Calculus of Variations

CrossMark

Boundedness of single layer potentials associated to divergence form parabolic equations with complex coefficients

Alejandro J. Castro[1] · Kaj Nyström[1] · Olow Sande[1]

Abstract We consider parabolic operators of the form

$$\partial_t + \mathcal{L}, \ \mathcal{L} := -\operatorname{div} A(X, t)\nabla,$$

in $\mathbb{R}_+^{n+2} := \{(X, t) = (x, x_{n+1}, t) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} : \ x_{n+1} > 0\}, n \geq 1$. We assume that A is a $(n + 1) \times (n + 1)$-dimensional matrix which is bounded, measurable, uniformly elliptic and complex, and we assume, in addition, that the entries of A are independent of the spatial coordinate x_{n+1} as well as of the time coordinate t. We prove that the boundedness of associated single layer potentials, with data in L^2, can be reduced to two crucial estimates (Theorem 1.1), one being a square function estimate involving the single layer potential. By establishing a local parabolic Tb-theorem for square functions we are then able to verify the two crucial estimates in the case of real, symmetric operators (Theorem 1.2). As part of this argument we establish a scale-invariant reverse Hölder inequality for the parabolic Poisson kernel (Theorem 1.3). Our results are important when addressing the solvability of the classical Dirichlet, Neumann and Regularity problems for the operator $\partial_t + \mathcal{L}$ in $\mathbb{R}_+^{n+2}$, with L^2-data on $\mathbb{R}^{n+1} = \partial\mathbb{R}_+^{n+2}$, and by way of layer potentials.

Mathematics Subject Classification 35K20 · 31B10

⊠ Kaj Nyström
  kaj.nystrom@math.uu.se

  Alejandro J. Castro
  alejandro.castro@math.uu.se

  Olow Sande
  olow.sande@math.uu.se

[1] Department of Mathematics, Uppsala University, 751 06 Uppsala, Sweden

⚫ Springer

1 Introduction and statement of main results

In this paper we establish certain estimates related to the solvability of the Dirichlet, Neumann and Regularity problems with data in L^2, in the following these problems are referred to as $(D2)$, $(N2)$ and $(R2)$, by way of layer potentials and for second order parabolic equations of the form

$$\mathcal{H}u := (\partial_t + \mathcal{L})u = 0, \tag{1.1}$$

where

$$\mathcal{L} := -\text{div } A(X,t)\nabla = -\sum_{i,j=1}^{n+1} \partial_{x_i}(A_{i,j}(X,t)\partial_{x_j})$$

is defined in $\mathbb{R}^{n+2} = \{(X,t) = (x_1,\ldots,x_{n+1},t) \in \mathbb{R}^{n+1} \times \mathbb{R}\}$, $n \geq 1$. $A = A(X,t) = \{A_{i,j}(X,t)\}_{i,j=1}^{n+1}$ is assumed to be a $(n+1) \times (n+1)$-dimensional matrix with complex coefficients satisfying the uniform ellipticity condition

$$\begin{aligned}
\text{(i)} \quad &\Lambda^{-1}|\xi|^2 \leq \text{Re}\left(\sum_{i,j=1}^{n+1} A_{i,j}(X,t)\xi_i\bar{\xi}_j\right), \\
\text{(ii)} \quad &|A\xi \cdot \zeta| \leq \Lambda|\xi||\zeta|,
\end{aligned} \tag{1.2}$$

for some Λ, $1 \leq \Lambda < \infty$, and for all $\xi, \zeta \in \mathbb{C}^{n+1}$, $(X,t) \in \mathbb{R}^{n+2}$. Here $u \cdot v = u_1 v_1 + \cdots + u_{n+1}v_{n+1}$, $\bar{u}$ denotes the complex conjugate of u and $u \cdot \bar{v}$ is the (standard) inner product on $\mathbb{C}^{n+1}$. In addition, we consistently assume that

$$A(x_1,\ldots,x_{n+1},t) = A(x_1,\ldots,x_n), \quad \text{i.e., } A \text{ is independent of } x_{n+1} \text{ and } t. \tag{1.3}$$

The solvability of $(D2)$, $(N2)$ and $(R2)$ for the operator $\mathcal{H}$ in $\mathbb{R}^{n+2}_+ = \{(x,x_{n+1},t) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} : x_{n+1} > 0\}$, with data prescribed on $\mathbb{R}^{n+1} = \partial\mathbb{R}^{n+2}_+ = \{(x,x_{n+1},t) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} : x_{n+1} = 0\}$ and by way of layer potentials, can roughly be decomposed into two steps: boundedness of layer potentials and invertibility of layer potentials. In this paper we first prove, in the case of equations of the form (1.1), satisfying (1.2) and (1.3) and the De Giorgi–Moser–Nash estimates stated in (2.6) and (2.7) below, that a set of key boundedness estimates for associated single layer potentials can be reduced to two crucial estimates (Theorem 1.1), one being a square function estimate involving the single layer potential. By establishing a local parabolic Tb-theorem for square functions, and by establishing a version of the main result in [15] for equations of the form (1.1), assuming in addition that A is real and symmetric, we are then subsequently able to verify the two crucial estimates in the case of real, symmetric operators (1.1) satisfying (1.2) and (1.3) (Theorem 1.2). As part of this argument we establish, and this is of independent interest, a scale-invariant reverse Hölder inequality for the parabolic Poisson kernel (Theorem 1.3). The invertibility of layer potentials, and hence the solvability of the Dirichlet, Neumann and Regularity problems L^2-data, is addressed in [33].

Jointly, this paper and [33] yield solvability for $(D2)$, $(N2)$ and $(R2)$, by way of layer potentials, when the coefficient matrix is either

    (i) a small complex perturbation of a constant (complex) matrix, or

    (ii) a real and symmetric matrix, or

    (iii) a small complex perturbation of a real and symmetric matrix.

In all cases the unique solutions can be represented in terms of layer potentials. We claim that the results established in this paper and in [33], and the tools developed, pave the way for important developments in the area of parabolic PDEs. In particular, it is interesting to generalize the present paper and [33] to the context of L^p and relevant endpoint spaces, and to challenge the assumption in (1.3).

The main results of this paper and [33] can jointly be seen as a parabolic analogue of the elliptic results established in [3] and we recall that in [3] the authors establish results concerning the solvability of the Dirichlet, Neumann and Regularity problems with data in L^2, i.e., (D2), (N2) and (R2), by way of layer potentials and for elliptic operators of the form $-\operatorname{div} A(X)\nabla$, in $\mathbb{R}^{n+1}_+ := \{X = (x, x_{n+1}) \in \mathbb{R}^n \times \mathbb{R} : x_{n+1} > 0\}$, $n \geq 2$, assuming that A is a $(n + 1) \times (n + 1)$-dimensional matrix which is bounded, measurable, uniformly elliptic and complex, and assuming, in addition, that the entries of A are independent of the spatial coordinate x_{n+1}. Moreover, if A is real and symmetric, (D2), (N2) and (R2) were solved in [27–29], but the major achievement in [3] is that the authors prove that the solutions can be represented by way of layer potentials. In [24] a version of [3], but in the context of L^p and relevant endpoint spaces, was developed and in [26] the structural assumption that A is independent of the spatial coordinate x_{n+1} is challenged. The core of the impressive arguments and estimates in [3] is based on the fine and elaborated techniques developed in the context of the proof of the Kato conjecture, see [4,5,20].

1.1 Notation

Based on (1.3) we let $\lambda = x_{n+1}$, and when using the symbol λ we will write the point $(X, t) = (x_1, \ldots, x_n, x_{n+1}, t)$ as $(x, t, \lambda) = (x_1, \ldots, x_n, t, \lambda)$. Using this notation,

$$\mathbb{R}^{n+2}_+ = \{(x, t, \lambda) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} : \lambda > 0\},$$

and

$$\mathbb{R}^{n+1} = \partial \mathbb{R}^{n+2}_+ = \{(x, t, \lambda) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} : \lambda = 0\}.$$

We write $\nabla := (\nabla_{||}, \partial_\lambda)$ where $\nabla_{||} := (\partial_{x_1}, \ldots, \partial_{x_n})$. We let $L^2(\mathbb{R}^{n+1}, \mathbb{C})$ denote the Hilbert space of functions $f : \mathbb{R}^{n+1} \to \mathbb{C}$ which are square integrable and we let $||f||_2$ denote the norm of f. We also introduce

$$||| \cdot ||| := \left(\int_0^\infty \int_{\mathbb{R}^{n+1}} | \cdot |^2 \frac{dx\,dt\,d\lambda}{\lambda} \right)^{1/2}. \tag{1.4}$$

Given $(x, t) \in \mathbb{R}^n \times \mathbb{R}$ we let $\|(x, t)\|$ be the unique positive solution ρ to the equation

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

Then $\|(\gamma x, \gamma^2 t)\| = \gamma \|(x, t)\|$, $\gamma > 0$, and we call $\|(x, t)\|$ the parabolic norm of (x, t). We define the parabolic first order differential operator $\mathbb{D}$ through the relation

$$\widehat{(\mathbb{D}f)}(\xi, \tau) := \|(\xi, \tau)\| \hat{f}(\xi, \tau),$$

where $\widehat{(\mathbb{D}f)}$ and $\hat{f}$ denote the Fourier transform of $\mathbb{D}f$ and f, respectively. We define the fractional (in time) differentiation operators $D^t_{1/2}$ through the relation

$$\widehat{(D^t_{1/2}f)}(\xi, \tau) := |\tau|^{1/2} \hat{f}(\xi, \tau).$$

We let H_t denote a Hilbert transform in the t-variable defined through the multiplier $i \operatorname{sgn}(\tau)$. We make the construction so that

$$\partial_t = D^t_{1/2} H_t D^t_{1/2}.$$

By applying Plancherel's theorem we have

$$\|\mathbb{D} f\|_2 \approx \|\nabla_{\|} f\|_2 + \|H_t D^t_{1/2} f\|_2 \approx \|\nabla_{\|} f\|_2 + \|D^t_{1/2} f\|_2,$$

with constants depending only on n.

1.2 Non-tangential maximal functions

Given $(x_0, t_0) \in \mathbb{R}^{n+1}$, and $\beta > 0$, we define the cone

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

Consider a function U defined on $\mathbb{R}^{n+2}_+$. The non-tangential maximal operator N^β_* is defined

$$N^\beta_*(U)(x_0, t_0) := \sup_{(x,t,\lambda) \in \Gamma^\beta(x_0,t_0)} |U(x, t, \lambda)|.$$

Given $(x, t) \in \mathbb{R}^{n+1}$, $\lambda > 0$, we let

$$Q_\lambda(x, t) := \{(y, s) : \ |x_i - y_i| < \lambda, \ |t - s| < \lambda^2\}$$

denote the parabolic cube on $\mathbb{R}^{n+1}$, with center (x, t) and side length λ. We let

$$W_\lambda(x, t) := \{(y, s, \sigma) : \ (y, s) \in Q_\lambda(x, t), \lambda/2 < \sigma < 3\lambda/2\}$$

be an associated Whitney type set. Using this notation we also introduce

$$\tilde{N}^\beta_*(U)(x_0, t_0) := \sup_{(x,t,\lambda) \in \Gamma^\beta(x_0,t_0)} \left(\fint_{W_\lambda(x,t)} |U(y, s, \sigma)|^2 \, dy ds d\sigma \right)^{1/2}.$$

We let

$$\Gamma(x_0, t_0) := \Gamma^1(x_0, t_0), \ N_*(U) := N^1_*(U), \ \tilde{N}_*(U) := \tilde{N}^1_*(U).$$

Furthermore, in many estimates it is necessary to increase the β in Γ^β as the estimate progresses. We will use the convention, when the exact β is not important, that $N_{**}(U), \tilde{N}_{**}(U)$, equal $N^\beta_*(U), \tilde{N}^\beta_*(U)$, for some $\beta > 1$. In fact, the L^p-norms of N_* and N^β_* are equivalent, for any $\beta > 0$ (see for example [16, Lemma 1, p. 166]).

1.3 Single layer potentials

Consider $\mathcal{H} = \partial_t + \mathcal{L} = \partial_t - \operatorname{div} A\nabla$ and $\mathcal{H}^* := -\partial_t + \mathcal{L}^*$, where $\mathcal{L}^*$ is the hermitian adjoint of $\mathcal{L}$, i.e., $\mathcal{L}^* = -\operatorname{div} A^*\nabla$. Assume that $\mathcal{H}, \mathcal{H}^*$, satisfy (1.2) and (1.3). Then $\mathcal{L} = -\operatorname{div} A\nabla$ defines, recall that A is independent of t, a maximal accretive operator on $L^2(\mathbb{R}^{n+1}, \mathbb{C})$ and $-\mathcal{L}$ generates a contraction semigroup on $L^2(\mathbb{R}^{n+1}, \mathbb{C})$, $e^{-t\mathcal{L}}$, for $t > 0$, see p. 28 in [6]. Let $K_t(X, Y)$ denote the distributional or Schwartz kernel of $e^{-t\mathcal{L}}$. In the statement of our main results, and hence throughout the paper, we will assume, in addition to (1.2) and (1.3), that $\mathcal{H}, \mathcal{H}^*$, both satisfy De Giorgi–Moser–Nash estimates stated in (2.6) and (2.7) below. This assumption implies, in particular, that $K_t(X, Y)$ is, for each $t > 0$, Hölder continuous in X

and Y and that $K_t(X, Y)$ satisfies the Gaussian (pointwise) estimates stated in Definition 2 on p. 29 in [6]. Under these assumptions we introduce

$$\Gamma(x, t, \lambda, y, s, \sigma) := \Gamma^{\mathcal{H}}(X, t, Y, s) := K_{t-s}(X, Y) = K_{t-s}(x, \lambda, y, \sigma)$$

whenever $t - s > 0$ and we put $\Gamma(x, t, \lambda, y, s, \sigma) = 0$ whenever $t - s < 0$. Based on (1.3) we in the following let

$$\Gamma_\lambda(x, t, y, s) := \Gamma(x, t, \lambda, y, s, 0),$$
$$\Gamma_\lambda^*(y, s, x, t) := \Gamma^*(y, s, 0, x, t, \lambda),$$

and we introduce associated single layer potentials

$$\mathcal{S}_\lambda^{\mathcal{H}} f(x, t) := \int_{\mathbb{R}^{n+1}} \Gamma_\lambda(x, t, y, s) f(y, s) \, dy ds,$$

$$\mathcal{S}_\lambda^{\mathcal{H}^*} f(x, t) := \int_{\mathbb{R}^{n+1}} \Gamma_\lambda^*(y, s, x, t) f(y, s) \, dy ds.$$

1.4 Statement of main results

The following are our main results.

Theorem 1.1 *Consider $\mathcal{H} = \partial_t - \mathrm{div}\, A\nabla$. Assume that $\mathcal{H}$, $\mathcal{H}^*$, satisfy (1.2) and (1.3) as well as the De Giorgi–Moser–Nash estimates stated in (2.6) and (2.7) below. Assume that there exists a constant C such that*

$$\text{(i)} \quad \sup_{\lambda>0} ||\partial_\lambda \mathcal{S}_\lambda^{\mathcal{H}} f||_2 + \sup_{\lambda>0} ||\partial_\lambda \mathcal{S}_\lambda^{\mathcal{H}^*} f||_2 \leq C||f||_2,$$

$$\text{(ii)} \quad |||\lambda \partial_\lambda^2 \mathcal{S}_\lambda^{\mathcal{H}} f||| + |||\lambda \partial_\lambda^2 \mathcal{S}_\lambda^{\mathcal{H}^*} f||| \leq C||f||_2, \tag{1.5}$$

whenever $f \in L^2(\mathbb{R}^{n+1}, \mathbb{C})$. Then there exists a constant c, depending at most on n, Λ, the De Giorgi–Moser–Nash constants and C, such that

$$\text{(i)} \quad ||N_*(\partial_\lambda \mathcal{S}_\lambda^{\mathcal{H}} f)||_2 + ||N_*(\partial_\lambda \mathcal{S}_\lambda^{\mathcal{H}^*} f)||_2 \leq c||f||_2,$$

$$\text{(ii)} \quad \sup_{\lambda>0} ||\mathbb{D}\mathcal{S}_\lambda^{\mathcal{H}} f||_2 + \sup_{\lambda>0} ||\mathbb{D}\mathcal{S}_\lambda^{\mathcal{H}^*} f||_2 \leq c||f||_2,$$

$$\text{(iii)} \quad ||\tilde{N}_*(\nabla_{||} \mathcal{S}_\lambda^{\mathcal{H}} f)||_2 + ||\tilde{N}_*(\nabla_{||} \mathcal{S}_\lambda^{\mathcal{H}^*} f)||_2 \leq c||f||_2,$$

$$\text{(iv)} \quad ||\tilde{N}_*(H_t D_{1/2}^t \mathcal{S}_\lambda^{\mathcal{H}} f)||_2 + ||\tilde{N}_*(H_t D_{1/2}^t \mathcal{S}_\lambda^{\mathcal{H}^*} f)||_2 \leq c||f||_2, \tag{1.6}$$

whenever $f \in L^2(\mathbb{R}^{n+1}, \mathbb{C})$.

Theorem 1.2 *Consider $\mathcal{H} = \partial_t - \mathrm{div}\, A\nabla$. Assume that $\mathcal{H}$ satisfies (1.2) and (1.3). Assume in addition that A is real and symmetric. Then there exists a constant C, depending at most on n, Λ, such that (1.5) holds with this C. In particular, the estimates in (1.6) all hold, with constants depending only on n, Λ, C, in the case when A is real, symmetric and satisfies (1.2) and (1.3).*

Theorem 1.3 *Assume that $\mathcal{H} = \partial_t - \mathrm{div}\, A\nabla$ satisfies (1.2) and (1.3). Suppose in addition that A is real and symmetric. Then the parabolic measure associated to $\mathcal{H}$, in $\mathbb{R}_+^{n+2}$, is absolutely continuous with respect to the measure $dxdt$ on $\mathbb{R}^{n+1} = \partial\mathbb{R}_+^{n+2}$. Moreover, let $Q \subset \mathbb{R}^{n+1}$ be a parabolic cube and let $K(A_Q, y, s)$ be the to $\mathcal{H}$ associated Poisson kernel at*

$A_Q := (x_Q, l(Q), t_Q)$ *where* (x_Q, t_Q) *is the center of the cube* Q *and* $l(Q)$ *defines its size. Then there exists* $c \geq 1$, *depending only on n and* Λ, *such that*

$$\int_Q |K(A_Q, y, s)|^2 \, dy ds \leq c|Q|^{-1}.$$

Remark 1.4 Note that (1.5) (i) is a uniform (in λ) L^2-estimate involving the first order partial derivative, in the λ-coordinate, of single layer potentials, while (1.5) (ii) is a square function estimate involving the second order partial derivatives, in the λ-coordinate, of single layer potentials. A relevant question is naturally in what generality the estimates in (1.5) can be expected to hold. In [33] it is proved, under additional assumptions, that these estimates are stable under small complex perturbations of the coefficient matrix. However, in the elliptic case and after [3] appeared, it was proved in [34], see [17] for an alternative proof, that if $-\text{div } A(X)\nabla$ satisfies the basic assumptions imposed in [3], then the elliptic version of (1.5) (ii) always holds. In fact, the approach in [34], which is based on functional calculus, even dispenses of the De Giorgi–Moser–Nash estimates underlying [3]. Furthermore, in the elliptic case (1.5) (ii) can be seen to imply (1.5) (i) by the results of [2]. Hence, in the elliptic case, and under the assumptions of [3], the elliptic version of (1.5) always holds. Based on this it is fair to pose the question whether or not a similar line of development can be anticipated in the parabolic case. Based on [32], this paper and [33], we anticipated that a parabolic version of [17] can be developed, To develop a parabolic version of [2] is a very interesting and potentially challenging project.

Theorem 1.3 is used in the proof of Theorem 1.2 and to our knowledge Theorems 1.1, 1.2 and 1.3 are all new. To put these results in the context of the current literature devoted to parabolic layer potentials and parabolic singular integrals, in C^1-regular or Lipschitz regular cylinders, it is fair to first mention [12–14] where a theory of singular integral operators with mixed homogeneity was developed and Theorem 1.1 (i)–(iv) were proved in the context of the heat operator and in the context of time-independent C^1-cylinders. These results were then extended in [7,8], still in the context of the heat operator, to the setting of time-independent Lipschitz domains. The more challenging setting of time-dependent Lipschitz type domains was considered in [18,21,30], see also [22]. In particular, in these papers the correct notion of time-dependent Lipschitz type domains, from the perspective of parabolic singular integral operators and parabolic layer potentials, was found. One major contribution of these papers, see [18,21,22] in particular, is the proof of Theorem 1.1 in the context of the heat operator in time-dependent Lipschitz type domains. Beyond these results the literature only contains modest contributions to the study of parabolic layer potentials associated to second order parabolic operators (in divergence form) with variable, bounded, measurable, uniformly elliptic (and complex) coefficients. Based on this we believe that our results will pave the way for important developments in the area of parabolic PDEs.

While Theorems 1.1 and 1.2 coincide, in the stationary case, with the set up and the corresponding results established in [3] for elliptic equations, we claim that our results, Theorem 1.1 in particular, are not, for at least two reasons, straightforward generalizations of the corresponding results in [3]. First, our result rely on [32] where certain square function estimates are established for second order parabolic operators of the form $\mathcal{H}$, and where, in particular, a parabolic version of the technology in [4] is developed. Second, in general the presence of the (first order) time-derivative forces one to consider fractional time-derivatives leading, as in [18,21,30], see also [22], to rather elaborate additional estimates. Theorem 1.3 gives a parabolic version of an elliptic result due to Jerison and Kenig [27] and a version of

the main result in [15] for equations of the form (1.1), assuming in addition that A is real and symmetric.

1.5 Proofs and organization of the paper

In general we will only supply the proof of our statements for $\mathcal{S}_\lambda := \mathcal{S}_\lambda^{\mathcal{H}}$. The corresponding results for $\mathcal{S}_\lambda^* := \mathcal{S}_\lambda^{\mathcal{H}^*}$ then follow readily by analogy. In Sect. 2, which is of preliminary nature, we introduce notation, weak solutions, state the De Giorgi–Moser–Nash estimates referred to in Theorem 1.1, we prove energy estimates, and we state/prove a few fact from Littlewood–Paley theory. In Sect. 3 we prove a set of important preliminary estimates related to the boundedness of single layer potentials: off-diagonal estimates and uniform (in λ) L^2-estimates. Section 4 is devoted to the proof of two important lemmas: Lemmas 4.1 and 4.2. To briefly describe these results we introduce $\Phi(f)$ where

$$\Phi(f) := \sup_{\lambda>0} ||\partial_\lambda \mathcal{S}_\lambda f||_2 + |||\lambda \partial_\lambda^2 \mathcal{S}_\lambda f|||. \tag{1.7}$$

Lemma 4.1 concerns estimates of non-tangential maximal functions and in this lemma we establish bounds of $||N_*(\partial_\lambda \mathcal{S}_\lambda f)||_2$, $||\tilde{N}_*(\nabla_{||} \mathcal{S}_\lambda f)||_2$ and $||\tilde{N}_*(H_t D_{1/2}^t \mathcal{S}_\lambda f)||_2$ in terms of a constant times

$$\Phi(f) + ||f||_2 + \sup_{\lambda>0} ||\mathbb{D} \mathcal{S}_\lambda f||_2.$$

In Lemma 4.2 we establish square function estimates of the form,

(i) $|||\lambda^{m+2l+4} \nabla \partial_\lambda \partial_t^{l+1} \partial_\lambda^{m+1} \mathcal{S}_\lambda f||| \leq c(\Phi(f) + ||f||_2)$,

(ii) $|||\lambda^{m+2l+4} \partial_t \partial_t^{l+1} \partial_\lambda^{m+1} \mathcal{S}_\lambda f||| \leq c(\Phi(f) + ||f||_2)$,

whenever $f \in L^2(\mathbb{R}^{n+1}, \mathbb{C})$, and for $m \geq -1$, $l \geq -1$. Using Lemma 4.1, the proof of Theorem 1.1 boils down to proving the estimate

$$\sup_{\lambda>0} ||\mathbb{D} \mathcal{S}_\lambda f||_2 \leq c(\Phi(f) + ||f||_2). \tag{1.8}$$

The estimate in (1.8), which is rather demanding, uses Lemma 4.2 and make extensive use of recent results concerning resolvents, square functions and Carleson measures, established in [32]. In Sect. 5 we collect the material from [32] needed in the proof of (1.8). In [32] a parabolic version of the main and hard estimate in [4] is established. In subsection 5.3, we also seize the opportunity to clarify some statements made in [32] concerning the Kato square root problem for parabolic operators. The conclusion is that in [32] the Kato square root problem for parabolic operators, with merely bounded and measurable coefficients, is solved for the first time in the literature. In Sect. 6 we prove (1.8) as a consequence of Lemmas 6.1, 6.2, and 6.3 stated below. For clarity, the final proof of Theorem 1.1, based on the estimates established in the previous sections, is summarized in Sect. 7. In Sect. 8 we prove Theorem 1.2 by first establishing a local parabolic Tb-theorem for square functions, see Theorem 8.1, and then by establishing Theorem 1.3. We believe that our proof of Theorem 1.3 adds to the clarity of the corresponding argument in [15].

2 Preliminaries

Let $x = (x_1, \ldots, x_n)$, $X = (x, x_{n+1})$, $(x, t) = (x_1, \ldots, x_n, t)$, $(X, t) = (x_1, \ldots, x_n, x_{n+1}, t)$. Given $(X, t) = (x, x_{n+1}, t)$, $r > 0$, we let $Q_r(x, t)$ and $\tilde{Q}_r(X, t)$ denote, respectively, the

parabolic cubes in $\mathbb{R}^{n+1}$ and $\mathbb{R}^{n+2}$, centered at (x, t) and (X, t), and of size r. By Q, $\tilde{Q}$ we denote any such parabolic cubes and we let $l(Q), l(\tilde{Q}), (x_Q, t_Q), (X_{\tilde{Q}}, t_{\tilde{Q}})$ denote their sizes and centers, respectively. Given $\gamma > 0$, we let γQ, $\gamma \tilde{Q}$ be the cubes which have the same centers as Q and $\tilde{Q}$, respectively, but with sizes defined by $\gamma l(Q)$ and $\gamma l(\tilde{Q})$. Given a set $E \subset \mathbb{R}^{n+1}$ we let $|E|$ denote its Lebesgue measure and by 1_E we denote the indicator function for E. Finally, by $||\cdot||_{L^2(E)}$ we mean $||\cdot 1_E||_2$. Furthermore, as mentioned and based on (1.3), we will frequently also use a different convention concerning the labeling of the coordinates: we let $\lambda = x_{n+1}$ and when using the symbol λ, the point $(X, t) = (x, x_{n+1}, t)$ will be written as $(x, t, \lambda) = (x_1, \ldots, x_n, t, \lambda)$. We write $\nabla = (\nabla_{||}, \partial_\lambda)$ where $\nabla_{||} = (\partial_{x_1}, \ldots, \partial_{x_n})$. The notation $L^2(\mathbb{R}^{n+1}, \mathbb{C})$, $||\cdot||_2$, $||(\cdot, \cdot)||$, $\mathbb{D}$, $D_{1/2}^t$, H_t, was introduced in Sect. 1.1 above. In the following we will, in addition to $\mathbb{D}$ and $D_{1/2}^t$, at instances also use the parabolic half-order time derivative

$$\widehat{\mathbb{D}_{n+1}f}(\xi, \tau) := \frac{\tau}{||(\xi, \tau)||} \hat{f}(\xi, \tau).$$

We let $\mathbb{H} := \mathbb{H}(\mathbb{R}^{n+1}, \mathbb{C})$ be the closure of $C_0^\infty(\mathbb{R}^{n+1}, \mathbb{C})$ with respect to

$$||f||_{\mathbb{H}} := ||\mathbb{D}f||_2. \tag{2.1}$$

By applying Plancherel's theorem we have

$$
\begin{aligned}
&\text{(i)} \quad ||f||_{\mathbb{H}} \approx ||\nabla_{||}f||_2 + ||H_t D_{1/2}^t f||_2 \approx ||\nabla_{||}f||_2 + ||D_{1/2}^t f||_2, \\
&\text{(ii)} \quad ||\mathbb{D}_{n+1}f||_2 \leq c||D_{1/2}^t f||_2,
\end{aligned}
\tag{2.2}
$$

with constants depending only on n. Furthermore, we let $\tilde{\mathbb{H}} := \tilde{\mathbb{H}}(\mathbb{R}^{n+2}, \mathbb{C})$ be the closure of $C_0^\infty(\mathbb{R}^{n+2}, \mathbb{C})$ with respect to

$$||F||_{\tilde{\mathbb{H}}} := \left(\int_{-\infty}^{\infty} \int_{\mathbb{R}^{n+1}} \left(|\partial_\lambda F|^2 + |\mathbb{D}F|^2 \right) dx dt d\lambda \right)^{1/2}.$$

Similarly, we let $\tilde{\mathbb{H}}_+ := \tilde{\mathbb{H}}_+(\mathbb{R}_+^{n+2}, \mathbb{C})$ be the closure of $C_0^\infty(\mathbb{R}_+^{n+2}, \mathbb{C})$ with respect to the expression in the last display but with integration over the interval $(-\infty, \infty)$ replaced by integration over the interval $(0, \infty)$.

2.1 Weak solutions

Let $\Omega \subset \{X = (x, x_{n+1}) \in \mathbb{R}^n \times \mathbb{R}_+\}$ be a domain and let, given $-\infty < t_1 < t_2 < \infty$, $\Omega_{t_1, t_2} = \Omega \times (t_1, t_2)$. We let $W^{1,2}(\Omega, \mathbb{C})$ be the Sobolev space of complex valued functions v, defined on Ω, such that v and ∇v are in $L^2(\Omega, \mathbb{C})$. $L^2(t_1, t_2, W^{1,2}(\Omega, \mathbb{C}))$ is the space of functions $u : \Omega_{t_1, t_2} \to \mathbb{C}$ such that

$$||u||_{L^2(t_1, t_2, W^{1,2}(\Omega, \mathbb{C}))} := \left(\int_{t_1}^{t_2} ||u(\cdot, t)||_{W^{1,2}(\Omega, \mathbb{C})}^2 dt \right)^{1/2} < \infty.$$

We say that $u \in L^2(t_1, t_2, W^{1,2}(\Omega, \mathbb{C}))$ is a weak solution to the equation

$$\mathcal{H}u = (\partial_t + \mathcal{L})u = 0, \tag{2.3}$$

in Ω_{t_1, t_2}, if

$$\int_{\mathbb{R}_+^{n+2}} \left(A\nabla u \cdot \nabla \bar{\phi} - u \partial_t \bar{\phi} \right) dX dt = 0, \tag{2.4}$$

whenever $\phi \in C_0^\infty(\Omega_{t_1,t_2}, \mathbb{C})$. Similarly, we say that u is a weak solution to (2.3) in $\mathbb{R}_+^{n+2}$ if $u\phi \in L^2(-\infty, \infty, W^{1,2}(\mathbb{R}^n \times \mathbb{R}_+, \mathbb{C}))$ whenever $\phi \in C_0^\infty(\mathbb{R}_+^{n+2}, \mathbb{C})$ and if (5.2) holds whenever $\phi \in C_0^\infty(\mathbb{R}_+^{n+2}, \mathbb{C})$. Assuming that $\mathcal{H}$ satisfies (1.2) and (1.3) as well as the De Giorgi–Moser–Nash estimates stated in (2.6) and (2.7) below, it follows that any weak solution is smooth as a function of t and in this case

$$\int_{\mathbb{R}_+^{n+2}} \left(A\nabla u \cdot \nabla \bar{\phi} + \partial_t u \bar{\phi} \right) \, dX dt = 0,$$

holds whenever $\phi \in C_0^\infty(\Omega_{t_1,t_2}, \mathbb{C})$. Furthermore, if u is globally defined in $\mathbb{R}_+^{n+2}$, and if $D_{1/2}^t u \overline{H_t D_{1/2}^t \phi}$ is integrable in $\mathbb{R}_+^{n+2}$, whenever $\phi \in C_0^\infty(\mathbb{R}_+^{n+2}, \mathbb{C})$, then

$$B_+(u, \phi) = 0 \quad \text{whenever} \quad \phi \in C_0^\infty(\mathbb{R}_+^{n+2}, \mathbb{C}), \tag{2.5}$$

where the sesquilinear form $B_+(\cdot, \cdot)$ is defined on $\tilde{\mathbb{H}}_+ \times \tilde{\mathbb{H}}_+$ as

$$B_+(u, \phi) := \int_0^\infty \int_{\mathbb{R}^{n+1}} \left(A\nabla u \cdot \nabla \bar{\phi} - D_{1/2}^t u \overline{H_t D_{1/2}^t \phi} \right) \, dx dt d\lambda.$$

In particular, whenever u is a weak solution to (2.3) in $\mathbb{R}_+^{n+2}$ such that $u \in \tilde{\mathbb{H}}_+$, then (2.5) holds. From now on, whenever we write that $\mathcal{H}u = 0$ in a bounded domain Ω_{t_1,t_2}, then we mean that (5.2) holds whenever $\phi \in C_0^\infty(\Omega_{t_1,t_2}, \mathbb{C})$, and when we write that $\mathcal{H}u = 0$ in $\mathbb{R}_+^{n+2}$, then we mean that (5.2) holds whenever $\phi \in C_0^\infty(\mathbb{R}_+^{n+2}, \mathbb{C})$.

2.2 De Giorgi–Moser–Nash estimates

We say that solutions to $\mathcal{H}u = 0$ satisfy De Giorgi–Moser–Nash estimates if there exist, for each $1 \le p < \infty$ fixed, constants c and $\alpha \in (0, 1)$ such that the following is true. Let $\tilde{Q} \subset \mathbb{R}^{n+2}$ be a parabolic cube and assume that $\mathcal{H}u = 0$ in $2\tilde{Q}$. Then

$$\sup_{\tilde{Q}} |u| \le c \left(\fint_{2\tilde{Q}} |u|^p \right)^{1/p}, \tag{2.6}$$

and

$$|u(X, t) - u(\tilde{X}, \tilde{t})| \le c \left(\frac{||(X - \tilde{X}, t - \tilde{t})||}{r} \right)^\alpha \left(\fint_{2\tilde{Q}} |u|^p \right)^{1/p}, \tag{2.7}$$

whenever $(X, t), (\tilde{X}, \tilde{t}) \in \tilde{Q}$, $r := l(\tilde{Q})$. The constant c and α will be referred to as the De Giorgi–Moser–Nash constants. It is well known that if (2.6) and (2.7) hold for one p, $1 \le p < \infty$, then these estimates hold for all p in this range.

2.3 Energy estimates

Lemma 2.1 *Assume that $\mathcal{H}$ satisfies* (1.2) *and* (1.3). *Let $\tilde{Q} \subset \mathbb{R}^{n+2}$ be a parabolic cube and let $\beta > 1$ be a fixed constant. Assume that $\mathcal{H}u = 0$ in $\beta\tilde{Q}$. Let $\phi \in C_0^\infty(\beta\tilde{Q})$ be a cut-off function for $\tilde{Q}$ such that $0 \le \phi \le 1$, $\phi = 1$ on $\tilde{Q}$. Then there exists a constant $c = c(n, \Lambda, \beta)$, $1 \le c < \infty$, such that*

$$\int |\nabla u(X, t)|^2 (\phi(X, t))^2 \, dX dt \le c \int |u(X, t)|^2 (|\nabla \phi(X, t)|^2 + \phi(X, t)|\partial_t \phi(X, t)|) \, dX dt.$$

Proof The lemma is a standard energy estimate. Indeed,

$$\int \left(A\nabla u \cdot \nabla(\bar{u}\phi^2) - u\partial_t(\bar{u}\phi^2) \right) dXdt = 0,$$

by the definition of weak solutions. Hence,

$$\int |\nabla u|^2 \phi^2 \, dxdt \leq c \int |u|^2 (|\nabla \phi|^2 + \phi|\partial_t\phi|) \, dXdt. \qquad \square$$

Lemma 2.2 *Assume that $\mathcal{H}$ satisfies* (1.2) *and* (1.3). *Let $Q \subset \mathbb{R}^{n+1}$ be a parabolic cube, $\lambda_0 \in \mathbb{R}$, and let $\beta_1 > 1$, $\beta_2 \in (0, 1]$ be fixed constants. Let $I = (\lambda_0 - \beta_2 l(Q), \lambda_0 + \beta_2 l(Q))$, $\gamma I = (\lambda_0 - \gamma\beta_2 l(Q), \lambda_0 + \gamma\beta_2 l(Q))$ for $\gamma \in (0, 1)$. Assume that $\mathcal{H}u = 0$ in $\beta_1^2 Q \times I$. Then there exists a constant $c = c(n, \Lambda, \beta_1, \beta_2)$, $1 \leq c < \infty$, such that*

$$\text{(i)} \quad \fint_Q |\nabla u(x, t, \lambda_0)|^2 \, dxdt \leq c \fint_{\beta_1 Q \times \frac{1}{4}I} |\nabla u(X, t)|^2 \, dXdt,$$

$$\text{(ii)} \quad \fint_Q |\nabla u(x, t, \lambda_0)|^2 \, dxdt \leq \frac{c}{l(Q)^2} \fint_{\beta_1^2 Q \times \frac{1}{2}I} |u(X, t)|^2 \, dXdt.$$

Proof It suffices to prove the lemma with $\beta_1 = 2$, $\beta_2 = 1$. Furthermore, we only prove (i) as (ii) follows from (i) and Lemma 2.1. For $\lambda_0 \in \mathbb{R}$ fixed, and with γI as above, we let

$$J_1 := \left(\fint_Q \left| \nabla u(x, t, \lambda_0) - \fint_{\frac{1}{16}I} \nabla u(x, t, \lambda) \, d\lambda \right|^2 dxdt \right)^{1/2},$$

$$J_2 := \left(\fint_Q \left| \fint_{\frac{1}{16}I} \nabla u(x, t, \lambda) \, d\lambda \right|^2 dxdt \right)^{1/2}.$$

Then

$$\left(\fint_Q |\nabla u(x, t, \lambda_0)|^2 \, dxdt \right)^{1/2} \leq J_1 + J_2.$$

Using the Hölder inequality

$$J_2 \leq c \left(\fint_{2Q \times \frac{1}{8}I} |\nabla u(X, t)|^2 \, dXdt \right)^{1/2}.$$

Using the fundamental theorem of calculus and the Hölder inequality,

$$J_1 \leq c l(Q) \left(\fint_{Q \times \frac{1}{16}I} |\nabla \partial_\lambda u(X, t)|^2 \, dXdt \right)^{1/2}.$$

Using that $\partial_\lambda u$ is a solution to the same equation as u it follows from Lemma 2.1 that

$$J_1 \leq c \left(\fint_{\frac{3}{2}Q \times \frac{1}{8}I} |\partial_\lambda u(X, t)|^2 \, dXdt \right)^{1/2}.$$

Hence the estimate in (i) follows. $\qquad \square$

Lemma 2.3 *Assume that $\mathcal{H}$ satisfies* (1.2) *and* (1.3). *Let $\tilde{Q} \subset \mathbb{R}^{n+2}$ be a parabolic cube and let $\beta > 1$ be a fixed constant. Assume that $\mathcal{H}u = 0$ in $\beta\tilde{Q}$. Then there exists a constant $c = c(n, \Lambda, \beta)$, $1 \le c < \infty$, such that*

$$\fint_{\tilde{Q}} |\partial_t u(X, t)|^2 \, dXdt \le \frac{c}{l(\tilde{Q})^4} \fint_{\beta\tilde{Q}} |u(X, t)|^2 \, dXdt.$$

Proof Let $\phi \in C_0^\infty(\beta\tilde{Q})$ be a cut-off function for $\tilde{Q}$ such that $0 \le \phi \le 1$, $\phi = 1$ on $\tilde{Q}$, $|\nabla\phi| \le c/l(\tilde{Q})$, $|\partial_t\phi| \le c/l(\tilde{Q})^2$. Let

$$J_1 := \int |\partial_t u|^2 \phi^4 \, dXdt,$$

and

$$J_2 := \int |\nabla u|^2 \phi^2 \, dXdt, \quad J_3 := \int |\nabla \partial_t u|^2 \phi^6 \, dXdt.$$

As $\partial_t u$ is a solution to the same equation as u,

$$\int \left(A\nabla\partial_t u \cdot \nabla(\bar{u}\phi^4) - \partial_t u \partial_t(\bar{u}\phi^4) \right) dXdt = 0.$$

Hence,

$$J_1 = \int \left((A\nabla\partial_t u \cdot \nabla\bar{u})\phi^4 + 4(A\nabla\partial_t u \cdot \nabla\phi)\bar{u}\phi^3 - 4(\partial_t u \partial_t \phi)\bar{u}\phi^3 \right) dXdt,$$

and

$$J_1 \le l(\tilde{Q})^2 \epsilon J_3 + \frac{c(\epsilon)}{l(\tilde{Q})^2} J_2 + \frac{c(\epsilon)}{l(\tilde{Q})^4} \fint_{\beta\tilde{Q}} |u(X, t)|^2 \, dXdt$$

where ϵ is a degree of freedom. Again using that $\partial_t u$ is a solution to the same equation as u, and essentially Lemma 2.1, we see that

$$J_3 \le c \int |\partial_t u|^2 \phi^4 (|\nabla\phi|^2 + |\partial_t\phi|) \, dXdt \le \frac{c}{l(\tilde{Q})^2} J_1.$$

Combining the above estimates, and again using Lemma 2.1, the lemma follows. □

2.4 Littlewood–Paley theory

We define a parabolic approximation of the identity, which will be fixed throughout the paper, as follows. Let $\mathcal{P} \in C_0^\infty(Q_1(0))$, $\mathcal{P} \ge 0$ be real-valued, $\int \mathcal{P} \, dxdt = 1$, where $Q_1(0)$ is the unit parabolic cube in $\mathbb{R}^{n+1}$ centered at 0. At instances we will also assume that $\int x_i \mathcal{P}(x, t) \, dxdt = 0$ for all $i \in \{1, \dots, n\}$. We set $\mathcal{P}_\lambda(x, t) = \lambda^{-n-2}\mathcal{P}(\lambda^{-1}x, \lambda^{-2}t)$ whenever $\lambda > 0$. We let $\mathcal{P}_\lambda$ denote the convolution operator

$$\mathcal{P}_\lambda f(x, t) = \int_{\mathbb{R}^{n+1}} \mathcal{P}_\lambda(x - y, t - s) f(y, s) \, dyds.$$

Similarly, by $\mathcal{Q}_\lambda$ we denote a generic approximation to the zero operator, not necessarily the same at each instance, but chosen from a finite set of such operators depending only on our

original choice of $\mathcal{P}_\lambda$. In particular, $\mathcal{Q}_\lambda(x, t) = \lambda^{-n-2}\mathcal{Q}(\lambda^{-1}x, \lambda^{-2}t)$ where $\mathcal{Q} \in C^\infty(\mathbb{R}^{n+1})$, $\int \mathcal{Q}\, dx dt = 0$. In addition we will, following [21], assume that $\mathcal{Q}_\lambda$ satisfies the conditions

$$|\mathcal{Q}_\lambda(x, t)| \leq \frac{c\lambda}{(\lambda + ||(x, t)||)^{n+3}},$$

$$|\mathcal{Q}_\lambda(x, t) - \mathcal{Q}_\lambda(y, s)| \leq \frac{c||(x - y, t - s)||^\alpha}{(\lambda + ||(x, t)||)^{n+2+\alpha}},$$

where the latter estimate holds for some $\alpha \in (0, 1)$ whenever $2||(x - y, t - s)|| \leq ||(x, t)||$. Under these assumptions it is well known that

$$\int_0^\infty \int_{\mathbb{R}^{n+1}} |\mathcal{Q}_\lambda f|^2 \frac{dx dt d\lambda}{\lambda} \leq c \int_{\mathbb{R}^{n+1}} |f|^2\, dx dt, \tag{2.8}$$

for all $f \in L^2(\mathbb{R}^{n+1}, \mathbb{C})$. In the following we collect a number of elementary observations used in the forthcoming sections.

Lemma 2.4 *Let $\mathcal{P}_\lambda$ be as above. Then*

$$|||\lambda \nabla \mathcal{P}_\lambda f||| + |||\lambda^2 \partial_t \mathcal{P}_\lambda f||| + |||\lambda \mathbb{D} \mathcal{P}_\lambda f||| \leq c||f||_2,$$

for all $f \in L^2(\mathbb{R}^{n+1}, \mathbb{C})$.

Proof This lemma essentially follows immediately from (2.8). For slightly more details we refer to the proof of Lemma 2.30 in [32]. $\qquad\qquad\square$

Consider a cube $Q \subset \mathbb{R}^{n+1}$. In the following we let $\mathcal{A}_\lambda^Q$ denote the dyadic averaging operator induced by Q, i.e., if $\hat{Q}_\lambda(x, t)$ is the minimal dyadic cube (with respect to the grid induced by Q) containing (x, t), with side length at least λ, then

$$\mathcal{A}_\lambda^Q f(x, t) := \fint_{\hat{Q}_\lambda(x,t)} f\, dy ds, \tag{2.9}$$

is the average of f over $\hat{Q}_\lambda(x, t)$.

Lemma 2.5 *Let $\mathcal{A}_\lambda^Q$ and $\mathcal{P}_\lambda$ be as above. Then*

$$\int_0^\infty \int_{\mathbb{R}^{n+1}} |(\mathcal{A}_\lambda^Q - \mathcal{P}_\lambda)f|^2 \frac{dx dt d\lambda}{\lambda} \leq c \int_{\mathbb{R}^{n+1}} |f|^2\, dx dt,$$

for all $f \in L^2(\mathbb{R}^{n+1}, \mathbb{C})$.

Proof The lemma follows by orthogonality estimates and we here include a sketch of the proof for completion. Let $F \in C_0^\infty(\mathbb{R}_+^{n+2}, \mathbb{C})$ be such that $|||F||| = 1$. It suffices to prove that

$$\int_0^\infty \int_{\mathbb{R}^{n+1}} F(x, t, \lambda)\overline{(\mathcal{A}_\lambda^Q - \mathcal{P}_\lambda)f(x, t)} \frac{dx dt d\lambda}{\lambda} \leq c||f||_2,$$

for all $f \in L^2(\mathbb{R}^{n+1}, \mathbb{C})$. To prove this we first note that $|(\mathcal{A}_\lambda^Q - \mathcal{P}_\lambda)f(x_0, t_0)| \leq cM(f)(x_0, t_0)$ whenever $(x_0, t_0) \in \mathbb{R}^{n+1}$ and where M is the parabolic Hardy–Littlewood maximal function. Hence,

$$\sup_{\lambda > 0} ||(\mathcal{A}_\lambda^Q - \mathcal{P}_\lambda)||_{2 \to 2} \leq c.$$

Let $\mathcal{Q}_\lambda$ be an approximation of the zero operator defined based on a function $\mathcal{Q}$ so normalized that $\mathcal{Q}_\lambda$ is a resolution of the identity, i.e.,

$$\int_0^\infty \mathcal{Q}_\lambda^2 g \, \frac{d\lambda}{\lambda} = g,$$

whenever $g \in C_0^\infty(\mathbb{R}^{n+1}, \mathbb{C})$. Then

$$||(\mathcal{A}_\lambda^Q - \mathcal{P}_\lambda)\mathcal{Q}_\sigma||_{2 \to 2} \le c \min\{(\lambda/\sigma)^\delta, (\sigma/\lambda)^\delta\}, \tag{2.10}$$

for some $\delta > 0$. Indeed, let $\mathcal{R}_\lambda(x, t, y, s)$ be the kernel associated to $\mathcal{A}_\lambda^Q - \mathcal{P}_\lambda$, i.e.,

$$\mathcal{R}_\lambda(x, t, y, s) = \frac{1}{|\hat{Q}_\lambda(x, t)|} 1_{\hat{Q}_\lambda(x,t)}(y, s) - \mathcal{P}_\lambda(x - y, t - s).$$

Then $\mathcal{R}_\lambda 1 = 0$ and it is easily seen that

(i) $|\mathcal{R}_\lambda(x, t, y, s)| \le \lambda^\delta(\lambda + ||(x, t)||)^{-n-2-\delta}$,

(ii) $\displaystyle\int_{\mathbb{R}^{n+1}} \sup_{\{(z,w): \, ||(z-y,w-s)|| \le \sigma\}} |\mathcal{R}_\lambda(x, t, z, w) - \mathcal{R}_\lambda(x, t, y, s)| \, dy \, ds \le c(\sigma/\lambda)^\delta$,

whenever $(x, t) \in \mathbb{R}^{n+1}$, $0 < \sigma \le \lambda < \infty$ and with $\delta = 1$. Note that there is an unfortunate statement in the corresponding proof in [32]: there (ii) was stated in a pointwise sense which can, obviously, not hold as the indicator function $1_{\hat{Q}_\lambda(x,t)}$ is not Hölder continuous. Using (i), (ii), one can, arguing as in the proof of display (3.7) and Remark 3.11 in [25], conclude the validity of (2.10). Let $h_\delta(\lambda, \sigma) := \min\{(\lambda/\sigma)^\delta, (\sigma/\lambda)^\delta\}$. We write

$$\left|\int_0^\infty \int_{\mathbb{R}^{n+1}} F(x, t, \lambda)\overline{(\mathcal{A}_\lambda^Q - \mathcal{P}_\lambda)f(x, t)} \, \frac{dx\,dt\,d\lambda}{\lambda}\right|$$
$$= \left|\int_0^\infty \int_0^\infty \int_{\mathbb{R}^{n+1}} F(x, t, \lambda)\overline{(\mathcal{A}_\lambda^Q - \mathcal{P}_\lambda)\mathcal{Q}_\sigma^2 f(x, t)} \, dx\,dt \frac{d\lambda}{\lambda} \frac{d\sigma}{\sigma}\right|,$$

Hence, using Cauchy–Schwarz we see that

$$\left|\int_0^\infty \int_{\mathbb{R}^{n+1}} F(x, t, \lambda)\overline{(\mathcal{A}_\lambda^Q - \mathcal{P}_\lambda)f(x, t)} \, \frac{dx\,dt\,d\lambda}{\lambda}\right| \le I_1^{1/2} I_2^{1/2},$$

where

$$I_1 := \int_0^\infty \int_0^\infty \int_{\mathbb{R}^{n+1}} |F(x, t, \lambda)|^2 h_\delta(\lambda, \sigma) \, dx\,dt \frac{d\lambda}{\lambda} \frac{d\sigma}{\sigma},$$
$$I_2 := \int_0^\infty \int_0^\infty \int_{\mathbb{R}^{n+1}} |(\mathcal{A}_\lambda^Q - \mathcal{P}_\lambda)\mathcal{Q}_\sigma^2 f(x, t)|^2 (h_\delta(\lambda, \sigma))^{-1} \, dx\,dt \frac{d\lambda}{\lambda} \frac{d\sigma}{\sigma}.$$

Integrating with respect to σ in I_1 we see that $I_1 \le c$. Furthermore, using (2.10) we see that

$$I_2 \le \int_0^\infty \int_0^\infty \int_{\mathbb{R}^{n+1}} |\mathcal{Q}_\sigma f(x, t)|^2 h_\delta(\lambda, \sigma) \, dx\,dt \frac{d\lambda}{\lambda} \frac{d\sigma}{\sigma}$$
$$\le c \int_0^\infty \int_{\mathbb{R}^{n+1}} |\mathcal{Q}_\sigma f(x, t)|^2 \, dx\,dt \frac{d\sigma}{\sigma} \le c||f||_2^2.$$

This completes the proof of the lemma. See also the proof of Lemma 4.3 in [25]. $\qquad\square$

3 Off-diagonal and uniform L^2-estimates for single layer potentials

We here establish a number of elementary and preliminary estimates for single layer potentials. We will consistently only formulate and prove results for $\mathcal{S}_\lambda := \mathcal{S}_\lambda^{\mathcal{H}}$ and for $\lambda > 0$, where $\mathcal{H} = \partial_t - \text{div}\, A\nabla$ is assumed to satisfy (1.2) and (1.3) as well as (2.6) and (2.7). The corresponding results for $\mathcal{S}_\lambda^* := \mathcal{S}_\lambda^{\mathcal{H}^*}$ follow by analogy. Here we will also use the notation $\text{div}_{||} = \nabla_{||}\cdot$, $D_i = \partial_{x_i}$ for $i \in \{1, \ldots, n+1\}$. We let

$$(\mathcal{S}_\lambda D_j)f(x,t) := \int_{\mathbb{R}^{n+1}} \partial_{y_j}\Gamma_\lambda(x,t,y,s)f(y,s)\,dyds, \ 1 \leq j \leq n,$$

$$(\mathcal{S}_\lambda D_{n+1})f(x,t) := \int_{\mathbb{R}^{n+1}} \partial_\sigma\Gamma(x,t,\lambda,y,s,\sigma)|_{\sigma=0}f(y,s)\,dyds.$$

We set

$$(\mathcal{S}_\lambda\nabla) := ((\mathcal{S}_\lambda D_1), \ldots, (\mathcal{S}_\lambda D_n), (\mathcal{S}_\lambda D_{n+1})),$$

$$(\mathcal{S}_\lambda\nabla\cdot)\mathbf{f} := \sum_{j=1}^{n+1}(\mathcal{S}_\lambda D_j)f_j,$$

whenever $\mathbf{f} = (f_1, \ldots, f_{n+1})$ and we note that

$$(\mathcal{S}_\lambda\nabla_{||})\cdot\mathbf{f}_{||} = -\mathcal{S}_\lambda(\text{div}_{||}\mathbf{f}_{||}), \quad (\mathcal{S}_\lambda D_{n+1})f_{n+1} = -\partial_\lambda\mathcal{S}_\lambda f_{n+1},$$

whenever $\mathbf{f} = (\mathbf{f}_{||}, f_{n+1}) \in C_0^\infty(\mathbb{R}^{n+1}, \mathbb{C}^{n+1})$ and by the translation invariance in the λ-variable. Given a function $f \in L^2(\mathbb{R}^{n+1}, \mathbb{C})$, and $h = (h_1, \ldots, h_{n+1}) \in \mathbb{R}^{n+1}$, we let $(\mathbb{D}^h f)(x,t) = f(x_1+h_1, \ldots, x_n+h_n, t+h_{n+1}) - f(x,t)$. Given $m \geq -1, l \geq -1$ we let

$$K_{m,\lambda}(x,t,y,s) := \partial_\lambda^{m+1}\Gamma_\lambda(x,t,y,s),$$

$$K_{m,l,\lambda}(x,t,y,s) := \partial_t^{l+1}\partial_\lambda^{m+1}\Gamma_\lambda(x,t,y,s), \tag{3.1}$$

and we introduce

$$d_\lambda(x,t,y,s) := |x-y| + |t-s|^{1/2} + \lambda.$$

Lemma 3.1 *Consider $m \geq -1$, $l \geq -1$. Then there exists constants $c_{m,l}$ and $\alpha \in (0,1)$, depending at most on n, Λ, the De Giorgi–Moser–Nash constants, m, l, such that*

(i) $|K_{m,l,\lambda}(x,t,y,s)| \leq c_{m,l}(d_\lambda(x,t,y,s))^{-n-m-2l-4}$,

(ii) $|(\mathbb{D}^h K_{m,l,\lambda}(\cdot,\cdot,y,s))(x,t)| \leq c_{m,l}||h||^\alpha(d_\lambda(x,t,y,s))^{-n-m-2l-4-\alpha}$,

(iii) $|(\mathbb{D}^h K_{m,l,\lambda}(x,t,\cdot,\cdot))(y,s)| \leq c_{m,l}||h||^\alpha(d_\lambda(x,t,y,s))^{-n-m-2l-4-\alpha}$,

whenever $2||h|| \leq ||(x-y,t-s)||$ or $2||h|| \leq \lambda$.

Proof Assume first that $l = -1$. Then $K_{m,l,\lambda} = K_{m,\lambda}$. In the case $m = -1$ the estimates in (i)–(iii) follow from (2.6) and (2.7), see also [1] and Section 1.4 in [6]. In the cases $m \geq 0$, the corresponding estimates follow by induction using (2.6), (2.7), Lemmas 2.1 and 2.2. This establishes the estimates in (i)–(iii) for $K_{m,-1,\lambda}$ whenever $m \geq -1$. We next consider the case of $K_{m,l,\lambda}, l \geq 0$. Fix $(y,s) \in \mathbb{R}^{n+1}$ and let $u = u(x,t,\lambda) = K_{m,l,\lambda}(x,t,y,s)$ for some $l \geq 0$. Given $(x,t,\lambda) \in \mathbb{R}_+^{n+2}$, let $\tilde{Q} \subset \mathbb{R}^{n+2}$ be the largest parabolic cube centered at (x,t,λ) such that $16\tilde{Q} \subset \mathbb{R}_+^{n+2}$ and such $\mathcal{H}u = 0$ in $16\tilde{Q}$. Then $l(\tilde{Q}) \approx \min\{\lambda, ||(x-y,t-s)||\}$, and

$$|\partial_t u(x,t,\lambda)| \leq c\left(\fint_{2\tilde{Q}}|\partial_t u|^2\,dXdt\right)^{1/2},$$

by (2.6) as $\partial_t u$ is a solution to the same equation as u. Using Lemma 2.3 we can therefore conclude that

$$|\partial_t u(x, t, \lambda)|^2 \leq \frac{c}{l(\tilde{Q})^4} \left(\fint_{8\tilde{Q}} |u|^2 \, dX dt \right).$$

Using this and induction, the estimate in (i) follows for $K_{m,l,\lambda}(x, t, y, s)$ whenever $l \geq -1$. Using (2.7), the estimates in (ii) and (iii) are proved similarly. $\qquad \square$

Lemma 3.2 *Consider $m \geq -1$, $l \geq -1$ and $\rho > 1$. Then there exist a constant $c_{m,l}$, depending at most on n, Λ, the De Giorgi–Moser–Nash constants, m, l, and a constant $c_{m,l,\rho}$, depending in addition on ρ, such that*

(i) $\displaystyle \int_{2^{k+1}Q \backslash 2^k Q} |(2^k l(Q))^{m+2l+3} \nabla_y K_{m,l,\lambda}(x, t, y, s)|^2 \, dy ds \leq c_{m,l} (2^k l(Q))^{-n-2}$,

(ii) $\displaystyle \int_{2Q} |(l(Q))^{m+2l+3} \nabla_y K_{m,l,\lambda}(x, t, y, s)|^2 \, dy ds \leq c_{m,l,\rho} (l(Q))^{-n-2}$,

$\quad if\ l(Q)/\rho \leq \lambda \leq \rho l(Q)$,

whenever $Q \subset \mathbb{R}^{n+1}$ is a parabolic cube, $k \geq 1$ is an integer and $(x, t) \in Q$.

Proof Fix $(x, t) \in Q$ and let

$$v(y, s, \lambda) := \overline{K_{m,l,\lambda}(x, t, y, s)}.$$

Then v is a solution to the adjoint equation. The lemma now follows from Lemma 2.2 (ii), applied to the adjoint equation, and Lemma 3.1 (i). Indeed, it is easy to see that Lemma 2.2 also is valid in when Q is replaced by the annular region $2^{k+1}Q \backslash 2^k Q$. $\qquad \square$

Lemma 3.3 *Consider $m \geq -1$, $l \geq -1$ and $\rho > 1$. Then there exist a constant $c_{m,l}$, depending at most on n, Λ, the De Giorgi–Moser–Nash constants, m, l, and a constant $c_{m,l,\rho}$, depending in addition on ρ, such that*

(i) $||\partial_t^{l+1} \partial_\lambda^{m+1} (\mathcal{S}_\lambda \nabla_{||} \cdot)(\mathbf{f} 1_{2^{k+1}Q \backslash 2^k Q})||_{L^2(Q)}^2$

$\quad \leq c_{m,l} 2^{-(n+2)k} (2^k l(Q))^{-2m-4l-6} ||\mathbf{f}||_{L^2(2^{k+1}Q \backslash 2^k Q)}^2$,

(ii) $||\partial_t^{l+1} \partial_\lambda^{m+1} (\mathcal{S}_\lambda \nabla_{||} \cdot)(\mathbf{f} 1_{2Q})||_{L^2(Q)}^2 \leq c_{m,l,\rho} (l(Q))^{-2m-4l-6} ||\mathbf{f}||_{L^2(2Q)}^2$,

$\quad if\ l(Q)/\rho \leq \lambda \leq \rho l(Q)$,

(iii) $||\partial_t^{l+1} \partial_\lambda^{m+1} (\mathcal{S}_\lambda)(f 1_{2^{k+1}Q \backslash 2^k Q})||_{L^2(Q)}^2 \leq c_{m,l} 2^{-(n+2)k}$

$\quad (2^k l(Q))^{-2m-4l-4} ||f||_{L^2(2^{k+1}Q \backslash 2^k Q)}^2$,

(iv) $||\partial_t^{l+1} \partial_\lambda^{m+1} (\mathcal{S}_\lambda)(f 1_{2Q})||_{L^2(Q)}^2 \leq c_{m,l,\rho} (l(Q))^{-2m-4l-4} ||f||_{L^2(2Q)}^2$,

$\quad if\ l(Q)/\rho \leq \lambda \leq \rho l(Q)$,

whenever $Q \subset \mathbb{R}^{n+1}$ is a parabolic cube, $k \geq 1$ is an integer, $\mathbf{f} \in L^2(\mathbb{R}^{n+1}, \mathbb{C}^n)$, and $f \in L^2(\mathbb{R}^{n+1}, \mathbb{C})$.

Proof Let $(x, t) \in Q$. To prove (i) we note that

$$|\partial_t^{l+1} \partial_\lambda^{m+1} (\mathcal{S}_\lambda \nabla_{||} \cdot)(\mathbf{f} 1_{2^{k+1} Q \backslash 2^k Q})(x, t)|^2$$

$$= \left| \int_{2^{k+1} Q \backslash 2^k Q} \nabla_y K_{m,l,\lambda}(x, t, y, s) \cdot \mathbf{f}(y, s) \, dy ds \right|^2$$

$$\leq ||\nabla_y K_{m,l,\lambda}(x, t, y, s)||_{L^2(2^{k+1} Q \backslash 2^k Q)}^2 ||\mathbf{f}||_{L^2(2^{k+1} Q \backslash 2^k Q)}^2$$

$$\leq c_{m,l} (2^k l(Q))^{-n-2m-4l-8} ||\mathbf{f}||_{L^2(2^{k+1} Q \backslash 2^k Q)}^2,$$

by Lemma 3.2 (i). Hence, integrating with respect to (x, t) we see that

$$||\partial_t^{l+1} \partial_\lambda^{m+1} (\mathcal{S}_\lambda \nabla_{||} \cdot)(\mathbf{f} 1_{2^{k+1} Q \backslash 2^k Q})||_{L^2(Q)}^2$$

$$\leq c_{m,l} (l(Q))^{n+2} (2^k l(Q))^{-n-2m-4l-8} ||\mathbf{f}||_{L^2(2^{k+1} Q \backslash 2^k Q)}^2$$

$$\leq c_{m,l} 2^{-(n+2)k} (2^k l(Q))^{-2m-4l-6} ||\mathbf{f}||_{L^2(2^{k+1} Q \backslash 2^k Q)}^2.$$

This completes the proof of (i). The proof of (ii) is similar. To prove (iii) we again consider $(x, t) \in Q$. Then

$$|\partial_t^{l+1} \partial_\lambda^{m+1} (\mathcal{S}_\lambda)(f 1_{2^{k+1} Q \backslash 2^k Q})(x, t)|^2$$

$$= \left| \int_{2^{k+1} Q \backslash 2^k Q} K_{m,l,\lambda}(x, t, y, s) f(y, s) \, dy ds \right|^2$$

$$\leq ||K_{m,l,\lambda}(x, t, y, s)||_{L^2(2^{k+1} Q \backslash 2^k Q)}^2 ||f||_{L^2(2^{k+1} Q \backslash 2^k Q)}^2$$

$$\leq c_{m,l} (2^k l(Q))^{-n-2m-4l-6} ||f||_{L^2(2^{k+1} Q \backslash 2^k Q)}^2.$$

We can now proceed as above to complete the proof of (iii). The proof of (iv) is similar. $\qquad \square$

Lemma 3.4 *Assume $m \geq -1$, $l \geq -1$, $m + 2l \geq -2$, Then there exists a constant $c_{m,l}$, depending at most on n, Λ, the De Giorgi–Moser–Nash constants, m, l, such that the following holds. Let $\mathbf{f} \in L^2(\mathbb{R}^{n+1}, \mathbb{C}^n)$ and $f \in L^2(\mathbb{R}^{n+1}, \mathbb{C})$. Then*

$$\text{(i)} \quad \sup_{\lambda>0} ||\lambda^{m+2l+3} \partial_t^{l+1} \partial_\lambda^{m+1} (\mathcal{S}_\lambda \nabla_{||} \cdot) \mathbf{f}||_2 \leq c_{m,l} ||\mathbf{f}||_2,$$

$$\text{(ii)} \quad \sup_{\lambda>0} ||\lambda^{m+2l+3} \partial_t^{l+1} \partial_\lambda^{m+1} (\nabla_{||} \mathcal{S}_\lambda f)||_2 \leq c_{m,l} ||f||_2.$$

Furthermore, if $m + 2l \geq -1$ then

$$\text{(iii)} \quad \sup_{\lambda>0} ||\lambda^{m+2l+2} \partial_t^{l+1} \partial_\lambda^{m+1} (\mathcal{S}_\lambda f)||_2 \leq c_{m,l} ||f||_2.$$

Proof We first note that to prove (ii) it suffices to only prove (i), as, by duality, (ii) follows from (i) applied to $\mathcal{S}_\lambda^*$. To prove (i), fix $\lambda > 0$ and consider $m \geq -1, l \geq -1$. Then

$$||\lambda^{m+2l+3} \partial_t^{l+1} \partial_\lambda^{m+1} (\mathcal{S}_\lambda \nabla_{||} \cdot) \mathbf{f}||_2^2 \leq \sum_Q \int_Q |\lambda^{m+2l+3} \partial_t^{l+1} \partial_\lambda^{m+1} (\mathcal{S}_\lambda \nabla_{||} \cdot) \mathbf{f}(x, t)|^2 \, dx dt,$$

where the sum runs over the dyadic grid of parabolic cubes with $l(Q) \approx \lambda$. With Q fixed we see that

$$
\int_Q |\lambda^{m+2l+3} \partial_t^{l+1} \partial_\lambda^{m+1} (\mathcal{S}_\lambda \nabla_{||} \cdot) \mathbf{f}(x,t)|^2 \, dxdt
$$

$$
\leq \int_Q |\lambda^{m+2l+3} \partial_t^{l+1} \partial_\lambda^{m+1} (\mathcal{S}_\lambda \nabla_{||} \cdot)(\mathbf{f} 1_{2Q})(x,t)|^2 \, dxdt
$$

$$
+ \sum_{k \geq 1} \int_Q |\lambda^{m+2l+3} \partial_t^{l+1} \partial_\lambda^{m+1} (\mathcal{S}_\lambda \nabla_{||} \cdot)(\mathbf{f} 1_{2^{k+1}Q \setminus 2^k Q})(x,t)|^2 \, dxdt
$$

$$
\leq c\lambda^{2m+4l+6}(l(Q))^{-2m-4l-6} ||\mathbf{f}||_{L^2(2Q)}^2
$$

$$
+ \sum_{k \geq 1} c 2^{-(n+2)k} \lambda^{2m+4l+6} (2^k l(Q))^{-2m-4l-6} ||\mathbf{f}||_{L^2(2^{k+1}Q \setminus 2^k Q)}^2
$$

$$
\leq c \left(||\mathbf{f}||_{L^2(2Q)}^2 + \sum_{k \geq 1} 2^{-(n+2)k} 2^{-(2m+4l+6)k} ||\mathbf{f}||_{L^2(2^{k+1}Q \setminus 2^k Q)}^2 \right),
$$

by Lemma 3.3 (i) and (ii), as $l(Q) \approx \lambda$. Hence,

$$
||\lambda^{m+2l+3} \partial_t^{l+1} \partial_\lambda^{m+1} (\mathcal{S}_\lambda \nabla_{||} \cdot) \mathbf{f}||_{L^2(\mathbb{R}^{n+1})}^2
$$

$$
\leq c||\mathbf{f}||_2^2 + c \sum_Q \sum_{k \geq 1} 2^{-(n+2)k} 2^{-(2m+4l+6)k} ||\mathbf{f}||_{L^2(2^{k+1}Q \setminus 2^k Q)}^2. \tag{3.2}
$$

To complete the proof of (i) we now note that there exists, given a point (x,t), at most $c_n 2^{(n+2)k}$ cubes Q such that $(x,t) \in 2^{k+1}Q \setminus 2^k Q$. Hence, using this, and the estimate in (3.2), we see that

$$
||\lambda^{m+2l+3} \partial_t^{l+1} \partial_\lambda^{m+1} (\mathcal{S}_\lambda \nabla_{||} \cdot) \mathbf{f}||_{L^2(\mathbb{R}^{n+1})}^2 \leq c||\mathbf{f}||_2^2 + c \sum_{k \geq 1} 2^{-(2m+4l+6)k} ||\mathbf{f}||_2^2
$$

$$
\leq c||\mathbf{f}||_2^2,
$$

as long as $m + 2l > -3$. This completes the proof of (i). Using Lemma 3.3 (iii) and (iv), the proof of (iii) is similar. We omit further details. $\qquad\square$

Lemma 3.5 *Let $f \in C_0^\infty(\mathbb{R}^{n+1}, \mathbb{C})$ and $\lambda_0 > 0$. Then $\mathcal{S}_{\lambda_0} f \in \mathbb{H}(\mathbb{R}^{n+1}, \mathbb{C}) \cap L^2(\mathbb{R}^{n+1}, \mathbb{C})$.*

Proof Given $f \in C_0^\infty(\mathbb{R}^{n+1}, \mathbb{C})$ we let $Q \subset \mathbb{R}^{n+1}$ be a parabolic cube, centered at $(0,0)$, such that the support of f is contained in Q. Let $\lambda_0 > 0$ be fixed. We have to prove that $||\nabla_{||} \mathcal{S}_{\lambda_0} f||_2 < \infty, ||H_t D_{1/2}^t \mathcal{S}_{\lambda_0} f||_2 < \infty$, and that $||\mathcal{S}_{\lambda_0} f||_2 < \infty$. To estimate $||\nabla_{||} \mathcal{S}_{\lambda_0} f||_2$ we see, by duality, that it suffices to bound

$$
\int_Q |(\mathcal{S}_{\lambda_0}^* \nabla_{||} \cdot) \mathbf{f}(x,t)|^2 \, dxdt \leq \int_Q |(\mathcal{S}_{\lambda_0}^* \nabla_{||} \cdot)(\mathbf{f} 1_{2Q})(x,t)|^2 \, dxdt
$$

$$
+ \sum_{k \geq 1} \int_Q |(\mathcal{S}_{\lambda_0}^* \nabla_{||} \cdot)(\mathbf{f} 1_{2^{k+1}Q \setminus 2^k Q})(x,t)|^2 \, dxdt,
$$

where $\mathbf{f} \in C_0^\infty(\mathbb{R}^{n+1}, \mathbb{C}^n)$, $||\mathbf{f}||_2 = 1$. However, now using the adjoint version of Lemma 3.3 (i), (ii) with $l = -1 = m$, we immediately see that

$$
\int_Q |(\mathcal{S}_{\lambda_0}^* \nabla_{||} \cdot) \mathbf{f}(x,t)|^2 \, dxdt \leq c(n, \Lambda, \lambda_0) < \infty,
$$

whenever $\mathbf{f} \in C_0^\infty(\mathbb{R}^{n+1}, \mathbb{C}^n)$, $||\mathbf{f}||_2 = 1$. To estimate $||H_t D_{1/2}^t \mathcal{S}_{\lambda_0} f||_2$ we first note that

$$||H_t D_{1/2}^t \mathcal{S}_{\lambda_0} f||_2^2 \leq ||\partial_t \mathcal{S}_{\lambda_0} f||_2 ||\mathcal{S}_{\lambda_0} f||_2.$$

Using Lemma 3.4 (iii) we see that $||\partial_t \mathcal{S}_{\lambda_0} f||_2 \leq c(n, \Lambda, \lambda_0)||f||_2 < \infty$. To estimate $||\mathcal{S}_{\lambda_0} f||_2$ we write

$$\int_{\mathbb{R}^{n+1}} |\mathcal{S}_{\lambda_0} f(x, t)|^2 \, dxdt \leq \int_{2Q} |\mathcal{S}_{\lambda_0} f(x, t)|^2 \, dxdt$$
$$+ \sum_{k \geq 1} \int_{2^{k+1}Q \setminus 2^k Q} |\mathcal{S}_{\lambda_0} f(x, t)|^2 \, dxdt.$$

Using this and Lemma 3.1 (i) we deduce that

$$\int_{\mathbb{R}^{n+1}} |\mathcal{S}_{\lambda_0} f(x, t)|^2 \, dxdt \leq c(n, \Lambda, \lambda_0) < \infty.$$

This completes the proof of the lemma. $\square$

4 Estimates of non-tangential maximal functions and square functions

Consider $\mathcal{S}_\lambda = \mathcal{S}_\lambda^{\mathcal{H}}$, for $\lambda > 0$, where $\mathcal{H} = \partial_t - \operatorname{div} A\nabla$ is assumed to satisfy (1.2) and (1.3) as well as (2.6) and (2.7). Recall the notation $||| \cdot |||$, $\Phi(f)$, introduced in (1.4), (1.7). This section is devoted to the proof of the following two lemmas.

Lemma 4.1 *Then there exists a constant c, depending at most on n, Λ, and the De Giorgi–Moser–Nash constants, such that*

(i) $||N_*(\partial_\lambda \mathcal{S}_\lambda f)||_2 \leq c \left(\sup_{\lambda > 0} ||\partial_\lambda \mathcal{S}_\lambda||_{2 \to 2} + 1 \right) ||f||_2,$

(ii) $||\tilde{N}_*(\nabla_{||} \mathcal{S}_\lambda f)||_2 \leq c \left(||f||_2 + \sup_{\lambda > 0} ||\nabla_{||} \mathcal{S}_\lambda f||_2 + ||N_{**}(\partial_\lambda \mathcal{S}_\lambda f)||_2 \right),$

(iii) $||\tilde{N}_*(H_t D_{1/2}^t \mathcal{S}_\lambda f)||_2 \leq c \left(||f||_2 + \sup_{\lambda > 0} ||H_t D_{1/2}^t \mathcal{S}_\lambda f||_2 \right)$
$$+ c \left(||\tilde{N}_{**}(\nabla_{||} \mathcal{S}_\lambda f)||_2 + ||N_{**}(\partial_\lambda \mathcal{S}_\lambda f)||_2 \right),$$

whenever $f \in L^2(\mathbb{R}^{n+1}, \mathbb{C})$.

Lemma 4.2 *Assume $m \geq -1$, $l \geq -1$. Let $\Phi(f)$ be defined as in (1.7). Assume that $\Phi(f) < \infty$ whenever $f \in L^2(\mathbb{R}^{n+1}, \mathbb{C})$. Then there exists a constant c, depending at most on n, Λ, the De Giorgi–Moser–Nash constants, and m, l, such that*

(i) $|||\lambda^{m+2l+4} \nabla \partial_\lambda \partial_t^{l+1} \partial_\lambda^{m+1} \mathcal{S}_\lambda f||| \leq c(\Phi(f) + ||f||_2),$

(ii) $|||\lambda^{m+2l+4} \partial_t \partial_t^{l+1} \partial_\lambda^{m+1} \mathcal{S}_\lambda f||| \leq c(\Phi(f) + ||f||_2),$

whenever $f \in L^2(\mathbb{R}^{n+1}, \mathbb{C})$.

4.1 Proof of Lemma 4.1

Throughout the proof we can, without loss of generality, assume that $f \in C_0^\infty(\mathbb{R}^{n+1}, \mathbb{C})$. We let $Q \subset \mathbb{R}^{n+1}$ be the (smallest) cube centered at $(0,0)$ such that the support of f is contained in $\frac{1}{2}Q$. Let $\delta > 0$ be small and let $1_{\lambda > 2\delta}$ denote the indicator function for the set $\{\lambda : \lambda > 2\delta\} \subset \mathbb{R}$.

Proof of Lemma 4.1 (i) We let $(x_0, t_0) \in \mathbb{R}^{n+1}$. Recall that the kernel of $\partial_\lambda S_\lambda$ is $K_{0,\lambda}(x, t, y, s)$ introduced in (3.1). $K_{0,\lambda}(x, t, y, s)$ is a (parabolic) Calderon–Zygmund kernel satisfying the Calderon–Zygmund type estimates stated in Lemma 3.1. Given $(x_0, t_0) \in \mathbb{R}^{n+1}$ we consider $(x, t, \lambda) \in \Gamma(x_0, t_0)$. Then

$$
|\partial_\lambda S_\lambda(f)(x, t) - \partial_\lambda S_\lambda(f)(x_0, t_0)|
$$
$$
\leq \int_{\mathbb{R}^{n+1}} |K_{0,\lambda}(x, t, y, s) - K_{0,\lambda}(x_0, t_0, y, s)| |f(y, s)| \, dy \, ds
$$
$$
\leq cM(f)(x_0, t_0),
$$

by Lemma 3.1 and where M is the parabolic Hardy–Littlewood maximal function. Hence

$$
N_*(1_{\lambda > 2\delta} \partial_\lambda S_\lambda f)(x_0, t_0) \leq \sup_{\lambda > 2\delta} |\partial_\lambda S_\lambda(f)(x_0, t_0)| + cM(f)(x_0, t_0),
$$

and we intend to estimate $|\partial_\lambda S_\lambda(f)(x_0, t_0)|$ for $\lambda > 2\delta$. To do this we fix $\lambda > 2\delta$ and we decompose $\partial_\lambda S_\lambda(f)(x_0, t_0)$ as

$$
\int_{||(x_0 - y, t_0 - s)|| > 5\lambda} (K_{0,\lambda}(x_0, t_0, y, s) - K_{0,\delta}(x_0, t_0, y, s)) f(y, s) \, dy \, ds
$$
$$
+ \int_{||(x_0 - y, t_0 - s)|| \leq 5\lambda} K_{0,\lambda}(x_0, t_0, y, s) f(y, s) \, dy \, ds
$$
$$
- \int_{\lambda < ||(x_0 - y, t_0 - s)|| < 5\lambda} K_{0,\delta}(x_0, t_0, y, s) f(y, s) \, dy \, ds
$$
$$
+ \int_{||(x_0 - y, t_0 - s)|| > \lambda} K_{0,\delta}(x_0, t_0, y, s) f(y, s) \, dy \, ds
$$
$$
=: I_1^\delta(x_0, t_0, \lambda) + I_2^\delta(x_0, t_0, \lambda) + I_3^\delta(x_0, t_0, \lambda) + I_4^\delta(x_0, t_0, \lambda).
$$

Using Lemma 3.1 we see that

$$
|I_1^\delta(x_0, t_0, \lambda) + I_2^\delta(x_0, t_0, \lambda) + I_3^\delta(x_0, t_0, \lambda)| \leq cM(f)(x_0, t_0).
$$

Furthermore,

$$
|I_4^\delta(x_0, t_0, \lambda)| \leq T_*^\delta f(x_0, t_0),
$$

where

$$
T_*^\delta f(x_0, t_0) = \sup_{\epsilon > 2\delta} |T_\epsilon^\delta f(x_0, t_0)|
$$

and

$$
T_\epsilon^\delta f(x_0, t_0) = \int_{||(x_0 - y, t_0 - s)|| > \epsilon} K_{0,\delta}(x_0, t_0, y, s) f(y, s) \, dy \, ds.
$$

We have to prove that $\mathcal{T}_*^\delta : L^2(\mathbb{R}^{n+1}, \mathbb{C}) \to L^2(\mathbb{R}^{n+1}, \mathbb{C})$ and we have to estimate $||\mathcal{T}_*^\delta||_{2\to 2}$. To do this we carry out an argument similar to the proof of Cotlar's inequality for Calderon–Zygmund operators. With $\epsilon > 0$ fixed, we let Q_ϵ be the the largest parabolic cube, centered at (x_0, t_0), which satisfies that $2Q_\epsilon \cap \{(y, s) \in \mathbb{R}^{n+1} : ||(x_0 - y, t_0 - s)|| > \epsilon\} = \emptyset$. Then $l(Q_\epsilon) \approx \epsilon$. Write $f = f 1_{2Q_\epsilon} + f 1_{\mathbb{R}^{n+1} \setminus 2Q_\epsilon}$. Then

$$|\mathcal{T}_\epsilon^\delta f(x_0, t_0)| = |\partial_\lambda \mathcal{S}_\delta (f 1_{\mathbb{R}^{n+1} \setminus 2Q_\epsilon})(x_0, t_0)|$$
$$\leq c M(f)(x_0, t_0) + |\partial_\lambda \mathcal{S}_\delta f(x, t)| + |\partial_\lambda \mathcal{S}_\delta (f 1_{2Q_\epsilon})(x, t)|,$$

whenever $(x, t) \in Q_\epsilon$ and where have used Lemma 3.1 once again. Let $r \in (0, 1)$. Taking a L^r average in the last display with respect to (x, t), we see that

$$|\mathcal{T}_\epsilon^\delta f(x_0, t_0)| \leq c M(f)(x_0, t_0) + (M(|\partial_\lambda \mathcal{S}_\delta f|^r)(x_0, t_0))^{1/r}$$
$$+ \left(\fint_{Q_\epsilon} |\partial_\lambda \mathcal{S}_\delta (f 1_{2Q_\epsilon})|^r \, dx dt \right)^{1/r}.$$

Hence,

$$|\mathcal{T}_\epsilon^\delta f(x_0, t_0)| \leq c M(f)(x_0, t_0) + (M(|\partial_\lambda \mathcal{S}_\delta f|^r)(x_0, t_0))^{1/r} + M(|\partial_\lambda \mathcal{S}_\delta f|)(x_0, t_0).$$

Furthermore, using an equality attributed to Kolmogorov, see Lemma 10 on p. 35 in [11] for example, and that the support of f is contained in Q, we see that

$$(M(|\partial_\lambda \mathcal{S}_\delta f|^r)(x_0, t_0))^{1/r} \leq c ||\partial_\lambda \mathcal{S}_\delta||_{L^1(Q) \to L^{1,\infty}(5Q)} M(f)(x_0, t_0),$$

where $L^{1,\infty}(5Q)$ is weak-L^1. Using that $\partial_\lambda \mathcal{S}_\delta$ is a Calderon–Zygmund operator one can deduce, by retracing, and localizing, the proof of weak estimates in Calderon–Zygmund theory based on L^2 estimates, that

$$||\partial_\lambda \mathcal{S}_\delta||_{L^1(Q) \to L^{1,\infty}(5Q)} \leq c \big(1 + ||\partial_\lambda \mathcal{S}_\delta||_{L^2(Q) \to L^2(\mathbb{R}^{n+1})}\big),$$

where c depends on the kernel $K_{0,\lambda}$ through the constants appearing in Lemma 3.1. For a detailed account of the dependence of the constant c, see [31]. Hence

$$\mathcal{T}_*^\delta f(x_0, t_0) \leq c \big(1 + ||\partial_\lambda \mathcal{S}_\delta||_{L^2(Q) \to L^2(\mathbb{R}^{n+1})}\big) M(f)(x_0, t_0) + M(|\partial_\lambda \mathcal{S}_\delta f|)(x_0, t_0)$$

and retracing the estimates we can conclude that we have proved that

$$N_*(1_{\lambda > 2\delta} \partial_\lambda \mathcal{S}_\lambda f)(x_0, t_0) \leq c \big(1 + ||\partial_\lambda \mathcal{S}_\delta||_{2\to 2}\big) M(f)(x_0, t_0) + M(|\partial_\lambda \mathcal{S}_\delta f|)(x_0, t_0)$$

whenever $(x_0, t_0) \in \mathbb{R}^{n+1}$ and $\delta > 0$. Hence,

$$||N_*(1_{\lambda > 2\delta} \partial_\lambda \mathcal{S}_\lambda f)||_2 \leq c \left(1 + \sup_{\lambda > 0} ||\partial_\lambda \mathcal{S}_\lambda||_{2\to 2} \right) ||f||_2,$$

whenever $f \in C_0^\infty(\mathbb{R}^{n+1}, \mathbb{C})$ and for a constant c, depending at most on n, Λ, and the De Giorgi–Moser–Nash constants, in particular c is independent of δ. Letting $\delta \to 0$ completes the proof of Lemma 4.1 (i). $\qquad\square$

Proof of Lemma 4.1 (ii) We let $(x_0, t_0) \in \mathbb{R}^{n+1}$. To estimate $\tilde{N}_*(1_{\lambda > 2\delta} \nabla_{||} \mathcal{S}_\lambda f)(x_0, t_0)$ it suffices to bound

$$\left(\fint_{W_\lambda(x_0, t_0)} |\nabla_{||} \mathcal{S}_\sigma f(y, s)|^2 \, dy ds d\sigma \right)^{1/2},$$

where

$$W_\lambda(x,t) := \{(y,s,\sigma) : \ (y,s) \in Q_\lambda(x,t), \lambda/2 < \sigma < 3\lambda/2\}$$

and for $\lambda > 4\delta/3$ which we from now on assume. In the following we let, for $m \in \{0, 1, \ldots, 4\}$

$$2^m W_\lambda(x,t) := \{(y,s,\sigma) : \ (y,s) \in Q_{2^m\lambda}(x,t), \lambda/2 - m\lambda 2^{-10} < \sigma < 3\lambda/2 + m\lambda 2^{-10}\}.$$

Then $2^0 W_\lambda(x,t) = W_\lambda(x,t)$. Using this notation and energy estimates, Lemma 2.1, we see that

$$\fint_{W_\lambda(x_0,t_0)} |\nabla_\| \mathcal{S}_\sigma f(y,s)|^2 \, dy ds d\sigma \le \frac{c}{\lambda^2} \fint_{2W_\lambda(x_0,t_0)} |\mathcal{S}_\sigma f(y,s) - A|^2 \, dy ds d\sigma,$$

where A is a constant which in the following is a degree of freedom. Furthermore, using (2.6) with $p = 1$ we see that

$$\left(\fint_{W_\lambda(x_0,t_0)} |\nabla_\| \mathcal{S}_\sigma f(y,s)|^2 \, dy ds d\sigma \right)^{1/2} \le \frac{c}{\lambda} \fint_{2^2 W_\lambda(x_0,t_0)} |\mathcal{S}_\sigma f(y,s) - A| \, dy ds d\sigma.$$

We write

$$\frac{1}{\lambda} \fint_{2^2 W_\lambda(x_0,t_0)} |\mathcal{S}_\sigma f(y,s) - A| \, dy ds d\sigma$$

$$\le \frac{1}{\lambda} \fint_{2^2 W_\lambda(x_0,t_0)} |\mathcal{S}_\sigma f(y,s) - \mathcal{S}_\sigma f(y,t_0)| \, dy ds d\sigma$$

$$+ \frac{1}{\lambda} \fint_{2^2 W_\lambda(x_0,t_0)} |\mathcal{S}_\sigma f(y,t_0) - A| \, dy ds d\sigma$$

$$=: I_1 + I_2.$$

By the fundamental theorem of calculus we have

$$I_1 \le \fint_{2^3 W_\lambda(x_0,t_0)} |\lambda \partial_t \mathcal{S}_\sigma f(y,s)| \, dy ds d\sigma.$$

Let $Q \subset \mathbb{R}^{n+1}$ be a parabolic cube centered at (x_0,t_0) and with side length 8λ. Then I_1 is bounded by

$$c \int_{\lambda/8}^{2\lambda} \int_Q |\lambda^{-n-2} \partial_t \mathcal{S}_\sigma(f \mathbf{1}_{2Q})(y,s)| \, dy ds d\sigma$$

$$+ c \int_{\lambda/8}^{2\lambda} \int_Q |\lambda^{-n-2} \left(\partial_t \mathcal{S}_\sigma(f \mathbf{1}_{\mathbb{R}^{n+1}\setminus 2Q})(y,s) - \partial_t \mathcal{S}_\sigma(f \mathbf{1}_{\mathbb{R}^{n+1}\setminus 2Q})(x_0,t_0) \right)| \, dy ds d\sigma$$

$$+ c \int_{\lambda/8}^{2\lambda} |\partial_t \mathcal{S}_\sigma(f \mathbf{1}_{\mathbb{R}^{n+1}\setminus 2Q})(x_0,t_0)| \, d\sigma$$

$$=: I_{11} + I_{12} + I_{13}.$$

Using Lemma 3.1 we see that

$$I_{11} + I_{12} \le c M(f)(x_0,t_0),$$

where M is the parabolic Hardy–Littlewood maximal function. Furthermore,

$$I_{13} \le c \sum_{k=1}^{\infty} \int_{\lambda/8}^{2\lambda} |\partial_t \mathcal{S}_{\sigma}(f \mathbf{1}_{2^{k+1}Q \setminus 2^k Q})(x_0, t_0)| \, d\sigma$$

$$\le c\lambda \sum_{k=1}^{\infty} (2^k \lambda)^{-n-3} \int_{2^{k+1}Q} |f(y, s)| \, dy ds \le cM(f)(x_0, t_0).$$

Hence, we can conclude that

$$I_1 \le cM(f)(x_0, t_0). \tag{4.1}$$

Focusing on I_2 we see that

$$\begin{aligned} I_2 &\le \frac{1}{\lambda} \fint_{2^2 W_\lambda(x_0, t_0)} |\mathcal{S}_\sigma f(y, t_0) - \mathcal{S}_{\delta/4} f(y, t_0)| \, dy ds d\sigma \\ &\quad + \frac{1}{\lambda} \fint_{2^2 W_\lambda(x_0, t_0)} |\mathcal{S}_{\delta/4} f(y, t_0) - A| \, dy ds d\sigma \\ &=: I_{21} + I_{22}. \end{aligned}$$

By the fundamental theorem of calculus

$$\begin{aligned} I_{21} &\le \frac{1}{\lambda} \fint_{2^3 W_\lambda(x_0, t_0)} \lambda |N_{**}^x(\partial_\lambda \mathcal{S}_\lambda f(\cdot, t_0))(y)| \, dy ds d\sigma \\ &\le M^x(N_{**}^x(\partial_\lambda \mathcal{S}_\lambda f(\cdot, t_0))(\cdot))(x_0), \end{aligned}$$

where M^x is the Hardy–Littlewood maximal function in x only and N_{**}^x is an elliptic non tangential maximal function on a fixed time slice. Finally, let A be the average of $\mathcal{S}_{\delta/4} f(y, t_0)$, with respect to y, on an spatial surface cube around x_0 with sidelength λ. Then, using the L^1-Poincare inequality we deduce that

$$I_{22} \le cM^x(\nabla_{||} \mathcal{S}_{\delta/4} f(\cdot, t_0))(x_0).$$

Retracing the argument we can conclude that

$$\begin{aligned} \tilde{N}_*(\mathbf{1}_{\lambda > 2\delta} \nabla_{||} \mathcal{S}_\lambda f)(x_0, t_0) &\le c\big(M(f)(x_0, t_0) + M^x(N_{**}^x(\partial_\lambda \mathcal{S}_\lambda f(\cdot, t_0))(\cdot))(x_0) \\ &\quad + M^x(\nabla_{||} \mathcal{S}_{\delta/4} f(\cdot, t_0))(x_0)\big). \end{aligned}$$

Hence

$$\begin{aligned} ||\tilde{N}_*(\mathbf{1}_{\lambda > 2\delta} \nabla_{||} \mathcal{S}_\lambda f)||_2^2 &\le c\big(||f||_2^2 + ||\nabla_{||} \mathcal{S}_{\delta/4} f||_2^2\big) \\ &\quad + \int_{-\infty}^{\infty} \int_{\mathbb{R}^n} |N_{**}^x(\partial_\lambda \mathcal{S}_\lambda f(\cdot, t))(x)|^2 \, dx dt. \end{aligned}$$

However,

$$N_{**}^x(\partial_\lambda \mathcal{S}_\lambda f(\cdot, t_0))(x_0) \le N_{**}(\partial_\lambda \mathcal{S}_\lambda f)(x_0, t_0)$$

and we can conclude that

$$||\tilde{N}_*(\mathbf{1}_{\lambda > 2\delta} \nabla_{||} \mathcal{S}_\lambda f)||_2 \le c\left(||f||_2 + \sup_{\lambda > 0} ||\nabla_{||} \mathcal{S}_\lambda f||_2 + ||N_{**}(\partial_\lambda \mathcal{S}_\lambda f)||_2\right).$$

This completes the proof of Lemma 4.1 (ii). □

Proof of Lemma 4.1 (iii) We again fix $(x_0, t_0) \in \mathbb{R}^{n+1}$ and we note that to estimate

$$\tilde{N}_*(1_{\lambda > 2\delta} H_t D^t_{1/2} \mathcal{S}_\lambda f)(x_0, t_0)$$

it suffices to bound

$$\left(\fint_{W_\lambda(x_0,t_0)} |H_t D^t_{1/2} \mathcal{S}_\sigma f(y,s)|^2 \, dy ds d\sigma \right)^{1/2}, \quad \lambda > 4\delta/3.$$

Consider $(y, s, \sigma) \in W_\lambda(x_0, t_0)$, $\lambda > 4\delta/3$, and let $K \gg 1$ be a degree of freedom to be chosen. Then

$$
\begin{aligned}
H_t D^t_{1/2}(\mathcal{S}_\sigma f)(y,s) &= \lim_{\epsilon \to 0} \int_{\epsilon \le |s-t| < 1/\epsilon} \frac{\mathrm{sgn}(s-t)}{|s-t|^{3/2}} (\mathcal{S}_\sigma f)(y,t) \, dt \\
&= \lim_{\epsilon \to 0} \int_{\epsilon \le |s-t| < (K\sigma)^2} \frac{\mathrm{sgn}(s-t)}{|s-t|^{3/2}} (\mathcal{S}_\sigma f)(y,t) \, dt \\
&\quad + \lim_{\epsilon \to 0} \int_{(K\sigma)^2 \le |s-t| < 1/\epsilon} \frac{\mathrm{sgn}(s-t)}{|s-t|^{3/2}} (\mathcal{S}_\sigma f)(y,t) \, dt \\
&=: g_1(y,s,\sigma) + g_2(y,s,\sigma).
\end{aligned}
$$

Let

$$g_3(x_0, t_0, \sigma) := \sup_{\{y:\, |y-x_0| \le 4\sigma\}} \sup_{\{\tau:\, |\tau-t_0| \le (4K\sigma)^2\}} |\partial_\tau (\mathcal{S}_\sigma f)(y,\tau)|.$$

Then, using the oddness about s of the kernel in the definition of g_1,

$$|g_1(y,s,\sigma)| \le cK\lambda g_3(x_0, t_0, \sigma),$$

whenever $(y, s, \sigma) \in W_\lambda(x_0, t_0)$. Hence,

$$\left(\fint_{W_\lambda(x_0,t_0)} |g_1(y,s,\sigma)|^2 \, dy ds d\sigma \right) \le c\lambda^2 \int_{\lambda/8}^{2\lambda} |g_3(x_0, t_0, \sigma)|^2 \, d\sigma.$$

To estimate the right hand side in the last display, let (y, τ) be such that $|y - x_0| \le 4\sigma$, $|\tau - t_0| \le (4K\sigma)^2$. Let $Q \subset \mathbb{R}^{n+1}$ be a parabolic cube centered at (x_0, t_0) and with side length $16K\sigma$. Then, for K large enough we see that

$$
\begin{aligned}
|\lambda \partial_\tau (\mathcal{S}_\sigma f)(y,\tau)| &\le \lambda |\partial_\tau \mathcal{S}_\sigma (f 1_{2Q})(y,\tau)| \\
&\quad + \lambda |\partial_\tau \mathcal{S}_\sigma (f 1_{\mathbb{R}^{n+1} \setminus 2Q})(y,\tau) - \partial_\tau \mathcal{S}_\sigma (f 1_{\mathbb{R}^{n+1} \setminus 2Q})(x_0, t_0)| \\
&\quad + \lambda |\partial_\tau \mathcal{S}_\sigma (f 1_{\mathbb{R}^{n+1} \setminus 2Q})(x_0, t_0)|.
\end{aligned}
$$

Basically repeating the proof of (4.1) we see that

$$\left(\fint_{W_\lambda(x_0,t_0)} |g_1(y,s,\sigma)|^2 \, dy ds d\sigma \right)^{1/2} \le cM(f)(x_0, t_0).$$

To estimate $g_2(y, s, \sigma)$, whenever $(y, s, \sigma) \in W_\lambda(x_0, t_0)$, we introduce the function

$$g_4(\bar{y}, \bar{s}, \sigma) := \lim_{\epsilon \to 0} \int_{(K\sigma)^2 \le |t-\bar{s}| < 1/\epsilon} \frac{\mathrm{sgn}(\bar{s}-t)}{|\bar{s}-t|^{3/2}} (\mathcal{S}_{\delta/4} f)(\bar{y}, t) \, dt.$$

Now

$$
\begin{aligned}
|g_2(y,s,\sigma) - g_4(x_0, t_0, \sigma)| &\le |g_2(y,s,\sigma) - g_2(x_0, s, \sigma)| \\
&\quad + |g_2(x_0, s, \sigma) - g_2(x_0, t_0, \sigma)| \\
&\quad + |g_2(x_0, t_0, \sigma) - g_4(x_0, t_0, \sigma)|.
\end{aligned}
$$

In particular,

$$
\begin{aligned}
|g_2(y, s, \sigma) - g_4(x_0, t_0, \sigma)| \;\leq\; & \int_{(K\sigma)^2 \leq |s-t|} \frac{|\mathcal{S}_\sigma f(y, t) - \mathcal{S}_\sigma f(x_0, t)|}{|t - s|^{3/2}} \, dt \\
& + \int_{(K\sigma)^2 \leq |\xi|} \frac{|\mathcal{S}_\sigma f(x_0, \xi + s) - \mathcal{S}_\sigma f(x_0, \xi + t_0)|}{|\xi|^{3/2}} \, d\xi \\
& + \int_{(K\sigma)^2 \leq |t-t_0|} \frac{|\mathcal{S}_\sigma f(x_0, t) - \mathcal{S}_{\delta/4} f(x_0, t)|}{|t_0 - t|^{3/2}} \, dt \\
=: \; & h_1(y, s, \sigma) + h_2(y, s, \sigma) + h_3(x_0, t_0, \sigma).
\end{aligned}
$$

We note that

$$
\begin{aligned}
h_2(y, s, \sigma) &\leq c\sigma^2 \int_{(K\sigma)^2 \leq |\xi|} \frac{N_*(\partial_t \mathcal{S}_\sigma f)(x_0, \xi + t_0)}{|\xi|^{3/2}} \, d\xi \\
&\leq c\sigma \int_{(K\sigma)^2 \leq |\xi|} \frac{M(f)(x_0, \xi + t_0)}{|\xi|^{3/2}} \, d\xi \leq c M^t(M(f)(x_0, \cdot))(t_0),
\end{aligned}
$$

where M^t is the Hardy–Littlewood maximal operator in the t-variable, as we see by arguing as above. Similarly,

$$
h_3(y, s, \sigma) \leq c M^t (N_*(\partial_\lambda \mathcal{S}_\sigma f)(x_0, \cdot))(t_0).
$$

We therefore focus on $h_1(y, s, \sigma)$. Let

$$
\tilde{h}_1(y, \sigma) := \int_{\lambda^2 \leq |t-t_0|} \frac{|\mathcal{S}_\sigma f(y, t) - \mathcal{S}_\sigma f(x_0, t)|}{|t - t_0|^{3/2}} \, dt.
$$

If K is large enough, then $h_1(y, s, \sigma) \leq c\tilde{h}_1(y, \sigma)$, whenever $(y, s, \sigma) \in W_\lambda(x_0, t_0)$. Hence we only have to estimate

$$
\left(\fint_{\hat{Q}_\lambda(x_0) \times I_{\lambda/2}(\lambda)} \tilde{h}_1^2 \, dy d\sigma \right)^{1/2} = \sup \left| \fint_{\hat{Q}_\lambda(x_0) \times I_{\lambda/2}(\lambda)} \tilde{h}_1 g \, dy d\sigma \right|,
$$

where $\hat{Q}_\lambda(x_0) \subset \mathbb{R}^n$ now is a (non-parabolic) cube with side length λ and center x_0, $I_{\lambda/2}(\lambda)$ is the interval $(\lambda/2, 3\lambda/2)$, and where the sup is taken with respect to all $g \in C_0^\infty(\mathbb{R}^{n+1}, \mathbb{R})$ such that

$$
\left(\fint_{\hat{Q}_\lambda(x_0) \times I_{\lambda/2}(\lambda)} g^2 \, dy d\sigma \right)^{1/2} = 1. \tag{4.2}
$$

Given g as in (4.2) we let

$$
E := \fint_{\hat{Q}_\lambda(x_0) \times I_{\lambda/2}(\lambda)} \tilde{h}_1 g \, dy d\sigma.
$$

Then

$$
\begin{aligned}
E &= \fint_{\hat{Q}_\lambda(x_0) \times I_{\lambda/2}(\lambda)} \left(\int_{\lambda^2 \leq |t-t_0|} \frac{|\mathcal{S}_\sigma f(y, t) - \mathcal{S}_\sigma f(x_0, t)|}{|t - t_0|^{3/2}} \, dt \right) g(y, \sigma) \, dy d\sigma \\
&\leq c \sum_{j=0}^{\infty} (\lambda^2 2^j)^{-3/2} \fint_{\hat{Q}_\lambda(x_0) \times I_{\lambda/2}(\lambda)} \left(\int_{I_j} |\mathcal{S}_\sigma f(y, t) - \mathcal{S}_\sigma f(x_0, t)| \, dt \right) g(y, \sigma) \, dy d\sigma,
\end{aligned}
$$

where $I_j = \{t : \lambda^2 2^j \leq |t - t_0| < \lambda^2 2^{j+1}\}$. Let $\eta \in (-\lambda^2/100, \lambda^2/100)$ be a degree of freedom. Given any integer $i \in \{2^{j-1}, \ldots, 2^{j+3}\}$ we let $t_{j,i}^{\pm} = t_0 \pm i\lambda^2$, $N_j = (2^{j+3} - 2^{j-1} + 1)$. Given η we let $I_{j,i}(t_{j,i}^{\pm} + \eta, \lambda^2)$ be the interval centered at $t_{j,i}^{\pm} + \eta$ and of length $2\lambda^2$. Then $\{I_{j,i}(t_{j,i}^{\pm} + \eta, \lambda^2)\}_i$ is, for each $\eta \in (-\lambda^2/100, \lambda^2/100)$, a covering of I_j and $\{I_{j,i}(t_{j,i} + \eta, \lambda^2/10^4)\}$ is a disjoint collection. Using this we see that $|E|$ can be bounded from above by

$$c\lambda^2 \sum_{j=0}^{\infty} (\lambda^2 2^j)^{-3/2} \sum_{i=1}^{N_j} \fint_{W_\lambda(x_0, t_{j,i}^{\pm} + \eta)} |\mathcal{S}_\sigma f(y, t) - \mathcal{S}_\sigma f(x_0, t)||g(y, \sigma)| \, dy dt d\sigma$$

$$\leq c\lambda^3 \sum_{j=0}^{\infty} (\lambda^2 2^j)^{-3/2} \sum_{i=1}^{N_j} \tilde{N}_{**}(\nabla_{||}\mathcal{S}_\lambda f)(x_0, t_{j,i}^{\pm} + \eta).$$

This estimate holds uniformly with respect to $\eta \in (-\lambda^2/100, \lambda^2/100)$. In particular, taking the average with respect to η we see that

$$|E| \leq c\lambda \sum_{j=0}^{\infty} (\lambda^2 2^j)^{-3/2} \int_{\{t: \lambda^2 2^{j-2} \leq |t-t_0| < \lambda^2 2^{j+4}\}} \tilde{N}_{**}(\nabla_{||}\mathcal{S}_\lambda f)(x_0, t) \, dt$$

$$\leq cM^t(\tilde{N}_{**}(\nabla_{||}\mathcal{S}_\lambda f)(x_0, \cdot))(t_0).$$

Putting the estimates together we can conclude, for $\lambda > 4\delta/3$, that

$$\left(\fint_{W_\lambda(x_0, t_0)} |H_t D_{1/2}^t \mathcal{S}_\sigma f(y, s)|^2 \, dy ds d\sigma\right)^{1/2}$$

is bounded by

$$cM^t(\tilde{N}_{**}(\nabla_{||}\mathcal{S}_\lambda f)(x_0, \cdot))(t_0) + cM^t(M(f)(x_0, \cdot))(t_0) + cM^t(N_*(\partial_\lambda \mathcal{S}_\sigma f)(x_0, \cdot))(t_0)$$

$$+ \left(\fint_{W_\lambda(x_0, t_0)} |g_4(x_0, t_0, \sigma)|^2 \, dy ds d\sigma\right)^{1/2},$$

where M^t is the Hardy–Littlewood maximal operator in the t-variable and M is the parabolic Hardy–Littlewood maximal function. Hence, letting

$$\psi(x_0, t_0) := \sup_{\sigma > 0} |g_4(x_0, t_0, \sigma)|$$

we see that

$$||\tilde{N}_*(1_{\lambda > 2\delta} H_t D_{1/2}^t \mathcal{S}_\lambda f)||_2 \leq c||f||_2 + c\left(||\tilde{N}_{**}(\nabla_{||}\mathcal{S}_\lambda f)||_2 + ||N_{**}(\partial_\lambda \mathcal{S}_\lambda f)||_2\right) + c||\psi||_2$$

where the constant c is independent of δ. Hence, to complete the proof of (iii) it remains to estimate $||\psi||_2$. To do this we first recall that $f \in C_0^\infty(\mathbb{R}^{n+1}, \mathbb{C})$. Hence, using Lemma 3.5 we know that $\mathcal{S}_{\delta/4} f \in \mathbb{H}(\mathbb{R}^{n+1}, \mathbb{C}) \cap L^2(\mathbb{R}^{n+1}, \mathbb{C})$. Using this it follows that

$$\mathcal{S}_{\delta/4} f(x, t) = cI_{1/2}^t(D_{1/2}^t \mathcal{S}_{\delta/4} f)(x, t) = cI_{1/2}^t h(x, t),$$

where $I_{1/2}^t$ is the (fractional) Riesz operator in t defined on the Fourier transform side through the multiplier $|\tau|^{-1/2}$ and $h(x, t) := (D_{1/2}^t \mathcal{S}_{\delta/4} f)(x, t)$. Using this we see that

$$\psi(x_0, t_0) = c \sup_{\epsilon > 0} |\tilde{V}_\epsilon h(x_0, t_0)| =: c\tilde{V}_* h(x_0, t_0),$$

where V_ϵ is defined on functions $k \in L^2(\mathbb{R}, \mathbb{R})$ by

$$V_\epsilon k(t) = \int_{\{|s-t|>\epsilon\}} \frac{\text{sgn}(t-s)I_{1/2}^t k(s)}{|s-t|^{3/2}} \, ds,$$

and $\tilde{V}_\epsilon h(x,t) = V_\epsilon h(x, \cdot)$ evaluated at t. However, using this notation we can apply Lemma 2.27 in [21] and conclude that

$$||\psi||_2 \le c||h||_2 = c||D_{1/2}^t \mathcal{S}_{\delta/4} f||_2 \le c \sup_{\lambda>0} ||H_t D_{1/2}^t \mathcal{S}_\lambda f||_2.$$

This completes the proof of Lemma 4.1 (iii). □

4.2 Proof of Lemma 4.2

We first note, using Lemmas 2.1, 2.3 and induction, that it suffices to prove

(i') $|||\lambda \nabla \partial_\lambda \mathcal{S}_\lambda f||| \le c\Phi(f) + c||f||_2,$

(ii') $|||\lambda \partial_t \mathcal{S}_\lambda f||| \le c\Phi(f) + c||f||_2,$

whenever $f \in L^2(\mathbb{R}^{n+1}, \mathbb{C})$. To prove (i') it suffices to estimate $|||\lambda \nabla_{||} \partial_\lambda \mathcal{S}_\lambda f|||$. Given $\epsilon > 0$ we let

$$A_1 := -\frac{1}{2} \int_\epsilon^{1/\epsilon} \int_{\mathbb{R}^{n+1}} \nabla_{||} \partial_\lambda^2 \mathcal{S}_\lambda f \cdot \overline{\nabla_{||} \partial_\lambda \mathcal{S}_\lambda f} \, \lambda^2 dx dt d\lambda,$$

$$A_2 := -\frac{1}{2} \int_\epsilon^{1/\epsilon} \int_{\mathbb{R}^{n+1}} \nabla_{||} \partial_\lambda \mathcal{S}_\lambda f \cdot \overline{\nabla_{||} \partial_\lambda^2 \mathcal{S}_\lambda f} \, \lambda^2 dx dt d\lambda,$$

$$A_3 := \int_{\mathbb{R}^{n+1}} \nabla_{||} \partial_\lambda \mathcal{S}_\lambda f \cdot \overline{\nabla_{||} \partial_\lambda \mathcal{S}_\lambda f} \, \lambda^2 dx dt \bigg|_{\lambda=1/\epsilon},$$

$$A_4 := \int_{\mathbb{R}^{n+1}} \nabla_{||} \partial_\lambda \mathcal{S}_\lambda f \cdot \overline{\nabla_{||} \partial_\lambda \mathcal{S}_\lambda f} \, \lambda^2 dx dt \bigg|_{\lambda=\epsilon}.$$

Using partial integration with respect to λ,

$$\int_\epsilon^{1/\epsilon} \int_{\mathbb{R}^{n+1}} \nabla_{||} \partial_\lambda \mathcal{S}_\lambda f \cdot \overline{\nabla_{||} \partial_\lambda \mathcal{S}_\lambda f} \, \lambda dx dt d\lambda = A_1 + A_2 + A_3 + A_4.$$

Furthermore, using Lemma 3.4 (ii),

$$|A_1| + |A_2| + |A_3| + |A_4| \le c|||\lambda^2 \nabla_{||} \partial_\lambda^2 \mathcal{S}_\lambda f|||^2 + c||f||_2^2,$$

with c independent of ϵ. Hence

$$|||\lambda \nabla_{||} \partial_\lambda \mathcal{S}_\lambda f|||^2 = \lim_{\epsilon \to 0} \int_\epsilon^{1/\epsilon} \int_{\mathbb{R}^{n+1}} \nabla_{||} \partial_\lambda \mathcal{S}_\lambda f \cdot \overline{\nabla_{||} \partial_\lambda \mathcal{S}_\lambda f} \, \lambda dx dt d\lambda$$
$$\le c|||\lambda^2 \nabla_{||} \partial_\lambda^2 \mathcal{S}_\lambda f|||^2 + c||f||_2^2. \tag{4.3}$$

(i′) now follows from an application of Lemma 2.1. To prove (ii′) we first introduce, for $\epsilon > 0$,

$$B_1 := -\frac{1}{2} \int_\epsilon^{1/\epsilon} \int_{\mathbb{R}^{n+1}} \partial_t \partial_\lambda \mathcal{S}_\lambda f \overline{\partial_t \mathcal{S}_\lambda f} \, \lambda^2 dx dt d\lambda,$$

$$B_2 := -\frac{1}{2} \int_\epsilon^{1/\epsilon} \int_{\mathbb{R}^{n+1}} \partial_t \mathcal{S}_\lambda f \overline{\partial_t \partial_\lambda \mathcal{S}_\lambda f} \, \lambda^2 dx dt d\lambda,$$

$$B_3 := \int_{\mathbb{R}^{n+1}} \partial_t \mathcal{S}_\lambda f \overline{\partial_t \mathcal{S}_\lambda f} \, \lambda^2 dx dt \Big|_{\lambda=1/\epsilon},$$

$$B_4 := -\int_{\mathbb{R}^{n+1}} \partial_t \mathcal{S}_\lambda f \overline{\partial_t \mathcal{S}_\lambda f} \, \lambda^2 dx dt \Big|_{\lambda=\epsilon}.$$

Then, using Lemma 3.4 (iii)

$$|B_1| + |B_2| + |B_3| + |B_4| \le c|||\lambda^2 \partial_t \partial_\lambda \mathcal{S}_\lambda f|||^2 + c||f||_2^2,$$

with c independent of ϵ. Hence, again by integration by parts with respect to λ,

$$|||\lambda \partial_t \mathcal{S}_\lambda f|||^2 = \lim_{\epsilon \to 0} \int_\epsilon^{1/\epsilon} \int_{\mathbb{R}^{n+1}} \partial_t \mathcal{S}_\lambda f \overline{\partial_t \mathcal{S}_\lambda f} \, \lambda dx dt d\lambda$$
$$\le c|||\lambda^2 \partial_t \partial_\lambda \mathcal{S}_\lambda f|||^2 + c||f||_2^2. \tag{4.4}$$

Furthermore, repeating the above argument it also follows that

$$|||\lambda^2 \partial_t \partial_\lambda \mathcal{S}_\lambda f|||^2 \le c|||\lambda^3 \partial_t \partial_\lambda^2 \mathcal{S}_\lambda f|||^2 + c||f||_2^2.$$

Finally, using Lemma 2.3 we can combine the above estimates and conclude that

$$|||\lambda \partial_t \mathcal{S}_\lambda f||| \le c\Phi(f) + c||f||_2.$$

This completes the proof of (ii′) and hence the proof of Lemma 4.2.

5 Resolvents, square functions and Carleson measures

In the following we collect some of the main results from [32] to be used in the proof of our main results. Throughout the section we assume that $\mathcal{H}, \mathcal{H}^*$ satisfy (1.2) and (1.3). We let

$$\mathcal{L}_{||} := -\operatorname{div}_{||} A_{||} \nabla_{||},$$

where $\operatorname{div}_{||}$ is the divergence operator in the variables $(\partial_{x_1}, \dots, \partial_{x_n})$. $A_{||}$ is the $n \times n$-dimensional sub matrix of A defined by $\{A_{i,j}\}_{i,j=1}^n$. We also let

$$\mathcal{H}_{||} := \partial_t + \mathcal{L}_{||}, \quad \mathcal{H}_{||}^* := -\partial_t + \mathcal{L}_{||}^*.$$

Using this notation the equation $\mathcal{H}u = 0$ can be written, formally, as

$$\mathcal{H}_{||}u - \sum_{j=1}^{n+1} A_{n+1,j} D_{n+1} D_j u - \sum_{i=1}^{n} D_i(A_{i,n+1} D_{n+1} u) = 0. \tag{5.1}$$

In the proof of Lemma 6.1 below we will use that (5.1) holds in an appropriate weak sense on cross sections $\lambda = $ constant. Indeed, let $\lambda \in (a, b)$ and let $\epsilon < \min(\lambda - a, b - \lambda)$. Set

$\varphi_\epsilon(\sigma) = \epsilon^{-1}\varphi(\sigma/\epsilon)$ where $\varphi \in C_0^\infty(-1/2, 1/2), 0 \leq \varphi, \int \varphi\, d\sigma = 1$. We let $\phi_{\lambda,\epsilon}(x, t, \sigma) = \psi(x, t)\varphi_\epsilon(\sigma)$ where $\psi \in C_0^\infty(\mathbb{R}^{n+1}, \mathbb{C})$. Then, by the notion of weak solutions we have

$$\int_{\mathbb{R}^{n+2}} \left(A_{||}(x)\nabla_{||}u(x, t, \sigma) \cdot \nabla_{||}\overline{\phi_{\lambda,\epsilon}(x, t, \sigma)} - u(x, t, \sigma)\partial_t\overline{\phi_{\lambda,\epsilon}(x, t, \sigma)} \right) dxdtd\sigma$$

$$= \sum_{j=1}^{n+1} \int_{\mathbb{R}^{n+2}} A_{n+1,j}(x)\partial_{x_j}\partial_\lambda u(x, t, \sigma)\overline{\phi_{\lambda,\epsilon}(x, t, \sigma)}\, dxdtd\sigma$$

$$- \sum_{i=1}^{n} \int_{\mathbb{R}^{n+2}} A_{i,n+1}(x)\partial_\lambda u(x, t, \sigma)\partial_{x_i}\overline{\phi_{\lambda,\epsilon}(x, t, \sigma)}\, dxdtd\sigma. \tag{5.2}$$

Hence, if

$$\nabla u, \ \nabla\partial_\lambda u \in L^2(\mathbb{R}^{n+1}, \mathbb{C}^{n+1}), \tag{5.3}$$

uniformly in $\lambda \in (a, b)$, with norms depending continuously on $\lambda \in (a, b)$, then we can conclude, by letting $\epsilon \to 0$ in (5.2), that

$$\int_{\mathbb{R}^{n+1}} \left(A_{||}(x)\nabla_{||}u(x, t, \lambda) \cdot \nabla_{||}\overline{\psi(x, t)} - u(x, t, \lambda)\partial_t\overline{\psi(x, t)} \right) dxdt$$

$$= \sum_{j=1}^{n+1} \int_{\mathbb{R}^{n+1}} A_{n+1,j}(x)\partial_{x_j}\partial_\lambda u(x, t, \lambda)\overline{\psi(x, t)}\, dxdt$$

$$- \sum_{i=1}^{n} \int_{\mathbb{R}^{n+1}} A_{i,n+1}(x)\partial_\lambda u(x, t, \lambda)\partial_{x_i}\overline{\psi(x, t)}\, dxdt. \tag{5.4}$$

In this sense, and under these assumptions, (5.1) holds on cross sections $\lambda = $ constant.

5.1 Resolvents and a parabolic Hodge decomposition associated to $\mathcal{H}_{||}$

Recall the function space $\mathbb{H} = \mathbb{H}(\mathbb{R}^{n+1}, \mathbb{C})$ introduced in (2.1). In the following we will consider, to ensure a Hilbertian structure, that this space is equipped with the equivalent semi norm stated on the right hand side in (2.2) (i). We let $\mathbb{H}^* = \mathbb{H}^*(\mathbb{R}^{n+1}, \mathbb{C})$ be the space dual to $\mathbb{H}$, with norm $||\cdot||_{\mathbb{H}^*}$, and we let $\langle\cdot, \cdot\rangle_{\mathbb{H}^*} : \mathbb{H}^* \times \mathbb{H} \to \mathbb{C}$ denote the duality pairing. We let $\bar{\mathbb{H}} = \bar{\mathbb{H}}(\mathbb{R}^{n+1}, \mathbb{C})$ be the closure of $C_0^\infty(\mathbb{R}^{n+1}, \mathbb{C})$ with respect to the norm

$$\|f\|_{\bar{\mathbb{H}}} := \|f\|_{\mathbb{H}} + \|f\|_2.$$

We let $\bar{\mathbb{H}}^* = \bar{\mathbb{H}}^*(\mathbb{R}^{n+1}, \mathbb{C})$ be the space dual to $\bar{\mathbb{H}}$, with norm $||\cdot||_{\bar{\mathbb{H}}^*}$, and we let $\langle\cdot, \cdot\rangle_{\bar{\mathbb{H}}^*} : \bar{\mathbb{H}}^* \times \bar{\mathbb{H}} \to \mathbb{C}$ denote the duality pairing. Let $B : \mathbb{H} \times \mathbb{H} \to \mathbb{R}$ be defined as

$$B(u, \phi) := \int_{\mathbb{R}^{n+1}} (A_{||}\nabla_{||}u \cdot \nabla_{||}\bar{\phi} - D_{1/2}^t u\overline{H_t D_{1/2}^t \phi})\, dxdt, \tag{5.5}$$

and let, for $\delta \in (0, 1)$, $B_\delta : \mathbb{H} \times \mathbb{H} \to \mathbb{R}$ be defined as

$$B_\delta(u, \phi) := \int_{\mathbb{R}^{n+1}} A_{||}\nabla_{||}u \cdot \overline{\nabla_{||}(I + \delta H_t)\phi}\, dxdt$$

$$- \int_{\mathbb{R}^{n+1}} D_{1/2}^t u\overline{H_t D_{1/2}^t(I + \delta H_t)\phi}\, dxdt. \tag{5.6}$$

Definition 5.1 Let $F \in \mathbb{H}^*(\mathbb{R}^{n+1}, \mathbb{C})$. We say that a function $u \in \mathbb{H}(\mathbb{R}^{n+1}, \mathbb{C})$ is a (weak) solution to the equation $\mathcal{H}_{||} u = F$, in $\mathbb{R}^{n+1}$, if

$$B(u, \phi) = \langle F, \phi \rangle_{\mathbb{H}^*},$$

whenever $\phi \in \mathbb{H}(\mathbb{R}^{n+1}, \mathbb{C})$.

Definition 5.2 Let $\lambda > 0$ be given. Let $F \in \bar{\mathbb{H}}^*(\mathbb{R}^{n+1}, \mathbb{C})$. We say that a function $u \in \bar{\mathbb{H}}(\mathbb{R}^{n+1}, \mathbb{C})$ is a (weak) solution to the equation $u + \lambda^2 \mathcal{H}_{||} u = F$, in $\mathbb{R}^{n+1}$, if

$$\int_{\mathbb{R}^{n+1}} u \bar{\phi} \, dx dt + \lambda^2 B(u, \phi) = \langle F, \phi \rangle_{\bar{\mathbb{H}}^*},$$

whenever $\phi \in \bar{\mathbb{H}}(\mathbb{R}^{n+1}, \mathbb{C})$.

Lemma 5.3 *Consider the operator $\mathcal{H}_{||} = \partial_t - \mathrm{div}_{||} A_{||} \nabla_{||}$ and assume that A satisfies (1.2), (1.3). Let $F \in \mathbb{H}^*(\mathbb{R}^{n+1}, \mathbb{C})$. Then there exists a weak solution to the equation $\mathcal{H}_{||} u = F$, in $\mathbb{R}^{n+1}$, in the sense of Definition 5.1. Furthermore,*

$$||u||_{\mathbb{H}} \leq c ||F||_{\mathbb{H}^*},$$

for some constant c depending only on n and Λ. The solution is unique up to a constant.

Proof This is essentially Lemma 2.6 in [32]. Let $\phi_\delta := (I + \delta H_t)\phi$, $\phi \in \mathbb{H}(\mathbb{R}^{n+1}, \mathbb{C})$, $\delta \in (0, 1)$. Then

$$|\langle F, \phi_\delta \rangle_{\mathbb{H}^*}| \leq c ||F||_{\mathbb{H}^*} ||\phi||_{\mathbb{H}}.$$

Consider the sesquilinear form $B_\delta(\cdot, \cdot)$ introduced in (5.6). If $\delta = \delta(n, \Lambda)$ is small enough, then $B_\delta(\cdot, \cdot)$ is a sesquilinear, bounded, coercive form on $\mathbb{H} \times \mathbb{H}$. Hence, using the Lax–Milgram theorem we see that there exists a unique $u \in \mathbb{H}$ such that

$$B(u, \phi_\delta) = B_\delta(u, \phi) = \langle F, \phi_\delta \rangle_{\mathbb{H}^*},$$

for all $\phi \in \mathbb{H}$. Using that $(I + \delta H_t)$ is invertible on $\mathbb{H}$, if $0 < \delta \ll 1$ is small enough, we can conclude that

$$B(u, \psi) = \langle F, \psi \rangle_{\mathbb{H}^*},$$

whenever $\psi \in \mathbb{H}$. The bound $||u||_{\mathbb{H}} \leq c ||F||_{\mathbb{H}^*}$ follows readily. This completes the existence and quantitative part of the lemma. The statement concerning uniqueness follows immediately. $\square$

Lemma 5.4 *Let $\lambda > 0$ be given. Consider the operator $\mathcal{H}_{||} = \partial_t - \mathrm{div}_{||} A_{||} \nabla_{||}$ and assume that A satisfies (1.2), (1.3). Let $F \in \bar{\mathbb{H}}^*(\mathbb{R}^{n+1}, \mathbb{C})$. Then there exists a weak solution to the equation $u + \lambda^2 \mathcal{H}_{||} u = F$, in $\mathbb{R}^{n+1}$, in the sense of Definition 5.2. Furthermore,*

$$||u||_2 + ||\lambda \nabla_{||} u||_2 + ||\lambda D^t_{1/2} u||_2 \leq c ||F||_{\bar{\mathbb{H}}^*},$$

for some constant c depending only on n and Λ. The solution is unique.

Proof See the proof of Lemma 2.7 in [32]. $\square$

Remark 5.5 Definitions 5.1, 5.2, Lemmas 5.3, and 5.4, all have analogous formulations for the operator $\mathcal{H}^*_{||}$.

Remark 5.6 Let $\lambda > 0$ be given. Consider the operator $\mathcal{H}_{||} = \partial_t - \text{div}_{||} A_{||} \nabla_{||}$. Let $F \in \bar{\mathbb{H}}^*(\mathbb{R}^{n+1}, \mathbb{C})$. By Lemma 5.4 the equation $u + \lambda^2 \mathcal{H}_{||} u = F$ has a unique weak solution $u \in \bar{\mathbb{H}}$. From now on we will denote this solution by $\mathcal{E}_\lambda F$. In the case of the operator $\mathcal{H}_{||}^*$ we denote the corresponding solution by $\mathcal{E}_\lambda^* F$. In this sense $\mathcal{E}_\lambda = (I + \lambda^2 \mathcal{H}_{||})^{-1}$ and $\mathcal{E}_\lambda^* = (I + \lambda^2 \mathcal{H}_{||}^*)^{-1}$.

Consider $\lambda > 0$ fixed, let $|h| \ll \lambda$ and consider $F \in \bar{\mathbb{H}}^*(\mathbb{R}^{n+1}, \mathbb{C})$. By definition,

$$\int_{\mathbb{R}^{n+1}} \mathcal{E}_{\lambda+h} F \bar{\phi} \, dx dt + (\lambda + h)^2 B(\mathcal{E}_{\lambda+h} F, \phi) = \langle F, \phi \rangle_{\bar{\mathbb{H}}^*},$$

$$\int_{\mathbb{R}^{n+1}} \mathcal{E}_\lambda F \bar{\phi} \, dx dt + \lambda^2 B(\mathcal{E}_\lambda F, \phi) = \langle F, \phi \rangle_{\bar{\mathbb{H}}^*}, \tag{5.7}$$

for all $\phi \in \bar{\mathbb{H}}(\mathbb{R}^{n+1}, \mathbb{C})$. We let $\mathcal{D}_\lambda^h F := \mathcal{E}_{\lambda+h} F - \mathcal{E}_\lambda F$. (5.7) implies

$$\int_{\mathbb{R}^{n+1}} \mathcal{D}_\lambda^h F \bar{\phi}_\delta \, dx dt + \lambda^2 B(\mathcal{D}_\lambda^h F, \phi_\delta) = -h(2\lambda + h) B(\mathcal{E}_{\lambda+h} F, \phi_\delta) \tag{5.8}$$

for all $\phi \in \bar{\mathbb{H}}(\mathbb{R}^{n+1}, \mathbb{C})$, $\phi_\delta := (I + \delta H_t)\phi$. Again, arguing as in the proof of Lemma 5.4 we see, if $\delta = \delta(n, \Lambda)$, $0 < \delta \ll 1$ is small enough and as $\mathcal{D}_\lambda^h F \in \bar{\mathbb{H}}(\mathbb{R}^{n+1}, \mathbb{C})$, that

$$||\mathcal{D}_\lambda^h F||_2 + ||\lambda \nabla_{||} \mathcal{D}_\lambda^h F||_2 + ||\lambda D_{1/2}^t \mathcal{D}_\lambda^h F||_2 \leq c|h| ||\mathcal{E}_{\lambda+h} F||_2 \leq c|h| ||F||_{\bar{\mathbb{H}}^*}, \tag{5.9}$$

where c is independent of h. Hence

$$\lim_{h \to 0} \mathcal{D}_\lambda^h F = \lim_{h \to 0} (\mathcal{E}_{\lambda+h} F - \mathcal{E}_\lambda F) = 0 \tag{5.10}$$

in the sense that

$$||\mathcal{D}_\lambda^h F||_2 + ||\lambda \nabla_{||} \mathcal{D}_\lambda^h F||_2 + ||\lambda D_{1/2}^t \mathcal{D}_\lambda^h F||_2 \to 0 \quad \text{as} \quad h \to 0. \tag{5.11}$$

Similarly,

$$\int_{\mathbb{R}^{n+1}} h^{-1} \mathcal{D}_\lambda^h F \bar{\phi}_\delta \, dx dt + \lambda^2 B(h^{-1} \mathcal{D}_\lambda^h F, \phi_\delta) = -(2\lambda + h) B(\mathcal{E}_{\lambda+h} F, \phi_\delta) \tag{5.12}$$

and hence

$$||h^{-1} \mathcal{D}_\lambda^h F||_2 + ||\lambda \nabla_{||} (h^{-1} \mathcal{D}_\lambda^h F)||_2 + ||\lambda D_{1/2}^t (h^{-1} \mathcal{D}_\lambda^h F)||_2 \leq c||F||_{\bar{\mathbb{H}}^*}, \tag{5.13}$$

where c is independent of h. Using (5.13), (5.12) and (5.11) we see, as λ is fixed, that

$$\lim_{h \to 0} h^{-1} \mathcal{D}_\lambda^h F =: \mathcal{G}_\lambda F \quad \text{weakly in} \quad \bar{\mathbb{H}}(\mathbb{R}^{n+1}, \mathbb{C}), \tag{5.14}$$

that (5.13) holds with $h^{-1} \mathcal{D}_\lambda^h F$ replaced by $\mathcal{G}_\lambda F$ and that

$$\int_{\mathbb{R}^{n+1}} \mathcal{G}_\lambda F \bar{\phi} \, dx dt + \lambda^2 B(\mathcal{G}_\lambda F, \phi) = -2\lambda B(\mathcal{E}_\lambda F, \phi) = -2\lambda \langle \mathcal{H}_{||} \mathcal{E}_\lambda F, \phi \rangle_{\bar{\mathbb{H}}^*} \tag{5.15}$$

whenever $\phi \in \bar{\mathbb{H}}(\mathbb{R}^{n+1}, \mathbb{C})$. We define

$$\partial_\lambda \mathcal{E}_\lambda F := \mathcal{G}_\lambda F \tag{5.16}$$

and hence

$$\partial_\lambda \mathcal{E}_\lambda F = -2\lambda \mathcal{E}_\lambda \mathcal{H}_{||} \mathcal{E}_\lambda F \tag{5.17}$$

in the sense of (5.15). Furthermore, if $F = f \in \mathbb{H}(\mathbb{R}^{n+1}, \mathbb{C})$ then

$$\langle \mathcal{H}_{||} \mathcal{E}_\lambda f, \phi \rangle_{\tilde{\mathbb{H}}^*} - \langle \mathcal{E}_\lambda \mathcal{H}_{||} f, \phi \rangle_{\tilde{\mathbb{H}}^*} = \langle \mathcal{H}_{||} \mathcal{E}_\lambda f, \phi \rangle_{\tilde{\mathbb{H}}^*} - \langle \mathcal{H}_{||} f, \mathcal{E}_\lambda^* \phi \rangle_{\tilde{\mathbb{H}}^*}$$
$$= B(\mathcal{E}_\lambda f, \phi) - B(f, \mathcal{E}_\lambda^* \phi) = 0, \tag{5.18}$$

and hence $\mathcal{H}_{||}$ and $\mathcal{E}_\lambda$ commute in this sense. Furthermore, as A is independent of t we can, by arguing similarly, conclude that if $f \in \mathbb{H}(\mathbb{R}^{n+1}, \mathbb{C})$, then

$$\langle \partial_t \mathcal{E}_\lambda f, \phi \rangle_{\tilde{\mathbb{H}}^*} - \langle \mathcal{E}_\lambda \partial_t f, \phi \rangle_{\tilde{\mathbb{H}}^*} = 0 = \langle \mathcal{L}_{||} \mathcal{E}_\lambda f, \phi \rangle_{\tilde{\mathbb{H}}^*} - \langle \mathcal{E}_\lambda \mathcal{L}_{||} f, \phi \rangle_{\tilde{\mathbb{H}}^*} \tag{5.19}$$

and hence ∂_t and $\mathcal{E}_\lambda$, and $\mathcal{L}_{||}$ and $\mathcal{E}_\lambda$, commute in this sense. In particular, if $F = f \in \mathbb{H}(\mathbb{R}^{n+1}, \mathbb{C})$ then

$$\partial_\lambda \mathcal{E}_\lambda f = -2\lambda \mathcal{E}_\lambda^2 \mathcal{H}_{||} f \tag{5.20}$$

in the sense of (5.15).

5.2 Estimates of resolvents

We here collect a set of the estimates for $\mathcal{E}_\lambda f$ and $\mathcal{E}_\lambda^* f$ to be used in the next section.

Lemma 5.7 *Let $\lambda > 0$ be given. Consider the operator $\mathcal{H}_{||} = \partial_t - \mathrm{div}_{||} A_{||} \nabla_{||}$ and assume that A satisfies (1.2), (1.3). Let Θ_λ denote any of the operators*

$$\mathcal{E}_\lambda, \lambda \nabla_{||} \mathcal{E}_\lambda, \lambda D_{1/2}^t \mathcal{E}_\lambda,$$

or

$$\lambda \mathcal{E}_\lambda D_{1/2}^t, \lambda^2 \nabla_{||} \mathcal{E}_\lambda D_{1/2}^t, \lambda^2 D_{1/2}^t \mathcal{E}_\lambda D_{1/2}^t,$$

and let $\tilde{\Theta}_\lambda$ denote any of the operators

$$\lambda \mathcal{E}_\lambda \, \mathrm{div}_{||}, \lambda^2 \nabla_{||} \mathcal{E}_\lambda \, \mathrm{div}_{||}, \lambda^2 D_{1/2}^t \mathcal{E}_\lambda \, \mathrm{div}_{||}.$$

Then there exist c, depending only on n, Λ, such that

$$\text{(i)} \quad \int_{\mathbb{R}^{n+1}} |\Theta_\lambda f(x, t)|^2 \, dx dt \leq c \int_{\mathbb{R}^{n+1}} |f(x, t)|^2 \, dx dt,$$

$$\text{(ii)} \quad \int_{\mathbb{R}^{n+1}} |\tilde{\Theta}_\lambda \mathbf{f}(x, t)|^2 \, dx dt \leq c \int_{\mathbb{R}^{n+1}} |\mathbf{f}(x, t)|^2 \, dx dt,$$

whenever $f \in L^2(\mathbb{R}^{n+1}, \mathbb{C})$, $\mathbf{f} \in L^2(\mathbb{R}^{n+1}, \mathbb{C}^n)$.

Proof This is Lemma 2.11 in [32]. □

Lemma 5.8 *Let $\lambda > 0$ be given. Consider the operator $\mathcal{H}_{||} = \partial_t - \mathrm{div}_{||} A_{||} \nabla_{||}$ and assume that A satisfies (1.2), (1.3). Let $A_{n+1}^{||} := (A_{1,n+1}, \ldots, A_{n,n+1})$,*

$$\mathcal{U}_\lambda := \lambda \mathcal{E}_\lambda \, \mathrm{div}_{||},$$

and let

$$\mathcal{R}_\lambda := \mathcal{U}_\lambda A_{n+1}^{||} - (\mathcal{U}_\lambda A_{n+1}^{||}) \mathcal{P}_\lambda,$$

where $\mathcal{P}_\lambda$ is a parabolic approximation of the identity. Then there exists a constant c, depending only on n, Λ, such that

$$||\mathcal{R}_\lambda f||_2 \leq c(||\lambda \nabla f||_2 + ||\lambda^2 \partial_t f||_2),$$

whenever $f \in C_0^\infty(\mathbb{R}^{n+1}, \mathbb{C})$.

Proof The lemma is a consequence of Lemma 2.27 in [32]. □

Lemma 5.9 *Let $\lambda > 0$ be given. Consider the operator $\mathcal{H}_{||} = \partial_t - \mathrm{div}_{||} A_{||} \nabla_{||}$ and assume that A satisfies* (1.2), (1.3). *Let $A_{n+1}^{||} := (A_{1,n+1}, \ldots, A_{n,n+1})$,*

$$\mathcal{U}_\lambda := \lambda \mathcal{E}_\lambda \, \mathrm{div}_{||},$$

and consider $\mathcal{U}_\lambda A_{n+1}^{||}$. Then there exists a constant c, depending only on n, Λ, such that

$$\int_0^{l(Q)} \int_Q |\mathcal{U}_\lambda A_{n+1}^{||}|^2 \frac{dx \, dt \, d\lambda}{\lambda} \leq c|Q|,$$

for all cubes $Q \subset \mathbb{R}^{n+1}$.

Proof This is Lemma 3.1 in [32]. □

Remark 5.10 For the details of the proof of Lemmas 5.8 and 5.9 we refer to [32]. We here simply note that for λ fixed, $(\mathcal{U}_\lambda A_{n+1}^{||})$ (and $\mathcal{R}_\lambda 1$) exists as an element in $L^2_{\mathrm{loc}}(\mathbb{R}^{n+1}, \mathbb{C})$. Indeed, let Q_R be the parabolic cube on $\mathbb{R}^{n+1}$ with center at $(0, 0)$ and with size determined by R. Writing

$$\mathcal{U}_\lambda A_{n+1}^{||} = \mathcal{U}_\lambda A_{n+1}^{||} 1_{2Q_R} + \mathcal{U}_\lambda A_{n+1}^{||} 1_{\mathbb{R}^{n+1} \setminus 2Q_R},$$

and using Lemma 5.7 we see that

$$||\mathcal{U}_\lambda (A_{n+1}^{||} 1_{2Q_R}) 1_{Q_R}||_2 \leq c||A||_\infty R^{(n+2)/2}.$$

Furthermore, by the off-diagonal estimates for $\mathcal{U}_\lambda$ proved in Lemma 2.17 in [32] it follows that also

$$||\mathcal{U}_\lambda (A_{n+1}^{||} 1_{\mathbb{R}^{n+1} \setminus 2Q_R}) 1_{Q_R}||_2 \leq c||A||_\infty R^{(n+2)/2}.$$

Theorem 5.11 *Consider the operators $\mathcal{H}_{||} = \partial_t + \mathcal{L}_{||} = \partial_t - \mathrm{div}_{||} A_{||} \nabla_{||}$, $\mathcal{H}_{||}^* = -\partial_t + \mathcal{L}_{||}^* = -\partial_t - \mathrm{div}_{||} A_{||}^* \nabla_{||}$, and assume that A satisfies* (1.2), (1.3). *Then there exists a constant c, $1 \leq c < \infty$, depending only on n, Λ, such that*

$$|||\lambda \mathcal{E}_\lambda \mathcal{H}_{||} f||| + |||\lambda \mathcal{E}_\lambda^* \mathcal{H}_{||}^* f||| \leq c||\mathbb{D}f||_2, \tag{5.21}$$

and

$$\begin{align}
\text{(i)} \quad & |||\partial_\lambda \mathcal{E}_\lambda f||| + |||\partial_\lambda \mathcal{E}_\lambda^* f||| \leq c||\mathbb{D}f||_2, \\
\text{(ii)} \quad & |||\lambda \partial_t \mathcal{E}_\lambda f||| + |||\lambda \partial_t \mathcal{E}_\lambda^* f||| \leq c||\mathbb{D}f||_2, \\
\text{(iii)} \quad & |||\lambda \mathcal{E}_\lambda \mathcal{L}_{||} f||| + |||\lambda \mathcal{E}_\lambda^* \mathcal{L}_{||}^* f||| \leq c||\mathbb{D}f||_2, \\
\text{(iv)} \quad & |||\lambda \mathcal{L}_{||} \mathcal{E}_\lambda f||| + |||\lambda \mathcal{L}_{||}^* \mathcal{E}_\lambda^* f||| \leq c||\mathbb{D}f||_2, \tag{5.22}
\end{align}$$

whenever $f \in \mathbb{H}(\mathbb{R}^{n+1}, \mathbb{C})$.

Proof (5.24) is Theorem 1.17 in [32], (5.22) (i)–(iv) is Corollary 1.18 in [32]. However, as the proof of Corollary 1.18 in [32] is presented in a slightly formal manner we here include the proof of the inequalities in (5.22) clarifying details. We only supply the proof in the case of $\mathcal{H}_{||}$. To prove (i) we note that $\partial_\lambda \mathcal{E}_\lambda f$ is defined as in (5.16) and that we have, using (5.20), $\partial_\lambda \mathcal{E}_\lambda f = -2\lambda \mathcal{E}_\lambda^2 \mathcal{H}_{||} f$ in the sense of (5.15). Hence (i) follows from (5.24). To prove (ii) we note that ∂_t and $\mathcal{E}_\lambda$ commute in the sense discussed above, see (5.19), and that

$$\lambda \mathcal{E}_\lambda \partial_t f = \lambda \mathcal{E}_\lambda \mathcal{H}_{||} f - \lambda \mathcal{E}_\lambda \mathcal{L}_{||} f.$$

Hence, using (5.24) we see that

$$|||\lambda \partial_t \mathcal{E}_\lambda f||| \leq c||\mathbb{D}f||_2 + |||\lambda \mathcal{E}_\lambda \mathcal{L}_{||} f|||.$$

Therefore, to prove (ii) it suffices to prove (iii). To prove (iii), we let $f \in \mathbb{H}(\mathbb{R}^{n+1}, \mathbb{C})$ and put $g = A_{||} \nabla_{||} f$. Using Lemma 5.3 we then see that there exists a weak solution u to the equation

$$\text{div}_{||}(g) = \mathcal{H}_{||} u \text{ such that } ||u||_\mathbb{H} \leq c||g||_2. \tag{5.23}$$

In particular,

$$\lambda \mathcal{E}_\lambda \mathcal{L}_{||} f = \lambda \mathcal{E}_\lambda \mathcal{H}_{||} u. \tag{5.24}$$

Hence, again using Theorem 5.11 we see that

$$|||\lambda \mathcal{E}_\lambda \mathcal{L}_{||} f||| \leq c||\mathbb{D}u||_2. \tag{5.25}$$

(iii) now follows by combining (5.23) and (5.25). To prove (iv) we simply note that $\mathcal{L}$ and $\mathcal{E}_\lambda$ commute in the sense of (5.19), and hence (iv) follows from the argument in (iii). This completes the proof of (5.22) (i)–(iv). □

5.3 Remark on the Kato problem for parabolic equations

In Section 5 in [32] implications of two of the results proved in [32], Theorem 1.17 and Theorem 1.19 in [32], for Kato square root problems related to the operator $\partial_t + \mathcal{L}_{||}$ (in [32] this operator is denoted $\partial_t + \mathcal{L}$), as well as generalizations of these results to operators $\partial_t - \text{div} A(x, t)\nabla$, i.e., to operators with time-dependent coefficients, are discussed. The discussion in the section is essentially flawless but the author neglects to properly state that the Kato square root problem for the operator $\partial_t + \mathcal{L}_{||}$ is in fact solved in [32]. Indeed, the core of the approach in [32] is the observation that $\partial_t + \mathcal{L}_{||}$ can be realized as an operator $\bar{\mathbb{H}} \to \bar{\mathbb{H}}^*$ via the sesquilinear form $B(u, \psi)$ introduced in (5.5):

$$\langle (\partial_t + \mathcal{L}_{||})u, \psi \rangle := B(u, \psi), \ u, \psi \in \bar{\mathbb{H}}.$$

By the arguments in [32] it follows, see also Lemma 5.4 above, that if $\theta \in \mathbb{C}$ with $\text{Re } \theta > 0$, then

$$\theta + \partial_t + \mathcal{L}_{||} : \mathcal{D}(\partial_t + \mathcal{L}_{||}) \to L^2(\mathbb{R}^{n+1}, \mathbb{C})$$

is bijective and the resolvent satisfies the estimate

$$\|(\theta + (\partial_t + \mathcal{L}_{||}))^{-1} f\|_2 \leq \frac{1}{\text{Re } \theta} \|f\|_2.$$

In particular, $\partial_t + \mathcal{L}_{||}$, with maximal domain $\mathcal{D}(\partial_t + \mathcal{L}_{||}) = \{u \in \bar{\mathbb{H}} : (\partial_t + \mathcal{L}_{||})u \in L^2(\mathbb{R}^{n+1}, \mathbb{C})\}$ in $L^2(\mathbb{R}^{n+1}, \mathbb{C})$, is maximal accretive and, see also the discussion in Section 5 in [32], $\partial_t + \mathcal{L}_{||}$ is sectorial and there is a square root $\sqrt{\partial_t + \mathcal{L}_{||}}$ abstractly defined by functional calculus. Furthermore, $\partial_t + \mathcal{L}_{||}$ has a bounded H^∞ calculus. This is an other way of formulating the discussion in Section 5 in [32] up to display (5.4) in [32]. Furthermore, the inequality

$$||\sqrt{\partial_t + \mathcal{L}_{||}} f||_2^2 \leq c \int_0^\infty \int_{\mathbb{R}^{n+1}} |(I + \lambda^2(\partial_t + \mathcal{L}_{||}))^{-1} \lambda(\partial_t + \mathcal{L}_{||})f|^2 \frac{dx dt d\lambda}{\lambda}, \tag{5.26}$$

does hold for all $f \in C_0^\infty(\mathbb{R}^{n+1}, \mathbb{C})$. In particular, the inequality in display (5.5) in [32] is valid and this was the only point left open in [32]. Based on this we can conclude, using the

main result proved in [32], that there exists a constant c, $1 \leq c < \infty$, depending only on n, Λ, such that

$$c^{-1}||\mathbb{D}f||_2 \leq ||\sqrt{\partial_t + \mathcal{L}_{||}}f||_2 \leq c||\mathbb{D}f||_2, \tag{5.27}$$

whenever $f \in \bar{\mathbb{H}}$.

6 Estimates in parabolic Sobolev spaces

Throughout this section we assume that $\mathcal{H}$, $\mathcal{H}^*$ satisfy (1.2) and (1.3) as well as (2.6) and (2.7). Using the estimates established and stated in Sects. 4 and 5 we in this section prove the following three lemmas.

Lemma 6.1 *Let* $\Phi(f)$ *be defined as in* (1.7). *Assume that* $\Phi(f) < \infty$ *whenever* $f \in L^2(\mathbb{R}^{n+1}, \mathbb{C})$. *Then there exists a constant* c, *depending at most on* n, Λ, *and the De Giorgi–Moser–Nash constants, such that*

$$||\nabla_{||}\mathcal{S}_{\lambda_0}f||_2 \leq c(\Phi(f) + ||f||_2 + ||N_{**}(\partial_\lambda \mathcal{S}_\lambda f)||_2),$$

whenever $f \in L^2(\mathbb{R}^{n+1}, \mathbb{C})$, $\lambda_0 > 0$.

Lemma 6.2 *Let* $\Phi(f)$ *be defined as in* (1.7). *Assume that* $\Phi(f) < \infty$ *whenever* $f \in L^2(\mathbb{R}^{n+1}, \mathbb{C})$. *Then there exists a constant* c, *depending at most on* n, Λ, *and the De Giorgi–Moser–Nash constants, such that*

$$||\mathbb{D}_{n+1}\mathcal{S}_{\lambda_0}f||_2^2 \leq c(\Phi(f) + ||f||_2),$$

whenever $f \in L^2(\mathbb{R}^{n+1}, \mathbb{C})$, $\lambda_0 > 0$.

Lemma 6.3 *There exists a constant* c, *depending at most on* n, *such that*

$$||H_t D_{1/2}^t \mathcal{S}_{\lambda_0}f||_2 \leq c(||\mathbb{D}_{n+1}\mathcal{S}_{\lambda_0}f||_2 + ||\nabla_{||}\mathcal{S}_{\lambda_0}f||_2),$$

whenever $f \in L^2(\mathbb{R}^{n+1}, \mathbb{C})$, $\lambda_0 > 0$.

The proofs of Lemmas 6.1–6.3 are given below.

6.1 Proof of Lemma 6.1

Throughout the proof we can, without loss of generality, assume that $f \in C_0^\infty(\mathbb{R}^{n+1}, \mathbb{C})$. Let $\lambda_0 > 0$ be fixed. To prove the lemma it suffices to estimate

$$I := \int_{\mathbb{R}^{n+1}} \bar{\mathbf{g}} \cdot \nabla_{||}\mathcal{S}_{\lambda_0}f \, dxdt,$$

where $\mathbf{g} \in C_0^\infty(\mathbb{R}^{n+1}, \mathbb{C}^n)$ and $||\mathbf{g}||_2 = 1$. Given $f \in C_0^\infty(\mathbb{R}^{n+1}, \mathbb{C})$, we note, see Lemma 3.5, that $\mathcal{S}_{\lambda_0}f \in \mathbb{H}(\mathbb{R}^{n+1}, \mathbb{C}) \cap L^2(\mathbb{R}^{n+1}, \mathbb{C})$. Hence, using Lemma 5.3,

$$I = \int_{\mathbb{R}^{n+1}} A_{||}\nabla_{||}\mathcal{S}_{\lambda_0}f \cdot \overline{\nabla_{||}v} \, dxdt + \int_{\mathbb{R}^{n+1}} H_t D_{1/2}^t(\mathcal{S}_{\lambda_0}f)\overline{D_{1/2}^t(v)} \, dxdt,$$

for a function $v \in \mathbb{H} = \mathbb{H}(\mathbb{R}^{n+1}, \mathbb{C})$ which satisfies

$$||v||_{\mathbb{H}} \leq c||\mathbf{g}||_2,$$

for some constant c depending only on n and Λ. Let

$$I_1 := \int_{\mathbb{R}^{n+1}} A_{||} \nabla_{||} \mathcal{S}_{\lambda_0} f \cdot \overline{\nabla_{||} v} \, dxdt,$$

$$I_2 := \int_{\mathbb{R}^{n+1}} H_t D_{1/2}^t (\mathcal{S}_{\lambda_0} f) \overline{D_{1/2}^t(v)} \, dxdt.$$

As $C_0^\infty(\mathbb{R}^{n+1}, \mathbb{C})$ is dense in $\mathbb{H}(\mathbb{R}^{n+1}, \mathbb{C})$ we can in the following also assume, without loss of generality, that $v \in C_0^\infty(\mathbb{R}^{n+1}, \mathbb{C})$. This reduction allows us to handle several boundary terms which appear when we integrate by parts.

We first estimate I_1. Recall the resolvents, $\mathcal{E}_\lambda = (I + \lambda^2 \mathcal{H}_{||})^{-1}$ and $\mathcal{E}_\lambda^* = (I + \lambda^2 \mathcal{H}_{||}^*)^{-1}$, introduced in Sect. 5. To start the estimate of I_1 we first note, applying Lemma 5.7, that

$$\left| \int_{\mathbb{R}^{n+1}} A_{||} \nabla_{||} \mathcal{E}_\lambda \mathcal{S}_{\lambda+\lambda_0} f \cdot \overline{\nabla_{||} \mathcal{E}_\lambda^* v} \, dxdt \right| \leq \frac{c}{\lambda^2} ||\mathcal{S}_{\lambda+\lambda_0} f||_2 ||v||_2. \tag{6.1}$$

Hence, using that

$$\mathcal{S}_{\lambda+\lambda_0} f - \mathcal{S}_{\lambda_0} f = \int_{\lambda_0}^{\lambda+\lambda_0} \partial_\sigma \mathcal{S}_\sigma f \, d\sigma, \tag{6.2}$$

the fact that $\Phi(f) < \infty$, Lemma 3.5 and that $f, v \in C_0^\infty(\mathbb{R}^{n+1}, \mathbb{C})$, we can use (6.1) to conclude that

$$\left| \int_{\mathbb{R}^{n+1}} A_{||} \nabla_{||} \mathcal{E}_\lambda \mathcal{S}_{\lambda+\lambda_0} f \cdot \overline{\nabla_{||} \mathcal{E}_\lambda^* v} \, dxdt \right| \longrightarrow 0 \quad \text{as } \lambda \to \infty. \tag{6.3}$$

Hence,

$$I_1 = -\int_0^\infty \partial_\lambda \left(\int_{\mathbb{R}^{n+1}} A_{||} \nabla_{||} \mathcal{E}_\lambda \mathcal{S}_{\lambda+\lambda_0} f \cdot \overline{\nabla_{||} \mathcal{E}_\lambda^* v} \, dxdt \right) d\lambda. \tag{6.4}$$

Consider $\lambda > 0$, $\lambda_0 > 0$ fixed, let $|h| \ll \min\{\lambda_0, \lambda\}$. Then

$$\int_{\mathbb{R}^{n+1}} A_{||} \nabla_{||} \mathcal{E}_{\lambda+h} \mathcal{S}_{\lambda+\lambda_0+h} f \cdot \overline{\nabla_{||} \mathcal{E}_{\lambda+h}^* v} \, dxdt$$

$$- \int_{\mathbb{R}^{n+1}} A_{||} \nabla_{||} \mathcal{E}_\lambda \mathcal{S}_{\lambda+\lambda_0} f \cdot \overline{\nabla_{||} \mathcal{E}_{\lambda+h}^* v} \, dxdt = T_1^h + T_2^h + T_3^h, \tag{6.5}$$

where

$$T_1^h := \int_{\mathbb{R}^{n+1}} A_{||} \nabla_{||} (\mathcal{E}_{\lambda+h} - \mathcal{E}_\lambda) \mathcal{S}_{\lambda+\lambda_0} f \cdot \overline{\nabla_{||} \mathcal{E}_{\lambda+h}^* v} \, dxdt,$$

$$T_2^h := \int_{\mathbb{R}^{n+1}} A_{||} \nabla_{||} \mathcal{E}_\lambda \mathcal{S}_{\lambda+\lambda_0} f \cdot \overline{\nabla_{||} (\mathcal{E}_{\lambda+h}^* v - \mathcal{E}_\lambda^* v)} \, dxdt,$$

$$T_3^h := \int_{\mathbb{R}^{n+1}} A_{||} \nabla_{||} \mathcal{E}_{\lambda+h} (\mathcal{S}_{\lambda+\lambda_0+h} f - \mathcal{S}_{\lambda+\lambda_0} f) \cdot \overline{\nabla_{||} \mathcal{E}_{\lambda+h}^* v} \, dxdt. \tag{6.6}$$

Using (5.7)–(5.16) we see that

$$\lim_{h \to 0} h^{-1} T_1^h = \int_{\mathbb{R}^{n+1}} A_{||} \nabla_{||} \partial_\lambda \mathcal{E}_\lambda \mathcal{S}_{\lambda+\lambda_0} f \cdot \overline{\nabla_{||} \mathcal{E}_\lambda^* v} \, dxdt,$$

$$\lim_{h \to 0} h^{-1} T_2^h = \int_{\mathbb{R}^{n+1}} A_{||} \nabla_{||} \mathcal{E}_\lambda \mathcal{S}_{\lambda+\lambda_0} f \cdot \overline{\nabla_{||} \partial_\lambda \mathcal{E}_\lambda^* v} \, dxdt,$$

$$\lim_{h \to 0} h^{-1} T_3^h = \int_{\mathbb{R}^{n+1}} A_{||} \nabla_{||} \mathcal{E}_\lambda \partial_\lambda \mathcal{S}_{\lambda+\lambda_0} f \cdot \overline{\nabla_{||} \mathcal{E}_\lambda^* v} \, dxdt. \tag{6.7}$$

Using these deductions we can conclude that

$$
\begin{aligned}
I_1 &= -\int_0^\infty \int_{\mathbb{R}^{n+1}} \left((A_{||}\nabla_{||}\partial_\lambda \mathcal{E}_\lambda \mathcal{S}_{\lambda+\lambda_0} f) \cdot \overline{\nabla_{||}\mathcal{E}_\lambda^* v} \right) dx dt d\lambda \\
&\quad -\int_0^\infty \int_{\mathbb{R}^{n+1}} \left((A_{||}\nabla_{||}\mathcal{E}_\lambda \mathcal{S}_{\lambda+\lambda_0} f) \cdot \overline{\nabla_{||}\partial_\lambda \mathcal{E}_\lambda^* v} \right) dx dt d\lambda \\
&\quad -\int_0^\infty \int_{\mathbb{R}^{n+1}} \left((A_{||}\nabla_{||}\mathcal{E}_\lambda \partial_\lambda \mathcal{S}_{\lambda+\lambda_0} f) \cdot \overline{\nabla_{||}\mathcal{E}_\lambda^* v} \right) dx dt d\lambda \\
&=: I_{11} + I_{12} + I_{13},
\end{aligned}
$$

and we emphasize that by our assumptions, and (5.7)–(5.16), $I_{11} - I_{13}$ are well defined. To proceed we first note that

$$
\begin{aligned}
I_{11} &= -\int_0^\infty \langle \mathcal{L}_{||}^* \mathcal{E}_\lambda^* v, \partial_\lambda \mathcal{E}_\lambda \mathcal{S}_{\lambda+\lambda_0} f \rangle_{\tilde{\mathbb{H}}^*} \, d\lambda = -\int_0^\infty \langle \mathcal{E}_\lambda^* \mathcal{L}_{||}^* v, \partial_\lambda \mathcal{E}_\lambda \mathcal{S}_{\lambda+\lambda_0} f \rangle_{\tilde{\mathbb{H}}^*} \, d\lambda, \\
I_{12} &= -\int_0^\infty \langle \mathcal{L}_{||}\mathcal{E}_\lambda \mathcal{S}_{\lambda+\lambda_0} f, \partial_\lambda \mathcal{E}_\lambda^* v \rangle_{\tilde{\mathbb{H}}^*} \, d\lambda = -\int_0^\infty \langle \mathcal{E}_\lambda \mathcal{L}_{||}\mathcal{S}_{\lambda+\lambda_0} f, \partial_\lambda \mathcal{E}_\lambda^* v \rangle_{\tilde{\mathbb{H}}^*} \, d\lambda,
\end{aligned}
$$

by (5.19). Let

$$
J := \int_0^\infty \int_{\mathbb{R}^{n+1}} |\mathcal{E}_\lambda \mathcal{L}_{||}\mathcal{S}_{\lambda+\lambda_0} f|^2 \, \lambda \, dx dt d\lambda.
$$

Then, using (5.17), the L^2-boundedness of $\mathcal{E}_\lambda$ and $\mathcal{E}_\lambda^*$, Lemma 5.7, and the square function estimates, Theorem 5.11, we see that

$$
\begin{aligned}
|I_{11}| + |I_{12}| &\leq c(|||\lambda \partial_t \mathcal{S}_{\lambda+\lambda_0} f||| + J^{1/2})||v||_{\mathbb{H}} \\
&\leq c(\Phi(f) + ||f||_2 + J^{1/2})||v||_{\mathbb{H}},
\end{aligned}
$$

where we on the last line have used Lemma 4.2. Next, referring to (5.4) we have

$$
\begin{aligned}
\mathcal{L}_{||}\mathcal{S}_{\lambda+\lambda_0} f &= \sum_{j=1}^{n+1} A_{n+1,j} D_{n+1} D_j \mathcal{S}_{\lambda+\lambda_0} f \\
&\quad + \sum_{i=1}^{n} D_i (A_{i,n+1} D_{n+1} \mathcal{S}_{\lambda+\lambda_0} f) + \partial_t \mathcal{S}_{\lambda+\lambda_0} f
\end{aligned}
$$

in a weak sense for almost every λ. Using this, and the L^2-boundedness of $\mathcal{E}_\lambda$, Lemma 5.7, we see that

$$
J \leq c(|||\lambda \nabla \partial_\lambda \mathcal{S}_{\lambda+\lambda_0} f|||^2 + |||\lambda \partial_t \mathcal{S}_{\lambda+\lambda_0} f|||^2 + \tilde{J}),
$$

where

$$
\tilde{J} := \int_0^\infty \int_{\mathbb{R}^{n+1}} \left| \mathcal{E}_\lambda \sum_{i=1}^{n} D_i (A_{i,n+1}\partial_\lambda \mathcal{S}_{\lambda+\lambda_0} f) \right|^2 \lambda \, dx dt d\lambda.
$$

In particular, again using Lemma 4.2 we see that

$$
J \leq c(\Phi(f) + ||f||_2 + \tilde{J}).
$$

To estimate $\tilde{J}$, let $A^{||}_{n+1} := (A_{1,n+1}, \ldots, A_{n,n+1})$. Then

$$\tilde{J} = \int_0^\infty \int_{\mathbb{R}^{n+1}} |\mathcal{E}_\lambda \operatorname{div}_{||}(A^{||}_{n+1} \partial_\lambda \mathcal{S}_{\lambda+\lambda_0} f)|^2 \, \lambda dx dt d\lambda$$

$$= \int_0^\infty \int_{\mathbb{R}^{n+1}} |\mathcal{U}_\lambda(A^{||}_{n+1} \partial_\lambda \mathcal{S}_{\lambda+\lambda_0} f)|^2 \, \frac{dx dt d\lambda}{\lambda},$$

where $\mathcal{U}_\lambda := \lambda \mathcal{E}_\lambda \operatorname{div}_{||}$. We write

$$\mathcal{U}_\lambda A^{||}_{n+1} = \mathcal{U}_\lambda A^{||}_{n+1} - (\mathcal{U}_\lambda A^{||}_{n+1})\mathcal{P}_\lambda + (\mathcal{U}_\lambda A^{||}_{n+1})\mathcal{P}_\lambda =: \mathcal{R}_\lambda + (\mathcal{U}_\lambda A^{||}_{n+1})\mathcal{P}_\lambda.$$

Then

$$\tilde{J} \leq \tilde{J}_1 + \tilde{J}_2,$$

where

$$\tilde{J}_1 := \int_0^\infty \int_{\mathbb{R}^{n+1}} |\mathcal{R}_\lambda \partial_\lambda \mathcal{S}_{\lambda+\lambda_0} f|^2 \, \frac{dx dt d\lambda}{\lambda},$$

$$\tilde{J}_2 := \int_0^\infty \int_{\mathbb{R}^{n+1}} |(\mathcal{U}_\lambda A^{||}_{n+1})\mathcal{P}_\lambda(\partial_\lambda \mathcal{S}_{\lambda+\lambda_0} f)|^2 \, \frac{dx dt d\lambda}{\lambda}.$$

Using Lemmas 5.8, and 4.2, we see that

$$\tilde{J}_1 \leq c \int_0^\infty \int_{\mathbb{R}^{n+1}} |\nabla \partial_\lambda \mathcal{S}_{\lambda+\lambda_0} f|^2 \, \lambda dx dt d\lambda$$

$$+ c \int_0^\infty \int_{\mathbb{R}^{n+1}} |\partial_t \partial_\lambda \mathcal{S}_{\lambda+\lambda_0} f|^2 \, \lambda^3 dx dt d\lambda$$

$$\leq c(\Phi(f)^2 + ||f||_2^2).$$

Furthermore, by the Carleson measure estimate in Lemma 5.9 we have

$$\tilde{J}_2 \leq c||N_*(\mathcal{P}_\lambda(\partial_\lambda \mathcal{S}_\lambda f))||_2^2.$$

Finally, we note that

$$||N_*(\mathcal{P}_\lambda(\partial_\lambda \mathcal{S}_\lambda f))||_2 \leq c||M(N_{**}(\partial_\lambda \mathcal{S}_\lambda f))||_2 \leq c||N_{**}(\partial_\lambda \mathcal{S}_\lambda f)||_2$$

where M is the parabolic Hardy–Littlewood maximal function. Putting all these estimates together we can conclude that

$$|I_{11}| + |I_{12}| \leq \big(\Phi(f) + ||f||_2 + ||N_{**}(\partial_\lambda \mathcal{S}_\lambda f)||_2\big)||v||_{\mathbb{H}},$$

which completes the estimate of $|I_{11}| + |I_{12}|$. We next estimate I_{13}. Integrating by parts with respect to λ we deduce, by repeating the argument above, that

$$I_{13} = -\int_0^\infty \int_{\mathbb{R}^{n+1}} \big(A_{||} \nabla_{||} \mathcal{E}_\lambda \partial_\lambda \mathcal{S}_{\lambda+\lambda_0} f \cdot \overline{\nabla_{||} \mathcal{E}_\lambda^* v}\big) \, dx dt d\lambda$$

$$= \int_0^\infty \int_{\mathbb{R}^{n+1}} \partial_\lambda \big(A_{||} \nabla_{||} \mathcal{E}_\lambda \partial_\lambda \mathcal{S}_{\lambda+\lambda_0} f \cdot \overline{\nabla_{||} \mathcal{E}_\lambda^* v}\big) \, \lambda dx dt d\lambda$$

$$= \int_0^\infty \int_{\mathbb{R}^{n+1}} \big((A_{||} \nabla_{||} \partial_\lambda \mathcal{E}_\lambda \partial_\lambda \mathcal{S}_{\lambda+\lambda_0} f) \cdot \overline{\nabla_{||} \mathcal{E}_\lambda^* v}\big) \, \lambda dx dt d\lambda$$

$$+ \int_0^\infty \int_{\mathbb{R}^{n+1}} \big((A_{||} \nabla_{||} \mathcal{E}_\lambda \partial_\lambda \mathcal{S}_{\lambda+\lambda_0} f) \cdot \overline{\nabla_{||} \partial_\lambda \mathcal{E}_\lambda^* v}\big) \, \lambda dx dt d\lambda$$

$$+ \int_0^\infty \int_{\mathbb{R}^{n+1}} \left((A_\| \nabla_\| \mathcal{E}_\lambda \partial_\lambda^2 \mathcal{S}_{\lambda+\lambda_0} f) \cdot \overline{\nabla_\| \mathcal{E}_\lambda^* v} \right) \lambda dx dt d\lambda$$

$$=: I_{131} + I_{132} + I_{133}.$$

By repeating the estimates above used to control $|I_{11}| + |I_{12}|$, we see that

$$(|I_{131}| + |I_{132}|)^2 \leq c \int_0^\infty \int_{\mathbb{R}^{n+1}} |\nabla \partial_\lambda^2 \mathcal{S}_{\lambda+\lambda_0} f|^2 \lambda^3 dx dt d\lambda,$$

$$+ c \int_0^\infty \int_{\mathbb{R}^{n+1}} |\partial_t \partial_\lambda \mathcal{S}_{\lambda+\lambda_0} f|^2 \lambda^3 dx dt d\lambda$$

$$+ c \int_0^\infty \int_{\mathbb{R}^{n+1}} |\partial_t \partial_\lambda^2 \mathcal{S}_{\lambda+\lambda_0} f|^2 \lambda^5 dx dt d\lambda + c ||N_*(\mathcal{P}_\lambda(\lambda \partial_\lambda^2 \mathcal{S}_\lambda f))||_2^2.$$

Furthermore,

$$I_{133} = \int_0^\infty \int_{\mathbb{R}^{n+1}} \left(A_\| \nabla_\| \mathcal{E}_\lambda \partial_\lambda^2 \mathcal{S}_{\lambda+\lambda_0} f \cdot \overline{\nabla_\| \mathcal{E}_\lambda^* v} \right) \lambda dx dt d\lambda$$

$$= - \int_0^\infty \int_{\mathbb{R}^{n+1}} \mathcal{E}_\lambda \partial_\lambda^2 \mathcal{S}_{\lambda+\lambda_0} f \overline{\mathcal{E}_\lambda^* \mathcal{L}_\|^* v} \lambda dx dt d\lambda,$$

by previous arguments. Using the L^2-boundedness of $\mathcal{E}_\lambda$, Lemma 5.7 and the square function estimate for $\mathcal{E}_\lambda^* \mathcal{L}_\|^*$, Theorem 5.11, we can conclude that

$$|I_{133}| \leq c \left(\int_0^\infty \int_{\mathbb{R}^{n+1}} |\partial_\lambda^2 \mathcal{S}_{\lambda+\lambda_0} f|^2 \lambda dx dt d\lambda \right)^{1/2} ||v||_{\mathbb{H}}.$$

Hence, again using Lemma 4.2 we see that

$$|I_{13}| \leq c \left(\Phi(f) + ||f||_2 + ||N_*(\mathcal{P}_\lambda(\lambda \partial_\lambda^2 \mathcal{S}_\lambda f))||_2 \right) ||v||_{\mathbb{H}},$$

Again

$$||N_*(\mathcal{P}_\lambda(\lambda \partial_\lambda^2 \mathcal{S}_\lambda f))||_2 \leq c ||M(N_{**}(\lambda \partial_\lambda^2 \mathcal{S}_\lambda f))||_2 \leq c ||N_{**}(\lambda \partial_\lambda^2 \mathcal{S}_\lambda f)||_2,$$

and using (2.6) and Lemma 2.1 we see that

$$||N_{**}(\lambda \partial_\lambda^2 \mathcal{S}_\lambda f)||_2 \leq c ||N_{**}(\partial_\lambda \mathcal{S}_\lambda f)||_2,$$

after a slight redefinition of the non-tangential maximal function on the right hand side. This completes the proof of I_1.

We next estimate I_2. To start the estimate of I_2 we first deduce, by arguing along the lines of (6.3)–(6.7), that

$$I_2 = - \int_0^\infty \int_{\mathbb{R}^{n+1}} \partial_\lambda \left(H_t D_{1/2}^t \mathcal{E}_\lambda \mathcal{S}_{\lambda+\lambda_0} f \cdot \overline{D_{1/2}^t \mathcal{E}_\lambda^* v} \right) dx dt d\lambda$$

$$= - \int_0^\infty \int_{\mathbb{R}^{n+1}} (H_t D_{1/2}^t \partial_\lambda \mathcal{E}_\lambda \mathcal{S}_{\lambda+\lambda_0} f) \cdot \overline{D_{1/2}^t \mathcal{E}_\lambda^* v} \, dx dt d\lambda$$

$$- \int_0^\infty \int_{\mathbb{R}^{n+1}} (H_t D_{1/2}^t \mathcal{E}_\lambda \mathcal{S}_{\lambda+\lambda_0} f) \cdot \overline{D_{1/2}^t \partial_\lambda \mathcal{E}_\lambda^* v} \, dx dt d\lambda$$

$$- \int_0^\infty \int_{\mathbb{R}^{n+1}} (H_t D_{1/2}^t \mathcal{E}_\lambda \partial_\lambda \mathcal{S}_{\lambda+\lambda_0} f) \cdot \overline{D_{1/2}^t \mathcal{E}_\lambda^* v} \, dx dt d\lambda$$

$$=: I_{21} + I_{22} + I_{23}.$$

Using the L^2-boundedness of $\mathcal{E}_\lambda$ and $\mathcal{E}_\lambda^*$, Lemma 5.7, and the square function estimates, Theorem 5.11, that $\mathcal{H}_{||}$ commutes with $\mathcal{E}_\lambda$, $D_{1/2}^t$, and $H_t D_{1/2}^t$, and that $\mathcal{H}_{||}^*$ commutes with $\mathcal{E}_\lambda^*$, $D_{1/2}^t$, and $H_t D_{1/2}^t$, in both cases in the sense described above, we can as in the estimate of $|I_{11}| + |I_{12}|$ deduce that

$$|I_{22}| \leq c|||\lambda\partial_t\mathcal{S}_{\lambda+\lambda_0}f|||\,||v||_{\mathbb{H}} \leq c(\Phi(f) + ||f||_2)||v||_{\mathbb{H}}. \tag{6.8}$$

At the final step of this deduction we have also used Lemma 4.2. Integrating by parts with respect to λ in I_{23}, and repeating the arguments used in the estimates of $|I_{21}|$ and $|I_{22}|$, it is easily seen, using Lemma 4.2, that

$$|I_{23}| \leq c(\Phi(f) + ||f||_2)||v||_{\mathbb{H}} + |\tilde{I}_{23}|,$$

where

$$\tilde{I}_{23} = \int_0^\infty \int_{\mathbb{R}^{n+1}} \left((H_t D_{1/2}^t \mathcal{E}_\lambda \partial_\lambda^2 \mathcal{S}_{\lambda+\lambda_0}f) \cdot \overline{D_{1/2}^t \mathcal{E}_\lambda^* v}\right) \lambda\,dxdtd\lambda.$$

However, again using Lemma 5.7 and Theorem 5.11

$$|\tilde{I}_{23}| \leq |||\lambda\partial_\lambda^2\mathcal{S}_{\lambda+\lambda_0}f|||\,|||\lambda\partial_t\mathcal{E}_\lambda^* v||| \leq c\Phi(f)||v||_{\mathbb{H}}.$$

This completes the proof of the lemma.

6.2 Proof of Lemma 6.2

To prove Lemma 6.2 it suffices to estimate

$$\int_{\mathbb{R}^{n+1}} (\mathbb{D}_{n+1}\mathcal{S}_{\lambda_0}f)\bar{g}\,dxdt$$

when $f, g \in C_0^\infty(\mathbb{R}^{n+1}, \mathbb{C})$, $||g||_2 = 1$. Let in the following $\mathcal{P}_\lambda$ be a parabolic approximation of the identity. Then, using (2.2) (ii) we see that

$$\left|\int_{\mathbb{R}^{n+1}} (\mathbb{D}_{n+1}\mathcal{S}_{\lambda+\lambda_0}f)\mathcal{P}_\lambda\bar{g}\,dxdt\right| \leq c||D_{1/2}^t\mathcal{S}_{\lambda+\lambda_0}f||_2||\mathcal{P}_\lambda\bar{g}||_2$$

$$\leq \frac{c}{\lambda^{n/2+1}}||\partial_t\mathcal{S}_{\lambda+\lambda_0}f||_2||\mathcal{S}_{\lambda+\lambda_0}f||_2.$$

Again using (6.2), Hölder's inequality, the fact that $\Phi(f) < \infty$, Lemmas 3.4 and 3.5 we deduce that

$$\left|\int_{\mathbb{R}^{n+1}} (\mathbb{D}_{n+1}\mathcal{S}_{\lambda+\lambda_0}f)\mathcal{P}_\lambda\bar{g}\,dxdt\right| \longrightarrow 0 \quad \text{as } \lambda \to \infty.$$

Hence,

$$-\int_{\mathbb{R}^{n+1}} (\mathbb{D}_{n+1}\mathcal{S}_{\lambda_0}f)\bar{g}\,dxdt = \int_0^\infty \int_{\mathbb{R}^{n+1}} \partial_\lambda((\mathbb{D}_{n+1}\mathcal{S}_{\lambda+\lambda_0}f)\mathcal{P}_\lambda\bar{g})\,dxdtd\lambda$$

$$= \int_0^\infty \int_{\mathbb{R}^{n+1}} (\mathbb{D}_{n+1}\partial_\lambda\mathcal{S}_{\lambda+\lambda_0}f)\mathcal{P}_\lambda\bar{g}\,dxdtd\lambda$$

$$+ \int_0^\infty \int_{\mathbb{R}^{n+1}} (\mathbb{D}_{n+1}\mathcal{S}_{\lambda+\lambda_0}f)\partial_\lambda(\mathcal{P}_\lambda\bar{g})\,dxdtd\lambda$$

$$=: I + II.$$

Note that $\mathbb{D}_{n+1} = i\mathbb{D}^{-1}\partial_t$ and that $\partial_\lambda\mathcal{P}_\lambda = \mathbb{D}\mathcal{Q}_\lambda$ where $\mathcal{Q}_\lambda$ is an approximation of the zero operator. To prove this one can use that the kernel of $\partial_\lambda\mathcal{P}_\lambda$ has not only zero mean but also

first order vanishing moments if $\mathcal{P}$ is an even function (see also [21, p. 366]). Using this we see that

$$|II|^2 \leq \left| \int_0^\infty \int_{\mathbb{R}^{n+1}} (\partial_t \mathcal{S}_{\lambda+\lambda_0} f) \mathcal{Q}_\lambda \bar{g} \, dx dt d\lambda \right|^2$$

$$\leq c \int_0^\infty \int_{\mathbb{R}^{n+1}} |\partial_t \mathcal{S}_{\lambda+\lambda_0} f|^2 \, \lambda dx dt d\lambda \leq c(\Phi(f) + ||f||_2)^2,$$

by (2.8) and Lemma 4.2. To handle I we again integrate by parts with respect to λ,

$$-I = \int_0^\infty \int_{\mathbb{R}^{n+1}} (\mathbb{D}_{n+1} \partial_\lambda^2 \mathcal{S}_{\lambda+\lambda_0} f) \mathcal{P}_\lambda \bar{g} \, \lambda dx dt d\lambda$$

$$+ \int_0^\infty \int_{\mathbb{R}^{n+1}} (\mathbb{D}_{n+1} \partial_\lambda \mathcal{S}_{\lambda+\lambda_0} f) \partial_\lambda (\mathcal{P}_\lambda \bar{g}) \, \lambda dx dt d\lambda$$

$$=: I_1 + I_2.$$

Arguing as above we immediately see that

$$|I_2|^2 \leq c \int_0^\infty \int_{\mathbb{R}^{n+1}} |\partial_t \partial_\lambda \mathcal{S}_{\lambda+\lambda_0} f|^2 \, \lambda^3 dx dt d\lambda \leq c(\Phi(f) + ||f||_2)^2.$$

Focusing on I_1, Lemma 2.4 implies

$$|I_1| \leq |||\lambda \partial_\lambda^2 \mathcal{S}_{\lambda+\lambda_0} f||| \, |||\lambda \mathbb{D}_{n+1} \mathcal{P}_\lambda g||| \leq c|||\lambda \partial_\lambda^2 \mathcal{S}_{\lambda+\lambda_0} f||| \, |||\lambda \mathbb{D} \mathcal{P}_\lambda g||| \leq c \Phi(f),$$

and the proof of the lemma is complete.

6.3 Proof of Lemma 6.3

Let $K \gg 2$ be a degree of freedom and let $\phi \in C_0^\infty(\mathbb{R})$ be an even function with $\phi = 1$ on $(-3/2, -2/K) \cup (2/K, 3/2)$ and with support in $(-2, -1/K) \cup (1/K, 2)$. Recall that the multiplier defining $D_{1/2}^t$ is $|\tau|^{1/2}$. We write

$$|\tau|^{1/2} = |\tau|^{1/2} \phi(\tau/||(\xi, \tau)||^2) + |\tau|^{1/2}(1 - \phi)(\tau/||(\xi, \tau)||^2)$$

$$= \text{sgn}(\tau) \frac{||(\xi, \tau)||}{|\tau|^{1/2}} \phi(\tau/||(\xi, \tau)||^2) \frac{\tau}{||(\xi, \tau)||}$$

$$- \sum_{j=1}^n |\tau|^{1/2} \frac{i\xi_j}{|\xi|^2} (1 - \phi)(\tau/||(\xi, \tau)||^2) i\xi_j.$$

Hence, introducing the multipliers

$$m_1(\xi, \tau) = \text{sgn}(\tau) \frac{||(\xi, \tau)||}{|\tau|^{1/2}} \phi(\tau/||(\xi, \tau)||^2),$$

$$m_{2,j}(\xi, \tau) = -|\tau|^{1/2} \frac{i\xi_j}{|\xi|^2} (1 - \phi)(\tau/||(\xi, \tau)||^2),$$

for $j \in \{1, \dots, n\}$ we can conclude the existence of kernels L_1, $L_{2,j}$, corresponding to m_1, $m_{2,j}$, such that

$$D_{1/2}^t = L_1 * \mathbb{D}_{n+1} + c \sum_{j=1}^n L_{2,j} * \partial_{x_j},$$

where $*$ denotes convolution. Choosing $K = K(n)$ large enough we see that the multipliers m_1 and $m_{2,j}$ are bounded, and hence L_1 and $L_{2,j}$ are bounded operators on $L^2(\mathbb{R}^{n+1}, \mathbb{C})$. This completes the proof of Lemma 6.3.

7 Proof of Theorem 1.1

Assume that $\mathcal{H}$, $\mathcal{H}^*$, satisfy (1.2) and (1.3) as well as the De Giorgi–Moser–Nash estimates stated in (2.6) and (2.7). Assume also that there exists a constant C such that (1.5) holds whenever $f \in L^2(\mathbb{R}^{n+1}, \mathbb{C})$. To prove Theorem 1.1 we need to prove that there exists a constant c, depending at most on n, Λ, the De Giorgi–Moser–Nash constants and C, such that the inequalities in (1.6) (i)–(iv) hold. Again, we only have to prove (1.6) (i)–(iv) for $\mathcal{S}_\lambda^{\mathcal{H}}$ as the corresponding results for $\mathcal{S}_\lambda^{\mathcal{H}^*}$ follow by analogy. To start the proof, we first note that (1.6) (i) is an immediate consequence of Lemma 4.1 (i) and the assumption in (1.5) (i). Using Lemmas 6.1, 6.2, and 6.3, we see that (1.6) (i) and the assumptions in (1.5) imply that

$$\sup_{\lambda > 0} ||\mathbb{D}\mathcal{S}_\lambda^{\mathcal{H}} f||_2 \leq c||f||_2.$$

This proves (1.6) (ii). (1.6) (iii), (iv), now follows immediately form these estimates and Lemma 4.1.

8 Proof of Theorems 1.2 and 1.3

Assume that $\mathcal{H} = \partial_t - \operatorname{div} A\nabla$ satisfies (1.2) and (1.3). Assume in addition that A is real and symmetric. Then (2.6) and (2.7) hold. To prove Theorem 1.2 we have to prove that there exists a constant C, depending at most on n, Λ, such that (1.5) holds with this C. We first focus on the estimate in (1.5) (ii). Consider

$$\psi_\lambda(x, t, y, s) := \lambda K_{1,\lambda}(x, t, y, s) = \lambda \partial_\lambda^2 \Gamma_\lambda(x, t, y, s). \tag{8.1}$$

Then, using Lemma 3.1 we see that $\psi_\lambda(x, t, y, s)$ satisfies the Calderon–Zygmund bounds

$$|\psi_\lambda(x, t, y, s)| \leq c|\lambda|(d_\lambda(x, t, y, s))^{-n-3}, \tag{8.2}$$

and

$$
\begin{aligned}
|\mathbb{D}^h(\psi_\lambda(\cdot, \cdot, y, s))(x, t)| &\leq c|\lambda|||h||^\alpha (d_\lambda(x, t, y, s))^{-n-3-\alpha} \\
&\leq c||h||^\alpha (d_\lambda(x, t, y, s))^{-n-2-\alpha},
\end{aligned}
\tag{8.3}
$$

for some $\alpha > 0$, whenever $2||h|| \leq (|x - y| + |t - s|^{1/2})$ or $2||h|| \leq |\lambda|$. Our proof of Theorem 1.2 is based on the following two theorems proved below.

Theorem 8.1 *Assume that ψ_λ satisfies (8.2) and (8.3). Let*

$$\theta_\lambda f(x, t) := \int_{\mathbb{R}^{n+1}} \psi_\lambda(x, t, y, s) f(y, s) \, dy ds,$$

whenever $f \in L^2(\mathbb{R}^{n+1}, \mathbb{C})$. Suppose that there exists a system $\{b_Q\}$ of functions, $b_Q : \mathbb{R}^{n+1} \to \mathbb{C}$, index by parabolic cubes $Q \subseteq \mathbb{R}^{n+1}$, and a constant c, independent of Q, such

that for each cube Q the following is true.

$$\text{(i)} \quad \int_{\mathbb{R}^{n+1}} |b_Q(x,t)|^2 \, dxdt \le c|Q|,$$

$$\text{(ii)} \quad \int_0^{l(Q)} \int_Q |\theta_\lambda b_Q(x,t)|^2 \, \frac{dxdtd\lambda}{\lambda} \le c|Q|,$$

$$\text{(iii)} \quad c^{-1}|Q| \le Re \int_Q b_Q(x,t) \, dxdt. \tag{8.4}$$

Then there exists a constant c such that

$$|||\theta_\lambda f||| = \left(\int_0^\infty \int_{\mathbb{R}^{n+1}} |\theta_\lambda f(x,t)|^2 \, \frac{dxdtd\lambda}{\lambda} \right)^{1/2} \le c||f||_2, \tag{8.5}$$

whenever $f \in L^2(\mathbb{R}^{n+1}, \mathbb{C})$.

The proofs of Theorems 1.3 and 8.1 are given below. We here use Theorems 1.3 and 8.1 to complete the proof of Theorem 1.2.

Proof of (1.5) *(ii)* We simply have to produce, using Theorem 8.1 and for θ_λ defined using the kernel in (8.1), a system $\{b_Q\}$ of functions satisfying (8.4) (i)–(iii). To do this we let

$$b_Q(y,s) := |Q| 1_Q \tilde{K}_-(A_Q^-, y, s),$$

whenever $(y,s) \in \mathbb{R}^{n+1}$, where 1_Q is the indicator function for the cube Q and where $\tilde{K}_-(A_Q^-, y, s)$ is the to $\mathcal{H}^* = -\partial_t + \mathcal{L}$ associated Poisson kernel, at $A_Q^- := (x_Q, -l(Q), t_Q)$, defined with respect to $\mathbb{R}_-^{n+2}$. Theorem 1.3 applies to $\tilde{K}_-(A_Q^-, \cdot, \cdot)$ modulo trivial modifications. To verify that b_Q satisfies (8.4) (i)–(iii), we first note that (i) is an immediate consequence of Theorem 1.3. Furthermore,

$$\int_{\mathbb{R}^{n+1}} b_Q(y,s) \, dyds = |Q|\tilde{\omega}_-^{A_Q^-}(Q) \ge c^{-1}|Q|,$$

by elementary estimates and where $\tilde{\omega}_-^{A_Q^-}$ is the associated parabolic measure at A_Q^- and defined with respect to $\mathbb{R}_-^{n+2}$. Hence (iii) follows and it only remains to establish (ii). Let $(x,t) \in Q, \lambda \in (0, l(Q))$ and note that

$$\begin{aligned}
\theta_\lambda b_Q(x,t) &= \int_{\mathbb{R}^{n+1}} \lambda \partial_\lambda^2 \Gamma_\lambda(x,t,y,s) b_Q(y,s) \, dyds \\
&= \lambda|Q| \int_Q \partial_\lambda^2 \Gamma_\lambda(x,t,y,s) \tilde{K}_-(A_Q^-, y, s) \, dyds \\
&= \lambda|Q| \left(\partial_\lambda^2 \Gamma(x,t,\lambda,x_Q,t_Q,-l(Q)) \right),
\end{aligned}$$

by the definition of A_Q^-, $\tilde{K}_-(A_Q^-, y, s)$, and as $\partial_\lambda^2 \Gamma(x,t,\lambda,x_Q,t_Q,-l(Q))$ solves $\mathcal{H}^* u = 0$ in $\mathbb{R}_-^{n+2}$. Using this, and (8.2), we see that (ii) follows by elementary manipulations. Hence, using Theorem 8.1 we can conclude the validity of (1.5) (ii). □

Proof of (1.5) *(i)* We first note, that we can throughout the proof assume, without loss of generality, that $f \in C_0^\infty(\mathbb{R}^{n+1}, \mathbb{R})$. Second, using Theorem 1.3 and the fact that if $\mathcal{H} = \partial_t - \text{div } A\nabla$ satisfies (1.2) and (1.3), and if A is real and symmetric, then the estimates of the non-tangential maximal function by the square function established in [9] for the heat equation, remain valid for solutions to $\mathcal{H}u = 0$. In particular, let $f \in C_0^\infty(\mathbb{R}^{n+1}, \mathbb{R})$ and

consider $\lambda > 0$ fixed. We let R and r be such that $\lambda \ll r \ll R$ and such that the support of f is contained in $Q_{R/4}(0,0)$. Then, using Theorem 1.3 and [9] we see that

$$||(\partial_\lambda \mathcal{S}_\lambda f)1_{Q_r(0,0)}||_2^2 \leq c|||\lambda \nabla \partial_\lambda \mathcal{S}_\lambda f|||^2 + cR^{n+2}|\partial_\lambda \mathcal{S}_{R/2} f(0,0)|^2,$$

for a constant c depending only on n, Λ. However,

$$R^{n+2}|\partial_\lambda \mathcal{S}_{R/2} f(0,0)|^2 \leq R^{-n-2}||f||_1^2.$$

Hence, first letting $R \to \infty$ and then letting $r \to \infty$ we can conclude that

$$||\partial_\lambda \mathcal{S}_\lambda f||_2 \leq c|||\lambda \nabla \partial_\lambda \mathcal{S}_\lambda f|||. \tag{8.6}$$

Using (4.3) we see that

$$|||\lambda \nabla \partial_\lambda \mathcal{S}_\lambda f||| \leq c|||\lambda \partial_\lambda^2 \mathcal{S}_\lambda f||| + c||f||_2. \tag{8.7}$$

(8.6), (8.7) and (1.5) (ii) now prove (1.5) (i). □

This completes the proof of Theorem 1.2 modulo Theorems 8.1 and 1.3.

8.1 Proof of Theorem 8.1

Though there are several references for this type of argument, see [10,19,25] and the references therein, we will, for completion, include a sketch/proof of the argument in our context. To start with, as ψ_λ satisfies (8.2) and (8.3) it is well-known, see [10], that to prove (8.5) it suffices to prove the Carleson measure estimate

$$\sup_{Q \subset \mathbb{R}^{n+1}} \frac{1}{|Q|} \int_0^{l(Q)} \int_Q |\theta_\lambda 1|^2 \frac{dx \, dt \, d\lambda}{\lambda} \leq c. \tag{8.8}$$

Using assumption (iii) in the statement of Theorem 8.1, and a by now well-known stopping time argument, see [19], one can conclude that

$$\sup_{Q \subset \mathbb{R}^{n+1}} \frac{1}{|Q|} \int_0^{l(Q)} \int_Q |\theta_\lambda 1|^2 \frac{dx \, dt \, d\lambda}{\lambda} \leq c \sup_{Q \subset \mathbb{R}^{n+1}} \frac{1}{|Q|} \int_0^{l(Q)} \int_Q |(\theta_\lambda 1)\mathcal{A}_\lambda^Q b_Q| \frac{dx \, dt \, d\lambda}{\lambda},$$

where $\mathcal{A}_\lambda^Q$ denotes the dyadic averaging operator induced by Q and introduced in (2.9). Hence, to prove (8.8) it suffices to prove that

$$\int_0^{l(Q)} \int_Q |(\theta_\lambda 1)\mathcal{A}_\lambda^Q b_Q| \frac{dx \, dt \, d\lambda}{\lambda} \leq c|Q|, \tag{8.9}$$

for all $Q \subset \mathbb{R}^{n+1}$. We write

$$(\theta_\lambda 1)\mathcal{A}_\lambda^Q b_Q = \mathcal{R}_\lambda^{(1)} b_Q + \mathcal{R}_\lambda^{(2)} b_Q + \theta_\lambda b_Q,$$

where

$$\mathcal{R}_\lambda^{(1)} b_Q := (\theta_\lambda 1)(\mathcal{A}_\lambda^Q - \mathcal{A}_\lambda^Q \mathcal{P}_\lambda) b_Q,$$
$$\mathcal{R}_\lambda^{(2)} b_Q := ((\theta_\lambda 1)\mathcal{A}_\lambda^Q \mathcal{P}_\lambda - \theta_\lambda) b_Q,$$

and where $\mathcal{P}_\lambda$ is a parabolic approximation of the identity. Using assumption (ii) in the statement of Theorem 8.1 we see that the contribution from the term $\theta_\lambda b_Q$ to the Carleson

measure in (8.9) is controlled. Hence we focus on the contributions from $\mathcal{R}_\lambda^{(1)} b_Q$ and $\mathcal{R}_\lambda^{(2)} b_Q$. Note that

$$\mathcal{R}_\lambda^{(1)} = (\theta_\lambda 1)(\mathcal{A}_\lambda^Q - \mathcal{A}_\lambda^Q \mathcal{P}_\lambda) = (\theta_\lambda 1)\mathcal{A}_\lambda^Q (\mathcal{A}_\lambda^Q - \mathcal{P}_\lambda).$$

Using (8.2), (8.3), and a version of Schur's lemma, we see that

$$||(\theta_\lambda 1)\mathcal{A}_\lambda^Q||_{2\to 2} \leq c.$$

Thus, by Lemma 2.5,

$$\int_0^{l(Q)} \int_Q |\mathcal{R}_\lambda^{(1)} b_Q(x,t)|^2 \frac{dxdtd\lambda}{\lambda} \leq c \int_0^\infty \int_{\mathbb{R}^{n+1}} |(\mathcal{A}_\lambda^Q - \mathcal{P}_\lambda) b_Q(x,t)|^2 \frac{dxdtd\lambda}{\lambda}$$

$$\leq c \int_{\mathbb{R}^{n+1}} |b_Q(x,t)|^2 \, dxdt \leq c|Q|.$$

It remains to estimate

$$\int_0^{l(Q)} \int_Q |\mathcal{R}_\lambda^{(2)} b_Q(x,t)|^2 \frac{dxdtd\lambda}{\lambda}.$$

However, using (8.2), (8.3), and that $\mathcal{R}_\lambda^{(2)} 1 = 0$, it follows by a well known orthogonality argument, and assumption (i) in the statement of Theorem 8.1, that

$$\int_0^{l(Q)} \int_Q |\mathcal{R}_\lambda^{(2)} b_Q(x,t)|^2 \frac{dxdtd\lambda}{\lambda} \leq \int_{\mathbb{R}^{n+1}} |b_Q(x,t)|^2 \, dxdt \leq c|Q|.$$

This completes the proof of Theorem 8.1.

8.2 Proof of Theorem 1.3

Under the assumptions of Theorem 1.3 there exists a Green's function $G = G(X,t,Y,s)$ to $\mathcal{H} = \partial_t + \mathcal{L} = \partial_t - \text{div } A\nabla$ in $\mathbb{R}_+^{n+2}$, and corresponding measures $\omega^{(X,t)}(\cdot)$, $\tilde{\omega}^{(X,t)}(\cdot)$, $(X,t) \in \mathbb{R}_+^{n+2}$ such that

$$\phi(X,t) = \int \left(A\nabla_Y G(X,t,Y,s) \cdot \nabla\phi(Y,s) + G(X,t,Y,s)\partial_s\phi(Y,s)\right) dYds$$

$$+ \int \phi(y,0,s) \, d\omega^{(X,t)}(y,s),$$

$$\phi(X,t) = \int \left(A\nabla_Y G(Y,s,X,t) \cdot \nabla\phi(Y,s) - G(Y,s,X,t)\partial_s\phi(Y,s)\right) dYds$$

$$+ \int \phi(y,0,s) \, d\tilde{\omega}^{(X,t)}(y,s), \tag{8.10}$$

whenever $\phi \in C_0^\infty(\mathbb{R}^{n+2})$ and where $(X,t) = (x,x_{n+1},t)$, $(Y,s) = (y,y_{n+1},s)$. In particular,

$$(\partial_t + \mathcal{L}_{X,t})G(X,t,Y,s) = \delta_{(0,0)}(X-Y,t-s),$$

and

$$(-\partial_s + \mathcal{L}_{Y,s})G(X,t,Y,s) = \delta_{(0,0)}(X-Y,t-s). \tag{8.11}$$

Furthermore, in this setting G has a number of well-known properties, see for example display (3.7) on p. 11 in [23], and given $f \in C(\mathbb{R}^{n+1}) \cap L^{\infty}(\mathbb{R}^{n+1})$,

$$u(X, t) = \int_{\mathbb{R}^{n+1}} f(y, s) \, d\omega^{(X,t)}(y, s),$$

gives the solution to the continuous Dirichlet problem $\mathcal{H}u = (\partial_t + \mathcal{L})u = (\partial_t - \operatorname{div} A\nabla)u = 0$ in $\mathbb{R}_+^{n+2}$, $u \in C(\mathbb{R}^{n+1} \times [0, \infty))$, and $u(x, 0, t) = f(x, t)$ whenever $(x, t) \in \mathbb{R}^{n+1}$. $\{\omega^{(X,t)} : (X, t) \in \mathbb{R}_+^{n+2}\}$ and $\{\tilde{\omega}^{(X,t)} : (X, t) \in \mathbb{R}_+^{n+2}\}$ are families of regular Borel measures on $\mathbb{R}^{n+1}$ which we call $\mathcal{H}$-caloric, or $\mathcal{H}$-parabolic measures, and $\mathcal{H}^*$-caloric, or $\mathcal{H}^*$-parabolic measures, respectively.

Given $\mathcal{H} = \partial_t - \operatorname{div} A\nabla$, satisfying (1.2) and (1.3) with constant Λ, A real and symmetric, let A_ϵ, $0 < \epsilon \ll 1$, be a smooth $(n + 1) \times (n + 1)$-matrix valued function, A_ϵ real and symmetric, such that $\mathcal{H}^\epsilon = \partial_t - \operatorname{div} A_\epsilon \nabla$ satisfies (1.2) and (1.3), with constants depending at most on n and Λ, and such that $|A_\epsilon - A| \leq \epsilon$ on $\mathbb{R}^{n+2}$. Let as above $G_\epsilon(X, t, Y, s)$, $\omega_\epsilon^{(X,t)}$, $\tilde{\omega}_\epsilon^{(X,t)}$, be the Green's function and boundary measures associated to $\mathcal{H}_\epsilon = \partial_t - \operatorname{div} A_\epsilon \nabla$, $\mathcal{H}_\epsilon^* = -\partial_t - \operatorname{div} A_\epsilon \nabla$. Extending G_ϵ and G to all of $\mathbb{R}^{n+2}$ by putting $G_\epsilon \equiv 0 \equiv G$ on $\mathbb{R}_-^{n+2}$ one can prove, by for instance following the argument in Lemma 3.37 in [23], that

$$\int \left(A_\epsilon \nabla_Y G_\epsilon(X, t, Y, s) \cdot \nabla\phi(Y, s) + G_\epsilon(X, t, Y, s)\partial_s\phi(Y, s) \right) dY ds$$

$$\rightarrow \int \left(A\nabla_Y G(X, t, Y, s) \cdot \nabla\phi(Y, s) + G(X, t, Y, s)\partial_s\phi(Y, s) \right) dY ds \quad (8.12)$$

and

$$\int \left(A_\epsilon \nabla_Y G_\epsilon(Y, s, X, t) \cdot \nabla\phi(Y, s) - G_\epsilon(Y, s, X, t)\partial_s\phi(Y, s) \right) dY ds$$

$$\rightarrow \int \left(A\nabla_Y G(Y, s, X, t) \cdot \nabla\phi(Y, s) - G(Y, s, X, t)\partial_s\phi(Y, s) \right) dY ds, \quad (8.13)$$

as $\epsilon \rightarrow 0$, whenever $(X, t) \in \mathbb{R}_+^{n+2}$ and $\phi \in C_0^\infty(K)$ where K is a compact subset of $\mathbb{R}^{n+2} \setminus \{(X, t)\}$. Hence, using (8.10), (8.12), (8.13) we can conclude that

$$\omega_\epsilon^{(X,t)} \rightarrow \omega^{(X,t)}, \quad \tilde{\omega}_\epsilon^{(X,t)} \rightarrow \tilde{\omega}^{(X,t)} \quad (8.14)$$

weakly as Radon measures on $\mathbb{R}^{n+1}$ as $\epsilon \rightarrow 0$.

Based on the above outline it follows that it suffices to prove Theorem 1.3 assuming that A is smooth. Indeed, consider, for $\epsilon > 0$ small, A_ϵ and assume that the parabolic measure associated to $\mathcal{H}_\epsilon$, in $\mathbb{R}_+^{n+2}$, is absolutely continuous with respect to the measure $dxdt$ on $\mathbb{R}^{n+1} = \partial\mathbb{R}_+^{n+2}$, let $Q \subset \mathbb{R}^{n+1}$ be a parabolic cube and let $K_\epsilon(A_Q, y, s)$ be the to $\mathcal{H}_\epsilon$ associated Poisson kernel at $A_Q := (x_Q, l(Q), t_Q)$ where (x_Q, t_Q) is the center of the cube Q and $l(Q)$ defines its size. Furthermore, assume that there exists $c \geq 1$, depending only on n and Λ, such that

$$\int_Q |K_\epsilon(A_Q, y, s)|^2 \, dyds \leq c|Q|^{-1}.$$

Then $K_\epsilon(A_Q, y, s) \rightarrow K(A_Q, y, s)$ weakly on Q as $\epsilon \rightarrow 0$ and

$$\int_Q |K(A_Q, y, s)|^2 \, dyds \leq c|Q|^{-1}.$$

Furthermore,

$$
\int \phi(y, s) d\omega^{A_Q}(y, s) = \lim_{\epsilon \to 0} \int \phi(y, s) d\omega_\epsilon^{A_Q}(y, s)
$$
$$
= \lim_{\epsilon \to 0} \int \phi(y, s) K_\epsilon(A_Q, y, s) \, dy ds = \int \phi(y, s) K(A_Q, y, s) \, dy ds,
$$

(8.15)

whenever $\phi \in C_0^\infty(Q \times (-l(Q)/2, l(Q)/2)$ and Theorem 1.3 follows.

In the following we prove Theorem 1.3 assuming that A is smooth. If A is smooth it follows that the solution to the Dirichlet problem $\mathcal{H}u = 0$ in $\mathbb{R}_+^{n+2}$, $u = f$ on $\mathbb{R}^{n+1}$, equals

$$
u(X, t) = \int_{\mathbb{R}^{n+1}} K(X, t, y, s) f(y, s) \, dy ds,
$$

where

$$
K(X, t, y, s) := \langle \nabla_Y G(X, t, Y, s), A(Y) e_{n+1} \rangle |_{y_{n+1}=0} = a_{n+1,n+1}(y) \partial_{y_{n+1}}
$$
$$
G(X, t, Y, s) |_{y_{n+1}=0}.
$$

Using (1.2) we see that $a_{n+1,n+1}$ is uniformly bounded from below. Let $Q \subset \mathbb{R}^{n+1}$ be a parabolic cube and let $A_Q := (X_Q, t_Q) := (x_Q, l(Q), t_Q)$, where (x_Q, t_Q) is the center of the cube and $l(Q)$ defines its size. We write $Q = \hat{Q} \times (t_Q - l(Q)^2/2, t_Q + l(Q)^2/2)$ where $\hat{Q} \subset \mathbb{R}^n$ is a (elliptic) cube in the space variables only. Then

$$
\int_Q (K(A_Q, y, s))^2 \, dy ds = \int_{t_Q - l(Q)^2/2}^{t_Q + l(Q)^2/2} \int_{\hat{Q}} (K(X_Q, t_Q, y, s))^2 \, dy ds
$$
$$
= \int_{-l(Q)^2/2}^{l(Q)^2/2} \int_{\hat{Q}} (K(X_Q, 0, y, s))^2 \, dy ds
$$
$$
= \int_{-l(Q)^2/2}^{l(Q)^2/2} \int_{\hat{Q}} (K(X_Q, 0, y, -s))^2 \, dy ds,
$$

(8.16)

by the translation invariance in the time-variable due to (1.3). Using the Harnack inequality we see that

$$
(K(X_Q, 0, y, -s))^2 \leq c K(X_Q, 0, y, -s) K(X_Q, 16l(Q)^2, y, s),
$$

(8.17)

whenever $(y, s) \in \hat{Q} \times [-l(Q)^2/2, l(Q)^2/2]$. Let

$$
\phi \in C_0^\infty(\mathbb{R}^{n+2} \setminus (\{(X_Q, 0)\} \cup \{(X_Q, 16l(Q)^2)\}))
$$

be such that

$$
\phi(y, y_{n+1}, s) = 1,
$$

(8.18)

whenever $(y, y_{n+1}, s) \in \hat{Q} \times [-l(Q)/16, l(Q)/16] \times [-l(Q)^2/2, l(Q)^2/2]$, and

$$
\phi(y, y_{n+1}, s) = 0,
$$

(8.19)

whenever $(y, y_{n+1}, s) \in \mathbb{R}^{n+2} \setminus (2\hat{Q} \times [-l(Q)/8, l(Q)/8] \times [-l(Q)^2, l(Q)^2])$. Furthermore, we choose ϕ so that

$$
|\nabla_Y \phi(Y, s)| \leq c l(Q)^{-1}, \quad |\partial_s \phi(Y, s)| \leq c l(Q)^{-2},
$$

(8.20)

whenever $(Y, s) \in \mathbb{R}^{n+2}$. Let $\Psi(Y, s) := \phi(Y, s)\partial_{y_{n+1}} v(Y, s)$, where

$$v(Y, s) := G(X_Q, 0, Y, -s),$$

and let

$$\tilde{v}(Y, s) := G(X_Q, 16l(Q)^2, Y, s).$$

Using (8.11) we see that

$$
\begin{aligned}
0 &= \int_{\mathbb{R}^{n+2}_+1} \left((-\partial_s + \mathcal{L}_{Y,s})G(X_Q, 16l(Q)^2, Y, s)\right)\Psi(Y, s)\, dY ds \\
&= \int_{\mathbb{R}^{n+2}_+} \left((-\partial_s + \mathcal{L}_{Y,s})\tilde{v}(Y, s)\right)\Psi(Y, s)\, dY ds.
\end{aligned}
$$

Using this identity, and integrating by parts, we see that

$$
\begin{aligned}
I &:= \int_{\mathbb{R}^{n+1}} \Psi(Y, s)|_{y_{n+1}=0} K(X_Q, 16l(Q)^2, y, s)\, dy ds \\
&= \int_{\mathbb{R}^{n+2}_+} \left((\partial_s + \mathcal{L}_{Y,s})\Psi(Y, s)\right)\tilde{v}(Y, s)\, dY ds. \quad (8.21)
\end{aligned}
$$

We will now use the identity in (8.21) to prove Theorem 1.3. Indeed,

$$
\begin{aligned}
(\partial_s + \mathcal{L}_{Y,s})\Psi &= \partial_s \Psi - \operatorname{div}(A\nabla_Y \Psi) \\
&= \partial_{y_{n+1}} v \partial_s \phi - \operatorname{div}((\partial_{y_{n+1}} v)A\nabla_Y \phi) - A\nabla_Y \partial_{y_{n+1}} v \cdot \nabla_Y \phi \\
&\quad + \phi(\partial_s \partial_{y_{n+1}} v - \operatorname{div}(A\nabla_Y \partial_{y_{n+1}} v)).
\end{aligned}
$$

The key observation is, as A is independent of y_{n+1}, that

$$
\begin{aligned}
\partial_s \partial_{y_{n+1}} v - \operatorname{div}(A\nabla_Y \partial_{y_{n+1}} v) &= \partial_{y_{n+1}}\left(\partial_s v - \operatorname{div}(A\nabla_Y v)\right) \\
&= \partial_{y_{n+1}}\left(((-\partial_s + \mathcal{L}_Y)G)(X_Q, 0, Y, -s)\right) = 0,
\end{aligned}
$$

on the support of ϕ. This is due to the presence of the minus sign in front of s in $G(X_Q, 0, Y, -s)$. Hence, using (8.21) and elementary manipulations, we see that

$$I = I_1 + I_2 - I_3.$$

where

$$I_1 := \int_{\mathbb{R}^{n+2}_+} \partial_{y_{n+1}} G(X_Q, 0, Y, -s)(\partial_s \phi(Y, s))G(X_Q, 16l(Q)^2, Y, s)\, dY ds,$$

$$I_2 := \int_{\mathbb{R}^{n+2}_+} \partial_{y_{n+1}} G(X_Q, 0, Y, -s)(A\nabla_Y \phi) \cdot \nabla_Y G(X_Q, 16l(Q)^2, Y, s)\, dY ds,$$

$$I_3 := \int_{\mathbb{R}^{n+2}_+} (A\nabla_Y \partial_{y_{n+1}} G(X_Q, 0, Y, -s) \cdot \nabla_Y \phi)G(X_Q, 16l(Q)^2, Y, s)\, dY ds.$$

Recall that ϕ satisfies (8.18)–(8.20) and let $E = \mathbb{R}^{n+2}_+ \cap \overline{\{(Y, s) : \phi(Y, s) \neq 0\}}$. Using this,

$$|I_1| \leq cl(Q)^{-2} \int_E |\partial_{y_{n+1}} G(X_Q, 0, Y, -s)||G(X_Q, 16l(Q)^2, Y, s)|\, dY ds,$$

$$|I_2| \leq cl(Q)^{-1} \int_E |\partial_{y_{n+1}} G(X_Q, 0, Y, -s)||\nabla_Y G(X_Q, 16l(Q)^2, Y, s)|\, dY ds,$$

$$|I_3| \leq cl(Q)^{-1} \int_E |\nabla_Y \partial_{y_{n+1}} G(X_Q, 0, Y, -s)||G(X_Q, 16l(Q)^2, Y, s)|\, dY ds.$$

Hence, using energy estimates and Gaussian bounds for the fundamental solution we deduce

$$|I| \leq |I_1| + |I_2| + |I_3| \leq c|Q|^{-1}.$$

Using this and (8.21) we see that

$$\int_{-l(Q)^2/2}^{l(Q)^2/2} \int_{\hat{Q}} K(X_Q, 0, y, -s) K(X_Q, 16l(Q)^2, y, s) \, dy ds \leq c|Q|^{-1}.$$

Hence, using (8.16) and (8.17) we can conclude that

$$\int_Q (K(A_Q, y, s))^2 \, dy ds \leq c|Q|^{-1}, \tag{8.22}$$

whenever $Q \subset \mathbb{R}^{n+1}$ is a parabolic cube, for a constant $c \geq 1$, depending only on n and Λ. Put together Theorem 1.3 follows.

References

1. Auscher, P.: Regularity theorems and heat kernel for elliptic operators. J. Lond. Math. Soc. (2) **54**(2), 284–296 (1996)
2. Auscher, P., Axelsson, A.: Weighted maximal regularity estimates and solvability of non-smooth elliptic systems I. Invent. Math. **184**(1), 47–115 (2011)
3. Alfonseca, M.A., Auscher, P., Axelsson, A., Hofmann, S., Kim, S.: Analyticity of layer potentials and L^2 solvability of boundary value problems for divergence form elliptic equations with complex L^∞ coefficients. Adv. Math. **226**, 4533–4606 (2011)
4. Auscher, P., Hofmann, S., Lacey, M., McIntosh, A., Tchamitchian, P.: The solution of the Kato square root problem for second order elliptic operators on R^n. Ann. Math. (2) **156**(2), 633–654 (2002)
5. Auscher, P., Hofmann, S., Lewis, J.L., Tchamitchian, P.: Extrapolation of Carleson measures and the analyticity of Kato's square root operators. Acta Math. **187**, 161–190 (2001)
6. Auscher, P., Tchamitchian, P.: Square root problem for divergence operators and related topics. Asterisque **249**, 1–172 (1998) (Societe Mathematique de France)
7. Brown, R.: The method of layer potentials for the heat equation in Lipschitz cylinders. Am. J. Math. **111**(2), 339–379 (1989)
8. Brown, R.: The initial-Neumann problem for the heat equation in Lipschitz cylinders. Trans. Am. Math. Soc. **320**(1), 1–52 (1990)
9. Brown, R.: Area integral estimates for caloric functions. Trans. Am. Math. Soc. **315**, 565–589 (1989)
10. Christ, M., Journe, J.-L.: Polynomial growth estimates for multilinear singular integral operators. Acta Math. **159**(1–2), 51–80 (1987)
11. Coifman, R., Meyer, Y.: Wavelets. Calderon–Zygmund and Multilinear Operators. Cambridge Studies in Advanced Mathematics, vol. 48. Cambridge University Press, Cambridge (1997)
12. Fabes, E.B., Rivière, N.M., Dirichlet and Neumann problems for the heat equation in C^1-cylinders. In: Harmonic Analysis in Euclidean Spaces (Proc. Sympos. Pure Math., Williams Coll., Williamstown, Mass., 1978) Proc. Sympos. Pure Math., XXXV, Part, Amer. Math. Soc., pp. 179–196, Providence, RI (1979)
13. Fabes, E.B., Rivière, N.M.: Symbolic calculus of kernels with mixed homogeneity. In: Calderon, A.P., (ed.) Singular Integrals. Proc. Symp. Pure. Math., vol. 10, A.M.S., Providence, pp. 106–127 (1967)
14. Fabes, E.B., Rivière, N.M.: Singular integrals with mixed homogeneity. Studia Math. **27**, 19–38 (1966)
15. Fabes, E.B., Salsa, S.: Estimates of caloric measure and the initial-Dirichlet problem for the heat equation in Lipschitz cylinders. Trans. Am. Math. Soc. **279**, 635–650 (1983)

16. Fefferman, C., Stein, E.M.: H^p spaces of several variables. Acta Math. **129**, 137–193 (1972)
17. Grau de la Herran, A., Hofmann, S.: Generalized local Tb theorems for square functions. Mathematika. **63**(1), 1–28 (2017)
18. Hofmann, S.: Parabolic singular integrals of Calderon-type, rough operators and caloric layer potentials. Duke Math. J. **90**(2), 209–259 (1997)
19. Hofmann, S.: Local T(b) theorems and applications in PDE. In: Harmonic Analysis and Partial Differential Equations. Contemp. Math., vol. 505, pp. 29–52. Amer. Math. Soc, Providence, RI (2010)
20. Hofmann, S., Lacey, M., McIntosh, A.: The solution of the Kato problem for divergence form elliptic operators with Gaussian heat kernel bounds. Ann. Math. **156**, 623–631 (2002)
21. Hofmann, S., Lewis, J.L.: L^2 solvability and representation by caloric layer potentials in time-varying domains. Ann. Math. (2) **144**(2), 349–420 (1996)
22. Hofmann, S., Lewis, J.L.: Square functions of Calderon type and applications. Rev. Mat. Iberoam. **17**(1), 1–20 (2001)
23. Hofmann, S., Lewis, J.L.: The Dirichlet Problem for Parabolic Operators with Singular Drift Terms, vol. 151. Mem. Amer. Math. Soc. 719 (2001)
24. Hofmann, S., Mitrea, M., Morris, A.: The method of layer potentials in L^p and endpoint spaces for elliptic operators with L^∞ coefficients. Proc. Lond. Math. Soc. **111**(3), 681–716 (2015)
25. Hofmann, S., McIntosh, A.: Boundedness and applications of singular integrals and square functions: a survey. Bull. Math. Sci. **1**(2), 201–244 (2011)
26. Hofmann, S., Mayboroda, S., Mourgoglou, M.: L^p and endpoint solvability results for divergence form elliptic equations with complex L^∞ coefficients. Adv. Math. **270**, 480–564 (2015)
27. Jerison, D., Kenig, C.: The Dirichlet problem in non-smooth domains. Ann. Math. **113**, 367–382 (1981)
28. Kenig, C., Pipher, J.: The Neumann problem for elliptic equations with nonsmooth coefficients. Invent. Math. **113**(3), 447–509 (1993)
29. Kenig, C., Pipher, J.: The Neumann problem for elliptic equations with nonsmooth coefficients. II. A celebration of John F. Nash, Jr. Duke Math. J. **81**(1), 227–250 (1996)
30. Lewis, J.L., Murray, M.: The method of layer potentials for the heat equation in time-varying domains. Mem. Am. Math. Soc. **114**(545), viii+157 (1995)
31. Nazarov, F., Treil, S., Volberg, A. Weak type estimates and Cotlar inequalities for Calderon–Zygmund operators on nonhomogeneous spaces. Int. Math. Res. Not. **9**, 463–487 (1998)
32. Nyström, K.: Square functions estimates and the Kato problem for second order parabolic operators in $\mathbb{R}^{n+1}$. Adv. Math. **293**, 1–36 (2016)
33. Nyström, K.: L^2 solvability of boundary value problems for divergence form parabolic equations with complex coefficients. Submittted December 2015
34. Rosen, A.: Layer potentials beyond singular integral operators. Publicacions Matematiques **57**(2), 429–454 (2013)