$ , i.e., for  $p \in (0, \infty)$ . A long term goal is to generalize the results in [DKPV] to a genuinely parabolic setting, i.e., to a setting of non-smooth and time-varying domains. Notable is that currently there are no results on higher order parabolic problems in a non-smooth and time-varying setting.

Note that in Theorem 1 we can only prove an  $L^p$  bound of the non-tangential maximal operator with the square function if  $p \in (2, \infty)$  instead of the full range  $(0, \infty)$ . Furthermore, in Theorem 2 the second inequality is much too weak compared to what one expects. The reason is that in our arguments we are not assuming anything about the  $L^p$ -solvability of the system. In particular all of the arguments in [DKPV] are based on the solvability of systems and higher order equations in  $L^2$  in Lipschitz domains with control of the non-tangential maximal operator using the data. In the case of systems of second order all arguments are based on the use

of singular integrals. In our case an inequality of the form

$$\sum_k \int_{\partial\Omega} |\tilde{N}_*(u^k)(X, t)|^2 d\sigma(X, t) \leq C \sum_s \int_{\partial\Omega} |u^s(X, t)|^2 d\sigma(X, t)$$

would immediately yield, using the arguments described in the bulk of this paper, the second inequality of Theorem 1 for  $p = 2$ . A localized version would similarly yield an extension of the second inequality of Theorem 2.

In [N] we proved solvability of Dirichlet, Neumann as well as regularity type problems in  $\Omega$  with data in  $L^p$  and  $L^p_{1,1/2}$  (the parabolic Sobolev space having tangential (spatial) gradients and half a time derivative in  $L^p$ ) for  $p \in (2 - \epsilon, 2 + \epsilon)$  assuming that  $A(x, \cdot)$  is uniformly Lipschitz with respect to the time variable and that  $\|D^t_{1/2} A\|_* \leq \epsilon_0 < \infty$  for  $\epsilon_0$  small enough. I.e., with this additional smallness assumption on  $\|D^t_{1/2} A\|_*$  we established the control, in  $L^2$ , of the non-tangential maximal operator using the data. In [HL] it is proven, in the situation of the heat operator, that the smallness condition imposed is in fact necessary for  $L^2$ -solvability of the Dirichlet problem. We therefore emphasize that Theorem 1 and Theorem 2 should therefore be seen as steps towards an analysis of systems depending less on singular integrals.

Concerning the arguments used in this paper we note that in [KKPT] a new method, based on a notion of  $\epsilon$ -approximability of harmonic functions introduced and proved by Dahlberg [D1], was developed for the analysis of elliptic measure. In fact the authors consider  $L^p$ -solvability, in Lipschitz domains  $\Omega$ , of the Dirichlet problem for divergence form operators  $L = \text{div}(A\nabla)$  just assuming that  $A = A(X) = \{a_{ij}(X)\}$  is real, bounded and uniformly elliptic. In particular no symmetry condition on the matrix  $A$  is imposed and in [KKPT] it is proven that if bounded solutions of  $Lu = 0$  are  $\epsilon$ -approximable, then  $d\omega_L \in A_\infty(d\sigma)$ . We define  $d\omega_L \in A_\infty(d\sigma)$  to mean that  $\omega_L$  is absolutely continuous with respect to  $d\sigma$  and that the Radon-Nikodym derivative satisfies a scale-invariant reverse Hölder inequality in  $L^p(d\sigma)$  for some  $p > 1$ . Here  $\omega_L$  is the  $L$ -elliptic measure defined by means of the Riesz representation theorem. Hence the deep question of whether  $d\omega_L \in A_\infty(d\sigma)$  can, for symmetric as well as non-symmetric operators in divergence form, be reduced to establishing the existence of  $\epsilon$ -approximability for bounded solutions. In [D1] Dahlberg based his proof of the  $\epsilon$ -approximability for bounded solutions on the  $L^2$ -equivalence between the non-tangential maximum of  $u$  and square functions which he had previously established in [D2], and in [KKPT] it is proven that the same approach can be used in more general situations. In fact the construction of  $\epsilon$ -approximants due to Dahlberg applies to functions  $u$  for which the De Giorgio, Moser, Nash theorems (see [G], [M], [Na]) are valid and for which the appropriate  $L^2$ -equivalence between the non-tangential maximal operator and square functions can be established. Hence the construction of  $\epsilon$ -approximants of solutions to  $Lu$  is essentially reduced to an analysis of the non-tangential maximal operator and square functions. It is important to point out that in [KKPT] the authors only prove  $\epsilon$ -approximability, in the case of non-symmetric operators, in  $\mathbf{R}^2$ , i.e., the higher dimensional case remains open.

Furthermore, in [LS], [LM], [HL] caloric measure was studied in a class of time-varying and non-smooth cylinders and the mutual absolute continuity of caloric measure and a naturally defined surface measure was proven, and the parabolic

analogue of the  $L^2$ -integrability of the Poisson kernel due to Dahlberg [D] established. The results in [LM], [HL] rely heavily on the use of singular integrals and layer potential techniques. Taking an appropriate pullback, caloric measure can also be analyzed through an analysis of operators of the form

$$\tilde{L} = \operatorname{div}(\tilde{A}\nabla) - \partial_t - \tilde{B} \cdot \nabla.$$

The conditions on the time-varying cylinders are then translated into appropriate conditions on the coefficients  $\tilde{A}$ ,  $\tilde{B}$  (see [HL1], [R]) and it is worth noting that  $\tilde{A}$  is symmetric. In [R] a parabolic version of the program of [KKPT] was presented for the operator  $\tilde{L}$ . In particular  $\epsilon$ -approximability for bounded solutions of  $\tilde{L}u = 0$  is proven and the mutual absolute continuity of  $\tilde{L}$ -caloric measure analyzed. As a consequence, the results in [LM] on caloric measure associated to the heat equation is reproven without the use of singular integrals. Again the analysis is reduced to an analysis of the non-tangential maximal operator and square functions.

The arguments used in Theorem 1 and Theorem 2 are based on real-variable techniques similar to those explored in [KKPT]. To prove Theorem 1 we first use a non-trivial argument based on partial integration to establish the first inequality of Theorem 2 and an even weaker form, in  $L^2$  though, of the second inequality. Based on a stopping time argument we then proceed and establish the second inequality of Theorem 2 in  $L^p$ , for  $p > 2$ , from which the second part of Theorem 1 can be recovered.

The rest of this paper is organized as follows. Section 2 is of preliminary nature while in section 3 we prove Theorem 1 and Theorem 2.

## 2. PRELIMINARIES

In this section, which is of a preliminary nature, we collect a few technical results on Carleson measures, the non-tangential maximal function, square functions, parabolic sawtooth domains and interior regularity for the system.

**2.1. Carleson measures and parabolic lifting.** Let  $P(z) \in C_0^\infty(\mathbf{R}^n)$ . We furthermore assume that  $P(z)$  is a non-negative even function and that  $\int_{\mathbf{R}^n} P(z) dz = 1$ . I.e., we assume that  $P(z)$  is a parabolic approximation of the identity. Let  $d = n + 1$  and define

$$P_\lambda(z) = \lambda^{-d} P(\lambda^{-\alpha} z) = \lambda^{-d} P\left(\frac{x}{\lambda}, \frac{t}{\lambda^2}\right).$$

For a locally integrable function  $f$  we denote by  $P_\lambda f$  the naturally defined operation of convolution. Define a ‘parabolic’ lifting  $\rho(\lambda, x, t)$  from  $\mathbf{R}_+^{n+1} = \mathbf{R}_+ \times \mathbf{R}^{n-1} \times \mathbf{R}$  onto  $\Omega = \{(x_0, x, t) \in \mathbf{R} \times \mathbf{R}^{n-1} \times \mathbf{R} : x_0 > A(x, t)\}$  in the following way:

$$\rho(\lambda, x, t) = (\lambda + P_{\gamma\lambda} A(x, t), x, t), \quad \rho(0, x, t) = (A(x, t), x, t).$$

Here  $\gamma$  is a small parameter and we can adjust  $\gamma$ , as  $\|\nabla_x A\|_\infty < \infty$ , so that

$$\frac{1}{2} \leq 1 + \frac{\partial P_{\gamma\lambda} A(z)}{\partial \lambda} \leq 3/2.$$

The following lemma is crucial and incorporates the geometry information in an analytic and quantitative way (for a proof see [HL, pp. 365-366]).

**Lemma 3.** *Let  $\sigma, \theta$  be non-negative integers and let  $\phi = (\phi_1, \dots, \phi_{n-1})$  be a multiindex. Define  $\ell = \sigma + |\phi| + \theta$ . If  $\|A\|_{comm} \leq \beta < \infty$  and if we define measures in the following way,*

$$dv = \left( \frac{\partial^\ell P_{\gamma\lambda} A(x, t)}{\partial \lambda^\sigma \partial x^\phi \partial t^\theta} \right)^2 \lambda^{2\ell+2\theta-3} dx dt d\lambda,$$

then this is a Carleson measure on  $\mathbf{R}_+^{n+1}$  if either  $\sigma + \theta \geq 1$  or  $|\phi| \geq 2$  with

$$(i) \quad v(B_r(z) \times (0, r)) \leq Cr^d \gamma^{(2-2\phi-4\theta)} b^2 (1 + \beta)^2.$$

Here  $b = \|\mathbf{D}_n A\|_*$  if  $\theta \geq 1$  and  $b = 1$ , if  $\theta = 0$ . Moreover if  $\ell \geq 1$ , then

$$(ii) \quad \left\| \frac{\partial^\ell P_{\gamma\lambda} A}{\partial \lambda^\sigma \partial x^\phi \partial t^\theta} \right\|_\infty \leq C_1 \gamma^{(1-|\phi|-2\theta)} \lambda^{1-\ell-\theta} b(1 + \beta).$$

**2.2. Maximal functions, square functions and parabolic sawtooths.** In the arguments of this paper we have to be very careful with the definition of the cones and the non-tangential maximal functions which appear in our inequalities. In the following we introduce the necessary notation assuming that  $\Omega$  is a time-varying domain  $TV(A, a_1, a_2)$  for some  $a_1$  and  $a_2$ . Hence we are working in a fixed coordinate system with coordinates  $(x_0, x, t)$  and  $\Omega = \{(x_0, x, t) \in \mathbf{R} \times \mathbf{R}^{n-1} \times \mathbf{R} : x_0 > A(x, t)\}$ . To continue we let  $a > 0$  and  $(X, t) = (x_0, x, t) \in \partial\Omega$ . We let  $\tilde{\Gamma}_a(X, t)$  be the parabolic cone

$$\tilde{\Gamma}_a(X, t) = \{(y_0, y, s) \in \Omega : \|(y - x, t - s)\| < a|y_0 - x_0|\}$$

and note that  $\tilde{\Gamma}_{a_1}(X, t) \subseteq \tilde{\Gamma}_{a_2}(X, t)$  in case  $a_1 \leq a_2$ . If  $h$  is a function defined on  $\Omega$  we define the non-tangential maximal function  $\tilde{N}_*(h) = \tilde{N}_*^a(h) : \partial\Omega \rightarrow \mathbf{R}$  as

$$\tilde{N}_*^a(h)(X, t) = \sup_{(Y, s) \in \tilde{\Gamma}_a(X, t)} |h|(Y, s).$$

We also introduce appropriate truncated versions in the following way. Let  $r > 0$  and let the parabolic cone, truncated at height  $r$  and centered at  $(X, t) = (x_0, x, t) \in \partial\Omega$ , be defined as

$$\tilde{\Gamma}_{a,r}(X, t) = \{(y_0, y, s) \in \Omega : \|(y - x, t - s)\| < a|y_0 - x_0|, \quad y_0 - x_0 < r\}.$$

Similarly we define the truncated non-tangential maximal function  $\tilde{N}_*^{a,r}(h) : \partial\Omega \rightarrow \mathbf{R}$  as

$$\tilde{N}_*^{a,r}(h)(X, t) = \sup_{(Y, s) \in \tilde{\Gamma}_{a,r}(X, t)} |h|(Y, s).$$

To continue we let  $B_{a\lambda}(x, t) = \{(y, s) : \|(x - y, t - s)\| < a\lambda\}$  and define

$$\begin{aligned} \Gamma_a(x, t) &:= \{(\lambda, y, s), 0 < \lambda, (y, s) \in B_{a\lambda}(x, t)\}, \\ \Gamma_{a,r}(x, t) &:= \{(\lambda, y, s), 0 < \lambda < r, (y, s) \in B_{a\lambda}(x, t)\}. \end{aligned}$$

For a function  $g$  defined on  $\mathbf{R}_+^{n+1}$  and for  $a \geq 1$  fixed we also introduce the following maximal function  $N_*(g) = N_*^a(g) : \mathbf{R}_+^{n+1} \rightarrow \mathbf{R}$ :

$$N_*^a(g)(x, t) = \sup_{\Gamma_a(x, t)} |g|(\lambda, y, s).$$

Similarly we introduce the truncated maximal function as

$$N_*^{a,r}(g)(x, t) = \sup_{\Gamma_{a,r}(x, t)} |g|(\lambda, y, s).$$

Let  $\rho(\lambda, x, t)$  be the parabolic lifting introduced in the previous subsection. Note that if  $a$  and  $r$  are fixed numbers, then one can easily prove that

$$\tilde{\Gamma}_{\tilde{a}, \tilde{r}}(\rho(0, x, t)) \subset \rho(\Gamma_{a,r}(x, t)),$$

provided that  $\tilde{a}$  is sufficiently small depending on  $a$  and  $\|A\|_{\text{comm}}$  and  $\tilde{r} = r + P_{\gamma r}A(x, t)$ . Hence choosing  $\gamma$  small we can make sure that  $|\tilde{r} - r|$  is small.

Using the notation just introduced we define, for functions  $u$  defined in  $\Omega$ , the associate square function as well as the associated truncated square function as

$$S^a(u)(X, t) = \left( \int_{\tilde{\Gamma}_a(X,t)} |\nabla u(Y, s)|^2 \delta(Y, s)^{-n} dY ds \right)^{1/2},$$

$$S^{a,r}(u)(X, t) = \left( \int_{\tilde{\Gamma}_{a,r}(X,t)} |\nabla u(Y, s)|^2 \delta(Y, s)^{-n} dY ds \right)^{1/2}.$$

Here  $\delta(Y, s)$  is the parabolic distance from  $(Y, s) \in \Omega$  to  $\partial\Omega$ . Note that if  $F \subset \partial\Omega$ , then by the theorem of Fubini,

$$\int_F |S^a(u)(X, t)|^2 d\sigma(X, t) \sim \int_{\cup_{(X,t) \in F} \tilde{\Gamma}_a(X,t)} |\nabla u(Y, s)|^2 \delta(Y, s) dY ds,$$

$$\int_F |S^{a,r}(u)(X, t)|^2 d\sigma(X, t) \sim \int_{\cup_{(X,t) \in F} \tilde{\Gamma}_{a,r}(X,t)} |\nabla u(Y, s)|^2 \delta(Y, s) dY ds.$$

We will in the bulk of the paper make use of the following lemma proved in [R].

**Lemma 4.** *Let  $F$  be an arbitrary subset of  $\mathbf{R}^n$ . Define the sawtooth region above  $F$  as*

$$\Phi(F) = \bigcup_{(x,t) \in F} \Gamma_a(x, t).$$

*Then  $\Phi(F)$  is a time-varying domain  $TV(A, a_1, a_2)$  for some triple  $(A, a_1, a_2)$ . Furthermore,  $\max\{a_1, a_2\} \leq f(a)$  where  $f(a) \rightarrow 0$  as  $a \rightarrow \infty$ .*

Let  $C_1(0, 0) := \{(x, t) \in \mathbf{R}^n \times \mathbf{R}; |x_j| < 1; |t| < 1\}$  and let  $C_r(X, t)$  be the set  $C_1(0, 0)$  with center translated to the point  $(X, t)$  and scaled, in the parabolic sense, to a size of  $r$ . The only result on interior regularity that we will use in this paper is the following. The proof follows from standard arguments.

**Lemma 5.** *Let  $\vec{u}$  be a solution in  $\Omega \subset \mathbf{R}^{n+1}$  to the system in (1) and assume that (2) and (3) hold. If  $C_{2r}(X, t) \subset \Omega$ , then for all  $(Z, \tau) \in C_r(X, t)$  we have that*

$$|u^k(Z, \tau) - u^k(X, t)| \leq C \left[ \sum_s \frac{1}{r^n} \int_{C_{2r}(X,t)} |\nabla u^s(Y, s)|^2 dY ds \right]^{1/2}.$$

### 3. EQUIVALENCE BETWEEN NON-TANGENTIAL MAXIMAL OPERATORS AND THE SQUARE FUNCTIONS

If  $\Delta_r$  is a cube on  $\mathbf{R}^{n-1} \times \mathbf{R}$  of dimensions  $r \times \dots \times r \times r^2$  we let

$$\tilde{\Delta}_r = \{(A(x, t), x, t) : (x, t) \in \Delta_r\}$$

be the associated surface cube on the boundary of our domain  $\Omega$ . We furthermore define an associated Carleson box above  $\tilde{\Delta}_r$ , denoted by  $\tilde{T}(\Delta_r)$ , as

$$\tilde{T}(\Delta_r) = \{(x_0, x, t) : (x, t) \in \Delta_r, A(x, t) \leq x_0 \leq A(x, t) + r\}.$$

**3.1. The main technical component.** In this section we will prove one technical but crucial lemma, Lemma 6 below. Let  $\Delta_r$  be a cube on  $\mathbf{R}^{n-1} \times \mathbf{R}$  of dimensions  $r \times \dots \times r \times r^2$ . Based on  $\Delta_r$  we define the following sets:

$$\begin{aligned} Q_{r,\beta_1}^+ &= \{(x_0, x, t) : (x, t) \in \Delta_r, A(x, t) \leq x_0 \leq A(x, t) + \beta_1 r\}, \\ Q_{r,\beta_1}^- &= \{(x_0, x, t) : (x, t) \in \Delta_r, A(x, t) \geq x_0 \geq A(x, t) - \beta_1 r\}, \\ Q_{r,\beta_1} &= Q_{r,\beta_1}^+ \cup Q_{r,\beta_1}^-, \\ W_{r,\beta_2,\beta_3}^+ &= \{(x_0, x, t) : (x, t) \in \Delta_r, A(x, t) + \beta_3 r \leq x_0 \leq A(x, t) + \beta_2 r\}, \\ W_{r,\beta_2,\beta_3}^- &= \{(x_0, x, t) : (x, t) \in \Delta_r, A(x, t) - \beta_3 r \geq x_0 \leq A(x, t) - \beta_2 r\}. \end{aligned}$$

In the last definitions  $\beta_3 < \beta_2$ . Let  $\alpha$  be the  $\text{Lip}(1,1/2)$  constant of  $A$ . In the following we let  $\beta_2 = 2\alpha$  and  $\beta_3 = 3\alpha/2$ . We furthermore let  $\phi_2 \in C_0^\infty(Q_{r,\beta_2})$  be such that  $\phi_2 = \phi_2(x_0)$  and such that  $\phi_2$  has the property that  $\phi_2(x_0) = 1$  for all  $(x_0, x, t) \in Q_{r,\beta_2}$ . Note that  $\phi_2$  just depends on the coordinate  $x_0$ . For  $\xi > 0$  we pick  $\phi_1 \in C_0^\infty(\mathbf{R}^n)$  such that  $\phi_1(x, t) = 1$  for all  $(x, t) \in \Delta_r$  and  $\phi_1(x, t) = 0$  for all  $(x, t) \in \mathbf{R}^{n-1} \times \mathbf{R} \setminus \Delta_{(1+\xi/2)r}$ . Based on this we define a test function  $\theta \in C_0^\infty(\mathbf{R}^{n+1})$  as  $\theta(x_0, x, t) = \phi_1(x, t)\phi_2(x_0)$ . Hence  $\theta$  has a well defined product structure. Let the ‘parabolic’ lifting  $\rho(\lambda, x, t)$  from  $\mathbf{R}_+^{n+1} = \mathbf{R}_+ \times \mathbf{R}^{n-1} \times \mathbf{R}$  onto  $\Omega = \{(x_0, x, t) \in \mathbf{R} \times \mathbf{R}^{n-1} \times \mathbf{R} : x_0 > A(x, t)\}$  be defined as in the previous section. For  $\theta$  as above we then define  $\tilde{\theta} = \theta \circ \rho$ .

We are now ready to state the lemma, and in the statement of the lemma we use summation convention and we want to emphasize, in order to avoid confusion, that  $u^r$  denotes the component  $r$  in the vector  $\vec{u}$ . I.e., this  $r$  is unrelated to the scale parameter  $r$  appearing in the definition of  $\tilde{\Delta}_r$ .

**Lemma 6.** *Let  $a > 0$  and let  $\xi > 0$  be a small number. Let  $\Delta_r \subset \mathbf{R}^{n-1} \times \mathbf{R}$  and let  $\theta$  be the test function defined above. Then there exists a constant  $C = C(\|A\|_{\text{comm}}, a, \xi)$  and a constant  $\delta = \delta(\|A\|_{\text{comm}})$  such that*

$$\int_0^\infty \int_{\mathbf{R}^n} |u_{x_j}^r \circ \rho|^2 \tilde{\theta}^2 \lambda dz d\lambda \leq C \int_{\tilde{\Delta}_{(1+\xi)r}} |\tilde{N}_*^{a,(1+\xi)\delta r}(u^s)(X, t)|^2 d\sigma(X, t).$$

Also for every  $\rho > 0$  small there exists  $C = C(\|A\|_{\text{comm}}, a, \xi, \rho)$  such that

$$\begin{aligned} \int_{\mathbf{R}^n} |u^r \circ \rho|^2 \tilde{\theta}^2 dz &\leq C \int_0^\infty \int_{\mathbf{R}^n} |u_{x_j}^r \circ \rho|^2 \tilde{\theta}^2 \lambda dz d\lambda + \frac{C}{r^2} \int_{K(\tilde{\Delta}_{(1+\xi)r}, \delta)} (u^s(Y, s))^2 dY ds \\ &+ \rho \int_{\tilde{\Delta}_{(1+\xi)r}} |\tilde{N}_*^{a,(1+\xi)\delta r}(u^s)(X, t)|^2 d\sigma(X, t). \end{aligned}$$

Here  $K(\tilde{\Delta}_{(1+\xi)r}, \delta)$  is the support of the function  $\theta_{x_0}$  that is located in  $\Omega$  and in particular if  $(Y, s) \in K(\tilde{\Delta}_{(1+\xi)r}, \delta)$ , then  $\delta(Y, s) \geq \beta r$  for a universal constant  $\beta = \beta(\|A\|_{\text{comm}}, a, \xi, \delta)$ .

*Proof.* Define

$$Q^{rs} = A_{00}^{rs} + \sum_{(i,j), i \neq 0, j \neq 0} A_{ij}^{rs} \frac{\partial P_{\gamma\lambda} A}{\partial x_i} \frac{\partial P_{\gamma\lambda} A}{\partial x_j} - 2 \sum_{j, j \neq 0} A_{0j}^{rs} \frac{\partial P_{\gamma\lambda} A}{\partial x_j}.$$

Then  $Q = \{Q^{rs}\}$  is an  $m \times m$  matrix, and by the Legendre-Hadamard condition  $Q$  is a positive definite matrix with eigenvalues bounded from below by  $C(1 + |\nabla P_{\gamma\lambda} A|^2)$ . In order to prove the lemma we will start by manipulating the expression

$$I = - \int_{\mathbf{R}^n} Q^{rs}(u^r \circ \rho)(u^s \circ \rho) \tilde{\theta}^2 dz.$$

Here  $\{r, s\}$  are initially assumed fixed, but we will at certain instances also sum over these indices. Integrating, in  $I$ , once by parts in the  $\lambda$ -direction we have

$$\begin{aligned} I &= \int_0^\infty \int_{\mathbf{R}^n} Q_\lambda^{rs}(u^r \circ \rho)(u^s \circ \rho) \tilde{\theta}^2 dz d\lambda + \int_0^\infty \int_{\mathbf{R}^n} Q^{rs}(u^r \circ \rho)_\lambda (u^s \circ \rho) \tilde{\theta}^2 dz d\lambda \\ &+ \int_0^\infty \int_{\mathbf{R}^n} Q^{rs}(u^r \circ \rho)(u^s \circ \rho)_\lambda \tilde{\theta}^2 dz d\lambda + \int_0^\infty \int_{\mathbf{R}^n} Q^{rs}(u^r \circ \rho)(u^s \circ \rho) (\tilde{\theta}^2)_\lambda dz d\lambda. \end{aligned}$$

Integrating by parts once more in the  $\lambda$ -direction in the three first terms we have

$$\begin{aligned} I &= - \int_0^\infty \int_{\mathbf{R}^n} Q_{\lambda\lambda}^{rs}(u^r \circ \rho)(u^s \circ \rho) \tilde{\theta}^2 \lambda dz d\lambda - \int_0^\infty \int_{\mathbf{R}^n} Q^{rs}(u^r \circ \rho)_{\lambda\lambda} (u^s \circ \rho) \tilde{\theta}^2 \lambda dz d\lambda \\ &- \int_0^\infty \int_{\mathbf{R}^n} Q^{rs}(u^r \circ \rho)(u^s \circ \rho)_{\lambda\lambda} \tilde{\theta}^2 \lambda dz d\lambda - 2 \int_0^\infty \int_{\mathbf{R}^n} Q_\lambda^{rs}(u^r \circ \rho)(u^s \circ \rho)_\lambda \tilde{\theta}^2 \lambda dz d\lambda \\ &- 2 \int_0^\infty \int_{\mathbf{R}^n} Q_\lambda^{rs}(u^r \circ \rho)_\lambda (u^s \circ \rho) \tilde{\theta}^2 \lambda dz d\lambda - 2 \int_0^\infty \int_{\mathbf{R}^n} Q^{rs}(u^r \circ \rho)_\lambda (u^s \circ \rho)_\lambda \tilde{\theta}^2 \lambda dz d\lambda \\ &- \int_0^\infty \int_{\mathbf{R}^n} Q_\lambda^{rs}(u^r \circ \rho)(u^s \circ \rho) (\tilde{\theta}^2)_\lambda \lambda dz d\lambda - \int_0^\infty \int_{\mathbf{R}^n} Q^{rs}(u^r \circ \rho)_\lambda (u^s \circ \rho) (\tilde{\theta}^2)_\lambda \lambda dz d\lambda \\ &- \int_0^\infty \int_{\mathbf{R}^n} Q^{rs}(u^r \circ \rho)(u^s \circ \rho)_\lambda (\tilde{\theta}^2)_\lambda \lambda dz d\lambda + \int_0^\infty \int_{\mathbf{R}^n} Q^{rs}(u^r \circ \rho)(u^s \circ \rho) (\tilde{\theta}^2)_\lambda dz d\lambda. \end{aligned}$$

We introduce the following notation for this decomposition:

$$I = T_1 + T_2 + T_3 + T_4 + T_5 + T_6 + T_7 + T_8 + T_9 + T_{10}.$$

We also introduce the notation

$$\begin{aligned} \tilde{G}_1 &= \left( \int_{\mathbf{R}^n} |N_*((u^r \circ \rho)\tilde{\theta})|^2 dz \right)^{1/2} \left( \int_0^\infty \int_{\mathbf{R}^n} |u_{x_j}^s \circ \rho|^2 \tilde{\theta}^2 \lambda dz d\lambda \right)^{1/2}, \\ \tilde{G}_2 &= \left( \int_0^\infty \int_{\mathbf{R}^n} |(u^r \circ \rho)(\tilde{\theta}_\lambda)|^2 \lambda dz d\lambda \right)^{1/2} \left( \int_0^\infty \int_{\mathbf{R}^n} |u_{x_j}^s \circ \rho|^2 \tilde{\theta}^2 \lambda dz d\lambda \right)^{1/2}, \\ \tilde{G}_3 &= \int_0^\infty \int_{\mathbf{R}^n} |u^r \circ \rho|^2 |(\tilde{\theta}^2)_\lambda| \lambda dz d\lambda, \\ \tilde{G}_4 &= \left( \int_{\mathbf{R}^n} |N_*((u^r \circ \rho)\tilde{\theta})|^2 dz \right)^{1/2} \left( \int_0^\infty \int_{\mathbf{R}^n} |(u^r \circ \rho)(\tilde{\theta}_\lambda)|^2 \lambda dz d\lambda \right)^{1/2}. \end{aligned}$$

We note that by using Lemma 3

$$|T_4| + |T_5| + |T_7| + |T_8| + |T_9| + |T_{10}| \leq C(\tilde{G}_1 + \tilde{G}_2 + \tilde{G}_3 + \tilde{G}_4).$$

Here the  $G$  in  $\tilde{G}_i$  stands for Good, and as discussed further below these are terms that we do not have to manipulate any further. Hence we are left with the terms  $T_1, T_2, T_3$  and  $T_6$ . We start by analyzing  $T_1$ . The term  $T_1$  decouples into a linear combination of terms of the type

$$\begin{aligned} T_{11} &= \int_0^\infty \int_{\mathbf{R}^n} \frac{\partial^3 P_{\gamma\lambda} A}{\partial x_i \partial \lambda^2} (u^r \circ \rho)(u^s \circ \rho) \tilde{\theta}^2 \lambda dz d\lambda, \\ T_{12} &= \int_0^\infty \int_{\mathbf{R}^n} \frac{\partial^2 P_{\gamma\lambda} A}{\partial x_i \partial \lambda} \frac{\partial^2 P_{\gamma\lambda} A}{\partial x_j \partial \lambda} (u^r \circ \rho)(u^s \circ \rho) \tilde{\theta}^2 \lambda dz d\lambda, \\ T_{13} &= \int_0^\infty \int_{\mathbf{R}^n} \frac{\partial P_{\gamma\lambda} A}{\partial x_i} \frac{\partial^3 P_{\gamma\lambda} A}{\partial x_j \partial \lambda^2} (u^r \circ \rho)(u^s \circ \rho) \tilde{\theta}^2 \lambda dz d\lambda. \end{aligned}$$

In these expressions  $i \neq 0$  and  $j \neq 0$ . To handle  $T_{11}$  and  $T_{13}$  the idea is to lift the derivative w.r.t.  $x_i$  in  $\frac{\partial^3 P_{\gamma\lambda} A}{\partial x_i \partial \lambda^2}$ , respectively the derivative w.r.t  $x_j$  in  $\frac{\partial^3 P_{\gamma\lambda} A}{\partial x_j \partial \lambda^2}$ , using partial integration. To carry this out we start by analyzing  $T_{11}$ . Using partial integration we get

$$\begin{aligned} T_{11} &= - \int_0^\infty \int_{\mathbf{R}^n} \frac{\partial^2 P_{\gamma\lambda} A}{\partial \lambda^2} (u^r \circ \rho)_{x_i} (u^s \circ \rho) \tilde{\theta}^2 \lambda dz d\lambda \\ &\quad - \int_0^\infty \int_{\mathbf{R}^n} \frac{\partial^2 P_{\gamma\lambda} A}{\partial \lambda^2} (u^r \circ \rho)(u^s \circ \rho)_{x_i} \tilde{\theta}^2 \lambda dz d\lambda \\ &\quad - \int_0^\infty \int_{\mathbf{R}^n} \frac{\partial^2 P_{\gamma\lambda} A}{\partial \lambda^2} (u^r \circ \rho)(u^s \circ \rho) (\tilde{\theta}^2)_{x_i} \lambda dz d\lambda = T_{111} + T_{112} + T_{113}. \end{aligned}$$

Again using Lemma 3 we have  $|T_{111}| + |T_{112}| \leq C\tilde{G}_1$ . Define

$$\tilde{G}_5 = \left( \int_{\mathbf{R}^n} |N_*((u^r \circ \rho)\tilde{\theta})|^2 dz \right)^{1/2} \left( \int_0^\infty \int_{\mathbf{R}^n} |(u^r \circ \rho)(\tilde{\theta}_{x_i})|^2 \lambda dz d\lambda \right)^{1/2}.$$

Then  $|T_{113}| \leq \tilde{\rho}\tilde{G}_5$  where the  $\tilde{\rho}$  indicates that the constant can be chosen to be arbitrarily small. We note that  $\tilde{\rho}$  is appearing through the parameter  $\gamma$  in the Carleson measure in Lemma 3. Hence in total  $|T_{11}| \leq |T_{111}| + |T_{112}| + |T_{113}| \leq C\tilde{G}_1 + \tilde{\rho}\tilde{G}_5$ .

For  $T_{13}$  we similarly get

$$\begin{aligned} T_{13} = & - \int_0^\infty \int_{\mathbf{R}^n} \frac{\partial^2 P_{\gamma\lambda} A}{\partial x_i \partial x_j} \frac{\partial^2 P_{\gamma\lambda} A}{\partial \lambda^2} (u^r \circ \rho)(u^s \circ \rho) \tilde{\theta}^2 \lambda dz d\lambda \\ & - \int_0^\infty \int_{\mathbf{R}^n} \frac{\partial P_{\gamma\lambda} A}{\partial x_i} \frac{\partial^2 P_{\gamma\lambda} A}{\partial \lambda^2} (u^r \circ \rho)_{x_j} (u^s \circ \rho) \tilde{\theta}^2 \lambda dz d\lambda \\ & - \int_0^\infty \int_{\mathbf{R}^n} \frac{\partial P_{\gamma\lambda} A}{\partial x_i} \frac{\partial^2 P_{\gamma\lambda} A}{\partial \lambda^2} (u^r \circ \rho)(u^s \circ \rho)_{x_j} \tilde{\theta}^2 \lambda dz d\lambda \\ & - \int_0^\infty \int_{\mathbf{R}^n} \frac{\partial P_{\gamma\lambda} A}{\partial x_i} \frac{\partial^2 P_{\gamma\lambda} A}{\partial \lambda^2} (u^r \circ \rho)(u^s \circ \rho)(\tilde{\theta}^2)_{x_j} \lambda dz d\lambda. \end{aligned}$$

In order to compactify our notation we also introduce

$$\tilde{G}_6 = \int_{\mathbf{R}^n} |N_*((u^r \circ \rho)\tilde{\theta})|^2 dz.$$

Arguing as above we have  $|T_{13}| \leq C\tilde{G}_1 + \rho\tilde{G}_5 + \rho\tilde{G}_6$ .

Finally we focus on  $T_{12}$ . Partial integration w.r.t.  $x_i$  gives

$$T_{12} = T_{121} - \int_0^\infty \int_{\mathbf{R}^n} \frac{\partial P_{\gamma\lambda} A}{\partial \lambda} \frac{\partial^3 P_{\gamma\lambda} A}{\partial x_j \partial x_i \partial \lambda} (u^r \circ \rho)(u^s \circ \rho) \tilde{\theta}^2 \lambda dz d\lambda$$

where, as above,  $|T_{121}| \leq C\tilde{G}_1 + \tilde{\rho}\tilde{G}_5 + \tilde{\rho}\tilde{G}_6$ . But by once again using using Lemma 3 we get  $|T_{12} - T_{121}| \leq \tilde{\rho}\tilde{G}_6$  and we can conclude that in total  $|T_1| \leq C\tilde{G}_1 + \tilde{\rho}\tilde{G}_5 + \tilde{\rho}\tilde{G}_6$ .

We continue by manipulating  $T_2$  and  $T_3$ . By symmetry we only have to treat  $T_2$ . From now on let  $D = (1 + \frac{\partial P_{\gamma\lambda} A}{\partial \lambda})$ . By carrying out the differentiation we get

$$\begin{aligned} T_2 = & - \int_0^\infty \int_{\mathbf{R}^n} Q^{rs}(u^r_{x_0 x_0} \circ \rho)(u^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda \\ & - \int_0^\infty \int_{\mathbf{R}^n} Q^{rs}(u^r_{x_0} \circ \rho)(u^s \circ \rho) \frac{\partial^2 P_{\gamma\lambda} A}{\partial \lambda^2} \tilde{\theta}^2 \lambda dz d\lambda. \end{aligned}$$

As  $u_i^s = A_{ij}^{rs} u^r_{x_i x_j}$  we have

$$u_i^p = A_{00}^{pp} u^r_{x_0 x_0} + \sum_{(i,j,r),(i,j) \neq (0,0)} A_{ij}^{rp} u^r_{x_i x_j}.$$

We now change variables. Assume initially that  $i \neq 0, j \neq 0$ . Then

$$\begin{aligned} (u_{x_i x_j}^r \circ \rho) &= (u_{x_j}^r \circ \rho)_{x_i} - (u_{x_j x_0}^r \circ \rho) \frac{\partial P_{\gamma\lambda A}}{\partial x_i} \\ &= (u_{x_j}^r \circ \rho)_{x_i} - (u_{x_0}^r \circ \rho)_{x_j} \frac{\partial P_{\gamma\lambda A}}{\partial x_i} + (u_{x_0 x_0}^r \circ \rho) \frac{\partial P_{\gamma\lambda A}}{\partial x_i} \frac{\partial P_{\gamma\lambda A}}{\partial x_j}. \end{aligned}$$

Suppose that  $i = 0, j \neq 0$ . Then

$$(u_{x_0 x_j}^r \circ \rho) = (u_{x_0}^r \circ \rho)_{x_j} - (u_{x_0 x_0}^r \circ \rho) \frac{\partial P_{\gamma\lambda A}}{\partial x_j}.$$

Combining these formulas,

$$\begin{aligned} u_t^p \circ \rho &= A_{ij}^{rp} u_{x_i x_j}^r \circ \rho = A_{00}^{rp} u_{x_0 x_0}^r \circ \rho \\ &+ \sum_{(i,j,r), i \neq 0, j \neq 0} A_{ij}^{rp} [(u_{x_j}^r \circ \rho)_{x_i} - (u_{x_0}^r \circ \rho)_{x_j} \frac{\partial P_{\gamma\lambda A}}{\partial x_i} \\ &+ (u_{x_0 x_0}^r \circ \rho) \frac{\partial P_{\gamma\lambda A}}{\partial x_i} \frac{\partial P_{\gamma\lambda A}}{\partial x_j}] \\ &+ 2 \sum_{(j,r), j \neq 0} A_{0j}^{rp} [(u_{x_0}^r \circ \rho)_{x_j} - (u_{x_0 x_0}^r \circ \rho) \frac{\partial P_{\gamma\lambda A}}{\partial x_j}]. \end{aligned}$$

Grouping the terms properly,

$$\begin{aligned} u_t^p \circ \rho &= \sum_r Q^{rp} u_{x_0 x_0}^r \circ \rho + \sum_{(i,j,r), i \neq 0, j \neq 0} A_{ij}^{rp} [(u_{x_j}^r \circ \rho)_{x_i} - (u_{x_0}^r \circ \rho)_{x_j} \frac{\partial P_{\gamma\lambda A}}{\partial x_i}] \\ &+ 2 \sum_{(j,r), j \neq 0} A_{0j}^{rp} (u_{x_0}^r \circ \rho)_{x_j}. \end{aligned}$$

Importing this formula into the formula for  $T_2$  and putting  $p = s$  we have,

$$T_2 = T_{21} + T_{22} + T_{23} + T_{24} + T_{25}$$

where

$$\begin{aligned} T_{21} &= - \int_0^\infty \int_{\mathbf{R}^n} (u_t^s \circ \rho)(u^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda, \\ T_{22} &= \sum_{(i,j,r), i \neq 0, j \neq 0} \int_0^\infty \int_{\mathbf{R}^n} A_{ij}^{rs} (u_{x_j}^r \circ \rho)_{x_i} (u^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda, \\ T_{23} &= - \sum_{(i,j,r), i \neq 0, j \neq 0} \int_0^\infty \int_{\mathbf{R}^n} A_{ij}^{rs} (u_{x_0}^r \circ \rho)_{x_j} \frac{\partial P_{\gamma\lambda A}}{\partial x_i} (u^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda, \\ T_{24} &= 2 \sum_{(j,r), j \neq 0} \int_0^\infty \int_{\mathbf{R}^n} A_{0j}^{rs} (u_{x_0}^r \circ \rho)_{x_j} (u^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda, \\ T_{25} &= - \int_0^\infty \int_{\mathbf{R}^n} Q^{rs} (u_{x_0}^r \circ \rho)(u^s \circ \rho) \frac{\partial^2 P_{\gamma\lambda A}}{\partial \lambda^2} \tilde{\theta}^2 \lambda dz d\lambda. \end{aligned}$$

We immediately see that  $|T_{25}| \leq C\tilde{G}_1$ . In fact by partial integration in the second, third and fourth term of this expression we get

$$\begin{aligned} T_2 - T_{21} &+ \sum_{(i,j,r), i \neq 0, j \neq 0} \int_0^\infty \int_{\mathbf{R}^n} A_{ij}^{rs}(u_{x_j}^r \circ \rho)(u^s \circ \rho)_{x_i} D^2 \tilde{\theta}^2 \lambda dz d\lambda \\ &- \sum_{(i,j,r), i \neq 0, j \neq 0} \int_0^\infty \int_{\mathbf{R}^n} A_{ij}^{rs}(u_{x_0}^r \circ \rho) \frac{\partial P_{\gamma\lambda} A}{\partial x_i}(u^s \circ \rho)_{x_j} D^2 \tilde{\theta}^2 \lambda dz d\lambda \\ &+ 2 \sum_{(j,r), j \neq 0} \int_0^\infty \int_{\mathbf{R}^n} A_{0j}^{rs}(u_{x_0}^r \circ \rho)(u^s \circ \rho)_{x_j} D^2 \tilde{\theta}^2 \lambda dz d\lambda \\ &= G. \end{aligned}$$

Here  $G$  denotes a sum of terms such that  $|G| \leq C(\tilde{G}_1 + \tilde{G}_7)$  where

$$\tilde{G}_7 = \left( \int_0^\infty \int_{\mathbf{R}^n} |u_{x_j}^s \circ \rho|^2 \tilde{\theta}^2 \lambda dz d\lambda \right)^{1/2} \left( \int_0^\infty \int_{\mathbf{R}^n} |(u^r \circ \rho)(\tilde{\theta}_{x_i})|^2 \lambda dz d\lambda \right)^{1/2}.$$

Let  $\Lambda$  be the set of index  $\{(i, j, r), i \neq 0, j \neq 0\}$ . By expanding each of the terms we can conclude that

$$T_2 - T_{21} - G = \tilde{T}_{21} + \tilde{T}_{22} + \tilde{T}_{23} + \tilde{T}_{24} + \tilde{T}_{25} + \tilde{T}_{26}$$

where

$$\begin{aligned} \tilde{T}_{21} &= - \sum_{\Lambda} \int_0^\infty \int_{\mathbf{R}^n} A_{ij}^{rs}(u_{x_j}^r \circ \rho)(u_{x_i}^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda, \\ \tilde{T}_{22} &= - \sum_{\Lambda} \int_0^\infty \int_{\mathbf{R}^n} A_{ij}^{rs}(u_{x_j}^r \circ \rho)(u_{x_0}^s \circ \rho) \frac{\partial P_{\gamma\lambda} A}{\partial x_i} D^2 \tilde{\theta}^2 \lambda dz d\lambda, \\ \tilde{T}_{23} &= \sum_{\Lambda} \int_0^\infty \int_{\mathbf{R}^n} A_{ij}^{rs}(u_{x_0}^r \circ \rho) \frac{\partial P_{\gamma\lambda} A}{\partial x_i}(u_{x_j}^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda, \\ \tilde{T}_{24} &= \sum_{\Lambda} \int_0^\infty \int_{\mathbf{R}^n} A_{ij}^{rs}(u_{x_0}^r \circ \rho) \frac{\partial P_{\gamma\lambda} A}{\partial x_i} \frac{\partial P_{\gamma\lambda} A}{\partial x_j}(u_{x_0}^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda, \\ \tilde{T}_{25} &= -2 \sum_{(j,r), j \neq 0} \int_0^\infty \int_{\mathbf{R}^n} A_{0j}^{rs}(u_{x_0}^r \circ \rho)(u_{x_j}^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda, \\ \tilde{T}_{26} &= -2 \sum_{(j,r), j \neq 0} \int_0^\infty \int_{\mathbf{R}^n} A_{0j}^{rs}(u_{x_0}^r \circ \rho)(u_{x_0}^s \circ \rho) \frac{\partial P_{\gamma\lambda} A}{\partial x_j} D^2 \tilde{\theta}^2 \lambda dz d\lambda. \end{aligned}$$

Note that if we sum over the indicies  $(r, s)$ , then  $\tilde{T}_{22} + \tilde{T}_{23} = 0$ . In the following we will therefore neglect these two terms and assume that we are summing over all

indices. Note that  $\tilde{T}_{21} + \tilde{T}_{24} + \tilde{T}_{25} + \tilde{T}_{26}$  equals

$$\begin{aligned} & \int_0^\infty \int_{\mathbf{R}^n} Q^{rs}(u_{x_0}^r \circ \rho)(u_{x_0}^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda - \int_0^\infty \int_{\mathbf{R}^n} A_{00}^{rs}(u_{x_0}^r \circ \rho)(u_{x_0}^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda \\ & - \int_0^\infty \int_{\mathbf{R}^n} A_{ij}^{rs}(u_{x_j}^r \circ \rho)(u_{x_i}^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda + \int_0^\infty \int_{\mathbf{R}^n} A_{00}^{rs}(u_{x_0}^r \circ \rho)(u_{x_0}^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda \\ & = \int_0^\infty \int_{\mathbf{R}^n} Q^{rs}(u_{x_0}^r \circ \rho)(u_{x_0}^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda - \int_0^\infty \int_{\mathbf{R}^n} A_{ij}^{rs}(u_{x_j}^r \circ \rho)(u_{x_i}^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda. \end{aligned}$$

Hence we can assume that

$$\begin{aligned} T_2 = -T_{21} - G & + \int_0^\infty \int_{\mathbf{R}^n} Q^{rs}(u_{x_0}^r \circ \rho)(u_{x_0}^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda \\ & - \int_0^\infty \int_{\mathbf{R}^n} A_{ij}^{rs}(u_{x_j}^r \circ \rho)(u_{x_i}^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda. \end{aligned}$$

Summarizing (assuming a similar derivation for  $T_3$ ) we have proved that

$$\begin{aligned} I & = T_1 + T_4 + T_5 + T_6 + G + T_7 + T_8 + T_9 + T_{10} \\ & + 2 \int_0^\infty \int_{\mathbf{R}^n} Q^{rs}(u_{x_0}^r \circ \rho)(u_{x_0}^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda \\ & - 2 \int_0^\infty \int_{\mathbf{R}^n} A_{ij}^{rs}(u_{x_j}^r \circ \rho)(u_{x_i}^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda \\ & - \int_0^\infty \int_{\mathbf{R}^n} (u_t^s \circ \rho)(u^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda - \int_0^\infty \int_{\mathbf{R}^n} (u_t^r \circ \rho)(u^r \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda. \end{aligned}$$

As

$$T_6 + 2 \int_0^\infty \int_{\mathbf{R}^n} Q^{rs}(u_{x_0}^r \circ \rho)(u_{x_0}^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda = 0$$

we can simplify one step further and conclude that

$$\begin{aligned} I & = T_1 + T_4 + T_5 + G + T_7 + T_8 + T_9 + T_{10} \\ & - 2 \int_0^\infty \int_{\mathbf{R}^n} A_{ij}^{rs}(u_{x_j}^r \circ \rho)(u_{x_i}^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda - 2 \int_0^\infty \int_{\mathbf{R}^n} (u_t^s \circ \rho)(u^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda. \end{aligned}$$

Using the deductions above,

$$\begin{aligned} I + 2 \int_0^\infty \int_{\mathbf{R}^n} A_{ij}^{rs}(u_{x_j}^r \circ \rho)(u_{x_i}^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda \\ = -2 \int_0^\infty \int_{\mathbf{R}^n} (u_t^s \circ \rho)(u^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda + E \end{aligned}$$

where

$$|E| \leq C(\tilde{G}_1 + \tilde{G}_2 + \tilde{G}_3 + \tilde{G}_4 + \tilde{G}_7) + \tilde{\rho}\tilde{G}_5 + \tilde{\rho}\tilde{G}_6.$$

Left is therefore essential to treat

$$T_{21} = - \int_0^\infty \int_{\mathbf{R}^n} (u_t^s \circ \rho)(u^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda.$$

But

$$\begin{aligned} T_{21} &= - \int_0^\infty \int_{\mathbf{R}^n} (u^s \circ \rho)_t (u^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda \\ &\quad + \int_0^\infty \int_{\mathbf{R}^n} (u^s \circ \rho)_\lambda (u^s \circ \rho) \frac{\partial P_{\gamma\lambda} A}{\partial t} D \tilde{\theta}^2 \lambda dz d\lambda = T_{211} + T_{212}. \end{aligned}$$

By familiar arguments  $|T_{212}| \leq C\tilde{G}_1$ . Integrating by parts w.r.t.  $\lambda$  in  $T_{211}$  we get

$$\begin{aligned} T_{211} &= \int_0^\infty \int_{\mathbf{R}^n} (u^s \circ \rho)_{\lambda t} (u^s \circ \rho) D^2 \tilde{\theta}^2 \lambda^2 dz d\lambda + \int_0^\infty \int_{\mathbf{R}^n} (u^s \circ \rho)_t (u^s \circ \rho)_\lambda D^2 \tilde{\theta}^2 \lambda^2 dz d\lambda \\ &\quad + 2 \int_0^\infty \int_{\mathbf{R}^n} (u^s \circ \rho)_t (u^s \circ \rho) D \frac{\partial^2 P_{\gamma\lambda} A}{\partial \lambda^2} \lambda^2 \tilde{\theta}^2 dz d\lambda \\ &\quad + \int_0^\infty \int_{\mathbf{R}^n} (u^s \circ \rho)_t (u^s \circ \rho) D^2 (\tilde{\theta}^2)_\lambda \lambda^2 dz d\lambda \\ &= T_{2111} + T_{2112} + T_{2113} + T_{2114}. \end{aligned}$$

By symmetrization we can assume that  $T_{2111} + T_{2112}$  has the form

$$\int_0^\infty \int_{\mathbf{R}^n} [(u^s \circ \rho)(u^s \circ \rho)]_{t\lambda} \left(1 + 2 \frac{\partial P_{\gamma\lambda} A}{\partial \lambda} + \left(\frac{\partial P_{\gamma\lambda} A}{\partial \lambda}\right)^2\right) \tilde{\theta}^2 \lambda^2 dz d\lambda.$$

Partial integration w.r.t.  $t$  in this formula we get (ignoring the minus sign)

$$\begin{aligned} &\int_0^\infty \int_{\mathbf{R}^n} [(u^s \circ \rho)(u^s \circ \rho)]_\lambda \frac{\partial}{\partial t} \left(1 + 2 \frac{\partial P_{\gamma\lambda} A}{\partial \lambda} + \left(\frac{\partial P_{\gamma\lambda} A}{\partial \lambda}\right)^2\right) \tilde{\theta}^2 \lambda^2 dz d\lambda \\ &+ \int_0^\infty \int_{\mathbf{R}^n} [(u^s \circ \rho)(u^s \circ \rho)]_\lambda \left(1 + 2 \frac{\partial P_{\gamma\lambda} A}{\partial \lambda} + \left(\frac{\partial P_{\gamma\lambda} A}{\partial \lambda}\right)^2\right) (\tilde{\theta}^2)_t \lambda^2 dz d\lambda. \end{aligned}$$

Using by now familiar arguments we can conclude that the first term is bounded by  $\tilde{G}_1$ . Introducing

$$\tilde{G}_8 = \left( \int_0^\infty \int_{\mathbf{R}^n} |(u^r \circ \rho)(\tilde{\theta}_t)|^2 \lambda^3 dz d\lambda \right)^{1/2} \left( \int_0^\infty \int_{\mathbf{R}^n} |u_{x_j}^s \circ \rho|^2 \tilde{\theta}^2 \lambda dz d\lambda \right)^{1/2}$$

we similarly get that the second term is bounded by  $C\tilde{G}_8$  and we can conclude that

$$|T_{2111} + T_{2112}| \leq C(\tilde{G}_1 + \tilde{G}_8).$$

We also use the equation and familiar manipulations

$$|T_{2113}| + |T_{2114}| \leq \tilde{\rho}\tilde{G}_9 + C\tilde{G}_{10}$$

where

$$\begin{aligned} \tilde{G}_9 &= \left( \int_0^\infty \int_{\mathbf{R}^n} |(u^s \circ \rho)_t|^2 \tilde{\theta}^2 \lambda^3 dz d\lambda \right)^{1/2} \left( \int_0^\infty \int_{\mathbf{R}^n} |N_*(\tilde{\theta}(u^s \circ \rho))|^2 dz \right)^{1/2}, \\ \tilde{G}_{10} &= \left( \int_0^\infty \int_{\mathbf{R}^n} |(u^s \circ \rho)_t|^2 \tilde{\theta}^2 \lambda^3 dz d\lambda \right)^{1/2} \left( \int_0^\infty \int_{\mathbf{R}^n} |(u^s \circ \rho)|^2 (\tilde{\theta}_\lambda)^2 \lambda dz \right)^{1/2}. \end{aligned}$$

Summarizing all of our estimates we have proved that

$$\begin{aligned} &\left| - \int_{\mathbf{R}^n} Q^{rs}(u^r \circ \rho)(u^s \circ \rho) \tilde{\theta}^2 dz + 2 \int_0^\infty \int_{\mathbf{R}^n} A_{ij}^{rs}(u_{x_j}^r \circ \rho)(u_{x_i}^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda \right| \\ &\leq C(\tilde{G}_1 + \tilde{G}_2 + \tilde{G}_3 + \tilde{G}_4 + \tilde{G}_7 + \tilde{G}_8 + \tilde{G}_{10}) + \tilde{\rho}\tilde{G}_5 + \tilde{\rho}\tilde{G}_6 + \tilde{\rho}\tilde{G}_9 \end{aligned}$$

where  $\tilde{G}_j, j \in \{1, 2, \dots, 10\}$ , is defined above. Before we can use the lengthy derivations given above to prove the two inequalities stated in the lemma we also need the following lemma. Summation convention applies.

**Lemma 7.**

$$\begin{aligned} &\int_0^\infty \int_{\mathbf{R}^n} |(u^s \circ \rho)_t|^2 \tilde{\theta}^2 \lambda^3 dz d\lambda \\ &\leq C \left( \int_0^\infty \int_{\mathbf{R}^n} |u_{x_j}^r \circ \rho|^2 \tilde{\theta}^2 \lambda dz d\lambda + \int_{\tilde{\Delta}_{(1+\xi)r}} |\tilde{N}_*^{a,(1+\xi)\delta r}(u^s)(X, t)|^2 d\sigma(X, t) \right). \end{aligned}$$

Using this lemma we will now complete the proof of Lemma 6. We first prove that the square function is dominated by the non-tangential maximum. By the strong ellipticity condition in (3) and the inequality derived above we get

$$\begin{aligned} &\int_0^\infty \int_{\mathbf{R}^n} |u_{x_j}^r \circ \rho|^2 \tilde{\theta}^2 \lambda dz d\lambda \leq C \int_0^\infty \int_{\mathbf{R}^n} A_{ij}^{rs}(u_{x_j}^r \circ \rho)(u_{x_i}^s \circ \rho) D^2 \tilde{\theta}^2 \lambda dz d\lambda \\ &\leq \int_{\mathbf{R}^n} |u^r \circ \rho|^2 \tilde{\theta}^2 dz + C(\tilde{G}_1 + \tilde{G}_2 + \tilde{G}_3 + \tilde{G}_4 + \tilde{G}_7 + \tilde{G}_8 + \tilde{G}_{10}) \\ &\quad + \tilde{\rho}\tilde{G}_5 + \tilde{\rho}\tilde{G}_6 + \tilde{\rho}\tilde{G}_9. \end{aligned}$$

Note that by simple estimates and by using the support properties of  $\tilde{\theta}_\lambda$  we have that

$$\tilde{G}_3 + \tilde{G}_4 + \tilde{G}_5 + \tilde{G}_6 \leq C \int_{\tilde{\Delta}_{(1+\xi)r}} |\tilde{N}_*^{a,(1+\xi)\delta r}(u^s)(X,t)|^2 d\sigma(X,t).$$

Furthermore, for arbitrary  $\epsilon_1 > 0$ ,

$$\begin{aligned} \tilde{G}_1 + \tilde{G}_2 + \tilde{G}_7 + \tilde{G}_8 &\leq C(\epsilon_1) \int_{\tilde{\Delta}_{(1+\xi)r}} |\tilde{N}_*^{a,(1+\xi)\delta r}(u^s)(X,t)|^2 d\sigma(X,t) \\ &\quad + C\epsilon_1 \int_0^\infty \int_{\mathbf{R}^n} |u_{x_j}^r \circ \rho|^2 \tilde{\theta}^2 \lambda dz d\lambda. \end{aligned}$$

Similarly for arbitrary  $\epsilon_2 > 0$  we get by using Lemma 7 and the support properties of  $\tilde{\theta}_\lambda$

$$\begin{aligned} \tilde{G}_9 + \tilde{G}_{10} &\leq C(\epsilon_2) \int_{\tilde{\Delta}_{(1+\xi)r}} |\tilde{N}_*^{a,(1+\xi)\delta r}(u^s)(X,t)|^2 d\sigma(X,t) \\ &\quad + C\epsilon_2 \int_0^\infty \int_{\mathbf{R}^n} |(u^s \circ \rho)_t|^2 \tilde{\theta}^2 \lambda^3 dz d\lambda \\ &\leq C(\epsilon_2) \int_{\tilde{\Delta}_{(1+\xi)r}} |\tilde{N}_*^{a,(1+\xi)\delta r}(u^s)(X,t)|^2 d\sigma(X,t) \\ &\quad + C\epsilon_2 \int_0^\infty \int_{\mathbf{R}^n} |u_{x_j}^r \circ \rho|^2 \tilde{\theta}^2 \lambda dz d\lambda. \end{aligned}$$

I.e.,

$$\begin{aligned} \int_0^\infty \int_{\mathbf{R}^n} |u_{x_j}^r \circ \rho|^2 \tilde{\theta}^2 \lambda dz d\lambda &\leq C(\epsilon_1, \epsilon_2) \int_{\tilde{\Delta}_{(1+\xi)r}} |\tilde{N}_*^{a,(1+\xi)\delta r}(u^s)(X,t)|^2 d\sigma(X,t) \\ &\quad + C\epsilon_1 \int_0^\infty \int_{\mathbf{R}^n} |u_{x_j}^r \circ \rho|^2 \tilde{\theta}^2 \lambda dz d\lambda + C\epsilon_2 \int_0^\infty \int_{\mathbf{R}^n} |u_{x_j}^r \circ \rho|^2 \tilde{\theta}^2 \lambda dz d\lambda. \end{aligned}$$

Appropriate choices of  $\epsilon_1$  and  $\epsilon_2$  complete the proof of the first part of Lemma 6.

Finally we prove the second inequality in Lemma 6. In this case the starting point is the estimate

$$\begin{aligned} \int_{\mathbf{R}^n} |u^r \circ \rho|^2 \tilde{\theta}^2 dz &\leq \int_0^\infty \int_{\mathbf{R}^n} |u_{x_j}^r \circ \rho|^2 \tilde{\theta}^2 \lambda dz d\lambda \\ &\quad + C(\tilde{G}_1 + \tilde{G}_2 + \tilde{G}_3 + \tilde{G}_4 + \tilde{G}_7 + \tilde{G}_8 + \tilde{G}_{10}) \\ &\quad + \tilde{\rho}\tilde{G}_5 + \tilde{\rho}\tilde{G}_6 + \tilde{\rho}\tilde{G}_9. \end{aligned}$$

Here we have to be a bit careful when we proceed, as we want to end up with, on the r.h.s., a term

$$\rho \int_{\tilde{\Delta}_{(1+\xi)r}} |\tilde{N}_*^{a,(1+\xi)\delta r}(u^s)(X, t)|^2 d\sigma(X, t)$$

where  $\rho$  can be chosen arbitrary small. Let  $\epsilon_1 > 0$  and  $\epsilon_2 > 0$  be arbitrary. Then arguing as above

$$\begin{aligned} \tilde{G}_1 + \tilde{G}_2 + \tilde{G}_7 + \tilde{G}_8 &\leq C\epsilon_1 \int_{\tilde{\Delta}_{(1+\xi)r}} |\tilde{N}_*^{a,(1+\xi)\delta r}(u^s)(X, t)|^2 d\sigma(X, t) \\ &\quad + C(\epsilon_1) \int_0^\infty \int_{\mathbf{R}^n} |u_{x_j}^r \circ \rho|^2 \tilde{\theta}^2 \lambda dz d\lambda. \end{aligned}$$

Also

$$\begin{aligned} \tilde{G}_3 + \tilde{G}_4 &\leq C\epsilon_2 \int_{\tilde{\Delta}_{(1+\xi)r}} |\tilde{N}_*^{a,(1+\xi)\delta r}(u^s)(X, t)|^2 d\sigma(X, t) \\ &\quad + \frac{C(\epsilon_2)}{r^2} \int_{K(\tilde{\Delta}_{(1+\xi)r, \delta})} (u^s(Y, s))^2 dY ds. \end{aligned}$$

Here  $K(\tilde{\Delta}_{(1+\xi)r, \delta})$  is the support of the function  $\theta_{x_0}$  that is located in  $\Omega$  and in particular if  $(Y, s) \in K(\tilde{\Delta}_{(1+\xi)r, \delta})$ , then  $\delta(Y, s) \geq \beta r$  for a universal constant  $\beta = \beta(\|A\|_{\text{comm}}, a, \xi, \delta)$ .

Continuing,

$$\tilde{G}_5 + \tilde{G}_6 \leq C \int_{\tilde{\Delta}_{(1+\xi)r}} |\tilde{N}_*^{a,(1+\xi)\delta r}(u^s)(X, t)|^2 d\sigma(X, t).$$

Now it only remains to estimate  $\tilde{G}_9$  and  $\tilde{G}_{10}$ . To do this we let  $\epsilon_3 > 0$  be arbitrary. Using Lemma 7 and the support properties of  $\tilde{\theta}_\lambda$

$$\tilde{G}_9 \leq C \int_{\tilde{\Delta}_{(1+\xi)r}} |\tilde{N}_*^{a,(1+\xi)\delta r}(u^s)(X, t)|^2 d\sigma(X, t) + C \int_0^\infty \int_{\mathbf{R}^n} |u_{x_j}^r \circ \rho|^2 \tilde{\theta}^2 \lambda dz d\lambda.$$

Also

$$\begin{aligned} \tilde{G}_{10} &\leq \frac{C(\epsilon_3)}{r^2} \int_{K(\tilde{\Delta}_{(1+\xi)r, \delta})} (u^s(Y, s))^2 dY ds + C\epsilon_3 \int_0^\infty \int_{\mathbf{R}^n} |(u^r \circ \rho)_t|^2 \tilde{\theta}^2 \lambda^3 dz d\lambda \\ &\leq \frac{C(\epsilon_3)}{r^2} \int_{K(\tilde{\Delta}_{(1+\xi)r, \delta})} (u^s(Y, s))^2 dY ds + C\epsilon_3 \int_{\tilde{\Delta}_{(1+\xi)r}} |\tilde{N}_*^{a,(1+\xi)\delta r}(u^s)(X, t)|^2 d\sigma(X, t) \\ &\quad + C(\epsilon_3) \int_0^\infty \int_{\mathbf{R}^n} |u_{x_j}^r \circ \rho|^2 \tilde{\theta}^2 \lambda dz d\lambda. \end{aligned}$$

Adding it all up we get

$$\int_{\mathbf{R}^n} |u^r \circ \rho|^2 \tilde{\theta}^2 dz \leq C \int_0^\infty \int_{\mathbf{R}^n} |u_{x_j}^r \circ \rho|^2 \tilde{\theta}^2 \lambda dz d\lambda + \frac{C}{r^2} \int_{K(\tilde{\Delta}_{(1+\xi)r}, \delta)} (u^s(Y, s))^2 dY ds$$

$$+ (\epsilon_1 + \epsilon_2 + \epsilon_3 + \tilde{\rho}) C \int_{\tilde{\Delta}_{(1+\xi)r}} |\tilde{N}_*^{a, (1+\xi)\delta r}(u^s)(X, t)|^2 d\sigma(X, t).$$

As  $\epsilon_1 + \epsilon_2 + \epsilon_3 + \tilde{\rho}$  can be made arbitrary small we can choose  $\rho = (\epsilon_1 + \epsilon_2 + \epsilon_3 + \tilde{\rho})C$  arbitrary small, and the proof is complete.  $\square$

*Proof of Lemma 7.* Unfortunately, in order to prove Lemma 7 it does seem inevitable to repeat many of the deductions carried out above but in a different order. Still we believe that it is more or less impossible for the reader to reconstruct the estimate unless we give a complete proof.

Let

$$I = \int_0^\infty \int_{\mathbf{R}^n} |(u^s \circ \rho)_t|^2 \tilde{\theta}^2 \lambda^3 dz d\lambda = \int_0^\infty \int_{\mathbf{R}^n} (u^s \circ \rho)_t (u^s \circ \rho)_t \tilde{\theta}^2 \lambda^3 dz d\lambda.$$

Then

$$I = \int_0^\infty \int_{\mathbf{R}^n} (u^s \circ \rho)_t (u_t^s \circ \rho) \tilde{\theta}^2 \lambda^3 dz d\lambda$$

$$+ \int_0^\infty \int_{\mathbf{R}^n} (u^s \circ \rho)_t (u_{x_0}^s \circ \rho) \frac{\partial P_{\gamma\lambda} A}{\partial t} \tilde{\theta}^2 \lambda^3 dz d\lambda = I_1 + I_2.$$

Using Lemma 3,

$$|I_2| \leq CI^{1/2} \left( \int_0^\infty \int_{\mathbf{R}^n} |u_{x_j}^r \circ \rho|^2 \tilde{\theta}^2 \lambda dz d\lambda \right)^{1/2}$$

so we only have to focus on  $I_1$ . But using the equation as in the proof of Lemma 6 we have

$$I_1 = I_{11} + I_{12} + I_{13},$$

$$I_{11} = \int_0^\infty \int_{\mathbf{R}^n} Q^{rs} (u^s \circ \rho)_t (u_{x_0 x_0}^r \circ \rho) \tilde{\theta}^2 \lambda^3 dz d\lambda,$$

$$I_{12} = \sum_{(i,j,r), i \neq 0, j \neq 0} \int_0^\infty \int_{\mathbf{R}^n} A_{ij}^{rs} (u^s \circ \rho)_t [(u_{x_j}^r \circ \rho)_{x_i} - (u_{x_0}^r \circ \rho)_{x_j} \frac{\partial P_{\gamma\lambda} A}{\partial x_i}] \tilde{\theta}^2 \lambda^3 dz d\lambda,$$

$$I_{13} = 2 \sum_{(j,r), j \neq 0} \int_0^\infty \int_{\mathbf{R}^n} A_{0j}^{rs} (u^s \circ \rho)_t (u_{x_0}^r \circ \rho)_{x_j} \tilde{\theta}^2 \lambda^3 dz d\lambda.$$

The way the proof proceeds is as follows. For each  $k \in \{1, 2, 3\}$  we, by partial integration, prove that  $I_{1k} = -I_{1k} + G$  where  $G$  is either  $\epsilon I$ , with  $\epsilon > 0$  arbitrary,

or a constant times one of the following:

$$\int_0^\infty \int_{\mathbf{R}^n} |u_{x_j}^r \circ \rho|^2 \tilde{\theta}^2 \lambda dz d\lambda, \quad \int_{\tilde{\Delta}_{(1+\xi)r}} |\tilde{N}_*^{a,(1+\xi)\delta r}(u^s)(X, t)|^2 d\sigma(X, t).$$

In fact we will only prove this for  $I_{13}$  and then state that the same argument is valid for  $I_{11}$  as well as  $I_{12}$ . In the following deduction we will ignore the summation sign in  $I_{13}$ . Starting out by lifting the  $x_j$  derivative we get

$$\begin{aligned} I_{13} &= -2 \int_0^\infty \int_{\mathbf{R}^n} A_{0j}^{rs}(u^s \circ \rho)_{tx_j}(u_{x_0}^r \circ \rho) \tilde{\theta}^2 \lambda^3 dz d\lambda \\ &\quad - 4 \int_0^\infty \int_{\mathbf{R}^n} A_{0j}^{rs}(u^s \circ \rho)_t(u_{x_0}^r \circ \rho) \tilde{\theta} \tilde{\theta}_{x_j} \lambda^3 dz d\lambda \\ &:= I_{131} + I_{132}. \end{aligned}$$

We note that

$$|I_{132}| \leq CI^{1/2} \left( \int_0^\infty \int_{\mathbf{R}^n} |(u_{x_0}^r \circ \rho)|^2 (\tilde{\theta}_{x_j})^2 \lambda^3 dz d\lambda \right)^{1/2}.$$

But

$$\int_0^\infty \int_{\mathbf{R}^n} |(u_{x_0}^r \circ \rho)|^2 (\tilde{\theta}_{x_j})^2 \lambda^3 dz d\lambda \leq \frac{C}{r^2} \int_0^\infty \int_{\mathbf{R}^n} |(u_{x_0}^r \circ \rho)|^2 \chi_{\{\tilde{\theta} > 0\}}(\lambda, z) \lambda^3 dz d\lambda$$

where  $\chi_{\{\tilde{\theta} > 0\}}(\lambda, z)$  is the indicator function of the support of  $\tilde{\theta}$ . Therefore by an interior regularity estimate and a simple geometric argument we can conclude that

$$\int_0^\infty \int_{\mathbf{R}^n} |(u_{x_0}^r \circ \rho)|^2 (\tilde{\theta}_{x_j})^2 \lambda^3 dz d\lambda \leq \int_{\tilde{\Delta}_{(1+\xi)r}} |\tilde{N}_*^{a,(1+\xi)\delta r}(u^s)(X, t)|^2 d\sigma(X, t).$$

We therefore only have to focus on  $I_{131}$ . Lifting the  $t$  derivative we get

$$\begin{aligned} I_{131} &= 2 \int_0^\infty \int_{\mathbf{R}^n} A_{0j}^{rs}(u^s \circ \rho)_{x_j}(u_{x_0}^r \circ \rho)_t \tilde{\theta}^2 \lambda^3 dz d\lambda \\ &\quad + 4 \int_0^\infty \int_{\mathbf{R}^n} A_{0j}^{rs}(u^s \circ \rho)_{x_j}(u_{x_0}^r \circ \rho) \tilde{\theta} \tilde{\theta}_t \lambda^3 dz d\lambda \\ &:= I_{1311} + I_{1312}. \end{aligned}$$

But

$$\begin{aligned} I_{1312} &= 4 \int_0^\infty \int_{\mathbf{R}^n} A_{0j}^{rs}(u^s \circ \rho)_{x_j}(u_{x_0}^r \circ \rho) \tilde{\theta}(\theta_t \circ \rho) \lambda^3 dz d\lambda \\ &\quad + 4 \int_0^\infty \int_{\mathbf{R}^n} A_{0j}^{rs}(u^s \circ \rho)_{x_j}(u_{x_0}^r \circ \rho) \tilde{\theta}(\theta_{x_0} \circ \rho) \frac{\partial P_{\gamma\lambda} A}{\partial t} \lambda^3 dz d\lambda := I_{13121} + I_{13122}. \end{aligned}$$

Using Lemma 3 and the properties of the support of  $\theta$  we get that

$$|I_{1312}| \leq C \int_0^\infty \int_{\mathbf{R}^n} |u_{x_j}^r \circ \rho|^2 \tilde{\theta}^2 \lambda dz d\lambda.$$

Therefore  $I_{1311}$  is left. But

$$\begin{aligned} I_{1311} &= 2 \int_0^\infty \int_{\mathbf{R}^n} A_{0j}^{rs}(u^s \circ \rho)_{x_j} (u^r \circ \rho)_t D^{-1} \tilde{\theta}^2 \lambda^3 dz d\lambda \\ &+ 2 \int_0^\infty \int_{\mathbf{R}^n} A_{0j}^{rs}(u^s \circ \rho)_{x_j} (u^r \circ \rho)_\lambda D^{-2} \frac{\partial^2 P_{\gamma\lambda} A}{\partial t \partial \lambda} \tilde{\theta}^2 \lambda^3 dz d\lambda := I_{13111} + I_{13112}. \end{aligned}$$

Using Lemma 3 we see that  $I_{13112}$  satisfies the same bound as  $I_{1312}$ . In the term  $I_{13111}$  we lift the  $\lambda$  derivative and we get

$$\begin{aligned} I_{13111} &= -2 \int_0^\infty \int_{\mathbf{R}^n} A_{0j}^{rs}(u^s \circ \rho)_{x_j \lambda} (u^r \circ \rho)_t D^{-1} \tilde{\theta}^2 \lambda^3 dz d\lambda \\ &- 2 \int_0^\infty \int_{\mathbf{R}^n} A_{0j}^{rs}(u^s \circ \rho)_{x_j} (u^r \circ \rho)_t D^{-2} \frac{\partial^2 P_{\gamma\lambda} A}{\partial \lambda^2} \tilde{\theta}^2 \lambda^3 dz d\lambda \\ &- 4 \int_0^\infty \int_{\mathbf{R}^n} A_{0j}^{rs}(u^s \circ \rho)_{x_j} (u^r \circ \rho)_t D^{-1} \tilde{\theta} \tilde{\theta}_\lambda \lambda^3 dz d\lambda \\ &- 6 \int_0^\infty \int_{\mathbf{R}^n} A_{0j}^{rs}(u^s \circ \rho)_{x_j} (u^r \circ \rho)_t D^{-1} \tilde{\theta}^2 \lambda^2 dz d\lambda \\ &- \lim_{\epsilon \rightarrow 0} 2 \int_{\mathbf{R}^n} A_{0j}^{rs}(u^s \circ \rho)_{x_j}(\epsilon, z) (u^r \circ \rho)_t(\epsilon, z) D^{-1} \tilde{\theta}^2(\epsilon, z) \epsilon^3 dz \\ &:= I_{131111} + I_{131112} + I_{131113} + I_{131114} + I_{131115}. \end{aligned}$$

We claim that by arguing as above

$$\begin{aligned} |I_{131112} + I_{131113} + I_{131114}| &\leq CI^{1/2} \left( \int_0^\infty \int_{\mathbf{R}^n} |(u_{x_j}^r \circ \rho)|^2 \tilde{\theta}^2 \lambda dz d\lambda \right)^{1/2} \\ &+ CI^{1/2} \left( \int_{\tilde{\Delta}_{(1+\xi)r}} |\tilde{N}_*^{a, (1+\xi)\delta r}(u^s)(X, t)|^2 d\sigma(X, t) \right)^{1/2}. \end{aligned}$$

Concerning  $I_{131115}$  we note, by interior regularity and the fact that  $u^r$  and  $u^s$  can be assumed to be bounded, that  $|I_{131115}|$  is finite. As we can assume, without loss of generality, that

$$\int_0^\infty \int_{\mathbf{R}^n} |(u_{x_j}^r \circ \rho)|^2 \tilde{\theta}^2 \lambda dz d\lambda < \infty,$$

we can conclude that  $\lim_{\lambda \rightarrow 0} \lambda(u_{x_j}^r \circ \rho(\lambda, z))(\tilde{\theta} \circ \rho(\lambda, z)) = 0$  for a.e  $z \in \mathbf{R}^{n-1} \times \mathbf{R}$ . Using this and interior regularity estimates we can conclude that  $I_{131115} = 0$ . We are therefore left with  $I_{131111}$ . But

$$I_{131111} = -2 \int_0^\infty \int_{\mathbf{R}^n} A_{0j}^{rs}(u_{x_0}^s \circ \rho)_{x_j}(u^r \circ \rho)_t \tilde{\theta}^2 \lambda^3 dz d\lambda$$

and we see that, by summing over all indices  $(r, s)$  and using the symmetry assumption of the system,  $I_{131111} = -I_{13}$ . Hence we have proved that  $I_{13} = -I_{13} + G$  where  $G$  is a sum of good terms in the sense defined above, and this means that we have established the appropriate bound on  $I_{13}$ . The arguments for  $I_{11}$  and  $I_{12}$  are similar and will be omitted. □

**3.2. Proof of Theorem 1 and Theorem 2.** We now intend to explore the inequalities proved in Lemma 6. In this section we will use the following lemma which essentially is a consequence of Lemma 6. In fact the lemma follows from Lemma 6 by a simple covering argument.

**Lemma 8.** *Let  $a > 0$ . For all cubes  $\Delta_r \subset \mathbf{R}^{n-1} \times \mathbf{R}$  there exists a constant  $C = C(\|A\|_{comm}, a)$  such that*

$$\int_{\tilde{\Delta}_r} |S^{a,r}(u^k)(X, t)|^2 d\sigma(X, t) \leq C \int_{\tilde{\Delta}_{2r}} |\tilde{N}_*^{a,2r}(u^s)(X, t)|^2 d\sigma(X, t).$$

Also for every  $\rho > 0$  small there exists  $C = C(\|A\|_{comm}, a, \rho)$  such that

$$\begin{aligned} \int_{\tilde{\Delta}_r} (u^k(X, t))^2 d\sigma(X, t) &\leq \rho \int_{\tilde{\Delta}_{2r}} |\tilde{N}_*^{a,2r}(u^s)(X, t)|^2 d\sigma(X, t) \\ &+ C \frac{1}{r^2} \int_{K(\tilde{\Delta}_{2r})} (u^s(Y, s))^2 dY ds \\ &+ C \int_{\tilde{\Delta}_{2r}} |S^{4a,2r}(u^s)(X, t)|^2 d\sigma(X, t). \end{aligned}$$

Here  $K(\tilde{\Delta}_{2r})$  is a compact set located in  $\Omega$  such that if  $(Y, s) \in K(\tilde{\Delta}_{2r})$ , then  $\delta(Y, s) \geq \beta r$  for a universal constant  $\beta = \beta(\|A\|_{comm}, a)$ .

Note that in the statement of the lemma we at this stage in fact do not have to enlarge the aperture of the cone defining the square function in the second inequality. Still this will be necessary later. In this section we mainly want to make progress in the direction of proving the second inequality of Theorem 2, as the first inequality follows from Lemma 8.

By scaling and translation there is no loss of generality to work in a neighbourhood of  $(A(0, 0), 0, 0) \in \partial\Omega$  and to consider a situation of unit size. Therefore we let  $(u^1, \dots, u^m)$  solve the system under consideration in  $\Omega$  and we let  $\Delta_0 = \{(x, t) \in \mathbf{R}^{n-1} \times \mathbf{R}; |x_j| < 1; |t| < 1\}$  denote the unit surface cube on  $\mathbf{R}^{n-1} \times \mathbf{R}$  centered at the origin  $(0, 0)$ . For functions  $f$  defined in  $\bar{\Omega}$  we define

$$M(f)(A(x, t), x, t) = \sup_{\Delta, \Delta \subset 3\Delta_0/2, (x,t) \in \Delta} \frac{1}{|\Delta|} \int_{\Delta} |f(A(y, s), y, s)| dy ds.$$

I.e.,  $M(f)$  is a localized Hardy-Littlewood maximal function defined relative to the graph of the function  $A(x, t)$ .

We intend to prove the following lemma.

**Lemma 9.** *Assume that  $\vec{u}(P_{\Delta_0}) = \vec{0}$  and let  $\beta > 0$  and  $0 < r_0 \ll 1$ . Let  $\Delta = \Delta_r \subset \Delta_0$ , with  $r < r_0$ , be a Whitney cube in the Whitney decomposition of the set*

$$\{(x, t) \in \Delta_0 : \tilde{N}_*^{a,1}(u^k)(A(x, t), x, t) > \beta\}.$$

Let  $A > 1$ ,  $\rho > 0$ ,  $\gamma > 0$  be degrees of freedom. Let  $E_{\Delta,\beta}$  denote the set

$$\left\{ (x, t) \in \Delta : \begin{aligned} &\tilde{N}_*^{a,1}(u^k)(A(x, t), x, t) > A\beta, \\ &\left[ \sum_s M([S^{4a,2}(u^s)]^2)(A(x, t), x, t) \right]^{1/2} \leq \gamma\beta, \\ &\left[ \sum_s M(\rho[\tilde{N}_*^{a,2}(u^s)]^2)(A(x, t), x, t) \right]^{1/2} \leq \gamma\beta \end{aligned} \right\}.$$

Then there exists  $\gamma_0$ , independent of  $\beta$ ,  $\Delta$  and  $\rho$  such that if  $\gamma < \gamma_0$ , then

$$|E_{\Delta,\beta}| \leq C\gamma^2|\Delta|.$$

Here  $C$  is independent of  $\beta$ ,  $\Delta$ ,  $\rho$  and  $\gamma$ .

We note that in the setting of the heat equation and appropriate pullbacks of the heat equation from time-varying domains a version of Lemma 9 is crucial to the arguments in [R]. Therefore it is important to point out that Lemma 5.4 in [R] and the subsequent argument are not correct as stated.

An immediate consequence of Lemma 9 is that if  $p > 2$  and if  $\vec{u}(P_{\Delta_0}) = \vec{0}$ , then

$$\begin{aligned} \sum_k \int_{\Delta_0} |\tilde{N}_*^{a,1}(u^k)(X, t)|^p d\sigma(X, t) &\leq C \sum_s \int_{2\Delta_0} |S^{4a,2}(u^s)(X, t)|^p d\sigma(X, t) \\ &+ C\rho \sum_s \int_{2\Delta_0} |\tilde{N}_*^{a,2}(u^s)(X, t)|^p d\sigma(X, t). \end{aligned}$$

Note that we can only, by the  $L^q$  continuity of the Hardy-Littlewood maximal function for  $q > 1$ , deduce this inequality for  $p > 2$ . In particular by the same arguments the following can be proven, as  $\rho$  can be chosen arbitrary small.

**Lemma 10.** *Let  $p > 2$ . Under the assumptions stated in Theorem 1 there exists a constant  $C$  independent of  $\vec{u}$  such that*

$$\sum_k \int_{\partial\Omega} |\tilde{N}_*(u^k)(X, t)|^p d\sigma(X, t) \leq C \sum_s \int_{\partial\Omega} |S(u^s)(X, t)|^p d\sigma(X, t).$$

We can conclude that Theorem 2 is proven by combining Lemma 8 and the consequence of Lemma 9. The second part of Theorem 1 follows from Lemma 10. To prove the first part of Theorem 1 we use the following real-variable argument. Let  $\lambda > 0$  and let  $\Delta = \Delta_r \subset \mathbf{R}^{n-1} \times \mathbf{R}$  be a Whitney cube, of scale  $r$ , in the Whitney decomposition of  $E_\lambda := \{(x, t) : S^{\vec{a}}(u^k)(A(x, t), x, t) > \lambda\}$  such that  $S^{\vec{a}}(u^k)(A(x^*, t^*), x^*, t^*) \leq \lambda$  for some point  $(x^*, t^*) \in 5\Delta$ .  $5\Delta$  is the cube having

the same center as  $\Delta$  but being scaled by a factor 5 and  $S^{\tilde{a}}(\cdot)$  is the square function operator previously introduced. Also let, for  $\eta \gg 1$  and  $\epsilon \ll 1$ ,

$$F_\lambda(\eta, \epsilon) := \{(x, t) \in \Delta : S^{\tilde{a}}(u^k)(A(x, t), x, t) > \eta\lambda, \sum_s \tilde{N}_*^a(u^s)(A(x, t), x, t) \leq \epsilon\lambda\}.$$

By a standard argument based on the interior regularity estimate stated in Lemma 5 we note that, if  $\eta$  is large enough and  $\epsilon$  is small enough, then  $S^{\tilde{a},r}(u^k)(A(x, t), x, t) > \eta\lambda/2$  if  $(x, t) \in F_\lambda(\eta, \epsilon)$ . If we define

$$\Omega(\Delta, \eta, \epsilon) := \bigcup_{(x,t) \in F_\lambda(\eta, \epsilon)} \tilde{\Gamma}_a(A(x, t), x, t),$$

then  $\Omega(\Delta, \eta, \epsilon)$  is the region defined as the union of parabolic cones with vertex on  $F_\lambda(\eta, \epsilon)$ . Hence

$$(\eta\lambda/2)^2 |F_\lambda(\eta, \epsilon)| \leq \int_{F_\lambda(\eta, \epsilon)} |S^{\tilde{a},r}(u^k)(A(x, t), x, t)|^2 dxdt.$$

Let  $\pi = \partial\Omega \rightarrow \mathbf{R}^n$  be the projection  $\pi(A(x, t), x, t) = (x, t)$ . Then trivially

$$\pi^{-1}(F_\lambda(\eta, \epsilon)) \subset \partial\Omega(\Delta, \eta, \epsilon)$$

and using Lemma 4 we can conclude that  $\Omega(\Delta, \eta, \epsilon) = \{(x_0, x, t) : x_0 > \hat{A}(x, t)\}$  where  $\|\hat{A}\|_{\text{comm}} < \beta$  with  $\beta$  independent of  $\Delta$ . Based on Lemma 8 we can assume that there exist constants  $C$  and  $\delta$ , independent of  $\eta, \epsilon$  and  $\Delta$  such that

$$\begin{aligned} & \int_{F_\lambda(\eta, \epsilon)} |S^{\tilde{a},r}(u^k)(A(x, t), x, t)|^2 dxdt \\ & \leq \int_{F_\lambda(\eta, \epsilon)} |S^{\tilde{a},r}(u^k)(\hat{A}(x, t), x, t)|^2 dxdt \\ & \leq C \sum_s \int_{(1+\xi)\bar{\Delta}} |\tilde{N}_{*,\Omega(\Delta, \eta, \epsilon)}^{a, (1+\xi)\delta r}(u^s)(\hat{A}(x, t), x, t)|^2 dxdt. \end{aligned}$$

The subscript in the maximal function indicates that it is defined w.r.t.  $\Omega(\Delta, \eta, \epsilon)$ . Hence by construction we can conclude that

$$(\eta\lambda/2)^2 |F_\lambda(\eta, \epsilon)| \leq C\epsilon^2\lambda^2|\Delta|$$

and in particular we have proved that  $|F_\lambda(\eta, \epsilon)| \leq C\epsilon^2/\eta^2|\Delta|$ . From this we can conclude by standard arguments that the first part of Theorem 1 is true and hence the proofs of Theorem 1 and Theorem 2 are complete modulo Lemma 9.

We will now set out to prove Lemma 9. To do this we define for  $k \in \{1, 2, \dots, m\}$  and  $(x_0, x, t) \in \Omega$

$$v^k(x_0, x, t) = u^k(x_0, x, t)\theta(x, t)\mu_\alpha(x_0 - A(x, t)).$$

Here  $\mu_\alpha \in C_0^\infty(\mathbf{R})$ ,  $\mu_\alpha(\lambda) = 1$  if  $0 < \lambda < \alpha/2$  and  $\mu_\alpha(\lambda) = 0$  if  $\alpha < \lambda$ .  $\alpha > 0$  will be determined later. We also let  $\theta \in C_0^\infty(\mathbf{R}^n)$ ,  $\theta(x, t) = 1$  if  $(x, t) \in \Delta_0$  and  $\theta(x, t) = 0$  if  $(x, t) \in \mathbf{R}^n \setminus 3\Delta_0/2$ . Recall that  $\tilde{\Gamma}_a(A(x, t), x, t)$  is the parabolic cone with aperture  $a$  and with vertex at  $(A(x, t), x, t) \in \partial\Omega$ . By  $\tilde{\Gamma}_a(A(x, t), x, t) + \lambda e_0$ , for  $\lambda >$

0, we mean the cone  $\tilde{\Gamma}_a(A(x, t), x, t)$  having its vertex translated to  $(A(x, t) + \lambda, x, t)$ . Using this notation we define, for each  $\beta \in (0, \infty)$ ,

$$h_\beta^k(x, t) = \sup\{\lambda \geq A(x, t) : \sup_{(y_0, y, s) \in [\tilde{\Gamma}_a(A(x, t), x, t) + \lambda e_0]} |v^k(y_0, y, s)| > \beta\}.$$

A simple geometric argument gives at hand that  $\{(h_\beta^k(x, t), x, t) : (x, t) \in \mathbf{R}^n\}$  is a Lip (1,1/2) graph with Lip (1,1/2) a constant determined by  $a$ , and this constant is by construction independent of  $\beta$ . We note that if  $\beta_1 < \beta_2$ , then  $h_{\beta_1}^k(x, t) > h_{\beta_2}^k(x, t)$  for all  $(x, t)$ . Furthermore it follows that for every  $(x, t) \in \Delta_0$  and for every  $\lambda \in (A(x, t), h_\beta^k(x, t))$ ,

$$\sup_{(y_0, y, s) \in [\tilde{\Gamma}_a(A(x, t), x, t) + \lambda e_0]} |v^k(y_0, y, s)| > \beta.$$

We define  $G_\beta^k = \{(h_\beta^k(x, t), x, t) : (x, t) \in \mathbf{R}^n\}$  as the graph of the function  $h_\beta^k(x, t)$ . We furthermore let  $\Omega_\beta^k = \{(x_0, x, t), x_0 > h_\beta^k(x, t)\}$ . As the function  $h_\beta^k(x, t)$  is only Lip(1,1/2) we need to construct an approximating function and an approximating graph which is time-varying in the sense of Definition 1. To do this we start by building a certain step function approximating  $G_\beta^k$ , but to formulate the construction we need to introduce some more notation. Let  $T(\Delta_0) = \{(\lambda, x, t); |x_j| < 1; |t| < 1; 0 < \lambda < 1\}$  be the Carleson box of unit size above  $\Delta_0$ . For each  $i \in \{0, 1, 2, \dots\}$  we divide  $\Delta_0$  into dyadic cubes  $\Delta_i^j$  each of dimensions  $2^{-i} \times \dots \times 2^{-i} \times 2^{-2i}$ . We denote the set of all such cubes by  $S_i$ . For each  $i \in \{0, 1, 2, \dots\}$  and to each  $\Delta_i^j \in S_i$  we associate two cubes, a Whitney cube  $Q(\Delta_i^j)$  and a Carleson box  $T(\Delta_i^j)$ . The Whitney cube is defined

$$Q(\Delta_i^j) = \{(\lambda, x, t), (x, t) \in \Delta_i^j; 2^{-i-1} < \lambda \leq 2^{-i}\}$$

and the Carleson box

$$T(\Delta_i^j) = \{(\lambda, x, t), (x, t) \in \Delta_i^j; 0 < \lambda \leq 2^{-i}\}.$$

The set of all Whitney cubes and the set of all Carleson boxes, defined by the size parameter  $i$ , are denoted by  $W_i$  and  $T_i$ , respectively. We also introduce a natural ordering of these geometric objects. Consider  $\Delta = \Delta_i^j$  for some  $i$  and  $j$  and let  $\tilde{\Delta} = \tilde{\Delta}_{\tilde{i}}^{\tilde{j}}$  for some  $\tilde{i}$  and  $\tilde{j}$ . Then  $\tilde{\Delta} \prec \Delta$  if  $\tilde{\Delta} \subset \Delta$  and  $\tilde{i} = i + 1$ . I.e., if  $\tilde{\Delta} \prec \Delta$ , then  $\tilde{\Delta}$  may be referred to as the son of  $\Delta$  or equivalently  $\Delta$  may be referred to as the father of  $\tilde{\Delta}$ .

To start the construction of the step function approximating  $G_\beta^k$  we consider a Whitney cube  $Q(\Delta_i^j)$  where  $\Delta_i^j \in S_i$  and we let  $N$  be a large integer. We consider for the moment  $Q(\Delta_i^j)$  as our universe and subdivide this Whitney cube dyadically until it is subdivided into a large number of cubes having dimensions  $2^{-i-N} \times 2^{-i-N} \times \dots \times 2^{-2i-2N}$ . We denote these new cubes by  $Q_l(\Delta_i^j)$  where  $l$  is the new parameter enumerating the cubes. Projecting  $Q_l(\Delta_i^j)$  in the  $(x, t)$ -plane we recover a dyadic cube  $\Delta_{i+N}^j \in S_{i+N}$ . We denote this projected cube  $\pi(Q_l(\Delta_i^j))$ . To be explicit we denote by  $\Delta_i^j(N)$  all subcubes of  $\Delta_i^j$  appearing as projections  $\pi(Q_l(\Delta_i^j))$  for some  $l$ . I.e., this is just the set of all dyadic subcubes of  $\Delta_i^j$  in  $S_{i+N}$ . We also let  $\Delta_i^j(N, k, \beta)$  denote the subset of all cubes in  $\Delta_i^j(N)$  having a

non-empty intersection with the projection of the graph  $G_\beta^k$ . Obviously  $\Delta_i^j(N, k, \beta)$  is the set of cubes such that there is a piece of graph above each such cube. Let  $\tilde{\Delta} \in \Delta_i^j(N, k, \beta)$  and let  $Q_{\tilde{l}_1}(\Delta_i^j), Q_{\tilde{l}_2}(\Delta_i^j), \dots, Q_{\tilde{l}_N}(\Delta_i^j)$  be an enumeration of all cubes having  $\tilde{\Delta}$  as its projection. We furthermore assume that these cubes are ordered, i.e.,  $Q_{\tilde{l}_1}(\Delta_i^j) \prec Q_{\tilde{l}_2}(\Delta_i^j) \prec \dots \prec Q_{\tilde{l}_N}(\Delta_i^j)$  where  $Q \prec \tilde{Q}$  in this situation means that  $\tilde{Q}$  is above  $Q$  in the  $x_0$ -direction. By  $Q_{\tilde{l}_k}^-(\Delta_i^j)$  we denote the bottom set (in the  $x_0$ -direction) of the set  $Q_{\tilde{l}_k}(\Delta_i^j)$  and let  $\lambda_{\tilde{l}_k}^-(\Delta_i^j)$  denote the  $x_0$ -coordinate of this set. Finally we let  $m(\tilde{\Delta}) \in \{1, 2, \dots, N\}$  be the largest  $m$  integer having the property  $Q_{\tilde{l}_m}^-(\Delta_i^j) \cap \Omega_\beta^k = \emptyset$ . Using this notation we define an auxiliary function as follows:

$$\tilde{g}_\beta^k(x, t) = \sum_{i,j} \sum_{\tilde{\Delta} \in \Delta_i^j(N, k, \beta)} \lambda_{\tilde{l}_m(\tilde{\Delta})}^-(\Delta_i^j) \chi_{\tilde{\Delta}}(x, t).$$

By definition  $\tilde{g}_\beta^k(x, t)$  is a step function and  $h_\beta^k(x, t) > \tilde{g}_\beta^k(x, t)$ . We now recall the following lemma proven in [R].

**Lemma 11.** *Let  $f : \Delta_0 \rightarrow \mathbf{R}$  be defined as  $f(x, t) = \sum_j a_j \chi_{A_j}(x, t)$  where the sets  $\{A_j\}$  and the positive constants  $\{a_j\}$  satisfy the following conditions:*

- *The sets  $\{A_j\}$  are parabolic dyadic cubes relative  $\Delta_0$ .*
- *There exist constants  $\tilde{k}_0$  and  $k_0$  such that  $\tilde{k}_0 \leq a_j / \text{diam} A_j \leq k_0$ .*
- *There exists a constant  $k_1$  such that for any  $A_j$  and  $A_i$ ,  $\overline{A_j} \cap \overline{A_i} \neq \emptyset$  we have that  $|a_j - a_i| \leq k_1 \min\{\text{diam} A_j, \text{diam} A_i\}$ .*

*Then for any  $0 < J \ll 1$  there exists a function  $\phi : \Delta_0 \rightarrow \mathbf{R}$  and a uniform constant  $c_1$  such that  $|\phi(x, t) - \phi(y, t)| < c_1$  for all  $x, y \in \mathbf{R}^{n-1}$  and  $t \in \mathbf{R}$  and  $D_{1/2}^t \phi \in BMO(\mathbf{R}^{n-1} \times \mathbf{R})$ . Furthermore,  $\phi$  approximates the function  $f$  in the sense that for each  $A_j$  there exists a concentric subcube  $\hat{\Delta}_j \subset A_j$  such that  $f = \phi$  on  $\hat{\Delta}_j$  and  $|A_j \setminus \hat{\Delta}_j| \leq CJ^{n+1}$ .*

One can easily verify that, by construction, the lemma applies to the function  $\tilde{g}_\beta^k$ . Hence we can conclude that given  $J, 0 < J \ll 1$ , there exists a function  $g_\beta^k(x, t)$  such that  $g_\beta^k(x, t) = \tilde{g}_\beta^k(x, t)$  for all  $(x, t) \in \hat{\Delta}$ , where  $\hat{\Delta} = \tilde{\Delta} / \hat{\gamma}$  for a universal constant  $\hat{\gamma}$  and  $|\tilde{\Delta} \setminus \hat{\Delta}| \leq CJ^{n+1}$ . This is true for all cubes  $\tilde{\Delta}$  in the construction above. Furthermore the function  $g_\beta^k$  defines a genuinely time-varying domain in the sense of Definition 1.

For functions  $f$  defined in  $\overline{\Omega}$  we also define

$$M_\beta(f)(g_\beta^k(x, t), x, t) = \sup_{\Delta, \Delta \subset 3\Delta_0/2, (x, t) \in \Delta} \frac{1}{|\Delta|} \int_{\Delta} |f(g_\beta^k(y, s), y, s)| dy ds.$$

I.e.,  $M_\beta(f)$  is a localized Hardy-Littlewood maximal function defined relative to the graph of the function  $g_\beta^k(x, t)$ . To prove Lemma 9 we need to establish two technical results.

**Lemma 12.** *Let  $\Delta = \Delta_r$  with  $r < r_0 \ll 1$  and assume that  $(x, t) \in E_{\Delta, \beta}$ . Then there exists  $\gamma_0$ , independent of  $\Delta$ , such that if  $\gamma \leq \gamma_0$ , then*

$$\tilde{N}_*^{a, r}(u^k)(A(x, t), x, t) \geq \beta/2.$$

*Proof.* By an appropriate choice of the Whitney decomposition we can assume that, by construction, there exists  $(\hat{x}, \hat{t}) \in 5\Delta \subset \Delta_0$  such that  $\tilde{N}_*^{a,1}(u^k)(A(\hat{x}, \hat{t}), \hat{x}, \hat{t}) \leq \beta$ . The cone  $\Gamma_a(A(x, t), x, t) \cap \{(x_0, x, t), A(x, t) \leq x_0 \leq A(x, t) + 1\}$  can be divided into the three sets

$$\begin{aligned} A_1 &= [\Gamma_a(A(x, t), x, t) \cap \Gamma_a(A(\hat{x}, \hat{t}), \hat{x}, \hat{t})] \cap \{(x_0, x, t), A(x, t) \leq x_0 \leq A(x, t) + 1\}, \\ A_2 &= (\Gamma_a(A(x, t), x, t) \setminus \Gamma_a(A(\hat{x}, \hat{t}), \hat{x}, \hat{t})) \cap \{(x_0, x, t), A(x, t) + r \leq x_0 < A(x, t) + 1\}, \\ A_3 &= (\Gamma_a(A(x, t), x, t) \setminus \Gamma_a(A(\hat{x}, \hat{t}), \hat{x}, \hat{t})) \cap \{(x_0, x, t), A(x, t) \leq x_0 \leq A(x, t) + r\}. \end{aligned}$$

We know that  $\tilde{N}_*^{a,1}(u^k)(A(x, t), x, t) > A\beta$  and that  $u^k(x_0, x, t) \leq \beta$  whenever  $(x_0, x, t) \in A_1$ . Let  $P \in A_2$  and let  $\hat{r} = \delta(P)$ . By a simple geometric argument and Lemma 5 we can conclude that there exists a point  $\tilde{P} \in \Gamma_a(A(\hat{x}, \hat{t}), \hat{x}, \hat{t}) \cap \{(x_0, x, t), A(x, t) \leq x_0 \leq A(x, t) + 1\}$  and a sequence of points  $P_0 = P, P_1, \dots, P_{l-1}, P_l = \tilde{P}$  with the following properties.  $l$  can be chosen independent of  $P$  and

$$|u^k(P_{j+1}) - u^k(P_j)| \leq C \left( \sum_s \frac{1}{(\hat{\gamma}\hat{r})^n} \int_{C_{\hat{\gamma}\hat{r}}(P_j)} |\nabla u^s(X, t)|^2 dX dt \right)^{1/2}$$

for some  $\hat{\gamma} > 0$ . As  $(x, t) \in E_{\Delta, \beta}$  we know, by definition, that

$$\sum_s M([S^{4a,2}(u^s)]^2)(A(x, t), x, t) \leq \gamma^2 \beta^2$$

and we can therefore conclude that

$$|u^k(P_{j+1}) - u^k(P_j)| \leq C\gamma\beta.$$

Using the triangle inequality  $|u^k(P)| \leq \beta + Cl\gamma\beta$ . As  $A > 1$  we hence know that there exists  $(x_0, x, t) \in A_2 \cup A_3$  such that  $u^k(x_0, x, t) \geq A\beta$ . If we therefore let  $\gamma$  be small enough we can conclude that there must exist a point  $(y_0, y, s) \in A_3$  such that  $u^k(y_0, y, s) \geq A\beta > \beta/2$ . This completes the proof of the lemma.  $\square$

**Lemma 13.** *Let  $\tilde{\gamma} > 1$ ,  $\Delta = \Delta_r$  with  $r < r_0 \ll 1$  and assume that  $(x, t) \in E_{\Delta, \beta}$ . Then there exists  $\gamma_0$ , independent of  $\tilde{\gamma}$ ,  $(x, t)$ , and  $\Delta$ , such that if  $\gamma \leq \gamma_0$ , then there exists a cube  $\tilde{\Delta}$  such that  $(x, t) \in 4\tilde{\Delta}$  and such that for all  $(\tilde{x}, \tilde{t}) \in \tilde{\Delta}/\tilde{\gamma}$*

$$u^k(g_\beta^k(\tilde{x}, \tilde{t}), \tilde{x}, \tilde{t}) \geq \beta/100.$$

*Proof.* We will prove that there exists a cube  $\tilde{\Delta}$  such that  $(x, t) \in 4\tilde{\Delta}$  and such that for all  $(\tilde{x}, \tilde{t}) \in \tilde{\Delta}$ ,

$$u^k(h_\beta^k(\tilde{x}, \tilde{t}), \tilde{x}, \tilde{t}) \geq \beta/100.$$

By construction this implies that

$$u^k(\tilde{g}_\beta^k(\tilde{x}, \tilde{t}), \tilde{x}, \tilde{t}) \geq \beta/100.$$

Using Lemma 11 we can then conclude that for all  $(\tilde{x}, \tilde{t}) \in \tilde{\Delta}/\hat{\gamma}$ ,

$$u^k(g_\beta^k(\tilde{x}, \tilde{t}), \tilde{x}, \tilde{t}) \geq \beta/100.$$

Let  $(x, t) \in E_{\Delta, \beta}$  and by applying Lemma 12 we know that  $N_*^{a,r} u^k(A(x, t), x, t) \geq \beta/2$ . We define  $\delta = h_\beta^k(x, t) - A(x, t)$ . We have that  $u^k(h_\beta^k(x, t), x, t) = \beta$ . Let  $\hat{\rho}$  be a small number and let  $\hat{P} = (h_\beta^k(x, t), x, t)$ . In the following we consider points  $P \in C_{\hat{\rho}\delta}(\hat{P})$ . Using Lemma 5

$$\begin{aligned} |u^k(P) - u^k(\hat{P})| &\leq C \left( \sum_s \frac{1}{(\hat{\rho}\delta)^n} \int_{C_{2\hat{\rho}\delta}(\hat{P})} |\nabla u^s(X, t)|^2 dX dt \right)^{1/2} \\ &\leq C \left[ \sum_s M([S^{4a,2}(u^s)]^2)(A(x, t), x, t) \right]^{1/2} \leq C\gamma\beta. \end{aligned}$$

Hence  $u^k(P) \geq u^k(\hat{P}) - C\gamma\beta \geq (1 - C\gamma)\beta$  for all  $P \in C_{\hat{\rho}\delta}(\hat{P})$ . This completes the proof of the lemma.  $\square$

We are now ready to prove Lemma 9.

*Proof.* We need to estimate  $|E_{\Delta, \beta}|$ , and to do so we let  $(x, t) \in E_{\Delta, \beta}$ . Using Lemma 13 we have that  $M_\beta(u^k \chi_{4\Delta})(g_\beta^k(x, t), x, t) \geq C\beta$ . Assume that the center of  $\Delta = \Delta_r$  is located at  $(\tilde{x}, \tilde{t})$ . To continue we define  $\tilde{u}(x_0, x, t) = u(x_0, x, t) - u(P_\Delta)$  and note that  $M_\beta(\tilde{u}^k \chi_{4\Delta})(g_\beta^k(x, t), x, t) \geq \tilde{C}\beta$ . By the maximal function theorem we get

$$|E_{\Delta, \beta}| \leq C\beta^{-2} \int_{4\Delta} (\tilde{u}^k)^2(g_\beta^k(x, t), x, t) dx dt \leq C\beta^{-2} \int_{4\tilde{\Delta}} (\tilde{u}^k)^2(X, t) d\sigma_\beta^k(X, t).$$

By the second part of Lemma 8 we have

$$\begin{aligned} &\int_{4\tilde{\Delta}_r} (\tilde{u}^k(X, t))^2 d\sigma_\beta^k(X, t) \\ &\leq \rho \int_{\tilde{\Delta}_{8r}} |\tilde{N}_*^{a,8r,\beta}(\tilde{u}^s)(X, t)|^2 d\sigma_\beta^k(X, t) + C \frac{1}{r^2} \int_{K(\tilde{\Delta}_{8r})} (\tilde{u}^s(Y, s))^2 dY ds \\ &\quad + C \int_{\tilde{\Delta}_{8r}} |S^{4a,8r,\beta}(\tilde{u}^s)(X, t)|^2 d\sigma_\beta^k(X, t). \end{aligned}$$

Here the  $\beta$  in  $\tilde{N}_*^{a,8r,\beta}$  and  $S^{4a,8r,\beta}$  indicate that these objects are defined with  $\Omega_\beta^k$  as the domain of reference. Using the normalization that  $\tilde{u}(P_\Delta) = 0$  we can assume, by using the interior regularity estimate of Lemma 5, that

$$\frac{1}{r^2} \int_{K(\tilde{\Delta}_{8r})} (\tilde{u}^s(Y, s))^2 dY ds \leq C \int_{\tilde{\Delta}_{8r}} |S^{4a,8r,\beta}(\tilde{u}^s)(X, t)|^2 d\sigma_\beta^k(X, t).$$

A crucial observation is that by construction

$$\begin{aligned} \tilde{N}_*^{a,8r,\beta}(\tilde{u}^s)(g_\beta^k(x, t), x, t) &\leq \tilde{N}_*^{a,2}(\tilde{u}^s)(A(x, t), x, t), \\ S^{4a,8r,\beta}(\tilde{u}^s)(g_\beta^k(x, t), x, t) &\leq S^{4a,2}(\tilde{u}^s)(A(x, t), x, t). \end{aligned}$$

Therefore

$$\sum_s \rho \int_{\tilde{\Delta}_{8r}} |\tilde{N}_*^{a,8r,\beta}(\tilde{u}^s)(X,t)|^2 d\sigma_\beta^k(X,t) \leq C \left[ \sum_s [M(\rho[\tilde{N}_*^{a,2}(\tilde{u}^s)]^2)(A(x,t),x,t)] |\Delta| \right],$$

$$\sum_s \int_{\tilde{\Delta}_{8r}} |S^{4a,8r,\beta}(\tilde{u}^s)(X,t)|^2 d\sigma_\beta^k(X,t) \leq C \left[ \sum_s [M([S^{4a,2}(u^s)]^2)(A(x,t),x,t))] |\Delta| \right].$$

Combining this with the definition of the set  $E_{\Delta,\beta}$  we can conclude that

$$|E_{\Delta,\beta}| \leq C\beta^{-2} \int_{4\tilde{\Delta}} (\tilde{u}^k)^2(X,t) d\sigma_\beta^k(X,t) \leq c\gamma^2 |\Delta|.$$

□

We note that in order to remove the restriction that  $p > 2$  in the second inequality of Theorem 1, a localized version of the inequality stated in Theorem 1 and a real-variable argument similar to the one used in the proof of the first part of Theorem 1 are needed. In particular we would like to be able to use Theorem 1 to prove that if  $p > 2$  and if  $\vec{u}(P_{\tilde{\Delta}_0}) = \vec{0}$ , then there exists a constant  $C$  independent of  $\vec{u}$  such that

$$\sum_k \int_{\tilde{\Delta}_0} |\tilde{N}_*^{a,1}(u^k)(X,t)|^p d\sigma(X,t) \leq C \sum_s \int_{2\tilde{\Delta}_0} |S^{4a,2}(u^s)(X,t)|^p d\sigma(X,t).$$

Currently it is not clear to us what estimates are needed for the system, on the atomic level, in order to establish this inequality. We therefore leave this question for future research.

## REFERENCES

- [D] B.E.J. Dahlberg, On estimates of harmonic measure, Arch. Rat. Mech. Anal. 65 (1977), 272-288. MR0466593 (57:6470)
- [D1] B.E.J. Dahlberg, Approximation of harmonic functions, Ann. Inst. Fourier Grenoble 30 (1980), 97-107. MR584274 (82i:31010)
- [D2] B.E.J. Dahlberg, Weighted norm inequalities for the Lusin area integral and the nontangential maximal function for functions harmonic in Lipschitz domains, Studia Math. 65 (1989), 297-314. MR592391 (82f:31003)
- [DKPV] B.E.J. Dahlberg, C.E. Kenig, J. Pipher, G.C. Verchota, Area integral estimates for higher order elliptic equations and systems, Annales de l'Inst. Fourier 47 (1997), 1425-1461. MR1600375 (98m:35045)
- [G] E. De Giorgi, Sulla differenziabilita e analiticita delle estremali degli integrali multipli regolari, Mem. Acad. Sci. Torino 3 (1957), 25-43. MR0093649 (20:172)
- [HL] S. Hofmann, J.L. Lewis,  $L^2$  solvability and representation by caloric layer potentials in time-varying domains, Annals of Math. 144 (2) (1996), 349-420. MR1418902 (97h:35072)
- [HL1] S. Hofmann, J.L. Lewis, The Dirichlet problem for parabolic operators with singular drift terms, Memoirs of the AMS, Volume 151, Providence 2001. MR1828387 (2002h:35108)
- [KKPT] C.E. Kenig, H. Koch, J. Pipher, T. Toro, A new approach to absolute continuity of elliptic measure with applications to non-symmetric equations, Adv. in Math. 153 (2000), 231-298. MR1770930 (2002f:35071)
- [LM] J.L. Lewis, M. Murray, The method of Layer potentials for the heat equation in time-varying domains, Memoir of the A.M.S. no. 545.
- [LS] J.L. Lewis, J. Silver, Parabolic measure and the Dirichlet problem for the heat equation in two dimensions, Indiana U. Math. J. 37 (1988), 801-839. MR982831 (90e:35079)
- [M] J. Moser, On Harnack's theorems for elliptic equations, Comm. Pure. Appl. Math. 14 (1961), 577-591. MR0159138 (28:2356)

- [Na] J. Nash, Continuity of solutions of parabolic and elliptic equations, Amer. J. Math. 80 (1958), 931-954. MR0100158 (20:6592)
- [N] K. Nyström, Boundary value problems and duality between  $L^p$  Dirichlet and regularity problems for second order parabolic systems in non-cylindrical domains, Collect. Math., Vol. 57, No. 2 (2006), 93-119. MR2206182 (2007g:35095)
- [R] J. Rivera-Noriega, Absolute continuity of parabolic measure and area integral estimates in non-cylindrical domains, Indiana University Mathematics Journal, Vol. 51, No. 2 (2003), 477-525. MR1976086 (2004d:35093)

DEPARTMENT OF MATHEMATICS, UMEÅ UNIVERSITY, S-90187 UMEÅ, SWEDEN  
*E-mail address:* `kaj.nystrom@math.umu.se`