# THE DIRICHLET PROBLEM FOR SECOND ORDER PARABOLIC OPERATORS IN DIVERGENCE FORM

Pascal Auscher, Moritz Egert, Kaj Nyström

# THE DIRICHLET PROBLEM FOR SECOND ORDER PARABOLIC OPERATORS IN DIVERGENCE FORM

PASCAL AUSCHER, MORITZ EGERT, AND KAJ NYSTRÖM

Abstract. We study parabolic operators $\mathcal{H} = \partial_t - \operatorname{div}_{\lambda,x} A(x,t)\nabla_{\lambda,x}$ in the parabolic upper half space $\mathbb{R}^{n+2}_+ = \{(\lambda,x,t) : \lambda > 0\}$. We assume that the coefficients are real, bounded, measurable, uniformly elliptic, but not necessarily symmetric. We prove that the associated parabolic measure is absolutely continuous with respect to the surface measure on $\mathbb{R}^{n+1}$ in the sense defined by $A_\infty(\mathrm{d}x\,\mathrm{d}t)$. Our argument also gives a simplified proof of the corresponding result for elliptic measure.

## 1. Introduction and statement of main results

A classical result due to Dahlberg [8] states in the context of Lipschitz domains that harmonic measure is absolutely continuous with respect to surface measure, and that the Poisson kernel (its Radon-Nykodim derivative) satisfies a scale-invariant reverse Hölder inequality in $\mathrm{L}^2$. Equivalently, the Dirichlet problem with $\mathrm{L}^2$-data can be solved with $\mathrm{L}^2$-control of a non-tangential maximal function. Ever since Dahlberg's original work the study of elliptic measure has been a very active area of research and a number of fine results have been established, see [1, 14, 19] for recent accounts of the state of the art.

In contrast to the study of elliptic measure, the fine properties of parabolic measure are considerably less understood. In [13] a parabolic version of Dahlberg's result was established for the heat equation in time-independent Lipschitz cylinders. A major contribution in the study of boundary value problems and parabolic measure for the heat equation in time-dependent Lipschitz type domains was achieved in [15, 20, 21]. In these papers the correct notion of time-dependent Lipschitz type cylinders, correct from the perspective of parabolic measure and parabolic layer potentials, was found. In particular, in [20, 21] the mutual absolute continuity of parabolic measure and surface measure and the $A_\infty$-property were established and in [15] the authors obtained a version of Dahlberg's result for parabolic measure associated to the heat equation in time-dependent Lipschitz-type domains. In this context the properties of parabolic measures were further analyzed in the influential work [16], parts of which have been simplified in [27].

Very recently, there have been advances in the theory of boundary value problems for second order parabolic equations (and systems) of the form

$$\mathcal{H}u := \partial_t u - \operatorname{div}_{\lambda,x} A(x,t)\nabla_{\lambda,x} u = 0, \tag{1.1}$$

in the upper-half parabolic space $\mathbb{R}^{n+2}_+ := \{(\lambda,x,t) \in \mathbb{R}\times\mathbb{R}^n\times\mathbb{R} : \lambda > 0\}$, $n \geq 1$, with boundary determined by $\lambda = 0$, assuming only bounded, measurable, uniformly elliptic and complex coefficients. In [6, 25, 26], the solvability for Dirichlet, regularity and Neumann problems with $\mathrm{L}^2$-data were established for the class of parabolic equations (1.1) under the additional assumptions that the elliptic part is also independent of the time variable $t$ and that it has either constant (complex) coefficients, real symmetric coefficients, or small perturbations thereof. Focusing on parabolic measure, a particular consequence of Theorem 1.3 in [6] is the generalization of [13] to equations of the form (1.1) but with $A$ real, symmetric and time-independent. This analysis was advanced further in [4], where a first order strategy to study boundary value problems of parabolic systems with second order

*Date*: November 2, 2016.

2010 *Mathematics Subject Classification.* Primary: 35K10, 35K20; Secondary: 26A33, 42B25.

*Key words and phrases.* Second order parabolic operator, non-symmetric coefficients, Dirichlet problem, parabolic measure, $A_\infty$-condition, Carleson measure estimate.

The authors were partially supported by the ANR project "Harmonic Analysis at its Boundaries", ANR-12-BS01-0013. The second author was also supported by a public grant as part of the FMJH.

elliptic part in the upper half-space was developed. The outcome of [4] was the possibility to address arbitrary parabolic equations (and systems) as in (1.1) with coefficients depending also on time and on the transverse variable with additional transversal regularity.

In this paper we advance the study of parabolic boundary value problems and parabolic measure even further. We consider parabolic equations as in (1.1), assuming that the coefficients are real, bounded, measurable, uniformly elliptic, but not necessarily symmetric. We prove that the associated parabolic measure is absolutely continuous with respect to the surface measure on $\mathbb{R}^{n+1}$ $(\mathrm{d}x\,\mathrm{d}t)$ in the sense defined by the Muckenhoupt class $A_\infty(\mathrm{d}x\,\mathrm{d}t)$. As consequences, the associated Poisson kernel exists, satisfies a scale-invariant reverse Hölder inequality in $\mathrm{L}^p$ for some $p \in (1,\infty)$, and the Dirichlet problem with $\mathrm{L}^q$-data, $q$ being the index dual to $p$, can be solved with appropriate control of non-tangential maximal functions. In particular, our main result, which is new already in the case when $A$ is symmetric and time-dependent, gives a parabolic analogue of the main result in [14] concerning elliptic measure. Our proof heavily relies on square function estimates and non-tangential estimates for parabolic operators with time-dependent coefficients that were only recently obtained by us in [4] as well as the reduction to a Carleson measure estimate proved in [9]. As we shall avoid the change of variables utilized in [14], this also gives a simpler and more direct proof of the $A_\infty$-property of elliptic measure.

1.1. **The coefficients.** We assume that $A = A(x,t) = \{A_{i,j}(x,t)\}_{i,j=0}^n$ is a real-valued $(n+1) \times (n+1)$-dimensional matrix, not necessarily symmetric, satisfying

$$\kappa|\xi|^2 \leq \sum_{i,j=0}^n A_{i,j}(x,t)\xi_i\xi_j, \qquad |A(x,t)\xi\cdot\zeta| \leq C|\xi||\zeta|, \tag{1.2}$$

for some $\kappa, C \in (0,\infty)$, which we refer to as the *ellipticity constants* of $A$, and for all $\xi,\zeta \in \mathbb{R}^{n+1}$, $(x,t) \in \mathbb{R}^{n+1}$. Here, given $u = (u_0,...,u_n)$, $v = (v_0,...,v_n) \in \mathbb{R}^{n+1}$ we write $u\cdot v := u_0v_0 + ... + u_nv_n$.

1.2. **Weak solutions.** If $\Omega$ is an open subset of $\mathbb{R}^{n+1}$, we let $\mathrm{H}^1(\Omega) = \mathrm{W}^{1,2}(\Omega)$ be the standard Sobolev space of complex valued functions $v$ defined on $\Omega$, such that $v$ and $\nabla v$ are in $\mathrm{L}^2(\Omega)$ and $\mathrm{L}^2(\Omega;\mathbb{C}^n)$, respectively. A subscripted 'loc' will indicate that these conditions hold locally. A function $u$ is called *weak solution* to the equation $\mathcal{H}u = 0$ on $\mathbb{R}^{n+1}_+ \times \mathbb{R}$ if it satisfies $u \in \mathrm{L}^2_{\mathrm{loc}}(\mathbb{R};\mathrm{W}^{1,2}_{\mathrm{loc}}(\mathbb{R}^{n+1}_+))$ and

$$\int_{\mathbb{R}}\iint_{\mathbb{R}^{n+1}_+} A\nabla_{\lambda,x}u\cdot\nabla_{\lambda,x}\phi\;\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda - \int_{\mathbb{R}}\iint_{\mathbb{R}^{n+1}_+} u\cdot\partial_t\phi\;\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda = 0$$

for all $\phi \in \mathrm{C}_0^\infty(\mathbb{R}^{n+2}_+)$.

1.3. **Parabolic measure.** Given $(x,t) \in \mathbb{R}^{n+1}$ and $r > 0$ we let $Q = Q_r(x) := B(x,r) \subset \mathbb{R}^n$ be the standard Euclidean ball centered at $x$ and of radius $r$, and we let $I = I_r(t) := (t-r^2, t+r^2)$. We let $\Delta = \Delta_r(x,t) = Q_r(x)\times I_r(t)$ and write $\ell(\Delta) := r$. We will use the convention that $cQ$ and $cI$ denote the dilates of balls and intervals, respectively, keeping the center fixed and dilating the radius by $c$ and we let $c\Delta := cQ \times c^2 I$.

Given $A$ real, satisfying (1.2), and $f$ continuous and compactly supported in $\mathbb{R}^{n+1}$, there exists a unique (weak) solution $u$ to the continuous Dirichlet problem $\mathcal{H}u = (\partial_t - \mathrm{div}_{\lambda,x}\,A(x,t)\nabla_{\lambda,x})u = 0$ in $\mathbb{R}^{n+2}_+$, $u$ continuous in $\overline{\mathbb{R}^{n+2}_+}$ and $u(0,x,t) = f(x,t)$ whenever $(x,t) \in \mathbb{R}^{n+1}$. Indeed, assume $f \geq 0$ and let $u_k$, $k \geq 1$, be the unique weak solution to $\mathcal{H}u = 0$ in $\Omega_k := (0,k)\times\Delta_k(0,0)$, with boundary values $f(x,t)\psi(||(x,t)||/k)$ on $\Delta_k(0,0t)$, and zero otherwise. Here, $||(x,t)|| := |x| + |t|^{1/2}$ and $\psi$ is a continuous decreasing function on $[0,\infty)$ such that $0 \leq \psi \leq 1$, $\psi(r) = 1$ for $0 \leq r \leq 1/2$, and $\psi(r) = 0$ for $r > 3/4$. Then $0 \leq u_k \leq u_{k+1} \leq ||f||_\infty$ in $\Omega_k$ and one can deduce, by the maximum principle and the Harnack inequality, see [24] for these estimates, that

$$\sup_{\Omega_l}|u_k - u_j| \leq c(u_k-u_j)(l,0,4l^2), \qquad \text{if } k > j \gg l.$$

In particular, $u$ can be constructed as the monotone and uniform limit of $\{u_k\}$ as $k\to\infty$ on the closure of $\Omega_l$ for each $l \geq 1$. Uniqueness follows from the maximum principle. Furthermore, by the maximum principle and the Riesz representation theorem we deduce

$$u(\lambda,x,t)=\iint_{\mathbb{R}^{n+1}} f(y,s)\,\mathrm{d}\omega(\lambda,x,t,y,s),\qquad \text{for all } (\lambda,x,t)\in\mathbb{R}^{n+2}_+,$$

where $\{\omega(\lambda,x,t,\cdot):\ (\lambda,x,t)\in\mathbb{R}^{n+2}_+\}$ is a family of regular Borel measures on $\mathbb{R}^{n+1}$ and we refer to $\omega(\lambda,x,t,\cdot)$ as $\mathcal{H}$-parabolic measure, or simply *parabolic measure* (at $(\lambda,x,t)$).

Given $r>0$ and $(x_0,t_0)\in\mathbb{R}^{n+1}$ we let

$$A_r^+(x_0,t_0):=(4r,x_0,t_0+16r^2).$$

Assume that $A$ satisfies (1.2). Then parabolic measure is a doubling measure in the sense that there exists a constant $c$, $1\le c<\infty$, depending only on $n$ and the ellipticity constants such that the following is true. Let $(x_0,t_0)\in\mathbb{R}^{n+1}$, $0<r_0<\infty$, $\Delta_0:=\Delta_{r_0}(x_0,t_0)$. Then

$$\omega(A_{4r_0}^+(x_0,t_0),2\Delta)\le c\,\omega(A_{4r_0}^+(x_0,t_0),\Delta)$$

whenever $\Delta\subset 4\Delta_0$. We refer to [11], [12] and [24] for details. The doubling property of parabolic measure serves as a starting point for further investigation. In this paper we are interested in scale invariant quantitative version of absolute continuity of parabolic measure with respect to the measure $\mathrm{d}x\,\mathrm{d}t$ on $\mathbb{R}^{n+1}$. Given a set $E\subset\mathbb{R}^{n+1}$ we let $|E|$ denote the Lebesgue measure of $E$.

**Definition 1.1.** Let $(x_0,t_0)\in\mathbb{R}^{n+1}$, $0<r_0<\infty$, $\Delta_0:=\Delta_{r_0}(x_0,t_0)$. We say that parabolic measure associated to $\mathcal{H}=\partial_t-\mathrm{div}_{\lambda,x}A(x,t)\nabla_{\lambda,x}$ at $A_{4r_0}^+(x_0,t_0)$ is in $A_\infty(\Delta_0,\,\mathrm{d}x\,\mathrm{d}t)$ if for every $\varepsilon>0$ there exists $\delta=\delta(\varepsilon)>0$ such that if $E\subset\Delta$ for some $\Delta\subset\Delta_0$, then

$$\frac{\omega(A_{4r_0}^+(x_0,t_0),E)}{\omega(A_{4r_0}^+(x_0,t_0),\Delta)}<\delta\quad\Longrightarrow\quad\frac{|E|}{|\Delta|}<\varepsilon.$$

Parabolic measure $\omega$ belongs to $A_\infty(\mathrm{d}x\,\mathrm{d}t)$ if $\omega(A_{4r_0}^+(x_0,t_0),\cdot)\in A_\infty(\Delta_0,\,\mathrm{d}x\,\mathrm{d}t)$ for all $\Delta_0$ as above and with uniform constants.

If $\omega$ belongs to $A_\infty(\mathrm{d}x\,\mathrm{d}t)$, then $\omega(A_{4r_0}^+(x_0,t_0),\cdot)$ and $\mathrm{d}x\,\mathrm{d}t$ are mutually absolutely continuous and hence one can write

$$\mathrm{d}\omega(A_{4r_0}^+(x_0,t_0),x,t)=K(A_{4r_0}^+(x_0,t_0),x,t)\,\mathrm{d}x\,\mathrm{d}t.$$

We refer to $K(A_{4r_0}^+(x_0,t_0),x,t)$ as the associated *Poisson kernel* (at $A_{4r_0}^+(x_0,t_0)$).

**Definition 1.2.** For $p\in(1,\infty)$ we say that $\omega$ belongs to the reverse Hölder class $B_p(\mathrm{d}x\,\mathrm{d}t)$ if there exists a constant $c$, $1\le c<\infty$, such that for all $\Delta_0:=\Delta_{r_0}(x_0,t_0)$ the Poisson kernel $K(A_{4r_0}^+(x_0,t_0),\cdot)$ satisfies the reverse Hölder inequality

$$\left(\fint\!\!\fint_{\Delta}(K(A_{4r_0}^+(x_0,t_0),x,t))^p\,\mathrm{d}x\,\mathrm{d}t\right)^{1/p}\le c\fint\!\!\fint_{\Delta}K(A_{4r_0}^+(x_0,t_0),x,t)\,\mathrm{d}x\,\mathrm{d}t$$

whenever $\Delta\subset\Delta_0$.

Note that as parabolic measure has the doubling property the statement that parabolic measure $\omega$ belongs to $A_\infty(\mathrm{d}x\,\mathrm{d}t)$ has several equivalent formulations. Furthermore, $A_\infty(\mathrm{d}x\,\mathrm{d}t)=\bigcup_{p>1}B_p(\mathrm{d}x\,\mathrm{d}t)$. We refer to [7] for more on $A_\infty$. For $(x,t)\in\mathbb{R}^{n+1}$, and a function $F$, we define the non-tangential maximal function

$$N_*F(x,t)=\sup_{\lambda>0}\sup_{\Lambda\times Q\times I}|F(\mu,y,s)|,\tag{1.3}$$

where $\Lambda=(\lambda/2,\lambda)$, $Q=B(x,\lambda)$ and $I=(t-\lambda^2,t+\lambda^2)$. Given $(x_0,t_0)\in\mathbb{R}^{n+1}$, $\eta>0$, we also introduce the parabolic cone

$$\Gamma^\eta(x_0,t_0):=\{(\lambda,x,t)\in\mathbb{R}^{n+2}_+:\ ||(x-x_0,t-t_0)||<\eta\lambda\}.\tag{1.4}$$

**Definition 1.3.** Let $q \in (1,\infty)$. We say that the Dirichlet problem for $\mathcal{H}$ in $\mathbb{R}^{n+2}_+$ with data in $\mathrm{L}^q(\mathbb{R}^{n+1})$, $D_q$ for short, is *uniquely solvable* if the following holds. Given $f \in \mathrm{L}^q(\mathbb{R}^{n+1})$ then there exists a unique weak solution $u$ such that

$$\begin{aligned} \mathcal{H}u &= 0 && \text{in } \mathbb{R}^{n+2}_+,\\ \lim_{\lambda\to 0} u(\lambda,\cdot,\cdot) &= f(\cdot,\cdot) && \text{in } \mathrm{L}^q(\mathbb{R}^{n+1}) \text{ and n.t.},\\ \|N_* u\|_q &< \infty. \end{aligned}$$

Here, n.t. is short for non-tangentially and means $u(\lambda,x,t) \to f(x_0,t_0)$ for almost every $(x_0,t_0) \in \mathbb{R}^{n+1}$ as $(\lambda,x,t) \to (x_0,t_0)$ through the parabolic cone $\Gamma^\eta(x_0,t_0)$ for some $\eta > 0$.

Assume that parabolic measure $\omega$ belongs to $A_\infty(\mathrm{d}x\,\mathrm{d}t)$ and, in particular, that $\omega$ belong to $B_p(\mathrm{d}x\,\mathrm{d}t)$ for some $p \in (1,\infty)$. This is equivalent to the solvability of $D_q$ for $\mathcal{H}$, $q$ being the dual index to $p$, see for example Theorem 6.2 in [24]. The results in [24] are derived under the assumption of symmetric coefficients. However, the lemmas underlying the proof of Theorem 6.2 in [24] do not rely on this assumption.

1.4. **Statement of the main result.** The following theorem is our main result.

**Theorem 1.4.** *Assume that $A$ satisfies (1.2). Then parabolic measure $\omega$ belongs to $A_\infty(\mathrm{d}x\,\mathrm{d}t)$ with constants depending only $n$ and the ellipticity constants. In particular, there exists $p \in (1,\infty)$ such that $\omega$ belongs to the reverse Hölder class $B_p(\mathrm{d}x\,\mathrm{d}t)$ with $p$ and the constant in the reverse Hölder inequality depending only $n$ and the ellipticity constants. Equivalently, $D_q$, where $q$ is the index dual to $p$, is solvable.*

Theorem 1.4 is new and gives the parabolic counterpart of the corresponding recent result for elliptic measure obtained in [14], with a simplified argument compared to [14]. As mentioned before, Theorem 1.4 is new even in the case when $A$ is symmetric and time-dependent. Note that in [18] the result of Dahlberg was proved for elliptic measure associated to the elliptic counterpart of (1.1) with symmetric $A$, that is, in this case the associated Poisson kernel exists and belongs to $B_2$. In contrast, in the parabolic case it is not clear if such a result holds true if we allow for time-dependent coefficients (the case of time-independent coefficients was treated in [6] and does give $B_2$).

Theorem 1.4 generalizes immediately to the setting of time-independent Lipschitz domains in the following sense. Consider the domain $\{(x_0,x,t) : x_0 > \varphi(x)\}$ above the graph of the time-independent Lipschitz function $\varphi$ and consider the equation

$$\partial_t u - \operatorname{div}_{x_0,x} A(x,t)\nabla_{x_0,x} u = 0$$

in this domain. Using the simple change of variables $(\lambda,x,t) \mapsto (\lambda+\varphi(x),x,t)$, this equation is equivalent to an equation in the upper parabolic half space to which Theorem 1.4 applies. In contrast, this argument does not apply to a time-dependent domain of the form $\{(x_0,x,t) : x_0 > \varphi(x,t)\}$ as the change of variables $(\lambda,x,t) \mapsto (\lambda+\varphi(x,t),x,t)$ with $\varphi$ Lipschitz in both $x$ and $t$ destroys the structure of the equations studied here. If $\varphi$ is only Lipschitz with respect to the parabolic metric, that is, Lipschitz continuous in $x$ and 1/2-Hölder continuous in $t$, then more elaborate changes of variables have to be employed but this changes the nature of the assumption on the coefficients, see [16] for details.

1.5. **Outline of the proof of Theorem 1.4.** The proof consists of three parts: a reduction to a Carleson measure estimate, the construction of a particular set $F$, and the proof of the Carleson measure estimate by partial integration. These three parts have four sources of insights [4,9,14,19]. In general, $c$ will denote a generic constant, not necessarily the same at each instance, which, unless otherwise stated, only depends on $n$ and the ellipticity constants. We often write $c_1 \lesssim c_2$ when we mean that $c_1/c_2$ is bounded by a constant depending only $n$ and the ellipticity constants.

*Reduction to a Carleson measure estimate.* The key insight in [19] is that the $A_\infty$-property of elliptic measure follows once a certain Carleson measure condition is verified. More recently, this idea has also been implemented in the parabolic context: On pp.19-22 in [9] it is shown that in order to conclude $\omega \in A_\infty(\mathrm{d}x\,\mathrm{d}t)$ it suffices to prove the following result, which we state here as our second main theorem.

**Theorem 1.5.** *Let $S \subset \mathbb{R}^{n+1}$ be a bounded Borel set and let $u(\lambda, x, t) := \omega(\lambda, x, t, S)$ be the corresponding weak solution to (1.1) created by the $\mathcal{H}$-parabolic measure $\omega$. Then $u$ satisfies the following Carleson measure estimate: for all parabolic cubes $\Delta \subset \mathbb{R}^{n+1}$,*

$$\int_0^{\ell(\Delta)} \iint_\Delta |\nabla_{\lambda,x} u|^2 \, \lambda \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\lambda \lesssim |\Delta|. \tag{1.5}$$

**Remark 1.6.** Theorem 1.5 is *a priori* equivalent to the statement that (1.5) holds for all parabolic cubes whenever $u$ is the unique solution to the continuous Dirichlet problem for $\mathcal{H}u = 0$ with continuous compactly supported boundary data $f$ satisfying $|f| \leq 1$, see Remark 5 in [9]. Note that in this case $|u| \leq 1$ by the maximum principle. This reformulation has the advantage that it allows one to assume that $A$ is smooth as long as all bounds depend on $A$ only through its ellipticity constants, see p. 20 in [16] for this type of reduction.

Based on Remark 1.6 we can assume *qualitatively* that $A$ is smooth and we are left with the task of proving the Carleson measure estimate (1.5) if $u$ is *any* weak solution to (1.1) bounded by $|u| \leq 1$. The fact that $u$ could be chosen continuous up to the boundary will not enter the argument.

Using that $\partial_\lambda u$ is a solution to $\mathcal{H}u = 0$ ($A$ is independent of $\lambda$), integration by parts in the integral in (1.5) and the standard Caccioppoli inequality on (parabolic) Whitney cubes, see Lemma 2.1 below, it follows that to prove (1.5) it suffices to prove for all parabolic cubes $\Delta$ the inequality

$$\int_0^{\ell(\Delta)} \iint_\Delta |\partial_\lambda u|^2 \, \lambda \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\lambda \lesssim |\Delta|. \tag{1.6}$$

Furthermore, as our equations have real and uniformly elliptic coefficients, the solution $\partial_\lambda u$ satisfies De Giorgi-Moser-Nash estimates, see for example Lemmas 3.4 and 3.5 in [16]. From a John-Nirenberg Lemma for Carleson measures, Lemma 2.14 in [5], it follows that for (1.6) it is enough to prove that the following holds: for each parabolic cube $\Delta \subset \mathbb{R}^{n+1}$, $r := \ell(\Delta)$, there is a Borel set $F \subset \Delta$ with $|\Delta| \lesssim |F|$, such that

$$\int_0^r \iint_F |\partial_\lambda u|^2 \, \lambda \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\lambda \lesssim |\Delta|. \tag{1.7}$$

This completes our reduction to a Carleson measure estimate. To avoid duplication with [9] and for the sake of brevity, we will not give more details concerning these facts. Instead we will simply prove Theorem 1.5 and Theorem 1.4 by verifying (1.7) for a properly constructed set $F$ and this is the main contribution of the paper.

*Construction of the set $F$.* In the context of elliptic measure the freedom of having a set $F \subset \Delta$ at one's disposal in (1.7) was cleverly brought into play in [14] via an adapted Hodge decomposition. Inspired by this, we look for a parabolic Hodge decomposition. To this end, we split the coefficient matrix $A$ as

$$A(x,t) = \begin{bmatrix} A_{\perp\perp}(x,t) & A_{\perp\|}(x,t) \\ A_{\|\perp}(x,t) & A_{\|\|}(x,t) \end{bmatrix}. \tag{1.8}$$

Then $A_{\perp\|}$ is an $n$-dimensional row vector and $A_{\|\perp}$ is an $n$-dimensional column vector. We have a similar decomposition of $A^*$, which is the transpose of $A$ since $A$ has real coefficients.

Introduce the parabolic operator $\mathcal{H}_\| := \partial_t - \operatorname{div}_x A_{\|\|} \nabla_x$ and its adjoint $\mathcal{H}_\|^* := -\partial_t - \operatorname{div}_x A_{\|\|}^* \nabla_x$ on $\mathbb{R}^{n+1}$. Let us recall that $\mathcal{H}_\|$ and $\mathcal{H}_\|^*$ admit the following *hidden coercivity* used systematically in [4, 6, 25, 26]. In fact, it appeared before in [17]. First, we define the homogeneous *energy space* $\dot{\mathsf{E}}(\mathbb{R}^{n+1})$ by taking the closure of test functions $v \in \mathrm{C}_0^\infty(\mathbb{R}^{n+1})$ with respect to the norm

$$\|v\|^2_{\dot{\mathsf{E}}(\mathbb{R}^{n+1})} := \iint_{\mathbb{R}^{n+1}} |\nabla_x v|^2 + |D_t^{1/2} v|^2 \, \mathrm{d}x \, \mathrm{d}t$$

and identifying functions that differ only by a constant. Here, the half-order $t$-derivative $D_t^{1/2}$ is defined via the Fourier symbol $\mathrm{i}|\tau|^{1/2}$. This closure can be realized in $\mathrm{L}^2(\mathbb{R}^{n+1}) + \mathrm{L}^\infty(\mathbb{R}^{n+1})$ and modulo constants $\dot{\mathsf{E}}(\mathbb{R}^{n+1})$ becomes a Hilbert space, see for example Section 3.2 in [4]. The corresponding inhomogeneous energy space $\mathsf{E}(\mathbb{R}^{n+1}) = \dot{\mathsf{E}}(\mathbb{R}^{n+1}) \cap \mathrm{L}^2(\mathbb{R}^{n+1})$ is equipped with the obvious Hilbertian norm. Denoting by $H_t$ the Hilbert transform with respect to the $t$-variable, we can factorize $\partial_t = D_t^{1/2} H_t D_t^{1/2}$ and this in turn allows us to define $\mathcal{H}_\parallel$ as a bounded operator from $\dot{\mathsf{E}}(\mathbb{R}^{n+1})$ into its (anti)-dual $\dot{\mathsf{E}}(\mathbb{R}^{n+1})^*$ via

$$(\mathcal{H}_\parallel u)(v) := \iint_{\mathbb{R}^{n+1}} D_t^{1/2} u \cdot \overline{H_t D_t^{1/2} v} + A_{\parallel\parallel} \nabla_x u \cdot \overline{\nabla_x v} \, \mathrm{d}x \, \mathrm{d}t. \tag{1.9}$$

The hidden coercivity of the sesquilinear form on the right-hand side now pays for this operator being invertible with operator norm depending only on $n$ and the ellipticity constants of $A_{\parallel\parallel}$, see Section 7 of [17] or Lemma 5.9 in [6]. An analogous construction applies to $\mathcal{H}_\parallel^*$. Considering a parabolic cube $\Delta = \Delta_r \subset \mathbb{R}^{n+1}$, we let $\chi_{8\Delta} = \chi_{8\Delta}(x,t)$ be a smooth cut off for $8\Delta$ which is 1 on $8\Delta$, vanishes outside of $16\Delta$ and satisfies $r|\nabla_x \chi_{8\Delta}| + r^2 |\partial_t \chi_{8\Delta}| \leq c$. Then, there exist $\varphi, \tilde{\varphi} \in \dot{\mathsf{E}}(\mathbb{R}^{n+1})$ solving

$$\mathcal{H}_\parallel^* \varphi = \operatorname{div}_x(A_{\perp\parallel} \chi_{8\Delta}), \qquad \mathcal{H}_\parallel \tilde{\varphi} = \operatorname{div}_x(A_{\parallel\perp} \chi_{8\Delta}), \tag{1.10}$$

and satisfying the *a priori* estimates

$$\begin{aligned} &\iint_{\mathbb{R}^{n+1}} |\nabla_x \varphi|^2 + |H_t D_t^{1/2} \varphi|^2 \, \mathrm{d}x \, \mathrm{d}t \lesssim \iint_{16\Delta} |A_{\perp\parallel}|^2 \, \mathrm{d}x \, \mathrm{d}t \lesssim |\Delta|, \\ &\iint_{\mathbb{R}^{n+1}} |\nabla_x \tilde{\varphi}|^2 + |H_t D_t^{1/2} \tilde{\varphi}|^2 \, \mathrm{d}x \, \mathrm{d}t \lesssim \iint_{16\Delta} |A_{\parallel\perp}|^2 \, \mathrm{d}x \, \mathrm{d}t \lesssim |\Delta|. \end{aligned} \tag{1.11}$$

As we can undo the factorization of $\partial_t$ leading to (1.9) if $v$ is a test function, (1.10) holds *a fortiori* in the usual weak sense. More in the spirit of operator theory, Lemma 4 in [3] shows that the part of $\mathcal{H}_\parallel$ in $\mathrm{L}^2(\mathbb{R}^{n+1})$ with maximal domain

$$\mathsf{D}(\mathcal{H}_\parallel) = \{u \in \mathsf{E}(\mathbb{R}^{n+1}) : \mathcal{H}_\parallel u \in \mathrm{L}^2(\mathbb{R}^{n+1})\}$$

is *maximal accretive*, that is, for every $\mu \in \mathbb{C}$ with $\operatorname{Re} \mu > 0$ the operator $\mu + \mathcal{H}_\parallel$ is invertible and $\|(\mu + \mathcal{H}_\parallel)^{-1}\|_{\mathrm{L}^2 \to \mathrm{L}^2} \leq (\operatorname{Re} \mu)^{-1}$ holds. The recent resolution of the Kato problem for parabolic operators identifies the domain of its unique maximal accretive square root as $\mathsf{D}(\mathcal{H}_\parallel^{1/2}) = \mathsf{E}(\mathbb{R}^{n+1})$ with a homogeneous estimate

$$\|\mathcal{H}_\parallel^{1/2} v\|_2 \sim \|\nabla_x v\|_2 + \|H_t D_t^{1/2} v\|_2 \qquad \text{for } v \in \mathsf{E}(\mathbb{R}^{n+1}),$$

see Theorem 2.6 in [4]. Thus, writing

$$(\mu + \mathcal{H}_\parallel)^{-1} v = \mathcal{H}_\parallel^{-1/2} (\mu + \mathcal{H}_\parallel)^{-1} \mathcal{H}_\parallel^{1/2} v,$$

we can extend $(\mu + \mathcal{H}_\parallel)^{-1}$ by density from $\mathsf{E}(\mathbb{R}^{n+1})$ to a bounded and invertible operator on $\dot{\mathsf{E}}(\mathbb{R}^{n+1})$. Again we also have the analogous results for $\mathcal{H}_\parallel^*$. In particular, for $m$ a natural number and $\lambda > 0$ we can introduce the higher order resolvents of $\varphi, \tilde{\varphi}$,

$$P_\lambda^* \varphi := (1 + \lambda^2 \mathcal{H}_\parallel^*)^{-m} \varphi, \qquad P_\lambda \tilde{\varphi} := (1 + \lambda^2 \mathcal{H}_\parallel)^{-m} \tilde{\varphi}, \tag{1.12}$$

within the homogeneous energy space $\dot{\mathsf{E}}(\mathbb{R}^{n+1})$. In the further course we will fix $m$ large enough (without trying to get optimal values) to have a number of estimates at our disposal.

Coming back to the actual construction of $F$, we also introduce the parabolic maximal differential operator

$$\mathbb{D} v(x,t) \ := \ \sup_{\varrho > 0} \fint\!\!\fint_{\Delta_\varrho(x,t)} \frac{|v(x,t) - v(y,s)|}{\|(x-y, t-s)\|} \, \mathrm{d}y \, \mathrm{d}s, \quad v \in \dot{\mathsf{E}}(\mathbb{R}^{n+1}), \tag{1.13}$$

which maps boundedly into $\mathrm{L}^2(\mathbb{R}^{n+1})$ as we shall prove later on in Lemma 2.3. Here, $\|\cdot\|$ indicates again the parabolic distance. In particular, (1.11) implies

$$\|\mathbb{D}\varphi\|_2 + \|\mathbb{D}\tilde{\varphi}\|_2 \lesssim |\Delta|^{\frac{1}{2}}. \tag{1.14}$$

The non-tangential maximal function operator $N_*$ acting on measurable functions $F$ on $\mathbb{R}^{n+2}_+$ was introduced in (1.3). For $(x,t) \in \mathbb{R}^{n+1}$ we also introduce the integrated non-tangential maximal function

$$\widetilde{N}_* F(x,t) = \sup_{\lambda>0} \left( \fint\!\!\fint\!\!\fint_{\Lambda\times Q\times I} |F(\mu,y,s)|^2 \,\mathrm{d}\mu\,\mathrm{d}y\,\mathrm{d}s \right)^{1/2}, \tag{1.15}$$

where $\Lambda = (\lambda/2,\lambda)$, $Q = B(x,\lambda)$ and $I = (t-\lambda^2, t+\lambda^2)$. If $g : \mathbb{R}^{n+1} \to \mathbb{R}$ and is locally integrable we let $\mathcal{M}(g)$ be the $(n+1)$-dimensional (parabolic) Hardy-Littlewood maximal function

$$\mathcal{M}(g)(x,t) = \sup_{\varrho>0} \fint\!\!\fint_{\Delta_\varrho(x,t)} |g| \,\mathrm{d}y\,\mathrm{d}s$$

and we let $\mathcal{M}_x$ and $\mathcal{M}_t$ denote the standard (euclidean) Hardy-Littlewood maximal operators in the $x$ and $t$ variables only. Our construction of $F$ is then done through the following definition.

**Definition 1.7.** Let $\Delta$ be fixed and also fix $m = n+1$. Given $\kappa_0 \gg 1$, we let $F \subset 16\Delta$ be the set of all $(x,t) \in 16\Delta$ such that the following requirements are met:

$$\begin{aligned}
&\text{(i)} && \mathcal{M}(|\nabla_x \varphi|^2)(x,t) + \mathcal{M}(|\nabla_x \tilde{\varphi}|^2)(x,t) \le \kappa_0^2,\\
&\text{(ii)} && \mathcal{M}_x\,\mathcal{M}_t(|H_t D_t^{1/2}\varphi|)(x,t) + \mathcal{M}_x\,\mathcal{M}_t(|H_t D_t^{1/2}\tilde{\varphi}|)(x,t) \le \kappa_0,\\
&\text{(iii)} && \mathbb{D}\varphi(x,t) + \mathbb{D}\tilde{\varphi}(x,t) \le \kappa_0,\\
&\text{(iv)} && N_*(\partial_\lambda P_\lambda^* \varphi)(x,t) + N_*(\partial_\lambda P_\lambda \tilde{\varphi})(x,t) \le \kappa_0,\\
&\text{(v)} && \widetilde{N}_*(\nabla_x P_\lambda^* \varphi)(x,t) + \widetilde{N}_*(\nabla_x P_\lambda \tilde{\varphi})(x,t) \le \kappa_0.
\end{aligned}$$

Given $\Delta$ and $\kappa_0 \gg 1$, let $F$ be defined as above. Then, using the weak type $(1,1)$ of $\mathcal{M}$, the strong type $(2,2)$ of $\mathcal{M}_x\,\mathcal{M}_t$, the estimates (1.11) and (1.14) and the $\mathrm{L}^2$-bounds for the non-tangential maximal functions that will later be obtained in Lemma 4.2 and Lemma 4.5, it follows that

$$|16\Delta \setminus F| \lesssim (\kappa_0^{-2} + \kappa_0^{-1})|\Delta|.$$

In particular, we can now choose $\kappa_0$, depending only on $n$ and the ellipticity constants, so that

$$\frac{|16\Delta \setminus F|}{|\Delta|} \le 1/1000. \tag{1.16}$$

This completes our construction of the set $F$ and from now on $\kappa_0$ is fixed as stated ensuring that (1.16) holds.

*Proof of the Carleson measure estimate.* Based on the previous steps, the proofs of Theorem 1.4 and Theorem 1.5 are reduced to verifying (1.7). To do this we construct, given $\Delta = \Delta_r$, $F \subset \Delta$ a Borel set and $\epsilon > 0$, a parabolic sawtooth region above $F$ using parabolic cones of aperture $0 < \eta \ll 1$. The parameter $\eta$ is an important degree of freedom in the argument. In (5.4) we will construct a (smooth) cut-off function $\Psi = \Psi_{\eta,\epsilon}$ such that $\Psi(\lambda,x,t) = 1$ on $F \times (2\epsilon, 2r)$ and $\Psi(\lambda,x,t) = 0$ if $\lambda \in (0,\epsilon) \cup (4r,\infty)$, and we let

$$J_{\eta,\epsilon} := \iiint_{\mathbb{R}^{n+2}_+} A\nabla_{\lambda,x} u \cdot \nabla_{\lambda,x} u\, \Psi^2 \lambda \,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda.$$

Then, by ellipticity of $A$,

$$\int_{2\epsilon}^{r} \iint_F |\partial_\lambda u|^2 \,\lambda \,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda \lesssim J_{\eta,\epsilon}. \tag{1.17}$$

Since $\Psi$ has compact support in the upper half space, we can ensure finiteness of $J_{\eta,\epsilon}$ and hence everything boils down to the following key lemma:

**Lemma 1.8** (Key Lemma)**.** *Let $\sigma, \eta \in (0,1)$ be given degrees of freedom. Then there exist a finite constant $c$ depending only on $n$ and the ellipticity constants, and a finite constant $\tilde{c}$ depending additionally on $\sigma$ and $\eta$, such that*

$$J_{\eta,\epsilon} \le (\sigma + c\eta) J_{\eta,\epsilon} + \tilde{c}|\Delta|.$$

Indeed, choosing $\sigma$ and $\eta$ small, both depending at most on $n$ and the ellipticity constants, we first derive

$$J_{\eta,\epsilon} \leq 2\tilde{c}|\Delta|,$$

where now $\eta$ is fixed but $\tilde{c}$ is still independent of $\epsilon$. On letting $\epsilon \to 0$, we see from (1.17) that the estimate (1.7) holds. As discussed before, this completes the proofs of Theorem 1.5 and Theorem 1.4.

1.6. **Organization of the paper.** Section 2 is partly of preliminary nature and we here prove (1.14). Section 3 is devoted to the important square function estimates underlying the proof of Theorem 1.8. These estimates rely on recent results established in [4]. In Section 4 we prove the non-tangential maximal function estimates underlying the statements in Definition 1.7 (iv)-(v). Based on the material of Sections 2-4 the set $F$ introduced in Definition 1.7 is well-defined and we can ensure (1.16). In particular, thereby the set $F \subset 16\Delta$ is fixed as we proceed into Section 5 and Section 6. In Section 5 we then introduce sawtooth domains above $F$, we define the cut-off function $\Psi = \Psi_{\eta,\epsilon}$ referred to above and we prove some auxiliary Carleson measure estimates. The proof of Lemma 1.8 is given in Section 6.

## 2. Technical tools

In this section we collect three technical lemmas that shall prove useful in the further course. We begin with standard Caccioppoli estimate which we here state without proof.

**Lemma 2.1** (Caccioppoli estimate). *Let $u$ be a weak solution to $\partial_t u - \operatorname{div}_{\lambda,x} A\nabla_{\lambda,x} u + \alpha u = 0$ on $\mathbb{R}^{n+2}_+$ where $\alpha \in \mathrm{L}^\infty(\mathbb{R}^{n+2}_+)$, $\alpha \geq 0$, and let $\psi \in \mathrm{C}_0^\infty(\mathbb{R}^{n+2}_+)$. Then*

$$\iiint |\nabla_{\lambda,x} u|^2 \psi^2 \, \mathrm{d}x \, \mathrm{d}\lambda \, \mathrm{d}t \leq c \iiint |u|^2 |\nabla_{\lambda,x}\psi|^2 \, \mathrm{d}x \, \mathrm{d}\lambda \, \mathrm{d}t$$

*for some finite constant $c$ depending on $n$ and the ellipticity constants of $A$.*

Next, we record a Poincaré-type estimate for functions in the homogeneous energy space $\dot{\mathsf{E}}(\mathbb{R}^{n+1})$. We use the standard notation for parabolic cubes introduced in Section 1.3.

**Lemma 2.2.** *Let $v \in \dot{\mathsf{E}}(\mathbb{R}^{n+1})$ and let $\Delta_\varrho = \Delta_\varrho(x_0, t_0) \subset \mathbb{R}^{n+1}$ be a parabolic cube. Then*

$$\frac{1}{\varrho} \fint\!\!\fint_{\Delta_\varrho} \Big| v - \fint\!\!\fint_{\Delta_\varrho} v \Big| \, \mathrm{d}x \, \mathrm{d}t \lesssim \mathcal{M}(|\nabla_x v|)(x_0,t_0) + \mathcal{M}_x \mathcal{M}_t(|H_t D_t^{1/2} v|)(x_0,t_0).$$

*Proof.* We write $\Delta_\varrho = Q_\varrho \times I_\varrho$ and we let

$$f(t) := \fint_{Q_\varrho} v(x,t) \, \mathrm{d}x,$$

noting that this function is contained in the homogeneous fractional Sobolev space $\dot{\mathrm{H}}^{1/2}(\mathbb{R})$, see Section 3.1 in [4]. Then

$$\fint\!\!\fint_{\Delta_\varrho} \Big| v - \fint\!\!\fint_{\Delta_\varrho} v \Big| \, \mathrm{d}x \, \mathrm{d}t \lesssim \varrho \mathcal{M}(|\nabla_x v|)(x_0,t_0) + \fint_{I_\varrho} \Big| f - \fint_{I_\varrho} f \Big| \, \mathrm{d}t$$

by Poincaré's inequality in the spatial variable $x$ only. Furthermore, for $f \in \dot{\mathrm{H}}^{1/2}(\mathbb{R})$ we have at hand the non-local Poincaré inequality

$$\fint_{I_\varrho} \Big| f - \fint_{I_\varrho} f \Big| \, \mathrm{d}t \leq \varrho \sum_{k\in\mathbb{Z}} \frac{1}{1+|k|^{3/2}} \fint_{k\varrho^2 + I_\varrho} |H_t D_t^{1/2} f| \, \mathrm{d}t,$$

see Lemma 8.3 in [4]. Rearranging the covering of the real line by translates of $I_\varrho$ into a covering by dyadic annuli, we obtain

$$\begin{aligned}\fint_{I_\varrho}\Big|f-\fint_{I_\varrho} f\Big|\,\mathrm{d}t &\leq \varrho\sum_{m\geq 0}2^{-m}\fint_{4^m I_\varrho}|H_tD_t^{1/2}f|\,\mathrm{d}\sigma\\ &\leq \varrho\sum_{m\geq 0}2^{-m}\fint\!\!\!\fint_{Q\times 4^m I_\varrho}|H_tD_t^{1/2}v|\,\mathrm{d}x\,\mathrm{d}t\\ &\leq 2\varrho\,\mathcal{M}_x(\mathcal{M}_t(|H_tD_t^{1/2}v|)(x_0,t_0),\end{aligned}$$

where the second step can rigorously be justified using Fubini's theorem, see Lemma 3.10 in [4]. $\square$

As a consequence, we obtain an important estimate for the parabolic maximal differential operator $\mathbb{D}$ defined in (1.13).

**Lemma 2.3.** *The operator $\mathbb{D}$ maps $\dot{\mathsf{E}}(\mathbb{R}^{n+1})$ boundedly into $\mathrm{L}^2(\mathbb{R}^{n+1})$.*

*Proof.* Let $v\in\dot{\mathsf{E}}(\mathbb{R}^{n+1})$. We first claim that

$$\begin{aligned}\frac{|v(x,t)-v(y,s)|}{\|(x-y,t-s)\|} &\lesssim \mathcal{M}(|\nabla_x v|)(x,t)+\mathcal{M}_x\,\mathcal{M}_t(|H_tD_t^{1/2}v|)(x,t)\\ &\quad+\mathcal{M}(|\nabla_x v|)(y,s)+\mathcal{M}_x\,\mathcal{M}_t(|H_tD_t^{1/2}v|)(y,s)\end{aligned}\tag{2.1}$$

holds for almost every $(x,t)$, $(y,s)\in\mathbb{R}^{n+1}$. Indeed, let $(x,t)$ be a Lebesgue point for $v$ and for $\varrho>0$ let $v_\varrho$ denote the average of $v$ over the parabolic cube $\Delta_\varrho:=\Delta_\varrho(x,t)$. Then, by a telescoping sum and an application of Lemma 2.2,

$$\begin{aligned}|v(x,t)-v_\varrho| &\leq \sum_{k=0}^{\infty}|v_{2^{-k-1}\varrho}-v_{2^{-k}\varrho}|\\ &\lesssim \sum_{k=0}^{\infty}2^{-k}\varrho\Big(\mathcal{M}(|\nabla_x v|)(x,t)+\mathcal{M}_x\,\mathcal{M}_t(|H_tD_t^{1/2}v|)(x,t)\Big)\\ &\leq 2\varrho\Big(\mathcal{M}(|\nabla_x v|)(x,t)+\mathcal{M}_x\,\mathcal{M}_t(|H_tD_t^{1/2}v|)(x,t)\Big).\end{aligned}$$

Furthermore, let also $(y,s)$ be a Lebesgue point for $v$ and assume that $(y,s)\in\Delta_\varrho(x,t)$. Then $\Delta_\varrho(x,t)\subset\Delta_{2\varrho}(y,s)$ and we obtain as above,

$$\begin{aligned}|v(y,s)-v_\varrho| &\leq \Big|v(y,s)-\fint\!\!\!\fint_{\Delta_{2\varrho}(y,s)}v\Big|+\Big|\fint\!\!\!\fint_{\Delta_{2\varrho}(y,s)}v-v_\varrho\Big|\\ &\lesssim \varrho\Big(\mathcal{M}(|\nabla_x v|)(y,s)+\mathcal{M}_x\,\mathcal{M}_t(|H_tD_t^{1/2}v|)(y,s)\Big).\end{aligned}$$

Now, for $(x,t)\neq(y,s)$ as above we can specify $\varrho:=\|(x-y,t-s)\|$ and (2.1) follows by adding up the previous two estimates. In particular, we obtain

$$\begin{aligned}\mathbb{D}v(x,t) &\lesssim \mathcal{M}(|\nabla_x v|)(x,t)+\mathcal{M}_x\,\mathcal{M}_t(|H_tD_t^{1/2}v|)(x,t)\\ &\quad+\mathcal{M}\,\mathcal{M}(|\nabla_x v|)(x,t)+\mathcal{M}\,\mathcal{M}_x\,\mathcal{M}_t(|H_tD_t^{1/2}\varphi|)(x,t)\end{aligned}$$

for almost every $(x,t)\in\mathbb{R}^{n+1}$ and since all occurring maximal operators are $\mathrm{L}^2$-bounded, we conclude $\|\mathbb{D}v\|_2\lesssim\|\nabla_x v\|_2+\|H_tD_t^{1/2}v\|_2$ as required. $\square$

## 3. Functional calculus and square function estimates

In this section we prove the important square function estimates for $\mathcal{H}_\parallel$ and $\mathcal{H}_\parallel^*$ underlying the proof of Lemma 1.8. Most of this material is taken from [4].

Given $\mu\in(0,\pi/2)$ we let

$$\mathrm{S}_\mu:=\{z\in\mathbb{C}:|\arg z|<\mu\text{ or }|\arg z-\pi|<\mu\}$$

denote the open double sector of angle $\mu$. We let

$$\Psi(\mathrm{S}_\mu) := \{\psi \in H^\infty(\mathrm{S}_\mu) : \exists\, \alpha > 0,\ C > 0,\ \text{such that } |\psi(z)| \le C \min\{|z|^\alpha, |z|^{-\alpha}\}\}$$

where $H^\infty(\mathrm{S}_\mu)$ is the set of all bounded holomorphic functions on $\mathrm{S}_\mu$. Furthermore, recall that an operator $T$ in a Hilbert space is *bisectorial* of *angle* $\omega \in (0, \pi/2)$ if its spectrum is contained in the closure of $\mathrm{S}_\omega$ and if, for each $\mu \in (\omega, \pi/2)$, the map $z \mapsto z(z-T)^{-1}$ is uniformly bounded on $\mathbb{C} \setminus \mathrm{S}_\mu$. In this case a bounded operator $\psi(T)$ is defined by the functional calculus for bisectorial operators and we refer to [23] or [10] for the few essentials of this theory used in this section. Turning to concrete operators, we represent vectors $h \in \mathbb{C}^{n+2}$ as

$$h = \begin{bmatrix} h_\perp \\ h_\parallel \\ h_\theta \end{bmatrix},$$

where the normal part $h_\perp$ is scalar valued, the tangential part $h_\parallel$ is valued in $\mathbb{C}^n$ and the time part $h_\theta$ is again scalar valued and let

$$P := \begin{bmatrix} 0 & \mathrm{div}_x & -D_t^{1/2} \\ -\nabla_x & 0 & 0 \\ -H_t D_t^{1/2} & 0 & 0 \end{bmatrix}, \qquad M := \begin{bmatrix} 1 & 0 & 0 \\ 0 & A_{\parallel\parallel} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Here, $M$ is considered as a bounded multiplication operator on $\mathrm{L}^2(\mathbb{R}^{n+1}; \mathbb{C}^{n+2})$ and the parabolic Dirac operator $P$ is an unbounded operator in $\mathrm{L}^2(\mathbb{R}^{n+1}; \mathbb{C}^{n+2})$ with maximal domain. The link with the parabolic operator $\mathcal{H}_\parallel$ is that $(PM)^2$ and $(MP)^2$ are operator matrices in block form

$$(PM)^2 = \begin{bmatrix} \mathcal{H}_\parallel & 0 & 0 \\ 0 & * & * \\ 0 & * & * \end{bmatrix}, \qquad (MP)^2 = \begin{bmatrix} \mathcal{H}_\parallel & 0 & 0 \\ 0 & * & * \\ 0 & * & * \end{bmatrix}, \tag{3.1}$$

where the entries $*$ do not play any role in the following but of course they could be computed explicitly. Note that taking adjoints in (3.1), hence using $(P^*M^*)^2$ or $(M^*P^*)^2$, allows to obtain $\mathcal{H}_\parallel^*$. The following theorem provides square function estimates.

**Theorem 3.1.** *The operator $PM$ is a bisectorial operator in $\mathrm{L}^2(\mathbb{R}^{n+1}; \mathbb{C}^{n+2})$ with angle $\omega$ of bisectoriality depending only upon $n$ and the ellipticity constants of $A$ and the same range as $P$, that is, $\mathsf{R}(PM) = \mathsf{R}(P)$. Let $\mu \in (\omega, \pi/2)$ and consider $\psi \in \Psi(\mathrm{S}_\mu)$ non vanishing on each connected component of $\mathrm{S}_\mu$. Then*

$$\int_0^\infty \|\psi(\lambda PM)h\|_2^2 \, \frac{\mathrm{d}\lambda}{\lambda} \sim \|h\|_2^2 \quad \text{if } h \in \overline{\mathsf{R}(PM)}$$

*and the implicit constants in this estimate depend only upon $n$, the ellipticity constants of $A$, $\mu$ and $\psi$. The same holds true for $MP$ on $\overline{\mathsf{R}(MP)} = M\overline{\mathsf{R}(P)}$ and with $PM$, $MP$, replaced by $P^*M^*$, $M^*P^*$.*

*Proof.* For $PM$, this is a mere consequence of Theorem 2.3 in [4]: Indeed, this theorem states all assertions apart from that only the quadratic estimate

$$\int_0^\infty \|\lambda PM(1 + \lambda^2 PMPM)^{-1}h\|_2^2 \, \frac{\mathrm{d}\lambda}{\lambda} \sim \|h\|_2^2 \qquad \text{for } h \in \overline{\mathsf{R}(PM)}$$

is mentioned. But due to a general result on quadratic estimates for bisectorial operators on Hilbert spaces, see [23] or Theorem 3.4.11 in [10], this quadratic estimate is in fact equivalent to the set of quadratic estimates stated above. The statement for $MP$ follows from the fact that this operator is similar to $PM$ on their respective ranges by $MP = M(PM)M^{-1}$. The statements for $P^*M^*$, $M^*P^*$ follow by duality, see again [10, 23]. □

Below, we single out some particular instances of the theorem above and reformulate them in terms of $\mathcal{H}_\parallel$ and $\mathcal{H}_\parallel^*$ to have direct references later on. Throughout, we let $\varphi$, $\tilde{\varphi}$ be as in (1.10), (1.11) and we recall that the resolvent operators $P_\lambda^*$, $P_\lambda$ were defined in (1.12) for the moment with $m$ unspecified.

**Lemma 3.2.** *There exists $c$, $1 \leq c < \infty$, depending only on $n$, the ellipticity constants and $m \geq 1$ such that*

$$\text{(i)} \qquad \iiint_{\mathbb{R}^{n+2}_+} |\partial_\lambda P_\lambda^* \varphi|^2 + |\partial_\lambda P_\lambda \tilde{\varphi}|^2 \, \frac{\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda}{\lambda} \leq c|\Delta|,$$

$$\text{(ii)} \qquad \iiint_{\mathbb{R}^{n+2}_+} |\lambda \nabla_x \partial_\lambda P_\lambda^* \varphi|^2 + |\lambda \nabla_x \partial_\lambda P_\lambda \tilde{\varphi}|^2 \, \frac{\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda}{\lambda} \leq c|\Delta|,$$

$$\text{(iii)} \qquad \iiint_{\mathbb{R}^{n+2}_+} |\lambda \mathcal{H}_{\|}^* P_\lambda^* \varphi|^2 + |\lambda \mathcal{H}_{\|} P_\lambda \tilde{\varphi}|^2 \, \frac{\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda}{\lambda} \leq c|\Delta|,$$

$$\text{(iv)} \qquad \iiint_{\mathbb{R}^{n+2}_+} |\lambda^2 \mathcal{H}_{\|}^* \partial_\lambda P_\lambda^* \varphi|^2 + |\lambda^2 \mathcal{H}_{\|} \partial_\lambda P_\lambda \tilde{\varphi}|^2 \, \frac{\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda}{\lambda} \leq c|\Delta|.$$

*Proof.* In the following we will only prove the estimates for $P_\lambda \tilde{\varphi}$, the estimates for $P_\lambda^* \varphi$ being proved similarly with $P^*$ and $M^*$ replacing $P$ and $M$. Note that $\tilde{\varphi} \in \dot{\mathsf{E}}(\mathbb{R}^{n+1})$ and hence the following calculations can be justified, for example, by approximating $\tilde{\varphi}$ by smooth and compactly supported functions in the semi-norm of $\dot{\mathsf{E}}(\mathbb{R}^{n+1})$. Keeping this in mind, we may directly argue with $\tilde{\varphi}$. We begin with (iii). Let

$$h := \begin{bmatrix} 0 \\ -\nabla_x \tilde{\varphi} \\ -H_t D_t^{1/2} \tilde{\varphi} \end{bmatrix} = P \begin{bmatrix} \tilde{\varphi} \\ 0 \\ 0 \end{bmatrix} \in \overline{\mathsf{R}(P)} = \overline{\mathsf{R}(PM)},$$

and note, using (3.1) and elementary manipulations of resolvents of $PM$ and $MP$, that

$$\begin{bmatrix} \lambda \mathcal{H}_{\|} P_\lambda \tilde{\varphi} \\ 0 \\ 0 \end{bmatrix} = \lambda (MP)^2 (1 + (\lambda MP)^2)^{-m} \begin{bmatrix} \tilde{\varphi} \\ 0 \\ 0 \end{bmatrix} = M(\lambda PM)(1 + (\lambda PM)^2)^{-m} P \begin{bmatrix} \tilde{\varphi} \\ 0 \\ 0 \end{bmatrix} = M\psi(\lambda PM) h$$

where $\psi(z) := z(1+z^2)^{-m}$. Hence,

$$\iiint_{\mathbb{R}^{n+2}_+} |\lambda \mathcal{H}_{\|} P_\lambda \tilde{\varphi}|^2 \, \frac{\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda}{\lambda} \lesssim \|h\|_2^2 = ||\nabla_x \tilde{\varphi}||_2^2 + ||H_t D_t^{1/2} \tilde{\varphi}||_2^2 \lesssim |\Delta|,$$

by an application of Theorem 3.1 and (1.11). This proves (iii). Likewise, (i) and (iv) follow with $\psi(z) = -2mz^2(1+z^2)^{-m-1}$ and $\psi(z) = -2mz^3(1+z^2)^{-m-1}$, respectively. Finally, to prove (ii) we write analogously

$$\begin{bmatrix} 0 \\ -\lambda \nabla_x \partial_\lambda P_\lambda \tilde{\varphi} \\ -\lambda H_t D_t^{1/2} \partial_\lambda P_\lambda \tilde{\varphi} \end{bmatrix} = P \begin{bmatrix} \lambda \partial_\lambda P_\lambda \tilde{\varphi} \\ 0 \\ 0 \end{bmatrix} = -2mP(\lambda MP)^2 (1 + (\lambda MP)^2)^{-m-1} \begin{bmatrix} \tilde{\varphi} \\ 0 \\ 0 \end{bmatrix} = \tilde{\psi}(\lambda PM) h$$

with $\tilde{\psi}(z) = -2mz^2(1+z^2)^{-m-1}$ and the claim follows by yet another application of Theorem 3.1. $\square$

**Lemma 3.3.** *There exists $c$, $1 \leq c < \infty$, depending only on $n$, the ellipticity constants and $m \geq 1$ such that*

$$\iiint_{\mathbb{R}^{n+2}_+} |(I - P_\lambda^*)\varphi|^2 + |(I - P_\lambda)\tilde{\varphi}|^2 \, \frac{\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda}{\lambda^3} \leq c|\Delta|.$$

*Proof.* We have

$$(I - P_\lambda)\tilde{\varphi} = \int_0^\lambda \partial_\sigma P_\sigma \tilde{\varphi} \, \mathrm{d}\sigma.$$

Applying Hardy's inequality and Lemma 3.2 (i) we see that

$$\iiint_{\mathbb{R}^{n+2}_+} |(I - P_\lambda)\tilde{\varphi}|^2 \, \frac{\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda}{\lambda^3} \lesssim \iiint_{\mathbb{R}^{n+2}_+} |\partial_\lambda P_\lambda \tilde{\varphi}|^2 \, \frac{\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda}{\lambda} \leq c|\Delta|.$$

The proof of the estimate for $(I - P_\lambda^*)\varphi$ is similar. $\square$

## 4. Non-tangential maximal function estimates

The pointwise non-tangential maximal operator $N_*$ was introduced in (1.3) and its integrated version $\widetilde{N}_*$ was defined in (1.15). In this section we use the previously obtained square function estimates to derive bounds for these maximal functions.

**Theorem 4.1.** *Let $h \in \overline{\mathsf{R}(PM)}$ and let $F(\lambda, x, t) = (\mathrm{e}^{-\lambda[PM]}h)(x,t)$, with $[PM] := \sqrt{(PM)^2}$. Then*

$$\|\widetilde{N}_* F\|_2 \sim \|h\|_2,$$

*where the implicit constants depend only on dimension and the ellipticity constants of $A$. The conclusion remains true also with $PM$ replaced by $P^*M^*$.*

*Proof.* For $PM$, this is Theorem 2.12 in [4]. The same statement can be proved for $P^*M^*$. □

In the following $P_\lambda^* \varphi$, $P_\lambda \tilde{\varphi}$ are again as defined in (1.12).

**Lemma 4.2.** *There exists $c$, $1 \leq c < \infty$, depending only on $n$, the ellipticity constants and $m \geq 1$ such that*

$$\|\widetilde{N}_*(\nabla_x P_\lambda^* \varphi)\|_2^2 + \|\widetilde{N}_*(\nabla_x P_\lambda \tilde{\varphi})\|_2^2 \leq c|\Delta|.$$

*Proof.* We only give the proof of the estimate of $\widetilde{N}_*(\nabla_x P_\lambda \tilde{\varphi})$. To start the proof we first note as in the proof of Lemma 3.2 (ii) that

$$\begin{bmatrix} 0 \\ -\nabla_x P_\lambda \tilde{\varphi} \\ -H_t D_t^{1/2} P_\lambda \tilde{\varphi} \end{bmatrix} = \vartheta(\lambda PM)h, \qquad h = \begin{bmatrix} 0 \\ -\nabla_x \tilde{\varphi} \\ -H_t D_t^{1/2} \tilde{\varphi} \end{bmatrix} = P \begin{bmatrix} \tilde{\varphi} \\ 0 \\ 0 \end{bmatrix} \in \overline{\mathsf{R}(PM)}$$

where now $\vartheta(z) = (1+z^2)^{-m}$. Thus, $-\nabla_x P_\lambda \tilde{\varphi} = (\vartheta(\lambda PM)h)_\parallel$ and we have to estimate $\|\widetilde{N}_*(\vartheta(\lambda PM)h)\|_2$. To this end, we first note

$$\|\widetilde{N}_*(\mathrm{e}^{-\lambda[PM]}h)\|_2^2 \lesssim \|h\|_2^2 = \|\nabla_x \tilde{\varphi}\|_2^2 + \|H_t D_t^{1/2} \tilde{\varphi}\|_2^2 \lesssim |\Delta|, \tag{4.1}$$

using Theorem 4.1, the construction of $h$ and (1.11). Now let $\psi(z) := \vartheta(z) - e^{-\sqrt{z^2}}$. Tonelli's theorem yields

$$\|\widetilde{N}_*(\psi(\lambda PM)h)\|_2^2 \lesssim \int_0^\infty \|\psi(\lambda PM)h\|_2^2 \, \frac{\mathrm{d}\lambda}{\lambda},$$

see for example Lemma 8.9 in [4] for an explicit proof. Since $\psi \in \Psi(\mathrm{S}_\mu)$ for every $\mu \in (0, \pi/2)$, we deduce from Theorem 3.1 that

$$\|\widetilde{N}_*(\psi(\lambda PM)h)\|_2^2 \lesssim \|h\|_2^2 \lesssim |\Delta|,$$

which in combination with (4.1) yields the claim. □

For the $\lambda$-derivatives of $P_\lambda^* \varphi$ and $P_\lambda \tilde{\varphi}$ we could get $\mathrm{L}^2$-bounds for the integrated non-tangential maximal function immediately from the square function estimate in Lemma 3.2 (i). However, this would not be enough for our purpose. To derive the required bounds for the *pointwise* non-tangential maximal function, we need the following lemma.

**Lemma 4.3.** *For $\lambda > 0$ and $m \geq 1$, the resolvent $P_\lambda = (1 + \lambda^2 \mathcal{H}_\parallel)^{-m}$, defined as a bounded operator on $\mathrm{L}^2(\mathbb{R}^{n+1})$, is represented by an integral kernel $K_{\lambda,m}$ with pointwise bounds*

$$|K_{\lambda,m}(x,t,y,s)| \leq \frac{C 1_{(0,\infty)}(t-s)}{\lambda^{2m}} (t-s)^{-n/2+m-1} \mathrm{e}^{-\frac{t-s}{\lambda^2}} \mathrm{e}^{-c\frac{|x-y|^2}{t-s}}, \tag{4.2}$$

*where $C, c > 0$ depend only on $n$, the ellipticity constants and $m$. An analogous representation holds for $(1 + \lambda^2 \mathcal{H}_\parallel^*)^{-m}$ with adjoint kernel $K_{\lambda,m}^*$.*

*Proof.* It suffices to do it when $m = 1$ as iterated convolution in $(x,t)$ of the estimate on the right hand side of (4.2) with $m = 1$ yields the result.

Let $f \in C_0^\infty(\mathbb{R}^{n+1})$. Let $u = (1 + \lambda^2 \mathcal{H}_\|)^{-1} f$ given by the functional calculus of $\mathcal{H}_\|$. Then $u \in L^2(\mathbb{R}^{n+1})$ and, in particular, $u$ is a weak solution to $\lambda^2 \partial_t u - \lambda^2 \operatorname{div}_x A_{\|\|} \nabla_x u + u = f$. On the other hand, by Aronson's result [2], the operator $\mathcal{H}_\|$ has a fundamental solution, denoted by $K(x,t,y,s)$, having bounds

$$|K(x,t,y,s)| \leq C 1_{(0,\infty)}(t-s) \cdot (t-s)^{-n/2} \mathrm{e}^{-c\frac{|x-y|^2}{t-s}} \qquad \text{for } x, y \in \mathbb{R}^n,\ t, s \in \mathbb{R}$$

with constants $C, c$ depending only on dimension and the ellipticity constants, and satisfying

$$\int_{\mathbb{R}^n} K(x,t,y,s)\, \mathrm{d}y = 1 \qquad \text{for } x \in \mathbb{R}^n,\ t, s \in \mathbb{R},\ t > s. \tag{4.3}$$

Set $K_{\lambda,1}(x,t,y,s) = \lambda^{-2} K(x,t,y,s) \mathrm{e}^{-\frac{t-s}{\lambda^2}}$ and $v(x,t) = \iint_{\mathbb{R}^{n+1}} K_{\lambda,1}(x,t,y,s) f(y,s)\, \mathrm{d}y\, \mathrm{d}s$. Aronson's estimate implies $v \in L^2(\mathbb{R}^{n+1})$ and a calculation shows that $v$ is a weak solution to the same equation as $u$. Thus, $w := u - v$ is a weak solution of $\partial_t w - \operatorname{div}_x A_{\|\|} \nabla_x w + \lambda^{-2} w = 0$ and we may use the Caccioppoli estimate of Lemma 2.1 in $\mathbb{R}^{n+1}$. Choosing test functions $\psi$ that converge to 1 reveals $\nabla_x w = 0$ as $w \in L^2(\mathbb{R}^{n+1})$. Hence $w$ depends only on $t$. Again, as $w \in L^2(\mathbb{R}^{n+1})$, $w$ must be 0. This shows that $P_\lambda f$ has the desired representation for all $f \in C_0^\infty(\mathbb{R}^{n+1})$ and we conclude by density. $\square$

**Remark 4.4.** The kernel representation from Lemma 4.3 can be extended from $L^2(\mathbb{R}^{n+1})$ to $\dot{\mathsf{E}}(\mathbb{R}^{n+1})$ since the latter embeds continuously into $L^2(\mathbb{R}^{n+1}) + L^\infty(\mathbb{R}^{n+1})$ modulo constants, see for example Lemma 3.11 in [4]. In this sense $(1 + \lambda^2 \mathcal{H}_\|)^{-m} 1 = 1$ holds due to (4.3).

**Lemma 4.5.** *Fix $m = n+1$ in the definitions of $P_\lambda^*$ and $P_\lambda$. There exists $c$, $1 \leq c < \infty$, depending only on $n$ and the ellipticity constants such that*

$$||N_*(\partial_\lambda P_\lambda^* \varphi)||_2^2 + ||N_*(\partial_\lambda P_\lambda \tilde{\varphi})||_2^2 \leq c|\Delta|.$$

*Proof.* By symmetry of definitions, we only have to prove one of the estimate and we do the one of $N_*(\partial_\lambda P_\lambda^* \varphi)$ for a change.

To start the proof, fix $(\mu, y, s) \in W(\lambda, x, t)$, where $W(\lambda, x, t) := \Lambda_\lambda \times Q_\lambda(x) \times I_\lambda(t) =$ and $\Lambda_\lambda = (\lambda/2, \lambda)$, $Q_\lambda(x) = B(x, \lambda)$ and $I_\lambda(t) = (t - \lambda^2, t + \lambda^2)$ is one of the Whitney regions used in the definition of $N_*$ and recall that $\Delta_\lambda(x,t) = Q_\lambda(x) \times I_\lambda(t)$. Let $\sigma \in \Lambda_\lambda$ be arbitrary for the moment. We note that within the functional calculus for $\mathcal{H}_\|^*$,

$$\partial_\mu P_\mu^* = -2m\mu \mathcal{H}_\|^* (1 + \mu^2 \mathcal{H}_\|^*)^{-m-1},$$

and we introduce $\tilde{P}_\mu^* := (1 + \mu^2 \mathcal{H}_\|^*)^{-1}$ to write

$$\partial_\mu P_\mu^* \varphi = -\frac{2m\mu}{\sigma} (1 + \mu^2 \mathcal{H}_\|^*)^{-m} (1 + \mu^2 \mathcal{H}_\|^*)^{-1} (1 + \sigma^2 \mathcal{H}_\|^*) \sigma \mathcal{H}_\|^* \tilde{P}_\sigma^* \varphi.$$

It is convenient to expand this identity as

$$\partial_\mu P_\mu^* \varphi = -\frac{2m\mu}{\sigma} (1 + \mu^2 \mathcal{H}_\|^*)^{-m} \left( \frac{\sigma^2}{\mu^2} + \left(1 - \frac{\sigma^2}{\mu^2}\right) (1 + \mu^2 \mathcal{H}_\|^*)^{-1} \right) \sigma \mathcal{H}_\|^* \tilde{P}_\sigma^* \varphi \tag{4.4}$$

since this reveals $\partial_\mu P_\mu^* \varphi = T(\sigma \mathcal{H}_\|^* \tilde{P}_\sigma^* \varphi)$, where the operator $T$ is given by a linear combination of the resolvent kernels $K_{\mu,m}^*$ and $K_{\mu,m+1}^*$ provided by Lemma 4.3. Setting $G_0(x,t) := \Delta_{2\lambda}(x,t)$ and $G_j(x,t) := \Delta_{2^{j+1}\lambda}(x,t) \setminus \Delta_{2^j\lambda}(x,t)$, $j \geq 1$, since $(\mu, y, s) \in \Lambda_\lambda \times \Delta_\lambda(x,t)$, we can infer pointwise estimates

$$|K_{\mu,m+k}^*(y,s,z,\tau)| \leq \frac{C}{\lambda^{n+2}} \mathrm{e}^{-c4^j} \qquad \text{if } (z,\tau) \in G_j(x,t),\ j \geq 0,\ m + k \geq n/2 + 1,$$

where $C, c > 0$ depend only on $n$, the ellipticity constants and $m + k$. Note that the bound for $j = 0$ only holds since $m + k \geq m = n + 1 \geq n/2 + 1$ guarantees that $K_{\mu,m+k}^*$ is *bounded*. As we have

$\lambda/2 < \sigma < \lambda$, the kernel $K^*$ of the operator acting on $\sigma \mathcal{H}_\parallel^* \tilde{P}_\sigma^* \varphi$ on the right-hand side of (4.4) has analogous bounds and we can eventually record

$$\begin{aligned} |\partial_\mu P_\mu^* \varphi(y,s)| &= \left| \iint_{\mathbb{R}^{n+1}} K^*(y,s,z,\tau) \sigma \mathcal{H}_\parallel^* \tilde{P}_\sigma^* \varphi(z,\tau) \, \mathrm{d}z \, \mathrm{d}\tau \right| \\ &\leq \sum_{j=0}^\infty C 2^{j(n+2)} \mathrm{e}^{-c4^j} \fint\!\!\fint_{G_j(x,t)} |\sigma \mathcal{H}_\parallel^* \tilde{P}_\sigma^* \varphi(z,\tau)| \, \mathrm{d}z \, \mathrm{d}\tau \end{aligned}$$

with $C, c > 0$ depending only on $n$ and the ellipticity constants. As $(\mu, y, s) \in W(\lambda, x, t)$ was arbitrary in this argument, we have in fact

$$\sup_{(\mu,y,s) \in W(\lambda,x,t)} |\partial_\mu P_\mu^* \varphi(y,s)|^2 \lesssim \sum_{j=0}^\infty e^{-c4^j} \fint\!\!\fint_{2^{j+1}Q_\lambda(x) \times 4^{j+2} I_\lambda(t)} |\sigma \mathcal{H}_\parallel^* \tilde{P}_\sigma^* \varphi(z,\tau)|^2 \, \mathrm{d}z \, \mathrm{d}\tau, \tag{4.5}$$

where we have also used Cauchy-Schwarz to switch to $\mathrm{L}^2$-averages and exploited the exponential decay. Since only the right-hand side depends on $\sigma \in \Lambda_\lambda$, we can average in $\sigma$ and take the supremum in $\lambda$ to find

$$N_*(\partial_\lambda P_\lambda^* \varphi)(x,t)^2 \lesssim \sum_{j=0}^\infty e^{-c4^j} \sup_{\lambda > 0} \int_{\lambda/2}^\lambda \fint\!\!\fint_{2^{j+1}Q_\lambda(x) \times 4^{j+2} I_\lambda(t)} |\sigma \mathcal{H}_\parallel^* \tilde{P}_\sigma^* \varphi(z,\tau)|^2 \, \frac{\mathrm{d}z \, \mathrm{d}\tau \, \mathrm{d}\sigma}{\sigma}.$$

By a direct application of Tonelli's theorem, see Lemma 8.9 in [4] for an explicit proof, this implies

$$\iint_{\mathbb{R}^{n+1}} |N_*(\partial_\lambda P_\lambda^* \varphi)(x,t)|^2 \, \mathrm{d}x \, \mathrm{d}t \lesssim \sum_{j=0}^\infty \mathrm{e}^{-c4^j} \iiint_{\mathbb{R}^{n+2}_+} |\sigma \mathcal{H}_\parallel^* \tilde{P}_\sigma^* \varphi(z,\tau)|^2 \, \frac{\mathrm{d}z \, \mathrm{d}\tau \, \mathrm{d}\sigma}{\sigma}$$

and hence the claim follows from Lemma 3.2 (i) applied with $m = 1$. □

## 5. Parabolic sawtooth domains associated with $F$

Throughout this section, let $\Delta$ and $\kappa_0 \gg 1$ be given and let $F \subset 16\Delta$ be the set introduced in Definition 1.7 with $P_\lambda = (1 + \lambda^2 \mathcal{H}_\parallel)^{-n-1}$, $P_\lambda^* = (1 + \lambda^2 \mathcal{H}_\parallel^*)^{-n-1}$ from now on. Let us recall that the non-tangential maximal operators $N_*$ and $\widetilde{N}_*$ at $(x_0, t_0) \in \mathbb{R}^{n+1}$ are defined with reference to the Whitney regions $\Lambda_\lambda \times Q_\lambda(x_0) \times I_\lambda(x_0)$, where $\Lambda_\lambda = (\lambda/2, \lambda)$, $Q_\lambda(x_0) = B(x_0, \lambda)$, $I_\lambda(t_0) = (t_0 - \lambda^2, t_0 + \lambda^2)$, and that $\Gamma(x_0, t_0)$ denotes the parabolic cone with vertex $(x_0, t_0)$ and aperture one, see (1.4). In particular, we have

$$\Gamma(x_0, t_0) \subset \bigcup_{\lambda > 0} \Lambda_\lambda \times Q_\lambda(x_0) \times I_\lambda(t_0).$$

Next, we introduce a sawtooth domain associated with $F$,

$$\Omega := \bigcup_{(x_0,t_0) \in F \cap \Delta} \Gamma(x_0, t_0),$$

and establish pointwise estimates for the differences

$$\theta_\lambda := \varphi - P_\lambda^* \varphi, \; \tilde{\theta}_\lambda := \tilde{\varphi} - P_\lambda \tilde{\varphi}.$$

**Lemma 5.1.** *We have*

$$\begin{aligned} &\text{(i)} && |\theta_\lambda(x,t)| + |\tilde{\theta}_\lambda(x,t)| \leq \kappa_0 \lambda && \textit{if } (\lambda,x,t) \in (0,\infty) \times F, \\ &\text{(ii)} && |\partial_\lambda \theta_\lambda(x,t)| + |\partial_\lambda \tilde{\theta}_\lambda(x,t)| = |\partial_\lambda P_\lambda^* \varphi(x,t)| + |\partial_\lambda P_\lambda \tilde{\varphi}(x,t)| \leq \kappa_0 && \textit{if } (\lambda,x,t) \in \Omega, \\ &\text{(iii)} && |\widetilde{N}_*(\nabla_x \theta_\lambda)(x,t)| + |\widetilde{N}_*(\nabla_x \tilde{\theta}_\lambda)(x,t)| \leq 2\kappa_0 && \textit{if } (x,t) \in F. \end{aligned}$$

*Proof.* If $(x,t) \in F$, then by the fundamental theorem of calculus and the construction of the set $F$, see Definition 1.7 (iv),

$$|\theta_\lambda(x,t)| + |\tilde{\theta}_\lambda(x,t)| = \int_0^\lambda |\partial_\sigma P_\sigma^* \varphi(x,t)| + |\partial_\sigma P_\sigma \tilde{\varphi}(x,t)| \, \mathrm{d}\sigma \leq \kappa_0 \lambda.$$

This proves (i). Similarly, consider $(\lambda, x, t) \in \Omega$. Then $(\lambda, x, t) \in \Gamma(x_0, t_0)$ for some $(x_0, t_0) \in F$ and since $\varphi$ and $\tilde{\varphi}$ are functions of $(x, t)$ only, we obtain

$$|\partial_\lambda \theta_\lambda(x,t)| = |\partial_\lambda P_\lambda^* \varphi(x,t)| \leq N_*(\partial_\sigma P_\sigma^* \varphi)(x_0,t_0)$$

together with an analogous estimate for $\partial_\lambda \tilde{\theta}_\lambda(x,t)$. Hence, (ii) is again a consequence of Definition 1.7 (iv). As $\varphi$ and $\tilde{\varphi}$ do not depend on $\lambda$, we also have

$$\begin{aligned}|\widetilde{N}_*(\nabla_x \theta_\lambda)(x,t)| + |\widetilde{N}_*(\nabla_x \tilde{\theta}_\lambda)(x,t)| \quad &\leq \quad \mathcal{M}(|\nabla_x \varphi|^2)(x,t)^{1/2} + \mathcal{M}(|\nabla_x \tilde{\varphi}|^2)(x,t)^{1/2} \\ &\quad + |\widetilde{N}_*(\nabla_x P_\lambda^* \varphi)(x,t)| + |\widetilde{N}_*(\nabla_x P_\lambda \tilde{\varphi})(x,t)|,\end{aligned}$$

showing that (iii) is a consequence of parts (i) and (v) in Definition 1.7. □

Our next lemma extends the bound in part (ii) above to the whole sawtooth region.

**Lemma 5.2.**

$$|\theta_\lambda(x,t)| + |\tilde{\theta}_\lambda(x,t)| \lesssim \kappa_0 \lambda \qquad \textit{for } (\lambda, x, t) \in \Omega.$$

*Proof.* By symmetry of the definitions it suffices to prove the bound for $\theta_\lambda$. As a preliminary observation note that if $(\lambda, x, t) \in \Omega$, then $(\lambda, x, t) \in \Gamma(x_0, t_0)$ for some $(x_0, t_0) \in F$ and in particular $(x, t) \in \Delta_\lambda(x_0, t_0)$. Since $\varphi$ is a weak solution to the equation $\mathcal{H}_\|^* \varphi = \mathrm{div}_x(A_{\perp \|} \chi_{8\Delta})$ on $\mathbb{R}^{n+1}$, we can then use the classical local estimates for weak solutions with real coefficients, see e.g. Theorem 6.17 in [22], to the effect that

$$\sup_{(x,t) \in \Delta_\lambda(x_0,t_0)} |\varphi(x,t) - \varphi(x_0,t_0)| \lesssim \lambda + \fint\!\!\fint_{\Delta_{2\lambda}(x_0,t_0)} |\varphi(x,t) - \varphi(x_0,t_0)| \, \mathrm{d}x \, \mathrm{d}t.$$

Hence, using the construction of the set $F$, see Definition 1.7 (iii), we deduce

$$\sup_{(x,t) \in \Delta_\lambda(x_0,t_0)} |\varphi(x,t) - \varphi(x_0,t_0)| \lesssim \lambda + \lambda \mathbb{D}\varphi(x_0,t_0) \leq \lambda + \kappa_0 \lambda. \tag{5.1}$$

To start with the actual proof of the estimate stated in the lemma, we let $(\lambda, x, t)$ and $(x_0, t_0)$ be fixed as above and we denote by $\varphi_{2\lambda}$ the average of $\varphi$ over the set $\Delta_{2\lambda}(x_0, t_0)$. Thinking of $P_\lambda^*$ as given by kernel representation from Lemma 4.3, see also Remark 4.4, we have $P_\lambda^* 1 = 1$ and consequently,

$$\begin{aligned}|\theta_\lambda(x,t)| = |(I - P_\lambda^*)\varphi(x,t)| \quad &\leq \quad |\varphi(x,t) - \varphi(x_0,t_0)| + |(I - P_\lambda^*)\varphi(x_0,t_0)| \\ &\quad + |P_\lambda^*(\varphi - \varphi_{2\lambda})(x_0,t_0)| + |P_\lambda^*(\varphi - \varphi_{2\lambda})(x,t)|.\end{aligned}$$

Hence, using (5.1) and Lemma 5.1 (i), we deduce

$$|\theta_\lambda(x,t)| \lesssim \kappa_0 \lambda + |P_\lambda^*(\varphi - \varphi_{2\lambda})(x_0,t_0)| + |P_\lambda^*(\varphi - \varphi_{2\lambda})(x,t)|. \tag{5.2}$$

To estimate the remaining two terms on the right, we bring in the kernel of $P_\lambda^*$ explicitly. Indeed, Lemma 4.3 yields

$$P_\lambda^*(\varphi - \varphi_{2\lambda})(x_0,t_0) = \iint_{\mathbb{R}^{n+1}} K_{\lambda,n+1}^*(x_0,t_0,z,\tau)(\varphi(z,\tau) - \varphi_{2\lambda}) \, \mathrm{d}z \, \mathrm{d}\tau$$

with a kernel enjoying the bound

$$\begin{aligned}|K_{\lambda,n+1}^*(x_0,t_0,z,\tau)| &\leq 1_{(0,\infty)}(\tau - t_0) \frac{C|t_0 - \tau|^{-n/2+n+1-1}}{\lambda^{2+2n}} \mathrm{e}^{-\frac{|t_0-\tau|}{\lambda^2}} \mathrm{e}^{-c\frac{|x_0-z|^2}{|t_0-\tau|}} \\ &\lesssim 1_{(0,\infty)}(\tau - t_0) \frac{C}{\lambda^{n+2}} \mathrm{e}^{-\frac{|t_0-\tau|}{\lambda^2}} \mathrm{e}^{-c\frac{|x_0-z|^2}{|t_0-\tau|}}\end{aligned}$$

for some constants $C, c > 0$ depending only on dimension and ellipticity. So, splitting $\mathbb{R}^{n+2}$ into $\Delta_{2\lambda}(x_0, t_0)$ and annuli $\Delta_{2^{j+1}\lambda}(x_0, t_0) \setminus \Delta_{2^j\lambda}(x_0, t_0)$, $j \geq 1$, we can infer that

$$|P_\lambda^*(\varphi - \varphi_{2\lambda})(x_0,t_0)| \lesssim \sum_{j=0}^{\infty} 2^{(n+2)j} \mathrm{e}^{-c4^j} \fint\!\!\fint_{\Delta_{2^{j+1}\lambda}(x_0,t_0)} |\varphi(z,\tau) - \varphi_{2\lambda}| \, \mathrm{d}z \, \mathrm{d}\tau. \tag{5.3}$$

Next, by a telescopic sum and Lemma 2.2 we deduce that

$$\begin{aligned}\fint\!\!\fint_{\Delta_{2^{j+1}\lambda}(x_0,t_0)} |\varphi(z,\tau) - \varphi_{2\lambda}| \,\mathrm{d}z\,\mathrm{d}\tau \ &\leq\ \sum_{k=1}^{j+1} \fint\!\!\fint_{\Delta_{2^k\lambda}(x_0,t_0)} |\varphi(z,\tau) - \varphi_{2^k\lambda}| \,\mathrm{d}z\,\mathrm{d}\tau \\ &\leq\ 2^{j+2}\lambda\Big(\mathcal{M}(|\nabla_x\varphi|)(x_0,t_0) + \mathcal{M}_x\,\mathcal{M}_t(|H_tD_t^{1/2}\varphi|)(x_0,t_0)\Big)\end{aligned}$$

and parts (i) and (ii) of Definition 1.7 guarantee that the last term is no larger than $2^{j+3}\lambda\kappa_0$. In particular, summing up in (5.3), we can conclude $|P_\lambda^*(\varphi - \varphi_{2\lambda})(x_0,t_0)| \lesssim \kappa_0\lambda$. The estimate of $|P_\lambda^*(\varphi - \varphi_{2\lambda})(x,t)|$ can be done similarly, taking into account $\Delta_\lambda(x_0,t_0) \subset \Delta_{2\lambda}(x,t)$ when writing out the telescopic sum of averages. Now, the claim follows from (5.2). $\square$

5.1. **An adapted cut-off and associated Carleson measures.** Here, we bring into play the degree of freedom $0 < \eta \ll 1$ and the parameter $0 < \epsilon \ll r$ that already appeared in the outline of Section 1.5.

Writing $\Gamma^\eta(x_0,t_0)$ for the parabolic cone with vertex $(x_0,t_0)$ and aperture $\eta$, we define the thinner sawtooth domains

$$\Omega_\eta := \bigcup_{(x_0,t_0)\in F\cap\Delta} \Gamma^{\eta/8}(x_0,t_0).$$

Then $\Omega_\eta \subset \Omega$. We are now going to define a smooth cut off adapted to $\Omega_\eta$.

Let $\Phi \in \mathrm{C}_0^\infty(\mathbb{R})$ be such that $\Phi(\varrho) = 1$ if $\varrho \leq 1/16$ and $\Phi(\varrho) = 0$ if $\varrho > 1/8$ and let $\Upsilon \in \mathrm{C}_0^\infty(\mathbb{R}^{n+2})$ be supported in $B(0,1/2048) \subset \mathbb{R}^{n+2}$ and satisfy $0 \leq \Upsilon \leq 1$ and $\int_{\mathbb{R}^{n+2}} \Upsilon(\mu,y,s)\,\mathrm{d}\mu\,\mathrm{d}y\,\mathrm{d}s = 1$. We then set

$$\begin{aligned}\Psi(\lambda,x,t) &:= \Psi_{\eta,\epsilon}(\lambda,x,t) \\ &:= \Phi\Big(\frac{\lambda}{32r}\Big)\Big(1 - \Phi\Big(\frac{\lambda}{32\epsilon}\Big)\Big) \int_{\Omega_{\eta/2}} \Upsilon\Big(\frac{\lambda-\mu}{\lambda}, \frac{x-y}{\lambda\eta}, \frac{s-t}{(\lambda\eta)^2}\Big) \frac{\mathrm{d}\mu\,\mathrm{d}y\,\mathrm{d}s}{\lambda(\lambda\eta)^{n+2}}.\end{aligned} \tag{5.4}$$

By construction, we have $\Psi \in \mathrm{C}_0^\infty(\mathbb{R}_+^{n+2})$, $0 \leq \Psi \leq 1$, $\Psi = 1$ on the open set $\Omega_{\eta/4} \cap ((2\epsilon,2r)\times\mathbb{R}^{n+1})$ and in particular on $(F\cap\Delta)\times(2\epsilon,2r)$ and the support of $\Psi$ is a subset of $\Omega_\eta \cap ((\epsilon,4r)\times 2\Delta)$. Making the link with the sawtooth domain $\Omega$ from the previous section, we note

$$(\lambda,x,t) \in \operatorname{supp}\Psi \quad\Longrightarrow\quad (\eta\lambda,x,t) \in \Omega.$$

Now, let $\delta(x,t)$ denote the parabolic distance from the point $(x,t) \in \mathbb{R}^{n+1}$ to $F \subset \mathbb{R}^{n+1}$ and let

$$\begin{aligned}E_1 &:= \{(\lambda,x,t) \in (0,4r)\times 2\Delta :\ \eta\lambda/32 \leq \delta(x,t) \leq \eta\lambda/8\}, \\ E_2 &:= \{(\lambda,x,t) \in (2r,4r)\times 2\Delta : \delta(x,t) \leq \eta\lambda/8\}, \\ E_3 &:= \{(\lambda,x,t) \in (\epsilon,2\epsilon)\times 2\Delta : \delta(x,t) \leq \eta\lambda/8\}.\end{aligned} \tag{5.5}$$

Let $(\alpha,\beta,\gamma) = (\alpha,\beta_1,\ldots,\beta_n,\gamma) \in \mathbb{N}^{n+2}\setminus\{0\}$ be a multiindex. Again by construction of $\Psi$ there exists a constant $\tilde{c}$ depending only on $\alpha$, $\beta$, $\gamma$, $\eta$ and $n$ such that

$$\Big|\frac{\partial^{\alpha+|\beta|+\gamma}}{\partial\lambda^\alpha\partial x^\beta\partial t^\gamma}\Psi(\lambda,x,t)\Big| \ \leq\ \frac{\tilde{c}}{|\lambda|^{\alpha+|\beta|+\gamma}} 1_{E_1\cup E_2\cup E_3}(\lambda,x,t). \tag{5.6}$$

The following Carleson lemma is also important in the next section.

**Lemma 5.3.** *It holds*

$$\iiint_{E_1\cup E_2\cup E_3} \frac{\mathrm{d}\lambda\,\mathrm{d}x\ dt}{\lambda} \leq \log(8)2^{n+2}|\Delta|.$$

*In particular, let $\eta$, $\epsilon$ and $\Psi = \Psi_{\eta,\epsilon}$ be as above, let $(\alpha,\beta,\gamma) = (\alpha,\beta_1,\ldots,\beta_n,\gamma) \in \mathbb{N}^{n+2}\setminus\{0\}$ be a multiindex and let $p \in (0,\infty)$. Then there exists $\tilde{c} = \tilde{c}(\alpha,\beta,\gamma,p,n,\eta) < \infty$, such that*

$$\iiint_{\mathbb{R}_+^{n+2}} \Big|\frac{\partial^{\alpha+|\beta|+\gamma}}{\partial\lambda^\alpha\partial x^\beta\partial t^\gamma}\Psi(\lambda,x,t)\Big|^p \lambda^{p(\alpha+|\beta|+2\gamma)-1}\,\mathrm{d}x\,\mathrm{d}t \leq \tilde{c}|\Delta|.$$

*Proof.* By definition of the sets $E_1$, $E_2$ and $E_3$, integration in $(x,t)$ takes place only on the cube $2\Delta$ and for $(x,t) \in 2\Delta$ fixed, integration in $\lambda$ is at most over the intervals $(\frac{8\delta(x,t)}{\eta}, \frac{32\delta(x,t)}{\eta})$, $(2r, 4r)$ and $(\epsilon, 2\epsilon)$, yielding a total contribution of $\log(8)$. Hence, the first claim follows from Tonelli's theorem and then the second one is a consequence of (5.6) $\square$

## 6. Proof of the Key Lemma

We are now ready to prove the Key Lemma, hence completing the proof of Theorem 1.4. As discussed in Section 1.5, throughout the proof we can *qualitatively* assume that $A$ is smooth. In that case, one can see that qualitatively $\varphi, \tilde{\varphi}, P_\lambda \tilde{\varphi}$ and $P_\lambda^* \varphi$ as well as $u$ are smooth by interior parabolic regularity. Furthermore, we will simply write $J$ for $J_{\eta,\epsilon}$ and we note – and this is a consequence of the introduction of $\epsilon$ – that no boundary terms will survive when we perform partial integration. Similarly, we will write $\Psi$ for $\Psi_{\eta,\epsilon}$ defined in Section 5.1. Throughout, $\sigma$ will denote a positive degree of freedom and $c$ will denote a generic constant (not necessarily the same at each instance), which depends only on the dimension $n$ and the ellipticity constants. In contrast, $\tilde{c}$ will denote a generic constant that may additionally depend on $\sigma$ and $\eta$. The fact that $|u| \leq 1$ will be used repeatedly in the proof.

To start the estimate of $J$ we first note that $u\Psi^2\lambda$ is a test function for the weak formulation of the equation for $u$. Hence,

$$0 = \iiint_{\mathbb{R}^{n+2}_+} A\nabla_{\lambda,x} u \cdot \nabla_{\lambda,x}(u\Psi^2\lambda) + \partial_t u(u\Psi^2\lambda) \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\lambda. \tag{6.1}$$

As

$$\nabla_{\lambda,x}(u\Psi^2\lambda) = (\nabla_{\lambda,x} u)\Psi^2\lambda + u\lambda\nabla_{\lambda,x}\Psi^2 + u\Psi^2\nabla_{\lambda,x}\lambda,$$

we have

$$\begin{aligned} J \;&=\; \iiint_{\mathbb{R}^{n+2}_+} A\nabla_{\lambda,x} u \cdot \nabla_{\lambda,x}(u\Psi^2\lambda) \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\lambda - \iiint_{\mathbb{R}^{n+2}_+} (A\nabla_{\lambda,x} u \cdot \nabla_{\lambda,x}\Psi^2) u \, \lambda \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\lambda \\ &\quad - \iiint_{\mathbb{R}^{n+2}_+} (A\nabla_{\lambda,x} u \cdot \nabla_{\lambda,x}\lambda) u\Psi^2 \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\lambda. \end{aligned}$$

Combining this with (6.1), we see that $J = J_1 + J_2 + J_3$, where

$$\begin{aligned} J_1 \;&:=\; -\iiint_{\mathbb{R}^{n+2}_+} (A\nabla_{\lambda,x} u \cdot \nabla_{\lambda,x}\Psi^2) u \, \lambda \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\lambda, \\ J_2 \;&:=\; -\iiint_{\mathbb{R}^{n+2}_+} (A\nabla_{\lambda,x} u \cdot \nabla_{\lambda,x}\lambda) u\Psi^2 \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\lambda, \\ J_3 \;&:=\; -\iiint_{\mathbb{R}^{n+2}_+} \partial_t u(u\Psi^2\lambda) \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\lambda. \end{aligned}$$

The estimates of $J_1$ and $J_3$ turn out to be straightforward: Indeed, by the Cauchy-Schwarz inequality

$$|J_1| \leq c\Big(\iiint_{\mathbb{R}^{n+2}_+} |\nabla_{\lambda,x} u|^2 \Psi^2 \, \lambda \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\lambda\Big)^{1/2} \Big(\iiint_{\mathbb{R}^{n+2}_+} |\nabla_{\lambda,x}\Psi|^2 \, \lambda \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\lambda\Big)^{1/2}$$

and hence, using the elementary Young's inequality, ellipticity of $A$ and the Carleson measure estimates in Lemma 5.3,

$$|J_1| \leq \sigma J + \tilde{c}|\Delta|.$$

Furthermore,

$$J_3 = -\frac{1}{2} \iiint_{\mathbb{R}^{n+2}_+} (\partial_t u^2)\Psi^2\lambda \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\lambda = \frac{1}{2} \iiint_{\mathbb{R}^{n+2}_+} u^2 \partial_t(\Psi^2)\lambda \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\lambda.$$

Thus, by the Carleson measure estimates in Lemma 5.3 and as $|u|, |\Psi| \leq 1$,

$$|J_3| \leq c \iiint_{\mathbb{R}^{n+2}_+} |\partial_t \Psi|\lambda \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\lambda \leq \tilde{c}|\Delta|.$$

As for $J_2$, we first use the decomposition (1.8) of the coefficients and split $J_2 = J_{21} + J_{22}$, where

$$\begin{aligned} J_{21} &:= -\iiint_{\mathbb{R}^{n+2}_+} (A_{\perp\|} \cdot \nabla_x u) u \Psi^2 \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\lambda, \\ J_{22} &:= -\iiint_{\mathbb{R}^{n+2}_+} (A_{\perp\perp} \partial_\lambda u) u \Psi^2 \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\lambda. \end{aligned}$$

Since $A$ does not depend on $\lambda$, integration by parts yields

$$J_{22} = -\frac{1}{2} \iiint_{\mathbb{R}^{n+2}_+} A_{\perp\perp} \partial_\lambda u^2 \Psi^2 \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\lambda = \frac{1}{2} \iiint_{\mathbb{R}^{n+2}_+} A_{\perp\perp} u^2 \partial_\lambda \Psi^2 \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\lambda$$

and hence $|J_{22}| \leq \tilde{c}|\Delta|$ follows again by Lemma 5.3. To estimate $J_{21}$ we use that

$$A_{\perp\|} \cdot \nabla_x \Big( \frac{u^2 \Psi^2}{2} \Big) = (A_{\perp\|} \cdot \nabla_x u) u \Psi^2 + (A_{\perp\|} \cdot \nabla_x \Psi) u^2 \Psi$$

and we write $J_{21} = J_{211} + J_{212}$, where

$$\begin{aligned} J_{211} &:= -\iiint_{\mathbb{R}^{n+2}_+} A_{\perp\|} \cdot \nabla_x \Big( \frac{u^2 \Psi^2}{2} \Big) \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\lambda, \\ J_{212} &:= \iiint_{\mathbb{R}^{n+2}_+} (A_{\perp\|} \cdot \nabla_x \Psi) u^2 \Psi \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\lambda. \end{aligned}$$

Once again, $|J_{212}| \leq \tilde{c}|\Delta|$ follows by Lemma 5.3. In order to handle $J_{211}$, we introduce $\varphi$ as in (1.10), that is, as the energy solution on $\mathbb{R}^{n+1}$ to the problem

$$\operatorname{div}_x(A_{\perp\|} \chi_{8\Delta}) = -\partial_t \varphi - \operatorname{div}_x(A_{\|\|}^* \nabla_x \varphi) = \mathcal{H}_\|^* \varphi.$$

The weak formulation with $\phi = u^2 \Psi^2(\lambda, \cdot, \cdot)/2$ as test function for $\lambda > 0$ fixed, which by construction of $\Psi$ is supported in $8\Delta$, yields

$$J_{211} = \int_{\mathbb{R}_+} \iint_{\mathbb{R}^{n+1}} \varphi \partial_t \Big( \frac{u^2 \Psi^2}{2} \Big) \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\lambda + \int_{\mathbb{R}_+} \iint_{\mathbb{R}^{n+1}} A_{\|\|}^* \nabla_x \varphi \cdot \nabla_x \Big( \frac{u^2 \Psi^2}{2} \Big) \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\lambda.$$

Recall that we write $\theta_{\eta\lambda} = \varphi - P_{\eta\lambda}^* \varphi$. Then, splitting $\varphi = \theta_{\eta\lambda} + P_{\eta\lambda}^* \varphi$ in both integrals, we may write

$$J_{211} = J_{2111} + J_{2112} + J_{2113} + J_{2114}, \tag{6.2}$$

where

$$\begin{aligned} J_{2111} &:= \iiint_{\mathbb{R}^{n+2}_+} (\theta_{\eta\lambda}) \partial_t \Big( \frac{u^2 \Psi^2}{2} \Big) \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\lambda, \\ J_{2112} &:= \iiint_{\mathbb{R}^{n+2}_+} (P_{\eta\lambda}^* \varphi) \partial_t \Big( \frac{u^2 \Psi^2}{2} \Big) \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\lambda, \\ J_{2113} &:= \iiint_{\mathbb{R}^{n+2}_+} A_{\|\|}^* \nabla_x \theta_{\eta\lambda} \cdot \nabla_x \Big( \frac{u^2 \Psi^2}{2} \Big) \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\lambda, \\ J_{2114} &:= \iiint_{\mathbb{R}^{n+2}_+} A_{\|\|}^* \nabla_x P_{\eta\lambda}^* \varphi \cdot \nabla_x \Big( \frac{u^2 \Psi^2}{2} \Big) \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\lambda. \end{aligned}$$

For the time being, let us concentrate on the second and fourth term in (6.2). Integrating by parts with respect to $\lambda$ leads us to

$$\begin{aligned} J_{2112}+J_{2114} &= -\iiint_{\mathbb{R}^{n+2}_+}(\partial_\lambda P^*_{\eta\lambda}\varphi)\partial_t\Big(\frac{u^2\Psi^2}{2}\Big)\,\lambda\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda \\ &\quad -\iiint_{\mathbb{R}^{n+2}_+}(P^*_{\eta\lambda}\varphi)\partial_t\partial_\lambda\Big(\frac{u^2\Psi^2}{2}\Big)\,\lambda\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda \\ &\quad -\iiint_{\mathbb{R}^{n+2}_+}(A^*_{\|\|}\nabla_x\partial_\lambda P^*_{\eta\lambda}\varphi)\cdot\nabla_x\Big(\frac{u^2\Psi^2}{2}\Big)\,\lambda\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda \\ &\quad -\iiint_{\mathbb{R}^{n+2}_+}(A^*_{\|\|}\nabla_x P^*_{\eta\lambda}\varphi)\cdot\nabla_x\partial_\lambda\Big(\frac{u^2\Psi^2}{2}\Big)\,\lambda\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda, \end{aligned}$$

where we have again used the $\lambda$-independence of the coefficients. We stress that throughout (and with a slight abuse of notation) $\partial_\lambda P^*_{\eta\lambda}\varphi$ denotes the derivative in $\lambda$ of the function $\lambda\mapsto P^*_{\eta\lambda}\varphi$, so that there is a factor $\eta$ showing up in front by the chain rule. A similar notational convention will apply to $\partial_\lambda\theta_{\eta\lambda}$. Taking into account the definition of the parabolic operator $\mathcal{H}^*_\|$, we can regroup these terms as

$$J_{2112}+J_{2114}=I_1+I_2+I_3,$$

where

$$\begin{aligned} I_1 &:= -\iiint_{\mathbb{R}^{n+2}_+}(\partial_\lambda P^*_{\eta\lambda}\varphi)\partial_t\Big(\frac{u^2\Psi^2}{2}\Big)\,\lambda\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda, \\ I_2 &:= -\iiint_{\mathbb{R}^{n+2}_+}(A^*_{\|\|}\nabla_x\partial_\lambda P^*_{\eta\lambda}\varphi)\cdot\nabla_x\Big(\frac{u^2\Psi^2}{2}\Big)\,\lambda\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda, \\ I_3 &:= \iiint_{\mathbb{R}^{n+2}_+}(\mathcal{H}^*_\| P^*_{\eta\lambda}\varphi)\partial_\lambda\Big(\frac{u^2\Psi^2}{2}\Big)\,\lambda\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda. \end{aligned}$$

By the Cauchy-Schwarz inequality and the square function estimates stated in parts (ii) and (iii) of Lemma 3.2 we first deduce that

$$\begin{aligned} |I_2|+|I_3| &\le \tilde{c}|\Delta|^{1/2}\Big(\iiint_{\mathbb{R}^{n+2}_+}\Big|\nabla_{\lambda,x}\Big(\frac{u^2\Psi^2}{2}\Big)\Big|^2\,\lambda\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda\Big)^{1/2} \\ &\le \tilde{c}|\Delta|^{1/2}\Big(\iiint_{\mathbb{R}^{n+2}_+}|\nabla_{\lambda,x}u|^2\Psi+|\nabla_{\lambda,x}\Psi|^2\,\lambda\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda\Big)^{1/2} \end{aligned}$$

and then, by Lemma 5.3 and Young's inequality, we can conclude $|I_2|+|I_3|\le\sigma J+\tilde{c}|\Delta|$. To estimate $I_1$, we write $I_1=I_{11}+I_{12}$, where

$$\begin{aligned} I_{11} &:= -\iiint_{\mathbb{R}^{n+2}_+}(\partial_\lambda P^*_{\eta\lambda}\varphi)u^2\partial_t\Big(\frac{\Psi^2}{2}\Big)\,\lambda\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda, \\ I_{12} &:= -\iiint_{\mathbb{R}^{n+2}_+}(\partial_\lambda P^*_{\eta\lambda}\varphi)u\partial_t u\Psi^2\,\lambda\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda. \end{aligned}$$

By a familiar argument relying on Cauchy-Schwarz, Lemma 3.2 and Lemma 5.3 we deduce $|I_{11}|\le\tilde{c}|\Delta|$. The estimate of $I_{12}$ is more involved. Here, we first use the equation $\partial_t u=\mathrm{div}_{\lambda,x}A\nabla_{\lambda,x}u$, which thanks to our smoothness assumption may be interpreted in the classical (pointwise) sense, in order to split $I_{12}=I_{121}+I_{122}$, where

$$\begin{aligned} I_{121} &:= -\iiint_{\mathbb{R}^{n+2}_+}(\partial_\lambda P^*_{\eta\lambda}\varphi)u\,\mathrm{div}_x(A\nabla_{\lambda,x}u)_\|\Psi^2\,\lambda\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda, \\ I_{122} &:= -\iiint_{\mathbb{R}^{n+2}_+}(\partial_\lambda P^*_{\eta\lambda}\varphi)u(A\nabla_{\lambda,x}\partial_\lambda u)_\perp\Psi^2\,\lambda\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda. \end{aligned}$$

Then,

$$\begin{aligned} I_{121} &= \iiint_{\mathbb{R}^{n+2}_+} \nabla_x(\partial_\lambda P^*_{\eta\lambda}\varphi)u \cdot (A\nabla_{\lambda,x}u)_{\|}\Psi^2\,\lambda\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda \\ &\quad + \iiint_{\mathbb{R}^{n+2}_+} (\partial_\lambda P^*_{\eta\lambda}\varphi)\nabla_x u \cdot (A\nabla_{\lambda,x}u)_{\|}\Psi^2\,\lambda\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda \\ &\quad + \iiint_{\mathbb{R}^{n+2}_+} (\partial_\lambda P^*_{\eta\lambda}\varphi)u(A\nabla_{\lambda,x}u)_{\|} \cdot \nabla_x\Psi^2\,\lambda\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda. \end{aligned}$$

For the first term on the right-hand side we can infer control by $\sigma J + \tilde{c}|\Delta|$ using Cauchy-Schwarz, Lemma 3.2 and Young's inequality in a by now familiar manner. For the other two terms we shall use for the first time the definition of the set $F$. More precisely, in virtue of Lemma 5.1 we can replace the resolvent by its pointwise upper bound $|\partial_\lambda P^*_{\eta\lambda}\varphi(x,t)| \le c\eta \le c$ noting that $(\lambda,x,t) \in \operatorname{supp}(\Psi)$ implies $(\eta\lambda,x,t) \in \Omega$ by construction, see Section 5.1. Having done this, the second integral on the right-hand side is bounded by $\eta J$ thanks to ellipticity of $A$ and for the third one we obtain a bound $\tilde{c}J^{1/2}|\Delta|^{1/2}$ by applying Cauchy-Schwarz and Lemma 5.3. Put together, we have

$$|I_{121}| \le (\sigma + c\eta)J + \tilde{c}|\Delta|.$$

Also, using Lemma 3.2 we immediately have

$$|I_{122}| \le \sigma I_{1221} + \tilde{c}|\Delta|, \tag{6.3}$$

where

$$I_{1221} := \iiint_{\mathbb{R}^{n+2}_+} |\nabla_{\lambda,x}\partial_\lambda u|^2\Psi^2\,\lambda^3\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda.$$

This term can be estimated using a Whitney type covering argument, the fact that $\partial_\lambda u$ is a solution and Caccioppoli's inequality: Indeed, let $\mathcal{W} = \{W_i\}$ denote a partitioning of $\mathbb{R}^{n+2}_+$ into (parabolic) Whitney cubes, that is, each $W_i$ has dyadic (parabolic) sidelength $\ell(W_i)$ and is located at distance $4\ell(W_i)$ to the boundary. Let $\phi_i \in \mathrm{C}_0^\infty(2W_i)$ be a standard cut-off for $W_i$ such that $0 \le \phi_i \le 1$, $|\nabla_{\lambda,x}\phi_i| + |\partial_t\phi_i|^{1/2} \le c/\ell(W_i)$ and $\sum_i \phi_i^2(\lambda,x,t) = 1$ for all $(\lambda,x,t) \in \mathbb{R}^{n+2}_+$. Then

$$\begin{aligned} I_{1221} &= \sum_i \iiint_{\mathbb{R}^{n+2}_+} |\nabla_{\lambda,x}\partial_\lambda u|^2\phi_i^2\Psi^2\,\lambda^3\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda \\ &\le c\sum_i \iiint_{\mathbb{R}^{n+2}_+} |\partial_\lambda u|^2|\nabla_{\lambda,x}(\phi_i\Psi)|^2\,\lambda^3\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda, \end{aligned}$$

by an application of Lemma 2.1 and hence, taking into account the finite overlap of the Whitney cubes,

$$I_{1221} \le c\iiint_{\mathbb{R}^{n+2}_+} |\partial_\lambda u|^2\Psi^2\,\lambda\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda + c\iiint_{\mathbb{R}^{n+2}_+} |\partial_\lambda u|^2|\nabla_{\lambda,x}\Psi|^2\,\lambda^3\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda.$$

Crudely employing ellipticity, the first integral on the right-hand side is under control by $cJ$. Since $\lambda|\partial_\lambda u| \le c$ in a pointwise fashion, as follows easily from DeGiorgi-Moser-Nash interior estimates, Caccioppoli's inequality and $|u| \le 1$, we can apply Lemma 5.3 to bound the second one by $\tilde{c}|\Delta|$. So, as to (6.3), we have

$$|I_{122}| \le \sigma J + \tilde{c}|\Delta|.$$

Put together we can conclude that the second and fourth term all the way back in (6.2) can be estimated by

$$|J_{2112} + J_{2114}| \le \sigma J + \tilde{c}|\Delta|.$$

At this stage of the proof it only remains to focus on $J_{2111} + J_{2113}$ and we note that by definition

$$\begin{aligned} J_{2111} + J_{2113} &= -\iiint_{\mathbb{R}^{n+2}_+} \theta_{\eta\lambda}\partial_t\Big(\frac{u^2\Psi^2}{2}\Big)\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda + \iiint_{\mathbb{R}^{n+2}_+} A^*_{\|\|}\nabla_x\theta_{\eta\lambda}\cdot\nabla_x\Big(\frac{u^2\Psi^2}{2}\Big)\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda \\ &= II_1 + II_2 + II_3, \end{aligned}$$

where

$$
\begin{aligned}
II_1 &:= \iiint_{\mathbb{R}^{n+2}_+} \partial_t u(\theta_{\eta\lambda} u\Psi^2)\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda + \iiint_{\mathbb{R}^{n+2}_+} \nabla_x\theta_{\eta\lambda}\cdot A_{\|\|}\nabla_x u(u\Psi^2)\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda,\\
II_2 &:= \frac{1}{2}\iiint_{\mathbb{R}^{n+2}_+} \theta_{\eta\lambda}u^2\partial_t\Psi^2\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda,\\
II_3 &:= \frac{1}{2}\iiint_{\mathbb{R}^{n+2}_+} \nabla_x\theta_{\eta\lambda}\cdot A_{\|\|}\nabla_x\Psi^2(u^2)\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda.
\end{aligned}
$$

Using Lemma 5.2, we have $|\theta_{\eta\lambda}(x,t)| \leq c\eta\lambda \leq c\lambda$ for $(\lambda,x,t)$ in the support of $\Psi$ and hence we can conclude, using Lemma 5.3, that $|II_2| \leq \tilde{c}|\Delta|$ holds. Similarly, for $II_3$ we would like to bring into play the (integrated) non-tangential control for $\nabla_x\theta_\lambda$ provided by Lemma 5.2 (iii). To this end, we use an "averaging trick" justified by Tonelli's theorem in order to write

$$
II_3 \leq \tilde{c}\iiint_{\mathbb{R}^{n+2}_+}\left(\fint\!\!\fint\!\!\fint_{W_{\eta/64}(\sigma,y,s)} |\nabla_x\theta_{\eta\lambda}\cdot A_{\|\|}\nabla_x\Psi^2(u^2)|\,\mathrm{d}\lambda\,\mathrm{d}x\,\mathrm{d}t\right)\mathrm{d}\sigma\,\mathrm{d}y\,\mathrm{d}s,
$$

where $W_{\eta/64}(\sigma,y,s)$ denotes the Whitney region $(\sigma/2,\sigma)\times Q_{\eta\sigma/64}(y)\times I_{\eta\sigma/64}(s)$. Now, letting

$$
\begin{aligned}
\tilde{E}_1 &:= \{(\sigma,y,s)\in(0,4r)\times 2\Delta:\ \eta\sigma/64 \leq \delta(y,s)\leq \eta\sigma/4\},\\
\tilde{E}_2 &:= \{(\sigma,y,s)\in(2r,8r)\times 4\Delta : \delta(y,s)\leq \eta\sigma/4\},\\
\tilde{E}_3 &:= \{(\sigma,y,s)\in(\epsilon,4\epsilon)\times 4\Delta : \delta(y,s)\leq \eta\sigma/4\},
\end{aligned}
$$

where again $\delta(x,t)$ denotes the parabolic distance from $(x,t)$ to the set $F$, we deduce from (5.5) and (5.6) that the integrand on the right-hand side vanishes outside of $\tilde{E}_1\cup\tilde{E}_2\cup\tilde{E}_3$ and that we have a bound

$$
\begin{aligned}
II_3 &\leq \tilde{c}\iiint_{\tilde{E}_1\cup\tilde{E}_2\cup\tilde{E}_3}\left(\fint\!\!\fint\!\!\fint_{W_{\eta/64}(\sigma,y,s)} |\nabla_x\theta_{\eta\lambda}|\,\mathrm{d}\lambda\,\mathrm{d}x\,\mathrm{d}t\right)\frac{\mathrm{d}\sigma\,\mathrm{d}y\,\mathrm{d}s}{\sigma}\\
&\leq \tilde{c}\iiint_{\tilde{E}_1\cup\tilde{E}_2\cup\tilde{E}_3}\left(\fint\!\!\fint\!\!\fint_{W_{\eta/64}(\eta\sigma,y,s)} |\nabla_x\theta_{\lambda}|^2\,\mathrm{d}\lambda\,\mathrm{d}x\,\mathrm{d}t\right)^{1/2}\frac{\mathrm{d}\sigma\,\mathrm{d}y\,\mathrm{d}s}{\sigma},
\end{aligned}
$$

the second step following from Cauchy-Schwarz and a simple change of variables. The definition of $\tilde{E}_1\cup\tilde{E}_2\cup\tilde{E}_3$ entails that for $(\sigma,y,s)$ in this union, the Whitney region $W_{\eta/64}(\eta\sigma,y,s)$ is contained in a cone $\Gamma^{1/2}(x_0,t_0)$ with vertex $(x_0,t_0)\in F$. In particular, $W_{\eta/64}(\eta\sigma,y,s)$ can be covered by a finite number (depending only on $n$) of Whitney regions $\Lambda\times Q\times I$ showing up in the definition of the integrated maximal function $\widetilde{N}_*$ on the set $F$. Consequently, Lemma 5.2 yields

$$
II_3 \leq \tilde{c}\iiint_{\mathbb{R}^{n+2}_+} 1_{\tilde{E}_1\cup\tilde{E}_2\cup\tilde{E}_3}(\sigma,y,s)\,\frac{\mathrm{d}\sigma\,\mathrm{d}y\,\mathrm{d}s}{\sigma}.
$$

Up to a change of parameters, the sets $\tilde{E}_1,\tilde{E}_2,\tilde{E}_3$ are similar to $E_1,E_2,E_3$ defined Subsection 5.1 and so we can rely on Lemma 5.3 to conclude $II_3\leq\tilde{c}|\Delta|$. To estimate $II_1$, we start out with the identity

$$
\nabla_x(\theta_{\eta\lambda}u\Psi^2) = (\nabla_x\theta_{\eta\lambda})u\Psi^2 + \theta_{\eta\lambda}(\nabla_x u)\Psi^2 + \theta_{\eta\lambda}u\nabla_x(\Psi^2)
$$

to see that $II_1 = II_{11} + II_{12}$, where

$$
\begin{aligned}
II_{11} &:= \iiint_{\mathbb{R}^{n+2}_+} \partial_t u(\theta_{\eta\lambda}u\Psi^2)\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda + \iiint_{\mathbb{R}^{n+2}_+} \nabla_x(\theta_{\eta\lambda}u\Psi^2)\cdot A_{\|\|}\nabla_x u\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda,\\
II_{12} &:= -\iiint_{\mathbb{R}^{n+2}_+} \nabla_x u\cdot A_{\|\|}\nabla_x u(\theta_{\eta\lambda}\Psi^2)\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda - \iiint_{\mathbb{R}^{n+2}_+} \nabla_x\Psi^2\cdot A_{\|\|}\nabla_x u(u\theta_{\eta\lambda})\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda.
\end{aligned}
$$

Using again the fact that $|\theta_{\eta\lambda}|\leq c\eta\lambda$ holds on the support of $\Psi$ along with Cauchy-Schwarz and Lemma 5.3, we deduce the estimate

$$
|II_{12}| \leq c\eta J + \tilde{c}|\Delta|^{1/2}J^{1/2} \leq (\sigma + c\eta)J + \tilde{c}|\Delta|.
$$

To estimate $II_{11}$, we capitalize again that the smoothness of our coefficients allows us to plug in the equation $\partial_t u = \operatorname{div}_{\lambda,x} A\nabla_{\lambda,x} u$ in the pointwise sense. Then, splitting $A$ according to (1.8), we can write $II_{11} = II_{111} + II_{112} + II_{113}$, where

$$\begin{aligned} II_{111} &:= -\iiint_{\mathbb{R}^{n+2}_+} A_{\|\perp} \cdot \nabla_x(\theta_{\eta\lambda} u \Psi^2)\partial_\lambda u \,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda, \\ II_{112} &:= -\iiint_{\mathbb{R}^{n+2}_+} A_{\perp\|} \cdot \nabla_x u \partial_\lambda(\theta_{\eta\lambda} u \Psi^2) \,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda, \\ II_{113} &:= -\iiint_{\mathbb{R}^{n+2}_+} A_{\perp\perp} \partial_\lambda(\theta_{\eta\lambda} u \Psi^2)\partial_\lambda u \,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda. \end{aligned}$$

Unwinding the derivative in $\lambda$ and using once more the bound $|\theta_{\eta\lambda}| \leq c\eta\lambda$ on the support of $\Psi$,

$$|II_{112} + II_{113}| \leq c \iiint_{\mathbb{R}^{n+2}_+} \eta\lambda|\nabla_{\lambda,x} u(\partial_\lambda u)\Psi^2| + \lambda|\nabla_{\lambda,x} u \partial_\lambda \Psi^2| + |\nabla_{\lambda,x} u \partial_\lambda \theta_{\eta\lambda}| \,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda.$$

Here, the first term gives a contribution $c\eta|\Delta|$, the second one can be treated by the familiar combination of Young's inequality and Lemma 5.3, whereas for the third term we make use of Lemma 3.2 (i) instead, noting that $\partial_\lambda \theta_{\eta\lambda} = \partial_\lambda P^*_{\eta\lambda}\varphi$ holds since $\varphi$ does not depend on $\lambda$. By these means, we find

$$|II_{112} + II_{113}| \leq (\sigma + c\eta)J + \tilde{c}|\Delta|.$$

Similarly, we obtain

$$|II_{111} - II_{1111}| \leq c\eta J + \tilde{c}|\Delta|,$$

where

$$II_{1111} := \iiint_{\mathbb{R}^{n+2}_+} A_{\|\perp} \cdot \nabla_x \theta_{\eta\lambda} u \Psi^2 \partial_\lambda u \,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda.$$

To estimate $II_{1111}$ we first integrate by parts in $\lambda$ and regroup derivatives to find

$$\begin{aligned} II_{1111} &= \frac{1}{2}\iiint_{\mathbb{R}^{n+2}_+} A_{\|\perp} \cdot \nabla_x \theta_{\eta\lambda} \Psi^2 \partial_\lambda u^2 \,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda \\ &= -\frac{1}{2}\iiint_{\mathbb{R}^{n+2}_+} A_{\|\perp} \cdot \nabla_x \theta_{\eta\lambda} \partial_\lambda \Psi^2 (u^2) \,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda \\ &\quad -\frac{1}{2}\iiint_{\mathbb{R}^{n+2}_+} A_{\|\perp} \cdot \nabla_x (\partial_\lambda P^*_{\eta\lambda}\varphi \Psi^2 u^2) \,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda \\ &\quad +\frac{1}{2}\iiint_{\mathbb{R}^{n+2}_+} A_{\|\perp} \cdot \partial_\lambda P^*_{\eta\lambda}\varphi \nabla_x (\Psi^2 u^2) \,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda. \end{aligned}$$

Note that the first term on the right-hand side has the same structure as $II_3$ with the only exception that we have a $\lambda$-derivative on $\Psi$ instead of an $x$-derivative. Hence, we can derive a bound $\tilde{c}|\Delta|$ by the very same methods. Also the third term on the right-hand side is of the same kind as a term we encountered earlier in the proof – $I_2$ in this case – which we already know how to bound by $\sigma J + \tilde{c}|\Delta|$.

All in all, we have reached a stage of the proof, where the only term that remains to be estimated is

$$III_1 := \iiint_{\mathbb{R}^{n+2}_+} A_{\|\perp} \cdot \nabla_x (\Psi^2 u^2 \partial_\lambda P^*_{\eta\lambda}\varphi) \,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda$$

and we remark that this final term resembles $J_{211}$ except that we have an additional factor $\partial_\lambda P^*_{\eta\lambda}\varphi$ acting to our favor. We now introduce $\tilde{\varphi}$ as in (1.10), that is, as the energy solution to the problem

$$\operatorname{div}_x(A_{\|\perp}\chi_{8\Delta}) = \partial_t \tilde{\varphi} - \operatorname{div}_x(A_{\|\|}\nabla_x \tilde{\varphi}) = \mathcal{H}_\| \tilde{\varphi}$$

on $\mathbb{R}^{n+1}$. We remark that $\Psi^2 u^2 \partial_\lambda P^*_{\eta\lambda}\varphi$ is qualitatively smooth and compactly supported, therefore it can be used as test function for the equation above in order to rewrite $III_1$. More precisely, we also recall $\tilde{\theta}_{\eta\lambda} = \tilde{\varphi} - P_{\eta\lambda}\tilde{\varphi}$ and write

$$III_1 = III_{11} + III_{12} + III_{13}, \tag{6.4}$$

where

$$\begin{aligned}
III_{11} &:= -\iiint_{\mathbb{R}^{n+2}_+} \partial_t\tilde{\theta}_{\eta\lambda}(\Psi^2 u^2 \partial_\lambda P^*_{\eta\lambda}\varphi)\, \mathrm{d}x\, \mathrm{d}t\, \mathrm{d}\lambda,\\
III_{12} &:= -\iiint_{\mathbb{R}^{n+2}_+} A_{\|\|}\nabla_x\tilde{\theta}_{\eta\lambda}\cdot\nabla_x(\Psi^2 u^2 \partial_\lambda P^*_{\eta\lambda}\varphi)\, \mathrm{d}x\, \mathrm{d}t\, \mathrm{d}\lambda,\\
III_{13} &:= -\iiint_{\mathbb{R}^{n+2}_+} \mathcal{H}_{\|}P_{\eta\lambda}\tilde{\varphi}(\Psi^2 u^2 \partial_\lambda P^*_{\eta\lambda}\varphi)\, \mathrm{d}x\, \mathrm{d}t\, \mathrm{d}\lambda.
\end{aligned}$$

The estimate $|III_{13}|\leq \tilde{c}|\Delta|$ is a consequence of the square function estimate in Lemma 3.2 (i) and (iii). To estimate $III_{12}$, we write $III_{12}=III_{121}+III_{122}+III_{123}$, where

$$\begin{aligned}
III_{121} &:= -\iiint_{\mathbb{R}^{n+2}_+} A_{\|\|}\nabla_x\tilde{\theta}_{\eta\lambda}\cdot\nabla_x(u^2)(\Psi^2 \partial_\lambda P^*_{\eta\lambda}\varphi)\, \mathrm{d}x\, \mathrm{d}t\, \mathrm{d}\lambda,\\
III_{122} &:= -\iiint_{\mathbb{R}^{n+2}_+} A_{\|\|}\nabla_x\tilde{\theta}_{\eta\lambda}\cdot\nabla_x(\Psi^2)(u^2 \partial_\lambda P^*_{\eta\lambda}\varphi)\, \mathrm{d}x\, \mathrm{d}t\, \mathrm{d}\lambda,\\
III_{123} &:= -\iiint_{\mathbb{R}^{n+2}_+} A_{\|\|}\nabla_x\tilde{\theta}_{\eta\lambda}\cdot\nabla_x \partial_\lambda P^*_{\eta\lambda}\varphi(u^2\Psi^2)\, \mathrm{d}x\, \mathrm{d}t\, \mathrm{d}\lambda.
\end{aligned}$$

Once having applied the pointwise bound $|\partial_\lambda P^*_{\eta\lambda}\varphi|\leq c\eta\leq c$ on the support of $\Psi$, see Lemma 5.2 (iii), the estimate of $III_{122}$ reduces to that of $II_3$ with $\tilde{\theta}_{\eta\lambda}$ in lieu of $\theta_{\eta\lambda}$. As the latter two functions share identical estimates, we can record $|III_{122}|\leq\tilde{c}|\Delta|$. To estimate $III_{121}$ we first note, using Cauchy-Schwarz' and Young's inequality, that

$$|III_{121}|\leq cIII_{1211}+\sigma J,$$

where

$$III_{1211}:=\iiint_{\mathbb{R}^{n+2}_+}\Psi^2|\partial_\lambda P^*_{\eta\lambda}\varphi|^2|\nabla_x\tilde{\theta}_{\eta\lambda}|^2\,\frac{\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda}{\lambda}.$$

Now, by the averaging trick already used in the estimate of $II_3$,

$$\begin{aligned}
III_{1211} &\leq \tilde{c}\iiint_{\mathbb{R}^{n+2}_+}\Big(\fint\!\!\!\iint_{W_{\eta/64}(\sigma,y,s)}\Big(\Psi^2|\partial_\lambda P^*_{\eta\lambda}\varphi|^2|\nabla_x\tilde{\theta}_{\eta\lambda}|^2\Big)\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda\Big)\frac{\mathrm{d}y\,\mathrm{d}s\,\mathrm{d}\sigma}{\sigma}\\
&\leq \tilde{c}\iiint_{\mathbb{R}^{n+2}_+}\Big(\sup_{(\lambda,x,t)\in W_{\eta/64}(\sigma,y,s)}|\partial_\lambda P^*_{\eta\lambda}\varphi(x,t)|\Big)^2\frac{\mathrm{d}y\,\mathrm{d}s\,\mathrm{d}\sigma}{\sigma},
\end{aligned}$$

where the second step follows again by Lemma 5.2 (iii) and elementary geometric considerations as in the estimate for $II_3$. As before, we write $W_{\eta/64}(\sigma,y,s):=\Lambda_\sigma\times Q_{\eta\sigma/64}(y)\times I_{\eta\sigma/64}(s)$. From (4.5) we obtain

$$\Big(\sup_{(\lambda,x,t)\in W_{\eta/64}(\sigma,y,s)}|\partial_\lambda P^*_{\eta\lambda}\varphi(x,t)|\Big)^2 \leq \tilde{c}\sum_{j=1}^\infty e^{-c4^j}\fint\!\!\!\int_{2^{j+1}Q_\sigma(y)\times 4^{j+1}I_\sigma(s)}|\sigma\mathcal{H}^*_{\|}\tilde{P}^*_\sigma\varphi(x,t)|^2\,\mathrm{d}x\,\mathrm{d}t,$$

where $\tilde{P}^*_\sigma=(1+\sigma^2\mathcal{H}^*_{\|})^{-1}$. Hence, using an averaging trick in the $(x,t)$-variables only,

$$\begin{aligned}
III_{1211} &\leq \tilde{c}\iiint_{\mathbb{R}^{n+2}_+}\sum_{j=1}^\infty e^{-c4^j}\fint\!\!\!\int_{2^{j+1}Q_\sigma(y)\times 4^{j+1}I_\sigma(s)}|\sigma\mathcal{H}^*_{\|}\tilde{P}^*_\sigma\varphi(x,t)|^2\,\mathrm{d}x\,\mathrm{d}t\frac{\mathrm{d}y\,\mathrm{d}s\,\mathrm{d}\sigma}{\sigma}\\
&= \tilde{c}\sum_{j=1}^\infty e^{-c4^j}\iiint_{\mathbb{R}^{n+2}_+}|\sigma\mathcal{H}^*_{\|}\tilde{P}^*_\sigma\varphi(x,t)|^2\,\frac{\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\sigma}{\sigma}\leq\tilde{c}|\Delta|,
\end{aligned}$$

where the final estimate follows from Lemma 3.2. So, we can conclude $|III_{121}| \leq \sigma J + \tilde{c}|\Delta|$. To estimate $III_{123}$ we use that $u$ is scalar-valued and write

$$\begin{aligned} III_{123} &= \iiint_{\mathbb{R}^{n+2}_+} \nabla_x(\tilde{\theta}_{\eta\lambda} u^2 \Psi^2) \cdot A^*_{\|\|} \nabla_x \partial_\lambda P^*_{\eta\lambda}\varphi \,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda \\ &\quad + \iiint_{\mathbb{R}^{n+2}_+} \tilde{\theta}_{\eta\lambda} \nabla_x(u^2\Psi^2) \cdot A^*_{\|\|} \nabla_x \partial_\lambda P^*_{\eta\lambda}\varphi \,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda. \end{aligned}$$

Note that so far we have neglected $III_{11}$ appearing in (6.4). Now, we come back to this term and combine it with the first integral on the right-hand side above to obtain

$$\begin{aligned} III_{11} + III_{123} &= -\iiint_{\mathbb{R}^{n+2}_+} \tilde{\theta}_{\eta\lambda} u^2 \Psi^2 (\mathcal{H}^*_{\|} \partial_\lambda P^*_{\eta\lambda}\varphi) \,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda \\ &\quad + \iiint_{\mathbb{R}^{n+2}_+} \tilde{\theta}_{\eta\lambda} \nabla_x(u^2\Psi^2) \cdot A^*_{\|\|} \nabla_x \partial_\lambda P^*_{\eta\lambda}\varphi \,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda \\ &\quad + \iiint_{\mathbb{R}^{n+2}_+} \tilde{\theta}_{\eta\lambda} \partial_t(\Psi^2 u^2) \partial_\lambda P^*_{\eta\lambda}\varphi \,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda. \end{aligned}$$

The first term on the right can be bounded by

$$\bigg(\iiint_{\mathbb{R}^{n+2}_+} |\tilde{\theta}_{\eta\lambda}|^2 \,\frac{\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda}{\lambda^3}\bigg)^{1/2} \bigg(\iiint_{\mathbb{R}^{n+2}_+} |\lambda^2 \mathcal{H}^*_{\|} \partial_\lambda P^*_{\eta\lambda}\varphi|^2 \,\frac{\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda}{\lambda}\bigg)^{1/2},$$

which in itself is bounded by $\tilde{c}|\Delta|$ by square function estimates, see Lemma 3.2 (iv) and Lemma 3.3. As for the second term on the right, having applied the pointwise bound $|\tilde{\theta}_{\eta\lambda}| \leq c\eta\lambda \leq c\lambda$ on the support of $\Psi$, we are left with the task of estimating $I_2$, which we have done before.

Altogether,

$$I_1^{\tilde{\theta}} := \iiint_{\mathbb{R}^{n+2}_+} \tilde{\theta}_{\eta\lambda} \partial_t(\Psi^2 u^2) \partial_\lambda P^*_{\eta\lambda}\varphi \,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda$$

is now the only term that remains to be estimated. It is instructive to observe that – upon replacing $|\tilde{\theta}_{\eta\lambda}|$ by its pointwise upper bound $c\lambda$ on the support of $\Psi$ – this is the same term as $2I_1$. Hence, we can follow the treatment of the latter almost *verbatim*, using the pointwise bound whenever feasible in order to reduce matters to estimates that have already been completed. So, we shall only outline the differences in this argument: To start the estimate we write $I_1^{\tilde{\theta}} = I_{11}^{\tilde{\theta}} + I_{12}^{\tilde{\theta}}$, where

$$\begin{aligned} I_{11}^{\tilde{\theta}} &:= \iiint_{\mathbb{R}^{n+2}_+} \tilde{\theta}_{\eta\lambda} \partial_t(\Psi^2) u^2 \partial_\lambda P^*_{\eta\lambda}\varphi \,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda, \\ I_{12}^{\tilde{\theta}} &:= 2\iiint_{\mathbb{R}^{n+2}_+} \tilde{\theta}_{\eta\lambda} u \partial_t u \Psi^2 \partial_\lambda P^*_{\eta\lambda}\varphi \,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda. \end{aligned}$$

The estimate for $I_{11}^{\tilde{\theta}}$ follows from that of $I_{11}$. To estimate $I_{12}^{\tilde{\theta}}$ we use the equation for $u$ and write $I_{12}^{\tilde{\theta}} = I_{121}^{\tilde{\theta}} + I_{122}^{\tilde{\theta}}$, where

$$\begin{aligned} I_{121}^{\tilde{\theta}} &:= 2\iiint_{\mathbb{R}^{n+2}_+} \tilde{\theta}_{\eta\lambda} u \,\mathrm{div}_x(A\nabla_{\lambda,x} u)_{\|} \Psi^2 \partial_\lambda P^*_{\eta\lambda}\varphi \,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda, \\ I_{122}^{\tilde{\theta}} &:= 2\iiint_{\mathbb{R}^{n+2}_+} \tilde{\theta}_{\eta\lambda} u (A\nabla_{\lambda,x} \partial_\lambda u)_{\perp} \Psi^2 \partial_\lambda P^*_{\eta\lambda}\varphi \,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda. \end{aligned}$$

Again, the estimate of $I_{122}^{\tilde\theta}$ follows from the bound for its counterpart $I_{122}$. Furthermore,

$$\begin{aligned} I_{121}^{\tilde\theta} &= 2\iiint_{\mathbb{R}^{n+2}_+} \tilde\theta_{\eta\lambda} u (A\nabla_{\lambda,x}u)_{\|}\Psi^2\cdot\nabla_x\partial_\lambda P^*_{\eta\lambda}\varphi \,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda \\ &\quad +2\iiint_{\mathbb{R}^{n+2}_+} \tilde\theta_{\eta\lambda}\nabla_x u\cdot (A\nabla_{\lambda,x}u)_{\|}\Psi^2\partial_\lambda P^*_{\eta\lambda}\varphi \,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda \\ &\quad +2\iiint_{\mathbb{R}^{n+2}_+} \tilde\theta_{\eta\lambda} u (A\nabla_{\lambda,x}u)_{\|}\cdot\nabla_x\Psi^2\partial_\lambda P^*_{\eta\lambda}\varphi \,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda \\ &\quad +2\iiint_{\mathbb{R}^{n+2}_+} \nabla_x\tilde\theta_{\eta\lambda} u (A\nabla_{\lambda,x}u)_{\|}\Psi^2\partial_\lambda P^*_{\eta\lambda}\varphi \,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\lambda, \end{aligned}$$

where the estimate for the first three terms follows from the bound for $I_{121}$ as before. Eventually, the fourth term, which shows up since unlike $\lambda$ the functions $\tilde\theta_{\eta\lambda}$ does also depend on $x$, can be bounded by $J^{1/2}|III_{1211}|^{1/2}\le \sigma J+\tilde c|\Delta|$ using Cauchy-Schwarz and the previously obtained bound for $III_{1211}$. Put together, this completes the proof of Theorem 1.8.

## References


[1] M.A. Alfonseca, P. Auscher, A. Axelsson, S. Hofmann and S. Kim. *Analyticity of layer potentials and $L^2$ solvability of boundary value problems for divergence form elliptic equations with complex $L^\infty$ coefficients.* Adv. Math. **226** (2011), 4533–4606. 1

[2] D. Aronson. *Non negative solutions of linear parabolic equations.* Ann. Scuola Norm. Sup. Pisa (3) **22** (1968), 607–694. 13

[3] P. Auscher and M. Egert. *On non-autonomous maximal regularity for elliptic operators in divergence form.* Arch. Math. **107** (2016), 271–284. 6

[4] P. Auscher, M. Egert and K. Nyström. *Boundary value problems for parabolic systems via a first order approach.* Available at https://arxiv.org/abs/1607.06246 1, 2, 4, 5, 6, 8, 9, 10, 12, 13, 14

[5] P. Auscher, S. Hofmann, J.L. Lewis and P. Tchamitchian. *Extrapolation of Carleson measures and the analyticity of Kato's square root operators.* Acta Math. **187** (2001), 161–190. 5

[6] A. Castro, K. Nyström and O. Sande. *Boundedness of single layer potentials associated to divergence form parabolic equations with complex coefficients.* To appear in Calc. Var. 1, 4, 5, 6

[7] R. Coifman and C. Fefferman. *Weighted norm inequalities for maximal function and singular integrals.* Studia Math. **51** (1974), 241–250. 3

[8] B. Dahlberg. *On estimates of harmonic measure*, Arch Rational Mech. Anal. **65** (1977), 275–288. 1

[9] M. Dindos, S. Petermichl and J. Pipher. *BMO solvability and the $A_\infty$ condition for second order parabolic operators.* Available at https://arviv.org/abs/1510.05813 2, 4, 5

[10] M. Egert. On Kato's conjecture and mixed boundary conditions. PhD Thesis, Sierke Verlag, Göttingen, 2015. 10

[11] E. Fabes and M. Safonov *Behaviour near the boundary of positive solutions of second order parabolic equations.* J. Fourier Anal. Appl. **3** (1997), 871–882. 3

[12] E. Fabes, M. Safonov and Y. Yuan. *Behavior near the boundary of positive solutions of second order parabolic equations.II.* Trans. Amer. Math. Soc. **351** (1999), 4947–4961. 3

[13] E. Fabes and S. Salsa. *Estimates of caloric measure and the initial-Dirichlet problem for the heat equation in Lipschitz cylinders.* Trans. Amer. Math. Soc. **279** (1983), 635–650. 1

[14] S. Hofmann, C. Kenig, S. Mayboroda and J. Pipher. *Square function/non-tangential maximal function estimates and the Dirichlet problem for non-symmetric elliptic operators.* J. Amer. Math. Soc. **28** (2015), 483–529. 1, 2, 4, 5

[15] S. Hofmann and J.L. Lewis. *$L^2$ solvability and representation by caloric layer potentials in time-varying domains.* Ann. of Math. (2) **144** (1996), 349–420. 1

[16] S. Hofmann and J.L. Lewis. *The Dirichlet problem for parabolic operators with singular drift terms.* Mem. Amer. Math. Soc. **151** (2001), no. 719, viii+113 pp. 1, 4, 5

[17] S. Hofmann and J.L. Lewis. *The $L^p$ Neumann problem for the heat equation in non-cylindrical domains.* J. Funct. Anal. **220** (2005), 1–54. 5, 6

[18] D. Jerison and C. Kenig. *The Dirichlet problem in nonsmooth domains.* Ann. of Math. (2) **113** (1981), 367–382. 4

[19] C. Kenig, B. Kirchheim, J. Pipher and T. Toro. *Square functions and absolute continuity of elliptic measure.* To appear in J. Geom. Anal. 1, 4, 5

[20] J. Lewis and M.A.M. Murray. *The method of layer potentials for the heat equation in time-varying domains.* Mem. Amer. Math. Soc. **114** (1995), no. 545, viii+157 pp. 1

[21] J. Lewis and J. Silver. *Parabolic measure and the Dirichlet problem for the heat equation in two dimensions.* Indiana Univ. Math. J. **37** (1988), 801–839. 1

[22] G.M. Lieberman. Second order parabolic differential equations. World Scientific Publishing Co., Inc., River Edge, NJ, 1996. 15

[23] A. McIntosh. *Operators which have an $H^\infty$ functional calculus*. In Miniconference on operator theory and partial differential equations, Proc. Centre Math. Anal. Austral. Nat. Univ., vol. 14, Austral. Nat. Univ., Canberra, 1986, 210–231. 10

[24] K. Nyström. *The Dirichlet problem for second order parabolic operators*. Indiana Univ. Math. J. **46** (1997), 183–245. 2, 3, 4

[25] K. Nyström. *Square functions estimates and the Kato problem for second order parabolic operators in $\mathbb{R}^{n+1}$*. Adv. Math. **293** (2016), 1–36. 1, 5

[26] K. Nyström. *$L^2$ Solvability of boundary value problems for divergence form parabolic equations with complex coefficients*. Available at http://arxiv.org/abs/1603.02823. 1, 5

[27] J. Rivera-Noriega. *Absolute continuity of parabolic measure and area integral estimates in non-cylindrical domains*. Indiana Univ. Math. J. **51** (2003), 477–525. 1

Laboratoire de Mathématiques d'Orsay, Univ. Paris-Sud, CNRS, Université Paris-Saclay, 91405 Orsay, France

*E-mail address*: pascal.auscher@math.u-psud.fr

*E-mail address*: moritz.egert@math.u-psud.fr

Department of Mathematics, Uppsala University, S-751 06 Uppsala, Sweden

*E-mail address*: kaj.nystrom@math.uu.se