ANALYSIS & PDE

Volume 18 No. 1 2025

ALIREZA ATAEI, MORITZ EGERT AND KAJ NYSTRÖM

THE KATO SQUARE ROOT PROBLEM FOR WEIGHTED PARABOLIC OPERATORS

msp

ANALYSIS AND PDE
Vol. 18 (2025), No. 1, pp. 141–169
DOI: 10.2140/apde.2025.18.141

THE KATO SQUARE ROOT PROBLEM FOR WEIGHTED PARABOLIC OPERATORS

ALIREZA ATAEI, MORITZ EGERT AND KAJ NYSTRÖM

We give a simplified and direct proof of the Kato square root estimate for parabolic operators with elliptic part in divergence form and coefficients possibly depending on space and time in a merely measurable way. The argument relies on the nowadays classical reduction to a quadratic estimate and a Carleson-type inequality. The precise organization of the estimates is different from earlier works. In particular, we succeed in separating space and time variables almost completely despite the nonautonomous character of the operator. Hence, we can allow for degenerate ellipticity dictated by a spatial A_2-weight, which has not been treated before in this context.

1. Introduction and main result

In the variables $(x, t) \in \mathbb{R}^n \times \mathbb{R} =: \mathbb{R}^{n+1}$, we consider parabolic operators of the form

$$\mathcal{H}u := \partial_t u - w^{-1} \operatorname{div}_x(A \nabla_x u), \tag{1-1}$$

where the weight $w = w(x)$ is time-independent and belongs to the spatial Muckenhoupt class $A_2(\mathbb{R}^n, \mathrm{d}x)$, and the coefficient matrix $A = A(x, t)$ is measurable with complex entries and possibly depends on all variables. Degeneracy is dictated by the same weight w in the sense that $w^{-1}A$ satisfies the classical uniform ellipticity condition (Section 2.3).

Weighted parabolic operators as in (1-1) occur in various contexts and applications, including the study of fractional powers [Litsgård and Nyström 2023] and heat kernels of Schrödinger equations with singular potential [Ishige et al. 2017]. For contributions to the study of local properties of solutions to $\mathcal{H}u = 0$ and Gaussian estimates, we refer to [Chiarenza and Serapioni 1985; Cruz-Uribe and Rios 2008].

The purpose of this paper is to establish the Kato (square root) estimate for $\mathcal{H}$, that is, to prove Theorem 1.1 stated below. We write $\mathrm{L}^2_\mu = \mathrm{L}^2(\mathbb{R}^{n+1}, \mathrm{d}w\, \mathrm{d}t)$, $\mathrm{d}\mu = \mathrm{d}w\, \mathrm{d}t = w(x)\, \mathrm{d}x\, \mathrm{d}t$, for the natural weighted Lebesgue space associated with $\mathcal{H}$, and E_μ for the corresponding first-order parabolic Sobolev space of functions u such that the spatial gradient $\nabla_x u$, as well as the half-order time derivative $D_t^{1/2} u$, is in L^2_μ. For the sake of the introduction, an intuitive interpretation of these objects suffices. We turn to rigorous definitions in Section 3.

MSC2020: primary 35K10, 42B25, 42B37; secondary 26A33, 47B44.
Keywords: Kato square root problem, second-order parabolic operators, Muckenhoupt weights, square function estimates, fractional derivatives.

© 2025 MSP (Mathematical Sciences Publishers). Distributed under the Creative Commons Attribution License 4.0 (CC BY).
Open Access made possible by subscribing institutions via Subscribe to Open.

Theorem 1.1. *The operator $\mathcal{H}$ can be defined as an m-accretive operator in L^2_μ associated with an accretive sesquilinear form with domain E_μ. The domain of its unique m-accretive square root is the same as the form domain, that is $\mathsf{D}(\sqrt{\mathcal{H}}) = \mathrm{E}_\mu$, and*

$$\|\sqrt{\mathcal{H}}u\|_{\mathrm{L}^2_\mu} \sim \|\nabla_x u\|_{\mathrm{L}^2_\mu} + \|D_t^{1/2}u\|_{\mathrm{L}^2_\mu}, \quad u \in \mathrm{E}_\mu,$$

holds with an implicit constant that depends on the dimension, the ellipticity parameters of A and the A_2-constant for w.

The time derivative ∂_t is a skew-adjoint operator, and hence there are no lower bounds for the formal pairing $\operatorname{Re}\langle \mathcal{H}u, u\rangle$ that contain derivatives in t. However, when the time variable describes the full real line, parabolic operators admit some "hidden coercivity" that can be revealed through the Hilbert transform H_t in the t-variable. Indeed, splitting $\partial_t = D_t^{1/2} H_t D_t^{1/2}$, the sesquilinear form associated with (1-1) over L^2_μ is given by

$$B(u, v) := \iint_{\mathbb{R}^{n+1}} w^{-1} A\nabla_x u \cdot \overline{\nabla_x v} + H_t D_t^{1/2} u \cdot \overline{D_t^{1/2} v} \,\mathrm{d}w\,\mathrm{d}t, \quad u, v \in \mathrm{E}_\mu, \tag{1-2}$$

and lower bounds including both time and space derivatives become available when taking $v = (1+\delta H_t)u$ with $\delta > 0$ small. This observation is originally due to Kaplan [1966] and has been rediscovered several times ever since; see [Dier and Zacher 2017; Hofmann and Lewis 2005; Nyström 2016] for example. M-accretivity of $\mathcal{H}$ essentially follows from this observation, but to the best of our knowledge an explicit statement, in the unweighted case $w = 1$, only appeared much later in [Auscher and Egert 2016]. For the reader's convenience, we reproduce the full argument in our weighted setting in Section 4. Being m-accretive, $\mathcal{H}$ admits a sectorial functional calculus and in particular a (unique) m-accretive square root $\sqrt{\mathcal{H}}$; see [Haase 2006; Kato 1966] for background. This is how our main result should be understood.

The pursuit of the solution of the Kato problem for unweighted elliptic operators (finally completed in [Auscher et al. 2002]) introduced new techniques that proved extremely viable for extensions and applications to other problems in harmonic analysis and partial differential equations [Amenta and Auscher 2018; Alfonseca et al. 2011; Auscher and Axelsson 2011; Auscher and Mourgoglou 2019; Auscher and Rosén 2012; Auscher et al. 2018; Castro et al. 2016; Escauriaza and Hofmann 2018; Hofmann et al. 2015; 2019; 2022; Nyström 2017]. For this reason, Kato-type estimates for different operators are desirable, and the results of this paper most surely have important implications for, and applications to, boundary value problems for weighted second-order parabolic operators.

Let us mention that the case of A_2-weighted elliptic operators was settled in [Cruz-Uribe and Rios 2015], see also [Cruz-Uribe et al. 2018] for an extension, and rediscovered in the more general framework of first-order Dirac operators in [Auscher et al. 2015]. The third author was first to develop the underlying harmonic analysis in the unweighted parabolic setting in [Nyström 2016], and in the same paper he proved the square function estimates that are essentially equivalent to Theorem 1.1 when $w \equiv 1$ and when the coefficients A are t-independent. Using a framework of parabolic Dirac operators, Auscher, together with the second and third authors, obtained the unweighted parabolic case when coefficients depend measurably on x *and* t [Auscher et al. 2020]. Our Theorem 1.1 completes this succession of results but there is more to it and that makes, as we shall discuss next, the present paper interesting even in the unweighted case.

Under the assumption $A = A(x)$ in [Nyström 2016], the operator $\mathcal{H}$ is an autonomous parabolic operator, and, in retrospect, the main result of that paper could have been obtained by interpolation from maximal regularity of the Cauchy problem for (1-1); see [Ouhabaz 2021]. (In fact, this argument requires only smoothness of order $\frac{1}{2}$ for the coefficients in the t-variable.) However, many of the techniques in [Nyström 2016], such as the parabolic off-diagonal estimates and the construction of Tb-type test functions, had already been strong enough for proving the parabolic Kato estimate in the presence of measurable t-dependence, and our proof of Theorem 1.1 shows exactly how, thereby making our result novel in at least two ways:

- We generalize all previous findings in the parabolic setting by combining measurable dependence of the coefficients on all variables with A_2-weighted degeneracy in space.
- We avoid the Dirac operator framework in [Auscher et al. 2020]. The resulting "second-order" approach for parabolic operators with time-dependent measurable coefficients has not appeared in the literature before, and, when restricted to the unweighted case $w \equiv 1$, it provides a significant simplification of the proof of [Auscher et al. 2020, Theorem 2.6].

Our ambition is to present an almost self-contained argument using only a minimal number of tools. We do not attempt to generalize all further results in [Auscher et al. 2020] to the weighted setting, which should be done by developing a parabolic weighted Dirac operator framework.

As is customary in the field, see [Auscher et al. 2002; Cruz-Uribe and Rios 2015; Hofmann et al. 2022; Nyström 2016], the first reduction in the proof of Theorem 1.1 is to use the bounded H^∞-calculus for m-accretive operators and a duality argument in order to reduce the matter to the one-sided quadratic estimate

$$\int_0^\infty \|\lambda\mathcal{H}(1+\lambda^2\mathcal{H})^{-1}u\|_{\mathrm{L}^2_\mu}^2 \frac{\mathrm{d}\lambda}{\lambda} \lesssim \|\nabla_x u\|_{\mathrm{L}^2_\mu}^2 + \|D_t^{1/2}u\|_{\mathrm{L}^2_\mu}^2, \quad u \in \mathrm{E}_\mu. \tag{1-3}$$

In contrast to the elliptic setting, this reduction does not follow immediately from classical results à la Kato and Lions [Lions 1962], since the sesquilinear form B in (1-2) is *not* closed due to the lack of lower bounds by half-order time derivatives. Some more care is needed but we settle the issue in Section 6. The quadratic estimate (1-3) is then achieved by slightly refining the techniques in [Nyström 2016], and the argument relies on (weighted) Littlewood–Paley theory in L^2 (Section 5), which eventually reduces matters to a Carleson measure estimate that can be proved through a Tb-procedure (Section 8).

It came as a surprise to us that, even though coefficients may depend measurably on all variables, the proof of (1-3) can be arranged in a way that almost completely separates time and space variables. This observation incarnates in three different stages of the proof:

- At the level of Littlewood–Paley theory, it suffices to use weighted elliptic theory in x and classical Fourier analysis in t. The required weighted theory has already been developed in detail by Cruz-Uribe and Rios [2012; 2015] in order to solve the weighted elliptic Kato problem.

- At the level of off-diagonal bounds (averaged "kernel" bounds, see Section 4.4), we only need estimates for operators involving differentiation in space. These estimates can be deduced directly from the equation and come with parabolic scaling. The much more involved off-diagonal decay and Poincaré inequalities for nonlocal derivatives such as $D_t^{1/2}$, which were fundamental novelties in [Auscher et al. 2020], can be avoided.
- At the level of the Tb-argument, the test functions can be constructed based on a product structure, which makes the argument more straightforward compared to the system of functions used in [Auscher et al. 2020].

These three observations have in common that we can regroup derivatives of resolvents of $\mathcal{H}$ in such a way that fine harmonic analysis estimates need only apply to the spatial parts, whereas the t-derivatives appear in blocks that are amenable for simple resolvent estimates in L^2_μ-norm. We shall indicate the most striking examples of this principle along with the proof of the Carleson measure estimate in the final section.

The next section contains some preliminary notation and conventions. The rest of the paper follows the outline above.

2. Preliminaries and basic assumptions

Given $(x, t) \in \mathbb{R}^n \times \mathbb{R}$, we let $\|(x, t)\| := \max\{|x|, |t|^{1/2}\}$. We call $\|(x, t)\|$ the parabolic norm of (x, t). Given a half-open cube $Q = \left(x - \frac{1}{2}r, x + \frac{1}{2}r\right]^n \subset \mathbb{R}^n$ parallel to the coordinate axes with sidelength r and an interval $I = \left(t - \frac{1}{2}r^2, t + \frac{1}{2}r^2\right]$, we call $\Delta := Q \times I \subset \mathbb{R}^{n+1}$ a parabolic cube of size r. Occasionally, we write $\Delta_r(x, t) = Q_r(x) \times I_r(t)$ and $r = \ell(\Delta)$ to indicate the center and size directly. A dyadic parabolic cube of size 2^j is by definition centered in $(2^j\mathbb{Z})^n \times (4^j\mathbb{Z})$. For every $c > 0$, and given Δ, we define $c\Delta$ as the parabolic cube with the same center as Δ and size $c\ell(\Delta)$.

2.1. ***Assumptions and notation concerning the weight.*** For general background and the results cited here, we refer to [Stein 1993, Chapter V]. The weight $w = w(x)$ is a real-valued function belonging to the Muckenhoupt class $A_2(\mathbb{R}^n, \mathrm{d}x)$, that is,

$$[w]_{A_2} := \sup_Q \left(\fint_Q w \,\mathrm{d}x \right)\left(\fint_Q w^{-1} \,\mathrm{d}x \right) < \infty, \tag{2-1}$$

where the supremum is taken with respect to all cubes $Q \subset \mathbb{R}^n$. We introduce the measure $\mathrm{d}w(x) := w(x)\,\mathrm{d}x$ on $\mathbb{R}^n$, and we write $w(E) = \int_E \mathrm{d}w$ for all Lebesgue measurable sets $E \subset \mathbb{R}^n$. For averages, we use the notation

$$(g)_{E,w} := \fint_E g(x)\,\mathrm{d}w(x) := \frac{1}{w(E)} \int_E g(x)w(x)\,\mathrm{d}x$$

if $w(E) \in (0, \infty)$ and g is locally integrable on $\mathbb{R}^n$ with respect to $\mathrm{d}w(x)$. It follows from (2-1) that there are constants $\eta \in (0, 1)$ and $\beta > 0$, depending only on n and $[w]_{A_2}$, such that

$$\beta^{-1}\left(\frac{|E|}{|Q|}\right)^{1/(2\eta)} \le \frac{w(E)}{w(Q)} \le \beta\left(\frac{|E|}{|Q|}\right)^{2\eta} \tag{2-2}$$

whenever $E \subset Q$ is measurable, and where $|\cdot|$ denotes Lebesgue measure in $\mathbb{R}^n$. In particular, there exists a constant D depending only on $[w]_{A_2}$ and n, called the doubling constant for w, such that

$$w(2Q) \le Dw(Q) \quad \text{for all cubes } Q \subset \mathbb{R}^n. \tag{2-3}$$

The measures

$$\begin{aligned} \mathrm{d}\mu = \mathrm{d}\mu(x,t) &:= w(x)\,\mathrm{d}x\,\mathrm{d}t, \\ \mathrm{d}\mu^{-1} = \mathrm{d}\mu^{-1}(x,t) &:= w(x)^{-1}\,\mathrm{d}x\,\mathrm{d}t \end{aligned} \tag{2-4}$$

are defined on $\mathbb{R}^{n+1}$. Naturally, μ and μ^{-1} can be seen as measures on $\mathbb{R}^{n+1}$ defined by $A_2(\mathbb{R}^{n+1}, \mathrm{d}x\,\mathrm{d}t)$ weights, and in the context of these measures we use similar notation as above. The doubling constant for μ with respect to parabolic scaling is $4D$.

2.2. ***Maximal functions.*** We introduce the maximal operators in the individual variables

$$\begin{aligned} \mathcal{M}^{(1)}(g_1)(x) &:= \sup_{r>0} \fint_{Q_r(x)} |g_1|\,\mathrm{d}x, \\ \mathcal{M}^{(2)}(g_2)(t) &:= \sup_{r>0} \fint_{I_r(t)} |g_2|\,\mathrm{d}t \end{aligned}$$

for all locally integrable functions g_1 and g_2 on $\mathbb{R}^n$ and $\mathbb{R}$, respectively. The operator $\mathcal{M}^{(1)}$ is bounded on the weighted space $\mathrm{L}^2(\mathbb{R}^n, \mathrm{d}w)$ with a bound depending on $[w]_{A_2}$ and n [Stein 1993, Theorem 1, p. 201]. Both maximal operators can be naturally extended to $\mathrm{L}^2(\mathbb{R}^{n+1}, \mathrm{d}\mu)$ by keeping one of the variables fixed, and they are bounded in this setting.

2.3. ***Assumptions on the coefficients.*** The matrix-valued function

$$A = A(x,t) = \{A_{i,j}(x,t)\}_{i,j=1}^n$$

is assumed to have complex measurable entries $A_{i,j}$ that satisfy

$$c_1|\xi|^2 w(x) \le \mathrm{Re}(A(x,t)\xi \cdot \bar{\xi}), \quad |A(x,t)\xi \cdot \zeta| \le c_2 w(x)|\xi||\zeta| \tag{2-5}$$

for some $c_1, c_2 \in (0,\infty)$ and for all $\xi, \zeta \in \mathbb{C}^n$, $(x,t) \in \mathbb{R}^{n+1}$. Here, $u \cdot v = u_1 v_1 + \cdots + u_n v_n$, and $\bar{u}$ denotes the complex conjugate of u so that $u \cdot \bar{v}$ is the standard inner product on $\mathbb{C}^n$. We refer to c_1, c_2 as the ellipticity constants of A. Assumption (2-5) is equivalent to saying that $w^{-1}A$ satisfies the classical uniform ellipticity condition.

2.4. ***Floating constants.*** We refer to n and the constants $[w]_{A_2}$, c_1, c_2, appearing in (2-1) and (2-5), as structural constants. For $A, B \in (0,\infty)$, the notation $A \lesssim B$ means that $A \le cB$ for some c depending at most on the structural constants. The notation $A \gtrsim B$ and $A \sim B$ should be interpreted similarly.

3. Weighted function spaces

In this section we give a brief account of the relevant weighted function spaces. We let $\mathrm{L}^2_w = \mathrm{L}^2(\mathbb{R}^n, \mathrm{d}w)$ be the Hilbert spaces of square integrable functions with respect to $\mathrm{d}w$. Its norm is denoted by $\|\cdot\|_{2,w}$, its inner product by $\langle \cdot , \cdot \rangle_{2,w}$, and the operator norm of linear operators on that space by $\| \cdot \|_{2\to 2,w}$. Thanks to the A_2-condition, we have

$$\mathrm{L}^2_w \subset \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^n, \mathrm{d}x), \tag{3-1}$$

and the class $\mathrm{C}^\infty_0(\mathbb{R}^n)$ of smooth and compactly supported test functions is dense in L^2_w via the usual truncation and convolution procedure [Kilpeläinen 1994, Section 1]. The same notation and properties apply to L^2_μ in $\mathbb{R}^{n+1}$.

Definition 3.1 (elliptic weighted Sobolev space). We write $\mathrm{H}^1_w := \mathrm{H}^1_w(\mathbb{R}^n)$ for the space of all $f \in \mathrm{L}^2_w$ for which the distributional gradient $\nabla_x f$ is (componentwise) in L^2_w, and we equip the space with the norm $\| \cdot \|_{\mathrm{H}^1_w} := (\| \cdot \|^2_{2,w} + \|\nabla_x \cdot \|^2_{2,w})^{1/2}$.

By construction H^1_w is a Hilbert space, and standard truncation and convolution techniques yield that $\mathrm{C}^\infty_0(\mathbb{R}^n)$ is dense in H^1_w; see [Kilpeläinen 1994, Theorem 2.5].

In order to define parabolic function spaces, we use the Fourier transform $\mathcal{F}$ in the time variable, keeping in mind that if $f \in \mathrm{L}^2(\mathbb{R}^{n+1}, \mathrm{d}\mu)$, then $f(x, \cdot) \in \mathrm{L}^2(\mathbb{R}, \mathrm{d}t)$ for a.e. $x \in \mathbb{R}^n$. The corresponding Fourier variable will be denoted by τ. Then,

$$H_t f := \mathcal{F}^{-1}(\mathrm{i}\,\mathrm{sgn}(\tau)\mathcal{F} f)$$

is our Hilbert transform. If $|\tau|^{1/2}\mathcal{F} f \in \mathrm{L}^2_\mu$, then we define the half-order time derivative

$$D^{1/2}_t f := \mathcal{F}^{-1}(|\tau|^{1/2}\mathcal{F} f),$$

and this is what we mean when we write $D^{1/2}_t f \in \mathrm{L}^2_\mu$. Using a classical formula for fractional Laplacians for a.e. fixed $x \in \mathbb{R}^n$, see [Di Nezza et al. 2012] for example, we obtain

$$\|D^{1/2}_t f\|^2_{2,\mu} = \frac{2}{\pi} \int_{\mathbb{R}^n} \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{|f(x,t) - f(x,s)|^2}{|t-s|^2} \,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}w(x), \tag{3-2}$$

with the right-hand side being finite precisely when $D^{1/2}_t f \in \mathrm{L}^2_\mu$.

Definition 3.2 (parabolic energy space). We write $\mathrm{E}_\mu := \mathrm{E}_\mu(\mathbb{R}^{n+1})$ for the space of all $f \in \mathrm{L}^2_\mu$ for which $\nabla_x f$, $D^{1/2}_t f \in \mathrm{L}^2{}_\mu$, and we equip the space with the norm

$$\| \cdot \|_{\mathrm{E}_\mu} := (\| \cdot \|^2_{2,\mu} + \|\nabla_x \cdot \|^2_{2,\mu} + \|D^{1/2}_t \cdot \|^2_{2,\mu})^{1/2}.$$

For $f \in \mathrm{E}_\mu$, we will refer to the vector $\mathbb{D} f := (\nabla_x f, D^{1/2}_t f)$ as the parabolic gradient of f.

Again, E_μ is a Hilbert space. Note that, in the unweighted setting of [Nyström 2016], the notation $\mathbb{D}$ has a slightly different meaning.

Lemma 3.3. *The following statements are true*:

(i) *The space* $\mathrm{C}^\infty_0(\mathbb{R}^{n+1})$ *is dense in* $\mathrm{E}_\mu(\mathbb{R}^{n+1})$.

(ii) *Multiplication by* $\mathrm{C}^1(\mathbb{R}^{n+1})$*-functions is bounded on* $\mathrm{E}_\mu(\mathbb{R}^{n+1})$.

Proof. We begin with (i). If $f \in \mathrm{E}_\mu$, then convolutions with smooth mollifiers, separately in x and t, provide smooth approximations in E_μ. For the convolution in space, this argument uses the A_2-condition on w as mentioned above. Hence, it suffices to approximate f by compactly supported functions in E_μ. To this end, we can follow the standard pattern of smooth truncation: We pick a sequence $(\eta_j)_j \subset \mathrm{C}_0^\infty(\mathbb{R}^{n+1})$ such that $\eta_j \to 1$ pointwise a.e. as $j \to \infty$, $\|\eta_j\|_\infty + j\|\nabla_x \eta_j\|_\infty + j\|\partial_t \eta_j\|_\infty \le c$ uniformly in j, and then we set $f_j := \eta_j f$. By dominated convergence, we obtain $f_j \to f$ and $\nabla_x f_j \to \nabla_x f$ in L^2_μ as $j \to \infty$. For the half-order derivative, we use (3-2) with $f_j - f$ in place of f. We first bound the integrand in (3-2) by

$$\begin{aligned}
&\frac{|(f_j - f)(x,t) - (f_j - f)(x,s)|^2}{|t-s|^2} \\
&\qquad \le 2\frac{|(\eta_j - 1)(x,t) - (\eta_j - 1)(x,s)|^2}{|t-s|^2}|f(x,t)|^2 + 2\frac{|f(x,t) - f(x,s)|^2}{|t-s|^2}|(\eta_j - 1)(x,s)|^2 \\
&\qquad \le 2\min\left\{c^2, \frac{4(c+1)^2}{|t-s|^2}\right\}|f(x,t)|^2 + 2(c+1)^2\frac{|f(x,t) - f(x,s)|^2}{|t-s|^2}.
\end{aligned} \tag{3-3}$$

The right-hand side is independent of j and integrable with respect to $\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}w(x)$. Since the middle term tends to 0 a.e. as $j \to \infty$, we conclude

$$\|D_t^{1/2}(f_j - f)\|_{2,\mu} \to 0$$

by dominated convergence. This completes the proof of (i).

As for (ii), we note that if $\eta \in \mathrm{C}^1(\mathbb{R}^{n+1})$ and $f \in \mathrm{E}_\mu$, then

$$\begin{aligned}
\|\eta f\|_{2,\mu} &\le \|\eta\|_\infty \|f\|_{2,\mu}, \\
\|\nabla_x(\eta f)\|_{2,\mu} &\le \|\eta\|_\infty \|\nabla_x f\|_{2,\mu} + \|\nabla_x \eta\|_\infty \|f\|_{2,\mu}, \\
\|D_t^{1/2}(\eta f)\|_{2,\mu} &\le \sqrt{8}\|\eta\|_\infty^{1/2}\|\partial_t \eta\|_\infty^{1/2}\|f\|_{2,\mu} + \|\eta\|_\infty \|D_t^{1/2} f\|_{2,\mu},
\end{aligned}$$

where the third line follows by the same splitting as in (3-3), but with η in place of $1 - \eta_j$. □

Lemma 3.3 (i) implies the chain of continuous and dense embeddings

$$\mathrm{E}_\mu \subset \mathrm{L}^2_\mu \simeq (\mathrm{L}^2_\mu)^* \subset (\mathrm{E}_\mu)^*, \tag{3-4}$$

where we use the upper star to denote (anti)-dual spaces. We have bounded operators

$$\begin{aligned}
D_t^{1/2} &: \mathrm{E}_\mu \to \mathrm{L}^2_\mu, \\
\nabla_x &: \mathrm{E}_\mu \to (\mathrm{L}^2_\mu)^n,
\end{aligned} \tag{3-5}$$

and we denote their adjoints with respect to (3-4) by

$$\begin{aligned}
D_t^{1/2} &: \mathrm{L}^2_\mu \to (\mathrm{E}_\mu)^*, \\
w^{-1}\operatorname{div}_x w &: (\mathrm{L}^2_\mu)^n \to (\mathrm{E}_\mu)^*.
\end{aligned} \tag{3-6}$$

Note carefully that $w^{-1}\operatorname{div}_x w$ is only a suggestive notation reflecting the formal action of this operator. In general, there is no guarantee that this operator splits into a composition of its three building blocks.

4. The parabolic operator

We continue by introducing the formal parabolic operator in (1-1) rigorously as an unbounded operator in the Hilbert space L^2_μ associated with a sesquilinear form.

Denoting by H_t the Hilbert transform in the t-variable and by $D_t^{1/2}$ the half-order time derivative as defined in Section 3, we can factorize

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

By (3-5) and (3-6), we have $\partial_t : \mathrm{E}_\mu \to (\mathrm{E}_\mu)^*$. We define $\mathcal{H}$ as a bounded operator $\mathrm{E}_\mu \to (\mathrm{E}_\mu)^*$ via

$$(\mathcal{H}u)(v) := B(u, v) := \iint_{\mathbb{R}^{n+1}} w^{-1} A\nabla_x u \cdot \overline{\nabla_x v} + H_t D_t^{1/2} u \cdot \overline{D_t^{1/2} v} \, \mathrm{d}\mu, \quad u, v \in \mathrm{E}_\mu. \tag{4-1}$$

In view of (3-4), it makes sense to consider the maximal restriction of $\mathcal{H}$ to an operator in L^2_μ, called the part of $\mathcal{H}$ in L^2_μ, with domain

$$\mathsf{D}(\mathcal{H}) := \{u \in \mathrm{E}_\mu(\mathbb{R}^{n+1}) : \mathcal{H}u \in \mathrm{L}^2_\mu(\mathbb{R}^{n+1})\}. \tag{4-2}$$

If $u \in \mathsf{D}(\mathcal{H})$, we have, for all $v \in \mathrm{E}_\mu$, that

$$(\mathcal{H}u)(v) = \iint_{\mathbb{R}^{n+1}} \mathcal{H}u \cdot \bar{v} \, \mathrm{d}\mu,$$

and a formal integration by parts in (4-1) reveals that it is indeed justified to say that the part of $\mathcal{H}$ in L^2_μ gives meaning to the formal expression in (1-1). More precisely, in terms of (3-5) and (3-6), we have that $\mathcal{H} : \mathrm{E}_\mu \to (\mathrm{E}_\mu)^*$ acts as the composition of operators

$$\mathcal{H} = D_t^{1/2} H_t D_t^{1/2} - (w^{-1} \operatorname{div}_x w)(w^{-1} A\nabla_x). \tag{4-3}$$

4.1. ***Hidden coercivity.*** The following lemma relies on the hidden coercivity (proved by Kaplan [1966]) of the parabolic sesquilinear form B in (4-1) that can be revealed through the Hilbert transform.

Lemma 4.1. *Let $\sigma \in \mathbb{C}$ with $\operatorname{Re}\sigma > 0$. For each $f \in (\mathrm{E}_\mu)^*$, there exists a unique $u \in \mathrm{E}_\mu$ such that $(\sigma + \mathcal{H})u = f$. Moreover,*

$$\|u\|_{\mathrm{E}_\mu} \le \sqrt{2} \max\left\{\frac{c_2 + 1}{c_1}, \frac{|\operatorname{Im}\sigma| + 1}{\operatorname{Re}\sigma}\right\} \|f\|_{(\mathrm{E}_\mu)^*}.$$

Proof. By Plancherel's theorem, the Hilbert transform H_t is isometric on E_μ. Hence, we can define a "twisted" sesquilinear form $B_{\delta,\sigma} : \mathrm{E}_\mu \times \mathrm{E}_\mu \to \mathbb{C}$ via

$$B_{\delta,\sigma}(u, v) := \iint_{\mathbb{R}^{n+1}} \Big(\sigma u \cdot \overline{(1 + \delta H_t)v} + w^{-1} A\nabla_x u \cdot \overline{\nabla_x (1 + \delta H_t)v} + H_t D_t^{1/2} u \cdot \overline{D_t^{1/2}(1 + \delta H_t)v}\Big) \, \mathrm{d}\mu, \tag{4-4}$$

where $\delta \in (0, 1)$ is to be chosen. Clearly $B_{\delta,\sigma}$ is bounded. Since H_t is skew-adjoint, we have

$$\operatorname{Re} \iint_{\mathbb{R}^{n+1}} H_t v \cdot \bar{v} \, \mathrm{d}\mu = 0, \quad v \in \mathrm{L}^2_\mu. \tag{4-5}$$

Expanding $B_{\delta,\sigma}(u,u)$ and using the above along with the weighted ellipticity of the coefficients A, we find

$$\operatorname{Re} B_{\delta,\sigma}(u,u) \geq \delta \|D_t^{1/2} u\|_{2,\mu}^2 + (c_1 - c_2\delta)\|\nabla_x u\|_{2,\mu}^2 + (\operatorname{Re}\sigma - \delta|\operatorname{Im}\sigma|)\|u\|_{2,\mu}^2. \tag{4-6}$$

Choosing $\delta = \min\{c_1/(c_2+1), \operatorname{Re}\sigma/(|\operatorname{Im}\sigma|+1)\}$, the factors in front of the second and third term in the last display become no less than δ. Hence, we obtain the coercivity estimate

$$\operatorname{Re} B_{\delta,\sigma}(u,u) \geq \min\left\{\frac{c_1}{c_2+1}, \frac{\operatorname{Re}\sigma}{|\operatorname{Im}\sigma|+1}\right\}\|u\|_{\mathrm{E}_\mu}^2, \quad v \in \mathrm{E}_\mu. \tag{4-7}$$

The Lax–Milgram lemma yields, for each $f \in (\mathrm{E}_\mu)^*$, a unique $u \in \mathrm{E}_\mu$ satisfying the estimate claimed in the lemma such that

$$B_{\delta,\sigma}(u,v) = f((1+\delta H_t)v), \quad v \in \mathrm{E}_\mu.$$

(Note that the additional factor $\sqrt{2}$ is an upper bound for the norm of $1+\delta H_t$ on E_μ.) Plancherel's theorem yields that $1+\delta H_t$ is an isomorphism on E_μ for all $\delta \in \mathbb{R}$. Thus,

$$\iint_{\mathbb{R}^{n+1}} \sigma u \cdot \bar{v} + w^{-1} A\nabla_x u \cdot \overline{\nabla_x v} + H_t D_t^{1/2} u \cdot \overline{D_t^{1/2} v} \, \mathrm{d}\mu = f(v), \quad v \in \mathrm{E}_\mu,$$

that is, $(\sigma + \mathcal{H})u = f$ as required. □

The proof above fails for $\delta = 0$ since $\operatorname{Re} B(\cdot\,,\cdot)$ does not control $\|D_t^{1/2} \cdot\|_{2,\mu}$ from above. As a consequence, B itself is *not* a closed sesquilinear form in the sense of Kato [1966] or, equivalently, $(\|\cdot\|_{2,\mu}^2 + \operatorname{Re} B(\cdot\,,\cdot))^{1/2}$ does *not* define an equivalent norm on E_μ. In [Auscher and Egert 2016, Lemma 4], it has been (essentially) shown that a parabolic analog of Kato's first representation theorem holds nonetheless. For convenience, we include the short argument with some minor improvements in the next section.

4.2. ***M-accretivity.*** Recall that an operator $\mathcal{H}$ in a Hilbert space such as L^2_μ is called *m-accretive* if it is closed and densely defined, with resolvent estimates

$$\|(\sigma+\mathcal{H})^{-1}\|_{2\to 2,\mu} \leq (\operatorname{Re}\sigma)^{-1}, \quad \sigma \in \mathbb{C}, \ \operatorname{Re}\sigma > 0.$$

Proposition 4.2. *The part of $\mathcal{H}$ in L^2_μ is m-accretive and $\mathsf{D}(\mathcal{H})$ is dense in E_μ.*

Proof. Fix $\sigma \in \mathbb{C}$ with $\operatorname{Re}\sigma > 0$. Lemma 4.1 yields that $\sigma + \mathcal{H} : \mathsf{D}(\mathcal{H}) \to \mathrm{L}^2_\mu$ is bijective. Given $f \in \mathrm{L}^2_\mu$, we set $u := (\sigma+\mathcal{H})^{-1} f$ and use ellipticity of the coefficients A and (4-5) to deduce

$$\begin{aligned}\operatorname{Re}\sigma\|u\|_{2,\mu}^2 &\leq \operatorname{Re}\iint_{\mathbb{R}^{n+1}} \sigma u \cdot \bar{u} + w^{-1} A\nabla_x u \cdot \overline{\nabla_x u} + H_t D_t^{1/2} u \cdot \overline{D_t^{1/2} u} \, \mathrm{d}\mu \\ &= \operatorname{Re}\iint_{\mathbb{R}^{n+1}} f \cdot \bar{u} \, \mathrm{d}\mu \leq \|f\|_{2,\mu}\|u\|_{2,\mu}.\end{aligned}$$

This gives the required resolvent bound $\|u\|_{2,\mu} \leq (\operatorname{Re}\sigma)^{-1}\|f\|_{2,\mu}$. Moreover, the part of $\mathcal{H}$ in L^2_μ is closed since it has a nonempty resolvent set, and the resolvent estimates for $\sigma > 0$ imply a dense domain [Haase 2006, Proposition 2.1.1]. This proves m-accretivity.

In order to prove that $\mathsf{D}(\mathcal{H})$ is dense in E_μ, we use the sesquilinear form $B_{\delta,1}$ in (4-4) with $\delta > 0$ chosen as in the proof of that lemma. Suppose $v \in \mathrm{E}_\mu$ is orthogonal to $\mathsf{D}(\mathcal{H})$ in E_μ. By the Lax–Milgram lemma, there is $w \in \mathrm{E}_\mu$ such that $\langle u, v\rangle_{\mathrm{E}_\mu} = B_{\delta,1}(u, w)$ for all $u \in \mathrm{E}_\mu$. For $u \in \mathsf{D}(H)$, this identity becomes $0 = \langle (1+\mathcal{H})u, (1+\delta H_t)w\rangle_{2,\mu}$, and since $1+\mathcal{H} : \mathsf{D}(\mathcal{H}) \to \mathrm{L}^2_\mu$ is bijective, we conclude that $(1+\delta H_t)w = 0$. Thus, we have $w = 0$ and therefore also $v = 0$. □

The adjoint $\mathcal{H}^*$ of $\mathcal{H}$ (seen as either a bounded operator $\mathrm{E}_\mu \to (\mathrm{E}_\mu)^*$ or an unbounded operator in L^2_μ) has the same properties as $\mathcal{H}$. Indeed, it can be checked by the very definition that it is associated with the sesquilinear form

$$B^*(u, v) = \overline{B(v, u)}$$

and that it formally corresponds to the backward-in-time operator

$$-\partial_t - w^{-1}(x)\,\mathrm{div}_x(A^*(x, t)\nabla_x).$$

Here A^* is the conjugate transpose of A.

4.3. ***Resolvent estimates.*** Using Proposition 4.2, we see that, for $\lambda > 0$, the resolvent operators

$$\begin{aligned} \mathcal{E}_\lambda &:= (I + \lambda^2\mathcal{H})^{-1}, \\ \mathcal{E}^*_\lambda &:= (I + \lambda^2\mathcal{H}^*)^{-1} \end{aligned} \tag{4-8}$$

are well defined as bounded operators $\mathrm{L}^2_\mu \to \mathrm{L}^2_\mu$ and $(\mathrm{E}_\mu)^* \to \mathrm{E}_\mu$. Moreover, they are adjoints of each other.

Lemma 4.3. *The following resolvent estimates hold uniformly for all $\lambda > 0$, all $f \in \mathrm{L}^2_\mu$ and all $\boldsymbol{f} \in (\mathrm{L}^2_\mu)^n$:*

(i) $\|\mathcal{E}_\lambda f\|_{2,\mu} + \|\lambda \mathbb{D}\mathcal{E}_\lambda f\|_{2,\mu} \lesssim \|f\|_{2,\mu}$,

(ii) $\|\lambda\mathcal{E}_\lambda D_t^{1/2} f\|_{2,\mu} + \|\lambda^2 \mathbb{D}\mathcal{E}_\lambda D_t^{1/2} f\|_{2,\mu} \lesssim \|f\|_{2,\mu}$,

(iii) $\|\lambda\mathcal{E}_\lambda w^{-1}\,\mathrm{div}_x(w\boldsymbol{f})\|_{2,\mu} + \|\lambda^2\mathbb{D}\mathcal{E}_\lambda w^{-1}\,\mathrm{div}_x(w\boldsymbol{f})\|_{2,\mu} \lesssim \|\boldsymbol{f}\|_{2,\mu}$.

The same estimates hold with $\mathcal{E}_\lambda$ replaced by $\mathcal{E}^*_\lambda$.

Proof. We first prove (i). Setting $u := (\lambda^{-2} + \mathcal{H})^{-1} f$, we have $\mathcal{E}_\lambda f = \lambda^{-2} u$, and by m-accretivity we obtain

$$\|\mathcal{E}_\lambda f\|_{2,\mu} \le \|f\|_{2,\mu}.$$

Next, we use the twisted sesquilinear form $B_{\delta,\sigma}$ as in (4-4) with parameter $\sigma = \lambda^{-2}$, so that by construction

$$B_{\delta,\sigma}(u, u) = \langle f, (1+\delta H_t)u\rangle_{2,\mu}. \tag{4-9}$$

With this choice for σ, we pick $\delta = c_1/(2c_2)$, use (4-6) on the left, and Cauchy–Schwarz on the right, in order to obtain

$$\|\mathbb{D}u\|^2_{2,\mu} \lesssim \|f\|_{2,\mu}\|u\|_{2,\mu} \le \lambda^2\|f\|^2_{2,\mu}.$$

This is the required uniform bound for $\lambda\mathbb{D}\mathcal{E}_\lambda f$. Since $\mathcal{H}$ is of the same type as $\mathcal{H}^*$ from the point of view of sesquilinear forms, the same estimates also hold for $\mathcal{E}^*_\lambda$ in place of $\mathcal{E}_\lambda$.

Next, we note that the estimates for the leftmost terms in (ii) and (iii) follow by duality from (i) applied to $\mathcal{E}_\lambda^*$.

In order to estimate the second term on the left in (ii), we set $u := (\lambda^{-2} + \mathcal{H})^{-1} D_t^{1/2} f$. Since $D_t^{1/2} f$ is now regarded as an element in $(\mathrm{E}_\mu)^*$, we get $\langle f, D_t^{1/2}(1 + \delta H_t)u\rangle_{2,\mu}$ on the right-hand side in (4-9), and from this we conclude

$$\|\mathbb{D}u\|_{2,\mu}^2 \lesssim \|f\|_{2,\mu}\|\mathbb{D}u\|_{2,\mu},$$

as required. The remaining term in (iii) is estimated in the same way upon replacing $D_t^{1/2} f$ by $w^{-1}\operatorname{div}_x(w\boldsymbol{f})$. □

4.4. *Off-diagonal estimates.* Given measurable subsets E and F of $\mathbb{R}^{n+1}$, we let

$$d(E, F) := \inf\{\|(x - y, t - s)\| : (x, t) \in E, (y, s) \in F\}$$

denote their parabolic distance. Lemma 4.4 below is an improvement of the uniform bounds in Lemma 4.3. We only state and prove Lemma 4.4 for the families of operators that will require it later. However, let us stress that such estimates are not to be expected in the presence of the nonlocal operator $D_t^{1/2}$, and one of the insights in [Auscher et al. 2020] was that in this case a nonlocal version of off-diagonal bounds should be used.

Lemma 4.4. *Assume that E and F are measurable subsets of $\mathbb{R}^{n+1}$, and let $d := d(E, F)$. Then, there exists a constant $c \in (0, \infty)$, depending only on the structural constants, such that*

(i) $\displaystyle\iint_F |\mathcal{E}_\lambda f|^2 + |\lambda \nabla_x \mathcal{E}_\lambda f|^2 \,\mathrm{d}\mu \lesssim e^{-d/(c\lambda)} \iint_E |f|^2 \,\mathrm{d}\mu,$

(ii) $\displaystyle\iint_F |\lambda \mathcal{E}_\lambda w^{-1}\operatorname{div}_x(w\boldsymbol{f})|^2 \,\mathrm{d}\mu \lesssim e^{-d/(c\lambda)} \iint_E |\boldsymbol{f}|^2 \,\mathrm{d}\mu$

for all $\lambda > 0$ and all $f \in \mathrm{L}^2_\mu$, $\boldsymbol{f} \in (\mathrm{L}^2_\mu)^n$ with support in E. The same statements are true when $\mathcal{E}_\lambda$ is replaced by $\mathcal{E}_\lambda^*$.

Proof. As in the proof of Lemma 4.3, it suffices to treat $\mathcal{E}_\lambda$. Based on Lemma 4.3, we see that it suffices to obtain the exponential estimate for $0 < \lambda \le \alpha d$, where for now $\alpha \in (0, 1)$ is a degree of freedom that will be determined later and which will only depend on the structural constants.

Let $u := \mathcal{E}_\lambda f$, and recall that

$$\iint_{\mathbb{R}^{n+1}} u\bar{v} + \lambda^2 w^{-1} A\nabla_x u \cdot \overline{\nabla_x v} + \lambda^2 H_t D_t^{1/2} u \cdot \overline{D_t^{1/2} v} \,\mathrm{d}\mu = \iint_{\mathbb{R}^{n+1}} f \cdot \bar{v} \,\mathrm{d}\mu \tag{4-10}$$

for all $v \in \mathrm{E}_\mu$. We can pick a real-valued $\tilde{\eta} \in \mathrm{C}^\infty(\mathbb{R}^{n+1})$ such that $\tilde{\eta} = 1$ on F, $\tilde{\eta} = 0$ on E, and such that

$$d|\nabla_x \tilde{\eta}| + d^2|\partial_t \tilde{\eta}| \le c$$

for some constant c only depending on n. The different scaling in the two terms is due to the definition of the parabolic distance. Next, we let

$$v := u\eta^2 \quad \text{with } \eta := e^{(\alpha d/\lambda)\tilde{\eta}} - 1. \tag{4-11}$$

For this choice of v, we rewrite the real part in (4-10) of the pairing containing half-order derivatives as follows. According to Lemma 3.3, there exists a sequence $\{u_i\} \subset \mathrm{C}_0^\infty(\mathbb{R}^{n+1})$ such that $u_i \to u$ in E_μ as $i \to \infty$. By the same lemma, $\eta^2 u_i \to \eta^2 u$ in E_μ, and therefore

$$\begin{aligned} \operatorname{Re} \iint_{\mathbb{R}^{n+1}} H_t D_t^{1/2} u \cdot \overline{D_t^{1/2} v} \,\mathrm{d}\mu &= \operatorname{Re} \lim_{i\to\infty} \iint_{\mathbb{R}^{n+1}} H_t D_t^{1/2} u_i \cdot \overline{D_t^{1/2}(u_i \eta^2)} \,\mathrm{d}t \,\mathrm{d}w \\ &= \lim_{i\to\infty} \operatorname{Re} \iint_{\mathbb{R}^{n+1}} \partial_t u_i \cdot \overline{u_i \eta^2} \,\mathrm{d}t \,\mathrm{d}w \\ &= \frac{1}{2} \lim_{i\to\infty} \iint_{\mathbb{R}^{n+1}} \partial_t |u_i|^2 \cdot \eta^2 \,\mathrm{d}t \,\mathrm{d}w \\ &= \frac{1}{2} \lim_{i\to\infty} - \iint_{\mathbb{R}^{n+1}} |u_i|^2 \cdot \partial_t(\eta^2) \,\mathrm{d}t \,\mathrm{d}w \\ &= -\frac{1}{2} \iint_{\mathbb{R}^{n+1}} |u|^2 \cdot \partial_t(\eta^2) \,\mathrm{d}\mu. \end{aligned}$$

Going back to (4-10) and using that $\eta = 0$ on E, we conclude that

$$\operatorname{Re} \iint_{\mathbb{R}^{n+1}} |u|^2 \eta^2 \,\mathrm{d}\mu + \lambda^2 w^{-1} A \nabla_x u \cdot \overline{\nabla_x (u\eta^2)} - \tfrac{1}{2} \lambda^2 |u|^2 \partial_t(\eta^2) \,\mathrm{d}\mu = 0.$$

Using this identity and ellipticity, we deduce

$$\begin{aligned} &\iint_{\mathbb{R}^{n+1}} |u|^2 \eta^2 \,\mathrm{d}\mu + c_1 \lambda^2 \iint_{\mathbb{R}^{n+1}} |\nabla_x u|^2 \eta^2 \,\mathrm{d}\mu \\ &\quad \le \lambda^2 \iint_{\mathbb{R}^{n+1}} |u|^2 |\eta| |\partial_t \eta| \,\mathrm{d}\mu + 2c_2 \lambda^2 \iint_{\mathbb{R}^{n+1}} |u| |\nabla_x u| |\eta| |\nabla_x \eta| \,\mathrm{d}\mu \\ &\quad \le \frac{1}{2} \iint_{\mathbb{R}^{n+1}} |u|^2 \eta^2 \,\mathrm{d}\mu + \frac{1}{2} \lambda^4 \iint_{\mathbb{R}^{n+1}} |u|^2 |\partial_t \eta|^2 \,\mathrm{d}\mu + \frac{1}{2} c_1 \lambda^2 \iint_{\mathbb{R}^{n+1}} |\nabla_x u|^2 \eta^2 \,\mathrm{d}\mu \\ &\qquad\qquad + 2 \frac{c_2^2}{c_1} \lambda^2 \iint_{\mathbb{R}^{n+1}} |u|^2 |\nabla_x \eta|^2 \,\mathrm{d}\mu. \end{aligned}$$

In conclusion,

$$\iint_{\mathbb{R}^{n+1}} |u|^2 \eta^2 \,\mathrm{d}\mu + c_1 \lambda^2 \iint_{\mathbb{R}^{n+1}} |\nabla_x u|^2 \eta^2 \,\mathrm{d}\mu \le \iint_{\mathbb{R}^{n+1}} |u|^2 \Bigl(\lambda^4 |\partial_t \eta|^2 + 4 \frac{c_2^2}{c_1} \lambda^2 |\nabla_x \eta|^2 \Bigr) \,\mathrm{d}\mu.$$

By the definition of η in (4-11) and since $\lambda \le \alpha d \le d$, we see that

$$|\partial_t \eta|^2 \le \frac{\alpha^2 d^2}{\lambda^2} |\eta + 1|^2 \frac{c^2}{d^4} \le c^2 \alpha^2 \lambda^{-4} |\eta + 1|^2$$

and

$$|\nabla_x \eta|^2 \le \frac{\alpha^2 d^2}{\lambda^2} |\eta + 1|^2 \frac{c^2}{d^2} = c^2 \alpha^2 \lambda^{-2} |\eta + 1|^2.$$

Thus, we get

$$\iint_{\mathbb{R}^{n+1}} |u|^2 \eta^2 \,\mathrm{d}\mu + c_1 \lambda^2 \iint_{\mathbb{R}^{n+1}} |\nabla_x u|^2 \eta^2 \,\mathrm{d}\mu \lesssim \alpha^2 \iint_{\mathbb{R}^{n+1}} |u|^2 |\eta + 1|^2 \,\mathrm{d}\mu.$$

At this point, we make our choice of α. Indeed, using the bound $|\eta+1|^2 \le 2(\eta^2+1)$, we choose α small enough to be able to absorb the part coming from η into the left-hand side. The conclusion is that

$$\iint_{\mathbb{R}^{n+1}} |u|^2\eta^2 \,\mathrm{d}\mu + \lambda^2 \iint_{\mathbb{R}^{n+1}} |\nabla_x u|^2\eta^2 \,\mathrm{d}\mu \lesssim \iint_{\mathbb{R}^{n+1}} |u|^2 \,\mathrm{d}\mu.$$

On the right-hand side, we can use Lemma 4.3 (i), and, on the left-hand side, we exploit that on F we have

$$\eta = e^{\alpha d/\lambda} - 1 \ge \tfrac{1}{2} e^{\alpha d/\lambda}$$

since we are assuming $\lambda \le \alpha d$. Consequently,

$$e^{2\alpha d/\lambda} \iint_F |u|^2 \,\mathrm{d}\mu + e^{2\alpha d/\lambda} \iint_F |\lambda \nabla_x u|^2 \,\mathrm{d}\mu \lesssim \iint_E |f|^2 \,\mathrm{d}\mu,$$

which proves (i).

The inequality in (ii) follows by a duality argument, using (i) for $\mathcal{E}_\lambda^*$ and interchanging the roles of E and F. In fact,

$$\begin{aligned}\iint_F |\lambda \mathcal{E}_\lambda w^{-1} \operatorname{div}_x(w\boldsymbol{f})|^2 \,\mathrm{d}\mu &= \sup_g \biggl(\iint_{\mathbb{R}^{n+1}} \lambda \mathcal{E}_\lambda w^{-1} \operatorname{div}_x(w\boldsymbol{f}) \cdot \bar{g} \,\mathrm{d}\mu \biggr)^2 \\ &= \sup_g \biggl(\iint_E -\boldsymbol{f} \cdot \overline{\lambda \nabla_x \mathcal{E}_\lambda^* g} \,\mathrm{d}\mu \biggr)^2,\end{aligned}$$

where the supremum is taken with respect to all $g \in \mathrm{L}^2_\mu$, with support in F, such that $\|g\|_{2,\mu} = 1$. We can now complete the proof by applying the Cauchy–Schwarz inequality and (i) of the lemma but for $\mathcal{E}_\lambda^*$. $\square$

5. Weighted Littlewood–Paley theory in the parabolic setting

We could develop a weighted parabolic Littlewood–Paley theory following the approach for singular integrals on spaces of homogeneous type [David et al. 1985]. However, since our weight w is time independent, we have decided to present a down-to-earth approach by combining weighted elliptic theory known in the field [Cruz-Uribe and Rios 2008; 2012; García-Cuerva and Rubio de Francia 1985] with Fourier analysis on the real line. Most of our estimates here are formulated using the square function norm

$$|||\cdot|||_{2,\mu} := \biggl(\int_0^\infty \iint_{\mathbb{R}^{n+1}} |\cdot|^2 \,\frac{\mathrm{d}\mu \,\mathrm{d}\lambda}{\lambda} \biggr)^{1/2}. \tag{5-1}$$

For the rest of the paper, $\mathcal{P} \in \mathrm{C}_0^\infty(\mathbb{R}^{n+1})$ is a fixed real-valued function in product form

$$\mathcal{P}(x,t) = \mathcal{P}^{(1)}(x)\mathcal{P}^{(2)}(t),$$

where $\mathcal{P}^{(1)}$ and $\mathcal{P}^{(2)}$ are both radial, nonnegative, and have integral 1. For all $x \in \mathbb{R}^n$, $t \in \mathbb{R}$, we set

$$\begin{aligned}\mathcal{P}_\lambda^{(1)}(x) &:= \lambda^{-n}\mathcal{P}^{(1)}(x/\lambda), \\ \mathcal{P}_\lambda^{(2)}(t) &:= \lambda^{-2}\mathcal{P}^{(2)}(t/\lambda^2), \\ \mathcal{P}_\lambda(x,t) &:= \mathcal{P}_\lambda^{(1)}(x)\mathcal{P}_\lambda^{(2)}(t) = \lambda^{-n-2}\mathcal{P}^{(1)}(x/\lambda)\mathcal{P}^{(2)}(t/\lambda^2)\end{aligned}$$

whenever $\lambda > 0$. With a slight abuse of notation, we let $\mathcal{P}_\lambda$ also denote the associated convolution operator

$$\mathcal{P}_\lambda f(x,t) = \mathcal{P}_\lambda * f(x,t) = \iint_{\mathbb{R}^{n+1}} \mathcal{P}_\lambda(x-y, t-s) f(y,s)\,\mathrm{d}y\,\mathrm{d}s,$$

and likewise for $\mathcal{P}_\lambda^{(1)}$ and $\mathcal{P}_\lambda^{(2)}$. We note that

$$\begin{aligned} |\mathcal{P}_\lambda^{(1)} f(x,t)| &\le \mathcal{M}^{(1)}(f(\,\cdot\,,t))(x), \\ |\mathcal{P}_\lambda^{(2)} f(x,t)| &\le \mathcal{M}^{(2)}(f(x,\cdot\,))(t), \\ |\mathcal{P}_\lambda f(x,t)| &\le \mathcal{M}^{(1)}(\mathcal{M}^{(2)} f(\,\cdot\,,t))(x) \end{aligned} \tag{5-2}$$

almost everywhere, for every $f \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^{n+1})$; see [Stein 1993, Section II.2.1]. In particular, these pointwise bounds hold for $f \in \mathrm{L}^2_\mu$. The boundedness of the maximal operators in L^2_μ implies

$$\sup_{\lambda>0} \|\mathcal{P}_\lambda\|_{2\to 2,\mu} \lesssim 1;$$

see Section 2.2.

Lemma 5.1. *For all* $f \in \mathrm{L}^2_\mu(\mathbb{R}^{n+1})$,

$$|||\lambda\nabla_x \mathcal{P}_\lambda f|||_{2,\mu} + |||\lambda^2 \partial_t \mathcal{P}_\lambda f|||_{2,\mu} + |||\lambda D_t^{1/2} \mathcal{P}_\lambda f|||_{2,\mu} \lesssim \|f\|_{2,\mu}.$$

Proof. Here, we write out in detail how the product structure of $\mathcal{P}_\lambda$ can be used to prove parabolic estimates in $\mathbb{R}^{n+1}$ through weighted elliptic theory and classical Fourier analysis. This motif will appear in all proofs of this section. Let $\hat{g}$ denote the Fourier transform in time of a function g on $\mathbb{R}^{n+1}$.

By uniform boundedness of $\mathcal{P}_\lambda^{(1)}$ in L^2_μ and Plancherel's theorem, we have

$$\begin{aligned} |||\lambda D_t^{1/2} \mathcal{P}_\lambda f|||_{2,\mu}^2 &= \int_0^\infty \iint_{\mathbb{R}^{n+1}} |\mathcal{P}_\lambda^{(1)} \lambda D_t^{1/2} \mathcal{P}_\lambda^{(2)} f|^2 \frac{\mathrm{d}\mu\,\mathrm{d}\lambda}{\lambda} \\ &\lesssim \int_{\mathbb{R}^n} \int_0^\infty \int_{\mathbb{R}} |\lambda D_t^{1/2} \mathcal{P}_\lambda^{(2)} f|^2 \frac{\mathrm{d}t\,\mathrm{d}\lambda}{\lambda}\,\mathrm{d}w \\ &= \int_{\mathbb{R}^n} \int_0^\infty \int_{\mathbb{R}} \bigl|\lambda|\tau|^{1/2} \widehat{\mathcal{P}^{(2)}}(\lambda^2\tau) \hat{f}(x,\tau)\bigr|^2 \frac{\mathrm{d}\tau\,\mathrm{d}\lambda}{\lambda}\,\mathrm{d}w(x). \end{aligned}$$

The integral in λ is finite and independent of τ since $\widehat{\mathcal{P}^{(2)}}$ is a radial Schwartz function. Applying Plancherel's theorem backwards, we get the desired bound by $\|f\|_{2,\mu}^2$. The same argument yields the bound for $|||\lambda^2 \partial_t \mathcal{P}_\lambda f|||_{2,\mu}$.

Finally, in order to bound $\lambda\nabla_x \mathcal{P}_\lambda f$, we use uniform boundedness of $\mathcal{P}_\lambda^{(2)}$ to get

$$|||\lambda\nabla_x \mathcal{P}_\lambda f|||_{2,\mu}^2 = \int_0^\infty \iint_{\mathbb{R}^{n+1}} |\mathcal{P}_\lambda^{(2)} \lambda\nabla_x \mathcal{P}_\lambda^{(1)} f|^2 \frac{\mathrm{d}\mu\,\mathrm{d}\lambda}{\lambda} \lesssim \int_{\mathbb{R}} \int_0^\infty \int_{\mathbb{R}^n} |\lambda\nabla_x \mathcal{P}_\lambda^{(1)} f|^2 \frac{\mathrm{d}w\,\mathrm{d}\lambda}{\lambda}\,\mathrm{d}t.$$

For fixed t, we now need weighted elliptic Littlewood–Paley theory. The operator $\lambda\nabla_x \mathcal{P}_\lambda^{(1)}$ acts by convolution with Ψ_λ, where $\Psi = \nabla_x \mathcal{P}^{(1)}$ has integral 0. Thus, we can use, e.g., [Cruz-Uribe and Rios 2012, Lemma 4.6] to control the integral in $\mathrm{d}w\,\mathrm{d}\lambda$ by $\|f(\,\cdot\,,t)\|_{2,w}^2$, and the proof is complete. □

Lemma 5.2. *For all* $f \in \mathrm{E}_\mu$,

$$|||\lambda^{-1}(I-\mathcal{P}_\lambda)f|||_{2,\mu} \lesssim \|\mathbb{D}f\|_{2,\mu}.$$

Proof. We first claim

$$|||\lambda^{-1}(I-\mathcal{P}_\lambda^{(1)})f|||_{2,\mu} \lesssim \|\nabla_x f\|_{2,\mu}, \quad |||\lambda^{-1}(I-\mathcal{P}_\lambda^{(2)})f|||_{2,\mu} \lesssim \|D_t^{1/2} f\|_{2,\mu}. \tag{5-3}$$

As in the proof of Lemma 5.1, this can be proved using Plancherel's theorem in t for the second term and weighted Littlewood–Paley theory with t fixed for the first term. The required weighted result is [Cruz-Uribe and Rios 2015, Proposition 2.3] (originally [Cruz-Uribe and Rios 2012, Proposition 4.7]) and the application to the concrete operator considered here is detailed in the lines following equation (4.3) in the same paper.

In order to complete the proof of the lemma, we simply write

$$(I-\mathcal{P}_\lambda) = \mathcal{P}_\lambda^{(2)}(1-\mathcal{P}_\lambda^{(1)}) + (1-\mathcal{P}_\lambda^{(2)}).$$

The result follows from (5-3) and the uniform boundedness of $\mathcal{P}_\lambda^{(2)}$ in L^2_μ. □

In the following we write $\Delta = Q \times I$ for parabolic cubes in $\mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}$.

Definition 5.3. We define $\mathcal{A}_\lambda^{(1)}$, $\mathcal{A}_\lambda^{(2)}$ and $\mathcal{A}_\lambda$ to be the dyadic averaging operators in x, t and (x,t) with respect to parabolic scaling, that is, if $\Delta = Q \times I$ is the dyadic parabolic cube with $\frac{1}{2}\ell(\Delta) < \lambda \le \ell(\Delta)$ containing (x,t), then

$$\begin{aligned}
\mathcal{A}_\lambda^{(1)} f(x,t) &:= \fint_Q f \,\mathrm{d}y, \\
\mathcal{A}_\lambda^{(2)} f(x,t) &:= \fint_I f \,\mathrm{d}s, \\
\mathcal{A}_\lambda f(x,t) &:= \fint\!\!\fint_\Delta f \,\mathrm{d}y\,\mathrm{d}s = \mathcal{A}_\lambda^{(1)}\mathcal{A}_\lambda^{(2)} f(x,t).
\end{aligned}$$

It follows from the bounds for the maximal operators in Section 2.2 and doubling that the dyadic averaging operators are bounded on L^2_μ, uniformly in λ.

Lemma 5.4. *Let* $\mathcal{P}_\lambda$ *and* $\mathcal{A}_\lambda$ *be as above. Then, for all* $f \in \mathrm{L}^2_\mu(\mathbb{R}^{n+1})$,

$$|||(\mathcal{A}_\lambda - \mathcal{P}_\lambda)f|||_{2,\mu} \lesssim \|f\|_{2,\mu}.$$

Proof. We follow our (general) strategy and write

$$\mathcal{A}_\lambda - \mathcal{P}_\lambda = \mathcal{A}_\lambda^{(2)}(\mathcal{A}_\lambda^{(1)} - \mathcal{P}_\lambda^{(1)}) + \mathcal{P}_\lambda^{(1)}(\mathcal{A}_\lambda^{(2)} - \mathcal{P}_\lambda^{(2)}),$$

where we have also used that $\mathcal{A}_\lambda^{(2)}$ and $\mathcal{P}_\lambda^{(1)}$ commute since they act in different variables. Since these operators are uniformly bounded on L^2_μ with respect to λ, we get

$$\begin{aligned}
|||(\mathcal{A}_\lambda - \mathcal{P}_\lambda)f|||_{2,\mu} \lesssim \int_{\mathbb{R}}\int_0^\infty \|(\mathcal{A}_\lambda^{(1)} - \mathcal{P}_\lambda^{(1)})f(\,\cdot\,,t)\|^2_{2,w} \frac{\mathrm{d}\lambda}{\lambda}\,\mathrm{d}t \\
+ \int_{\mathbb{R}^n}\int_0^\infty \|(\mathcal{A}_\lambda^{(2)} - \mathcal{P}_\lambda^{(2)})f(x,\cdot\,)\|^2_{2,\mathrm{d}t} \frac{\mathrm{d}\lambda}{\lambda}\,\mathrm{d}w(x).
\end{aligned}$$

For the first term on the right, we can rely on the weighted elliptic version of the lemma [Cruz-Uribe and Rios 2012, Lemma 5.2]. For the second term on the right, we can make a change of variables $\lambda' = \lambda^2$ and use the unweighted one-dimensional version of the lemma, which of course follows from the same reference or the classical proof in [Auscher and Tchamitchian 1998, Appendix C, (4)]. □

6. Reduction to a quadratic estimate

The purpose of this short section is to reduce our main result, Theorem 1.1, to the quadratic estimate

$$|||\lambda \mathcal{H} \mathcal{E}_\lambda f|||_{2,\mu} \lesssim \|\nabla_x f\|_{2,\mu} + \|H_t D_t^{1/2} f\|_{2,\mu}, \quad f \in \mathrm{E}_\mu. \tag{6-1}$$

Recall that $\mathcal{E}_\lambda = (1 + \lambda^2 \mathcal{H})^{-1}$. Since the sesquilinear form associated with $\mathcal{H}$ is not closed, see Section 4, classical results à la Lions [1962] as in the elliptic case do not apply, and here we give full details of this reduction.

At this point, we require some essentials from functional calculus. We give references along the way, and we refer the reader to [Haase 2006; McIntosh 1986] for further background. Since $\mathcal{H}$ is m-accretive (Proposition 4.2), it has a unique m-accretive square root $\sqrt{\mathcal{H}}$ defined by the functional calculus for sectorial operators, and the same is true for the adjoint $\mathcal{H}^*$ with $\sqrt{\mathcal{H}^*} = (\sqrt{\mathcal{H}})^*$.

In order to see the reduction alluded to above, we start out with the Calderón reproducing formula in [Haase 2006, Theorem 5.2.6], and we write

$$\sqrt{\mathcal{H}} f = \frac{16}{\pi} \int_0^\infty \lambda^3 \mathcal{H}^2 (1 + \lambda^2 \mathcal{H})^{-3} f \frac{\mathrm{d}\lambda}{\lambda}, \tag{6-2}$$

where $f \in \mathsf{D}(\sqrt{\mathcal{H}})$ and the integral is understood as an improper Riemann integral in L^2_μ. Testing this identity against $g \in \mathrm{L}^2_\mu$ and applying Cauchy–Schwarz, we obtain

$$|\langle \sqrt{\mathcal{H}} f, g \rangle_{2,\mu}| \le \frac{16}{\pi} |||\lambda \mathcal{H}(1 + \lambda^2 \mathcal{H})^{-1} f|||_{2,\mu} \times |||\lambda^2 \mathcal{H}^* (1 + \lambda^2 \mathcal{H}^*)^{-2} g|||_{2,\mu}.$$

The second term is controlled by a structural constant times $\|g\|_{2,\mu}$ since $\mathcal{H}^*$ is m-accretive in L^2_μ — more precisely, this follows from von Neumann's inequality [Haase 2006, Theorem 7.1.7] and the characterization of the emerging functional calculus due to McIntosh [Haase 2006, Theorem 7.3.1]. Taking the supremum over all g yields

$$\|\sqrt{\mathcal{H}} f\|_{2,\mu} \lesssim |||\lambda \mathcal{H}(I + \lambda^2 \mathcal{H})^{-1} f|||_{2,\mu}.$$

Let us now suppose that (6-1) holds. Then, we obtain

$$\|\sqrt{\mathcal{H}} f\|_{2,\mu} \lesssim \|\nabla_x f\|_{2,\mu} + \|H_t D_t^{1/2} f\|_{2,\mu}, \tag{6-3}$$

when f is in $\mathrm{E}_\mu \cap \mathsf{D}(\sqrt{\mathcal{H}}) \supset \mathsf{D}(\mathcal{H})$. However, since this space is dense in E_μ, by Proposition 4.2, and as $\sqrt{\mathcal{H}}$ is closed, the estimate extends to all $f \in \mathrm{E}_\mu$. Next, we note that $\mathcal{H}^*$ is similar to an operator in the same class as $\mathcal{H}$ under conjugation with the "time reversal" $f(t, x) \mapsto f(-t, x)$ and conjugation of A. Hence, we also have

$$\|\sqrt{\mathcal{H}^*} g\|_{2,\mu} \lesssim \|\nabla_x g\|_{2,\mu} + \|H_t D_t^{1/2} g\|_{2,\mu} \tag{6-4}$$

whenever $g \in \mathrm{E}_\mu$. Using (4-6) with $\sigma = 0$ and δ small enough depending on the structural constants, we obtain, for all $f \in \mathsf{D}(\mathcal{H})$, that

$$\begin{aligned}\delta \|\nabla_x f\|_{2,\mu}^2 + \delta \|D_t^{1/2} f\|_{2,\mu}^2 &\le |\langle \mathcal{H} f, (1+\delta H_t) f\rangle_{2,\mu}| \\ &\le \|\sqrt{\mathcal{H}} f\|_{2,\mu} \|\sqrt{\mathcal{H}^*}(1+\delta H_t) f\|.\end{aligned}$$

Now, (6-4) with $g := (1+\delta H_t) f \in \mathrm{E}_\mu$ implies

$$\|\nabla_x f\|_{2,\mu} + \|D_t^{1/2} f\|_{2,\mu} \lesssim \|\sqrt{\mathcal{H}} f\|_{2,\mu}. \tag{6-5}$$

Since $\mathsf{D}(\mathcal{H})$ is dense in $\mathsf{D}(\sqrt{\mathcal{H}})$ for the graph norm [Haase 2006, Proposition 3.1.1(h)], the estimate extends to all $f \in \mathsf{D}(\sqrt{\mathcal{H}})$.

In conclusion, we have seen that (6-1) implies the statement of Theorem 1.1 through the estimates (6-3) and (6-5). Therefore, the rest of the paper is devoted to the task of proving (6-1).

7. Principal part approximation

In order to prove the square function estimate (6-1), we will eventually split $\mathcal{H}$ into its elliptic and parabolic parts and perform the "hard" analysis only on the elliptic part. This will lead us to the operators

$$\mathcal{U}_\lambda := \lambda \mathcal{E}_\lambda w^{-1} \operatorname{div}_x w, \quad \lambda > 0. \tag{7-1}$$

These operators appeared in Lemma 4.4 on off-diagonal estimates and in particular they are uniformly bounded on $(\mathrm{L}^2_\mu)^n$. Here, we continue their analysis.

Given a cube $Q = Q_r(x) \subset \mathbb{R}^n$ and an interval $I = I_r(t)$, we let $\Delta := Q \times I$ and set

$$\begin{aligned}C_k(\Delta) = C_k(Q \times I) &:= 2^{k+1}\Delta \setminus 2^k \Delta, \quad k = 1, 2, \dots, \\ C_0(\Delta) &:= 2\Delta.\end{aligned}$$

In the following, we denote the characteristic function of a set E by 1_E. We use off-diagonal estimates to define $\mathcal{U}_\lambda$ on $(\mathrm{L}^\infty)^n$.

Definition 7.1. For $\boldsymbol{b} \in (\mathrm{L}^\infty)^n$, we define

$$\mathcal{U}_\lambda \boldsymbol{b} =: \lim_{k\to\infty} \mathcal{U}_\lambda(\boldsymbol{b} 1_{2^k\Delta}), \tag{7-2}$$

with convergence locally in $(\mathrm{L}^2_\mu)^n$, where on the right Δ is any parabolic cube.

Definition 7.1 is meaningful and independent of the choice of Δ as we shall see next. To start, if Δ' is any parabolic cube, then for $m > l$ large enough to guarantee that $\Delta' \subset 2^{l-1}\Delta$, applying Lemma 4.4 with $E = C_j(\Delta)$ and $F = \Delta'$ for $j = l, \dots, m-1$ yields

$$\begin{aligned}\|\mathcal{U}_\lambda(\boldsymbol{b} 1_{2^m\Delta \setminus 2^l \Delta})\|_{\mathrm{L}^2_\mu(\Delta')} &\le \sum_{j=l}^{m-1} \|\mathcal{U}_\lambda(\boldsymbol{b} 1_{C_j(\Delta)})\|_{\mathrm{L}^2_\mu(\Delta')} \\ &\lesssim \mu(\Delta)^{1/2} \|\boldsymbol{b}\|_\infty \sum_{j=l}^{m-1} e^{-\ell(\Delta)2^{j-1}/c\lambda} (4D)^{j+1}.\end{aligned}$$

Recall that $4D$ is the doubling constant for μ; see (2-3). The right-hand side converges to 0 as $m, l \to \infty$. In conclusion, $\{\mathcal{U}_\lambda(\boldsymbol{b}1_{2^l\Delta})\}_l$ is a Cauchy sequence locally in $(\mathrm{L}^2_\mu)^n$. By the same argument, Definition 7.1 is independent of the particular choice Δ. Taking $\Delta' = \Delta$ and $l = 1$, we get

$$\begin{aligned}\|\mathcal{U}_\lambda \boldsymbol{b}\|_{\mathrm{L}^2_\mu(\Delta)} &\le \|\mathcal{U}_\lambda(\boldsymbol{b}1_{2\Delta})\|_{\mathrm{L}^2_\mu(\Delta)} + \Big\|\lim_{m\to\infty} \mathcal{U}_\lambda(\boldsymbol{b}1_{2^m\Delta\setminus 2\Delta})\Big\|_{\mathrm{L}^2_\mu(\Delta)} \\ &\lesssim \mu(\Delta)^{1/2}\|\boldsymbol{b}\|_\infty\Big(1+\sum_{j=1}^\infty e^{-\ell(\Delta)2^{j-1}/c\lambda}(4D)^{j+1}\Big).\end{aligned} \tag{7-3}$$

Lemma 7.2. *Let $\boldsymbol{b} \in (\mathrm{L}^\infty)^n$ and $f \in \mathrm{L}^2_\mu$. Then,*

$$\|(\mathcal{U}_\lambda \boldsymbol{b})\cdot \mathcal{A}_\lambda f\|_{2,\mu} \lesssim \|\boldsymbol{b}\|_\infty \|f\|_{2,\mu}.$$

Proof. If $\Delta \subset \mathbb{R}^{n+1}$ is a parabolic cube such that $\frac{1}{2}\ell(\Delta) < \lambda \le \ell(\Delta)$, then by (7-3) we have

$$\iint_\Delta |\mathcal{U}_\lambda \boldsymbol{b}|^2 \,\mathrm{d}\mu \lesssim \mu(\Delta)\|\boldsymbol{b}\|_\infty^2.$$

Since $\mathcal{A}_\lambda f$ is constant on each such Δ, we obtain

$$\iint_\Delta |(\mathcal{U}_\lambda \boldsymbol{b})\cdot \mathcal{A}_\lambda f|^2 \,\mathrm{d}\mu \le \iint_\Delta |\mathcal{U}_\lambda \boldsymbol{b}|^2 \,\mathrm{d}\mu \cdot ⨌_\Delta |\mathcal{A}_\lambda f|^2 \,\mathrm{d}\mu \lesssim \|\boldsymbol{b}\|_\infty^2 \iint_\Delta |\mathcal{A}_\lambda f|^2 \,\mathrm{d}\mu.$$

The claim follows by summing in Δ and using that $\mathcal{A}_\lambda$ is uniformly bounded on L^2_μ with respect to λ; see Section 5. □

Writing $A = (A_1, \dots, A_n)$ with A_j the j-th column of A, we can use Definition 7.1 to define the action of $\mathcal{U}_\lambda$ on the bounded matrix-valued function $w^{-1}A$ by

$$(\mathcal{U}_\lambda w^{-1}A) := \mathcal{U}_\lambda(w^{-1}A) := (\mathcal{U}_\lambda(w^{-1}A_1), \dots, \mathcal{U}_\lambda(w^{-1}A_n)).$$

We will approximate $\mathcal{U}_\lambda w^{-1}A$ by operators that act via multiplication on the maximal dyadic cubes of size at most λ. To be precise, we will consider

$$\mathcal{R}_\lambda f := \mathcal{U}_\lambda(w^{-1}Af) - (\mathcal{U}_\lambda w^{-1}A)\cdot \mathcal{A}_\lambda f. \tag{7-4}$$

This is nowadays called the "principal part approximation". Using Lemmas 4.3 and 7.2, we see that the $\mathcal{R}_\lambda$ are uniformly bounded on L^2_μ for $\lambda > 0$. Moreover, we prove the following bound.

Proposition 7.3. *Let $f \in \mathrm{L}^2_\mu \cap \mathrm{C}^\infty$. Then,*

$$\|\mathcal{R}_\lambda f\|_{2,\mu} \lesssim \|\lambda \nabla_x f\|_{2,\mu} + \|\lambda^2 \partial_t f\|_{2,\mu}.$$

For the proof, we need the following weighted Poincaré-type inequality. In the following we abbreviate $(f)_\Delta = (f)_{\Delta,\mathrm{d}x\,\mathrm{d}t}$.

Lemma 7.4. *Let $f \in \mathrm{C}^\infty$. Then, for all parabolic cubes Δ and all nonnegative integers k,*

$$\iint_{C_k(\Delta)} |(f-(f)_\Delta)|^2 \,\mathrm{d}\mu \le c2^{k(n+2)} \iint_{2^{k+1}\Delta} \ell(\Delta)^2|\nabla_x f|^2 + \ell(\Delta)^4|\partial_t f|^2 \,\mathrm{d}\mu,$$

where c depends only on n and $[w]_{A_2}$.

Proof. Let $\Delta = Q \times I$ be a parabolic cube. We set $g := (f)_{Q,\mathrm{d}x}$, which is a function of t, and we split

$$f - (f)_\Delta = (f - (f)_{Q,\mathrm{d}x}) + (g - (g)_{I,\mathrm{d}t}).$$

To the first term we can apply the weighted Poincaré inequality in the x-variable from (the proof of) [Heinonen et al. 1993, Theorem 15.26] and to the second term the standard Poincaré inequality in the t-variable. The result is

$$\biggl(\iint_\Delta |(f - (f)_\Delta)|^2 \,\mathrm{d}\mu\biggr)^{1/2} \le c\biggl(\iint_\Delta \ell(\Delta)^2 |\nabla_x f|^2 + \ell(\Delta)^4 |\partial_t f|^2 \,\mathrm{d}\mu\biggr)^{1/2}.$$

Note that in [Heinonen et al. 1993] balls are used instead of cubes, but doubling allows us to switch between one and the other. For the general result it suffices to write

$$f - (f)_\Delta = (f - (f)_{2^{k+1}\Delta}) + ((f)_{2^{k+1}\Delta} - (f)_{2^k\Delta}) + \cdots + ((f)_{2\Delta} - (f)_\Delta)$$

and to use the estimate above on the cubes $2^{k+1}\Delta$ and then on $2^{k+1}\Delta, \ldots, 2\Delta$. □

Proof of Proposition 7.3. We note that if $(x, t) \in \mathbb{R}^{n+1}$ and $\lambda > 0$, then

$$\mathcal{R}_\lambda f(x, t) = \mathcal{U}_\lambda(w^{-1} A(f - (f)_\Delta))(x, t),$$

where Δ is the unique dyadic parabolic cube with $\frac{1}{2}\ell(\Delta) < \lambda \le \ell(\Delta)$ that contains (x, t). Thus,

$$\begin{aligned}\|\mathcal{R}_\lambda f\|_{2,\mu}^2 &= \sum_\Delta \iint_\Delta |\mathcal{U}_\lambda(w^{-1} A(f - (f)_\Delta))|^2 \,\mathrm{d}\mu \\ &\le \sum_\Delta \biggl(\sum_{k=0}^\infty \biggl(\iint_\Delta |\mathcal{U}_\lambda(w^{-1} A \cdot 1_{C_k(\Delta)}(f - (f)_\Delta))|^2 \,\mathrm{d}\mu\biggr)^{1/2}\biggr)^2,\end{aligned}$$

and therefore

$$\begin{aligned}\|\mathcal{R}_\lambda f\|_{2,\mu}^2 &\lesssim \sum_\Delta \biggl(\sum_{k=0}^\infty e^{-2^k/c} \biggl(\iint_{C_k(\Delta)} |(f - (f)_\Delta)|^2 \,\mathrm{d}\mu\biggr)^{1/2}\biggr)^2 \\ &\lesssim \sum_\Delta \sum_{k=0}^\infty e^{-2^k/c} \iint_{C_k(\Delta)} |(f - (f)_\Delta)|^2 \,\mathrm{d}\mu \\ &\lesssim \sum_\Delta \sum_{k=0}^\infty e^{-2^k/c} 2^{k(n+2)} \iint_{2^{k+1}\Delta} \lambda^2 |\nabla_x f|^2 + \lambda^4 |\partial_t f|^2 \,\mathrm{d}\mu \\ &\le \biggl(\sum_{k=0}^\infty e^{-2^k/c} 2^{(2k+1)(n+2)}\biggr) \iint_{\mathbb{R}^{n+1}} \lambda^2 |\nabla_x f|^2 + \lambda^4 |\partial_t f|^2 \,\mathrm{d}\mu,\end{aligned}$$

where we used, in succession, the off-diagonal estimates, Cauchy–Schwarz inequality, Lemma 7.4, and the fact that each point in $\mathbb{R}^{n+1}$ is contained in exactly $2^{(k+1)(n+2)}$ of the cubes $2^{k+1}\Delta$. The sum in k is still finite, and the proof is complete. □

8. Proof of Theorem 1.1

After the reduction in Section 6, it remains to prove the quadratic estimate (6-1) that we now write in the form

$$|||\lambda\mathcal{E}_\lambda\mathcal{H}f|||_{2,\mu} \lesssim \|\mathbb{D}f\|_{2,\mu}, \quad f \in \mathrm{E}_\mu. \tag{8-1}$$

In the following we will use the operators $\mathcal{P}_\lambda$, $\mathcal{A}_\lambda$, $\mathcal{U}_\lambda$, $\mathcal{R}_\lambda$ that have been introduced in Sections 4, 5 and 7. Collecting the estimates from these sections, we can at this stage prove the following.

Proposition 8.1. *Let $f \in \mathrm{E}_\mu$. Then,*

$$|||(\lambda\mathcal{E}_\lambda\mathcal{H} + (\mathcal{U}_\lambda w^{-1}A)\cdot\mathcal{A}_\lambda\nabla_x)f|||_{2,\mu} \lesssim \|\mathbb{D}f\|_{2,\mu}.$$

Proof. We begin by writing

$$\lambda\mathcal{E}_\lambda\mathcal{H}f = \lambda\mathcal{E}_\lambda\mathcal{H}\mathcal{P}_\lambda f + \lambda\mathcal{H}\mathcal{E}_\lambda(I-\mathcal{P}_\lambda)f. \tag{8-2}$$

Using the identity

$$\lambda\mathcal{H}\mathcal{E}_\lambda = \lambda^{-1}(I-\mathcal{E}_\lambda),$$

the uniform L^2_μ-boundedness of $\mathcal{E}_\lambda$, and Lemma 5.2, we see that

$$|||\lambda\mathcal{H}\mathcal{E}_\lambda(I-\mathcal{P}_\lambda)f|||_{2,\mu} \lesssim |||\lambda^{-1}(I-\mathcal{P}_\lambda)f|||_{2,\mu} \lesssim \|\mathbb{D}f\|_{2,\mu}.$$

Next, we use (4-3) to write

$$\lambda\mathcal{E}_\lambda\mathcal{H}\mathcal{P}_\lambda f = -\mathcal{U}_\lambda w^{-1}A\nabla_x\mathcal{P}_\lambda f + \lambda\mathcal{E}_\lambda D_t^{1/2}H_tD_t^{1/2}\mathcal{P}_\lambda f. \tag{8-3}$$

Using Lemma 4.3 (i) and then Lemma 5.1, we see that

$$\begin{aligned} |||\lambda\mathcal{E}_\lambda D_t^{1/2}H_tD_t^{1/2}\mathcal{P}_\lambda f|||_{2,\mu} &= |||\lambda\mathcal{E}_\lambda D_t^{1/2}\mathcal{P}_\lambda D_t^{1/2}H_tf|||_{2,\mu} \\ &\lesssim |||\lambda D_t^{1/2}\mathcal{P}_\lambda D_t^{1/2}H_tf|||_{2,\mu} \\ &\lesssim \|D_t^{1/2}f\|_{2,\mu}. \end{aligned} \tag{8-4}$$

Finally, we bring the principal part approximation into play. We use $\mathcal{U}_\lambda$ and $\mathcal{R}_\lambda$ to write

$$\begin{aligned} \mathcal{U}_\lambda w^{-1}A\nabla_x\mathcal{P}_\lambda f &= \mathcal{U}_\lambda w^{-1}A\mathcal{P}_\lambda\nabla_x f \\ &= \mathcal{R}_\lambda\mathcal{P}_\lambda\nabla_x f + (\mathcal{U}_\lambda w^{-1}A)\cdot\mathcal{A}_\lambda(\mathcal{P}_\lambda-\mathcal{A}_\lambda)\nabla_x f + (\mathcal{U}_\lambda w^{-1}A)\cdot\mathcal{A}_\lambda\nabla_x f, \end{aligned} \tag{8-5}$$

where we have also used that $(\mathcal{A}_\lambda)^2 = \mathcal{A}_\lambda$ for the last term. Applying Proposition 7.3 and Lemma 5.1, we have

$$|||\mathcal{R}_\lambda\mathcal{P}_\lambda\nabla_x f|||_{2,\mu} \lesssim |||\lambda\nabla_x\mathcal{P}_\lambda\nabla_x f|||_{2,\mu} + |||\lambda^2\partial_t\mathcal{P}_\lambda\nabla_x f|||_{2,\mu} \lesssim \|\mathbb{D}f\|_{2,\mu}.$$

Also, by Lemmas 5.4 and 7.2, we have

$$|||(\mathcal{U}_\lambda w^{-1}A)\cdot\mathcal{A}_\lambda(\mathcal{P}_\lambda-\mathcal{A}_\lambda)\nabla_x f|||_{2,\mu} \lesssim |||(\mathcal{A}_\lambda-\mathcal{P}_\lambda)\nabla_x f|||_{2,\mu} \lesssim \|\nabla_x f\|_{2,\mu}.$$

Looking back at the successive splittings in (8-2), (8-3) and (8-5), we see that the only term that has not been treated in the square function norm is $(\mathcal{U}_\lambda w^{-1}A)\cdot\mathcal{A}_\lambda\nabla_x f$. This proves the claim. □

To conclude the square function estimate for the final term $(\mathcal{U}_\lambda w^{-1}A)\cdot\mathcal{A}_\lambda\nabla_x f$, we establish Lemma 8.3 below. The lemma states that

$$|\mathcal{U}_\lambda w^{-1}A|^2\,\frac{\mathrm{d}\mu\,\mathrm{d}\lambda}{\lambda}$$

is a Carleson measure and that we have good control of the constants. Hence,

$$|||(\mathcal{U}_\lambda w^{-1}A)\cdot\mathcal{A}_\lambda\nabla_x f|||_{2,\mu}\lesssim\|\nabla_x f\|_{2,\mu}$$

follows by Carleson's inequality for parabolic cubes; see Lemma 8.2. This completes the proof of the estimate in (8-1), and hence the proof of Theorem 1.1 modulo Lemma 8.3. The reader should observe that, in our proof of (8-1), we have split off the time derivative ∂_t from $\mathcal{H}$ and we have controlled the part coming from ∂_t by a standard Littlewood–Paley estimate in (8-4).

For convenience, we include a proof of the version of Carleson's inequality that is used above. We adapt the elegant dyadic argument found in [Morris 2012, Theorem 4.3].

Lemma 8.2. *Let ν be a Borel measure on $\mathbb{R}^{n+1}\times\mathbb{R}^+$ that satisfies*

$$\|\nu\|_{\mathcal{C}}:=\sup_{\Delta}\frac{\nu(\Delta\times(0,\ell(\Delta)])}{\mu(\Delta)}<\infty,$$

where the supremum is taken over all dyadic parabolic cubes $\Delta\subset\mathbb{R}^{n+1}$. Then there is a constant c that only depends on n and $[w]_{A_2}$ such that, for every $f\in\mathrm{L}^2_\mu$,

$$\int_0^\infty\iint_{\mathbb{R}^{n+1}}|\mathcal{A}_\lambda f(x,t)|^2\,\mathrm{d}\nu(x,t,\lambda)\le c\|\nu\|_{\mathcal{C}}\iint_{\mathbb{R}^{n+1}}|f|^2\,\mathrm{d}\mu.$$

Proof. For $i\in\mathbb{Z}$, let $\{\Delta_i^j\}_j$ be the partition of $\mathbb{R}^{n+1}$ into dyadic parabolic cubes such that $\ell(\Delta_i^j)=2^i$. We have

$$\int_0^\infty\iint_{\mathbb{R}^{n+1}}|\mathcal{A}_\lambda f(x,t)|^2\,\mathrm{d}\nu(x,t,\lambda)=\sum_{i=-\infty}^{\infty}\sum_j\left|\fint\!\!\fint_{\Delta_i^j}f\,\mathrm{d}y\,\mathrm{d}s\right|^2\nu(\Delta_i^j\times(2^{i-1},2^i])=\sum_{i=-\infty}^{\infty}\sum_j|f_i^j|^2\nu_i^j,$$

where we have introduced $\nu_i^j:=\nu(\Delta_i^j\times(2^{i-1},2^i])$ and $f_i^j:=\fint\!\!\fint_{\Delta_i^j}f\,\mathrm{d}y\,\mathrm{d}s$. For $r>0$, let $\{\Delta_k(r)\}_k$ be the collection of maximal dyadic parabolic cubes Δ_i^j such that $|f_i^j|>r$. Note that these cubes are pairwise disjoint and contained in $\{\mathcal{M}^{(1)}\mathcal{M}^{(2)}f>r\}$. Hence,

$$\begin{aligned}\sum_{i=-\infty}^{\infty}\sum_j|f_i^j|^2\nu_i^j&=\int_0^\infty 2r\sum_{i=-\infty}^{\infty}\sum_j 1_{\{|f_i^j|>r\}}\nu_i^j\,\mathrm{d}r\le\int_0^\infty 2r\sum_k\sum_{\Delta\subset\Delta_k(r)}\nu\big(\Delta\times\big(\tfrac12\ell(\Delta),\ell(\Delta)\big]\big)\,\mathrm{d}r\\&=\int_0^\infty 2r\sum_k\nu(\Delta_k(r)\times(0,\ell(\Delta_k(r))])\,\mathrm{d}r\\&\le\|\nu\|_{\mathcal{C}}\int_0^\infty 2r\sum_k\mu(\Delta_k(r))\,\mathrm{d}r\le\|\nu\|_{\mathcal{C}}\int_0^\infty 2r\mu(\{\mathcal{M}^{(1)}\mathcal{M}^{(2)}f>r\})\,\mathrm{d}r\\&=\|\nu\|_{\mathcal{C}}\|\mathcal{M}^{(1)}\mathcal{M}^{(2)}f\|_{2,\mu}^2.\end{aligned}$$

Now, the claim follows from the Hardy–Littlewood–Muckenhoupt inequality. □

The rest of the section is devoted to the proof of the following lemma.

Lemma 8.3. *For all dyadic parabolic cubes $\Delta = Q \times I \subset \mathbb{R}^{n+1}$,*

$$\int_0^{\ell(\Delta)} \iint_\Delta |\mathcal{U}_\lambda w^{-1} A|^2 \frac{\mathrm{d}\mu\, \mathrm{d}\lambda}{\lambda} \lesssim \mu(\Delta).$$

The proof of Lemma 8.3 is based on the use of appropriate local Tb-type test functions.

8.1. ***Construction of appropriate local Tb-type test functions.*** Let $\zeta \in \mathbb{C}^n$ with $|\zeta| = 1$, and let ζ_i denote the i-th component of ζ for $1 \le i \le n$. We let χ and η be smooth functions on $\mathbb{R}^n$ and $\mathbb{R}$, respectively, whose values are in [0, 1]. The function χ is equal to 1 on $\left[-\frac{1}{2}, \frac{1}{2}\right]^n$ and has support in $(-1, 1)^n$, and η is equal to 1 on $\left[-\frac{1}{2}, \frac{1}{2}\right]$ with support in $(-1, 1)$. We fix a parabolic dyadic cube Δ and denote its center by (x_Δ, t_Δ). We first introduce

$$\chi_\Delta(x, t) := \chi\left(\frac{x - x_\Delta}{\ell(\Delta)}\right)\eta\left(\frac{t - t_\Delta}{\ell(\Delta)^2}\right).$$

Based on ζ and χ_Δ, we introduce

$$L_\Delta^\zeta(x, t) := \chi_\Delta(x, t)(\Phi_\Delta(x) \cdot \bar{\zeta}), \qquad \Phi_\Delta(x) := (x - x_\Delta).$$

Clearly, $L_\Delta^\zeta \in \mathrm{E}_\mu$. Using the function L_Δ^ζ and $0 < \epsilon \ll 1$, we define the test function

$$f_{\Delta,\epsilon}^\zeta := \mathcal{E}_{\epsilon\ell(\Delta)} L_\Delta^\zeta = (I + (\epsilon\ell(\Delta))^2\mathcal{H})^{-1} L_\Delta^\zeta. \tag{8-6}$$

Lemma 8.4. *Let $\zeta \in \mathbb{C}^n$ with $|\zeta| = 1$, and let $0 < \epsilon \ll 1$ be a degree of freedom. Given a parabolic dyadic cube Δ, define $f_{\Delta,\epsilon}^\zeta$ as in* (8-6)*. Then,*

(i) $\|f_{\Delta,\epsilon}^\zeta - L_\Delta^\zeta\|_{2,\mu}^2 \lesssim (\epsilon\ell(\Delta))^2\mu(\Delta)$,

(ii) $\|\mathbb{D}(f_{\Delta,\epsilon}^\zeta - L_\Delta^\zeta)\|_{2,\mu}^2 \lesssim \mu(\Delta)$,

(iii) $\|\mathbb{D} f_{\Delta,\epsilon}^\zeta\|_{2,\mu}^2 \lesssim \mu(\Delta)$.

Proof. Note that

$$\begin{aligned} f_{\Delta,\epsilon}^\zeta - L_\Delta^\zeta &= -(\epsilon\ell(\Delta))^2\mathcal{E}_{\epsilon\ell(\Delta)}\mathcal{H}L_\Delta^\zeta \\ &= -(\epsilon\ell(\Delta))^2\mathcal{E}_{\epsilon\ell(\Delta)} D_t^{1/2} H_t D_t^{1/2} L_\Delta^\zeta + (\epsilon\ell(\Delta))^2\mathcal{E}_{\epsilon\ell(\Delta)} w^{-1}\operatorname{div}_x w(w^{-1}A\nabla_x L_\Delta^\zeta). \end{aligned}$$

Hence, using the uniform L_μ^2-boundedness of $(\epsilon\ell(\Delta))\mathcal{E}_{\epsilon\ell(\Delta)} D_t^{1/2}$ and $(\epsilon\ell(\Delta))\mathcal{E}_{\epsilon\ell(\Delta)} w^{-1}\operatorname{div}_x w$, see Lemma 4.3, we get

$$\iint_{\mathbb{R}^{n+1}} |f_{\Delta,\epsilon}^\zeta - L_\Delta^\zeta|^2 \,\mathrm{d}\mu \lesssim \iint_{\mathbb{R}^{n+1}} |(\epsilon\ell(\Delta))\mathbb{D}L_\Delta^\zeta|^2 \,\mathrm{d}\mu.$$

Furthermore,

$$\iint_{\mathbb{R}^{n+1}} |\mathbb{D}L_\Delta^\zeta|^2 \,\mathrm{d}\mu = \iint_{\mathbb{R}^{n+1}} |\nabla_x L_\Delta^\zeta|^2 \,\mathrm{d}\mu + \iint_{\mathbb{R}^{n+1}} |D_t^{1/2} L_\Delta^\zeta|^2 \,\mathrm{d}\mu \lesssim \mu(\Delta) \tag{8-7}$$

by the construction of L_Δ^ζ (to estimate $D_t^{1/2} L_\Delta^\zeta$ we use the homogeneity of the Fourier symbol). Similarly, we deduce that

$$\iint_{\mathbb{R}^{n+1}} |\mathbb{D}(f_{\Delta,\epsilon}^\zeta - L_\Delta^\zeta)|^2 \,\mathrm{d}\mu \lesssim \mu(\Delta).$$

This proves (i) and (ii). To prove (iii), we simply use (ii) and (8-7). □

Lemma 8.5. *Given a parabolic dyadic cube* $\Delta = Q \times I$, *let* $f^{\zeta}_{\Delta,\epsilon}$ *be defined as in* (8-6). *There exist* $\epsilon \in (0, 1)$, *depending only on the structural constants, and a finite set* W *of unit vectors in* $\mathbb{C}^n$, *whose cardinality depends on* ϵ *and* n, *such that*

$$\sup_{\Delta} \frac{1}{|\Delta|} \int_0^{\ell(\Delta)} \iint_{\Delta} |\mathcal{U}_\lambda w^{-1} A|^2 \frac{\mathrm{d}\mu\,\mathrm{d}\lambda}{\lambda} \lesssim \sum_{\zeta \in W} \sup_{\Delta} \frac{1}{|\Delta|} \int_0^{\ell(\Delta)} \iint_{\Delta} |(\mathcal{U}_\lambda w^{-1} A) \cdot \mathcal{A}_\lambda \nabla_x f^{\zeta}_{\Delta,\epsilon}|^2 \frac{\mathrm{d}\mu\,\mathrm{d}\lambda}{\lambda},$$

where the supremum is taken over all dyadic parabolic cubes $\Delta \subset \mathbb{R}^{n+1}$.

Proof. Consider a degree of freedom $\epsilon > 0$. Given a unit vector ζ in $\mathbb{C}^n$, we introduce the cone

$$C^{\epsilon}_{\zeta} := \{u \in \mathbb{C}^n : |u - (u \cdot \bar{\zeta})\zeta| \le \epsilon |u \cdot \bar{\zeta}|\}.$$

We note that we can cover $\mathbb{C}^n$ by a finite number of such cones $\{C^{\epsilon}_{\zeta}\}$. The number of cones that are needed depends on ϵ and n. In the following, we fix one C^{ϵ}_{ζ}. We let

$$\gamma^{\epsilon}_{\lambda,\zeta}(x, t) := 1_{C^{\epsilon}_{\zeta}}(\mathcal{U}_\lambda w^{-1} A(x, t)) \cdot \mathcal{U}_\lambda w^{-1} A(x, t)$$

and consider a fixed dyadic parabolic cube $\Delta = Q \times I \subset \mathbb{R}^{n+1}$.

Step 1: Estimate of the test function along $\bar{\zeta}$. We first estimate

$$\iint_{\Delta} (1 - \nabla_x f^{\zeta}_{\Delta,\epsilon} \cdot \zeta)\,\mathrm{d}x\,\mathrm{d}t. \tag{8-8}$$

To start the estimate, we write

$$1 - \nabla_x f^{\zeta}_{\Delta,\epsilon} \cdot \zeta = \nabla_x g^{\zeta}_{\Delta,\epsilon} \cdot \zeta + (1 - \nabla_x L^{\zeta}_{\Delta} \cdot \zeta),$$

where $g^{\zeta}_{\Delta,\epsilon} := L^{\zeta}_{\Delta} - f^{\zeta}_{\Delta,\epsilon}$. By construction, we have $\nabla_x L^{\zeta}_{\Delta}(x, t) = \bar{\zeta}$ whenever $(x, t) \in \Delta$. Hence,

$$\iint_{\Delta} (1 - \nabla_x L^{\zeta}_{\Delta} \cdot \zeta)\,\mathrm{d}x\,\mathrm{d}t = 0.$$

We have to estimate the contribution to the integral in (8-8) coming from $\nabla_x g^{\zeta}_{\Delta,\epsilon} \cdot \zeta$. To do this, let $s \in (0, 1)$ yet to be chosen, and let $\varphi\colon \mathbb{R}^{n+1} \to [0, 1]$ be a smooth function which is 1 on $\Delta_s := (1-s)Q \times (1-s^2)I$, supported on Δ, and satisfies $\|\nabla_x \varphi\|_\infty \le c(s\ell(\Delta))^{-1}$, $\|\partial_t \varphi\|_\infty \le c(s\ell(\Delta))^{-2}$ for a dimensional constant $c > 0$. Using φ, we see that

$$\iint_{\Delta} \nabla_x g^{\zeta}_{\Delta,\epsilon} \cdot \zeta\,\mathrm{d}x\,\mathrm{d}t = \iint_{\Delta} (1 - \varphi) \nabla_x g^{\zeta}_{\Delta,\epsilon} \cdot \zeta\,\mathrm{d}x\,\mathrm{d}t + \iint_{\Delta} \varphi \nabla_x g^{\zeta}_{\Delta,\epsilon} \cdot \zeta\,\mathrm{d}x\,\mathrm{d}t =: \mathrm{I} + \mathrm{II}.$$

Using the Cauchy–Schwarz inequality, Lemma 8.4 (ii) and (2-2) for the A_2-weight $\mu^{-1}(x, t) = w^{-1}(x)$, we obtain

$$\begin{aligned} |\mathrm{I}| &\le \Bigl(\iint_{\Delta} |1 - \varphi|^2\,\mathrm{d}\mu^{-1}\Bigr)^{1/2} \Bigl(\iint_{\Delta} |\nabla_x g^{\zeta}_{\Delta,\epsilon}|^2\,\mathrm{d}\mu\Bigr)^{1/2} \\ &\lesssim \mu^{-1}(\Delta \setminus \Delta_s)^{1/2} \mu(\Delta)^{1/2} \lesssim s^{\eta} \mu^{-1}(\Delta)^{1/2} \mu(\Delta)^{1/2} \le s^{\eta} [w]_{A_2} |\Delta|. \end{aligned}$$

To estimate II, we integrate by parts to get

$$\mathrm{II} = -\iint_{\mathbb{R}^{n+1}} g^{\zeta}_{\Delta,\epsilon} \nabla_x \varphi \cdot \zeta \, \mathrm{d}x \, \mathrm{d}t,$$

and using the Cauchy–Schwarz inequality and Lemma 8.4 (i), we obtain similarly

$$\begin{aligned} |\mathrm{II}| &\le \biggl(\iint_{\mathbb{R}^{n+1}} |\nabla_x \varphi|^2 \, \mathrm{d}\mu^{-1} \biggr)^{1/2} \biggl(\iint_{\mathbb{R}^{n+1}} |g^{\zeta}_{\Delta,\epsilon}|^2 \, \mathrm{d}\mu \biggr)^{1/2} \\ &\lesssim (s\ell(\Delta))^{-1} \mu(\Delta)^{1/2} \epsilon \ell(\Delta) \mu^{-1}(\Delta)^{1/2} \le \epsilon s^{-1} [w]_{A_2} |\Delta|. \end{aligned}$$

We now choose $s = \epsilon^{1/(\eta+1)}$, so that the estimates for I and II come with the same power of ϵ. Putting the estimates together, we obtain, for the integral in (8-8), that

$$\frac{1}{|\Delta|} \biggl| \iint_{\Delta} 1 - \nabla_x f^{\zeta}_{\Delta,\epsilon} \cdot \zeta \, \mathrm{d}x \, \mathrm{d}t \biggr| \lesssim \epsilon^{\eta/(\eta+1)}. \tag{8-9}$$

Using Lemma 8.4 (iii) and the Cauchy–Schwarz inequality, we also see that

$$\frac{1}{|\Delta|} \iint_{\Delta} |\nabla_x f^{\zeta}_{\Delta,\epsilon}| \, \mathrm{d}x \, \mathrm{d}t \le \frac{1}{|\Delta|} \biggl(\iint_{\Delta} |\nabla_x f^{\zeta}_{\Delta,\epsilon}|^2 \, \mathrm{d}\mu \biggr)^{1/2} \mu^{-1}(\Delta)^{1/2} \lesssim 1. \tag{8-10}$$

Step 2: Choice of ϵ. Using the estimates in the last two displays, we see, if ϵ is chosen small enough, that

$$\frac{1}{|\Delta|} \iint_{\Delta} \mathrm{Re}(\nabla_x f^{\zeta}_{\Delta,\epsilon} \cdot \zeta) \, \mathrm{d}x \, \mathrm{d}t \ge \tfrac{7}{8}$$

and

$$\frac{1}{|\Delta|} \iint_{\Delta} |\nabla_x f^{\zeta}_{\Delta,\epsilon}| \, \mathrm{d}x \, \mathrm{d}t \le c$$

for some large constant c depending only on the structural constants. We now perform a stopping-time decomposition as in [Auscher et al. 2002] to select a collection $\mathcal{S}'_{\zeta} = \{\Delta'\}$ of dyadic parabolic subcubes of Δ, which are maximal with respect to the property that either

$$\frac{1}{|\Delta'|} \iint_{\Delta'} \mathrm{Re}(\nabla_x f^{\zeta}_{\Delta,\epsilon} \cdot \zeta) \, \mathrm{d}x \, \mathrm{d}t \le \tfrac{3}{4} \tag{8-11}$$

or

$$\frac{1}{|\Delta'|} \iint_{\Delta'} |\nabla_x f^{\zeta}_{\Delta,\epsilon}| \, \mathrm{d}x \, \mathrm{d}t \ge (4\epsilon)^{-2} \tag{8-12}$$

holds. In other words, we parabolically dyadically subdivide Δ and stop the first time either (8-11) or (8-12) hold. Then, $\mathcal{S}'_{\zeta} = \{\Delta'\}$ is a disjoint set of the parabolic dyadic subcubes of Δ. Let $\mathcal{S}''_{\zeta} = \{\Delta''\}$ be the collection of all the parabolic dyadic subcubes of Δ not contained in any $\Delta' \in \mathcal{S}'_{\zeta}$. Then, each $\Delta'' \in \mathcal{S}''_{\zeta}$ satisfies

$$\begin{aligned} \frac{1}{|\Delta''|} \iint_{\Delta''} \mathrm{Re}(\nabla_x f^{\zeta}_{\Delta,\epsilon} \cdot \zeta) \, \mathrm{d}x \, \mathrm{d}t &\ge \tfrac{3}{4}, \\ \frac{1}{|\Delta''|} \iint_{\Delta''} |\nabla_x f^{\zeta}_{\Delta,\epsilon}| \, \mathrm{d}x \, \mathrm{d}t &\le (4\epsilon)^{-2}. \end{aligned} \tag{8-13}$$

At this stage, we claim that, by the same type of argument as in the proof of statement (i) in Proposition 5.7 in [Auscher et al. 2002], there exists $\epsilon \in (0, 1)$ even smaller and depending only on the structural constants and $\eta' = \eta'(\epsilon) \in (0, 1)$ such that

$$\Big|\bigcup_{\Delta' \in \mathcal{S}'_\zeta} \Delta'\Big| \leq (1-\eta')|\Delta|. \tag{8-14}$$

In particular, from now on ϵ is fixed. For completeness and the convenience of the reader, we include a proof here.

Let E_1 and E_2 be the unions of all parabolic cubes in $\mathcal{S}'_\zeta$ which satisfy (8-11) and (8-12), respectively. Then,

$$\Big|\bigcup_{\Delta' \in \mathcal{S}'_\zeta} \Delta'\Big| \leq |E_1| + |E_2|.$$

For $|E_2|$, we have

$$|E_2| \leq (4\epsilon)^2 \sum_{\Delta' \in \mathcal{S}'_\zeta} \iint_{\Delta'} |\nabla_x f^\zeta_{\Delta,\epsilon}| \,\mathrm{d}x\,\mathrm{d}t \leq (4\epsilon)^2 \iint_{\Delta} |\nabla_x f^\zeta_{\Delta,\epsilon}| \,\mathrm{d}x\,\mathrm{d}t \leq (4\epsilon)^2 c|\Delta|,$$

where we used (8-10) in the last step. To control $|E_1|$, we let $h := 1 - \mathrm{Re}(\nabla_x f^\zeta_{\Delta,\epsilon} \cdot \zeta)$ and write

$$|E_1| \leq 4 \sum_{\Delta'} \iint_{\Delta'} h \,\mathrm{d}x\,\mathrm{d}t = 4 \iint_{\Delta} h \,\mathrm{d}x\,\mathrm{d}t - 4 \iint_{\Delta \setminus E_1} h \,\mathrm{d}x\,\mathrm{d}t, \tag{8-15}$$

where the sum is taken over all parabolic subcubes of E_1. By (8-9), the first term on the right is controlled by $\epsilon^{\eta/(\eta+1)}|\Delta|$ times a constant depending on the structural constants. Using in succession the Cauchy–Schwarz inequality, Lemma 8.4 (iii), the A_2-property and Young's inequality, the second term on the right is controlled by

$$\begin{aligned} 4|\Delta \setminus E_1| + 4\mu^{-1}(\Delta \setminus E_1)^{1/2}\Big(\iint_{\Delta} |\nabla_x f^\zeta_{\Delta,\epsilon}|^2 \,\mathrm{d}\mu\Big)^{1/2} &\leq 4|\Delta \setminus E_1| + 4\tilde{c}\mu^{-1}(\Delta \setminus E_1)^{1/2}\mu(\Delta)^{1/2} \\ &\leq 4|\Delta \setminus E_1| + 4\tilde{c}|\Delta \setminus E_1|^{\eta}|\Delta|^{1-\eta} \\ &\leq (4 + \tilde{c}\epsilon^{-1/\eta})|\Delta \setminus E_1| + \tilde{c}\epsilon^{1-\eta}|\Delta|, \end{aligned}$$

where $\tilde{c}$ depends on the structural constants and changes from line to line. Going back to (8-15) and rearranging terms, we find

$$|E_1| \leq \frac{4 + \tilde{c}\epsilon^{-1/\eta} + \tilde{c}(\epsilon^{\eta/(\eta+1)} + \epsilon^{1-\eta})}{5 + \tilde{c}\epsilon^{-1/\eta}}|\Delta|,$$

and, taking ϵ small enough, we conclude (8-14).

Since μ is an A_2-weight, we obtain from (8-14) — and upon taking η' smaller depending on the structural constants and ϵ — that

$$\mu\Big(\bigcup_{\Delta' \in \mathcal{S}'_\zeta} \Delta'\Big) \leq (1-\eta')\mu(\Delta); \tag{8-16}$$

see, for example, [Stein 1993, p. 196] for this A_∞-property of A_2-weights.

Step 3: Reintroducing the averaging operator. Given Δ, we consider $\Delta'' \in \mathcal{S}''_\zeta$ as above. Set

$$v := \frac{1}{\mu(\Delta'')} \iint_{\Delta''} \nabla_x f^\zeta_{\Delta,\epsilon} \,\mathrm{d}x\,\mathrm{d}t \in \mathbb{C}^n. \tag{8-17}$$

If $(x,t) \in \Delta''$ and $\frac{1}{2}\ell(\Delta'') < \lambda \le \ell(\Delta'')$, then $v = (\mathcal{A}_\lambda \nabla_x f^\zeta_{\Delta,\epsilon})(x,t)$. Assume that

$$u := (\mathcal{U}_\lambda w^{-1} A)(x,t) \in C^\epsilon_\zeta.$$

The pair of vectors (u, v) satisfies the estimates in (8-13). Thus, we can apply [Auscher et al. 2002, Lemma 5.10] with $w := \zeta$ and conclude that $|u| \le 4|u \cdot v|$; that is,

$$|\gamma^\epsilon_{\lambda,\zeta}(x,t)| \le 4|(\mathcal{U}_\lambda w^{-1} A(x,t)) \cdot (\mathcal{A}_\lambda \nabla_x f^\zeta_{\Delta,\epsilon})(x,t)|. \tag{8-18}$$

We next observe that, by construction, the Carleson box $\Delta \times (0, \ell(\Delta)]$ can be partitioned into Carleson boxes $\Delta' \times (0, \ell(\Delta')]$, with $\Delta' \in \mathcal{S}'_\zeta$, and Whitney boxes $\Delta'' \times \big(\frac{1}{2}\ell(\Delta''), \ell(\Delta'')\big]$, with $\Delta'' \in \mathcal{S}''_\zeta$. In particular,

$$\frac{1}{\mu(\Delta)} \int_0^{\ell(\Delta)} \iint_\Delta |\gamma^\epsilon_{\lambda,\zeta}(x,t)|^2 \frac{\mathrm{d}\mu\,\mathrm{d}\lambda}{\lambda} =: \mathrm{I} + \mathrm{II},$$

where

$$\mathrm{I} := \frac{1}{\mu(\Delta)} \sum_{\Delta' \in \mathcal{S}'_\zeta} \int_0^{\ell(\Delta')} \iint_{\Delta'} |\gamma^\epsilon_{\lambda,\zeta}(x,t)|^2 \frac{\mathrm{d}\mu\,\mathrm{d}\lambda}{\lambda},$$

$$\mathrm{II} := \frac{1}{\mu(\Delta)} \sum_{\Delta'' \in \mathcal{S}''_\zeta} \int_{\ell(\Delta'')/2}^{\ell(\Delta'')} \iint_{\Delta''} |\gamma^\epsilon_{\lambda,\zeta}(x,t)|^2 \frac{\mathrm{d}\mu\,\mathrm{d}\lambda}{\lambda}.$$

Using (8-16), we obtain

$$\mathrm{I} \le \frac{1}{\mu(\Delta)} \sum_{\Delta' \in \mathcal{S}'_\zeta} A^\epsilon_\zeta \mu(\Delta') \le (1-\eta') A^\epsilon_\zeta,$$

where

$$A^\epsilon_\zeta := \sup_{\tilde{\Delta}} \frac{1}{\mu(\tilde{\Delta})} \int_0^{\ell(\tilde{\Delta})} \iint_{\tilde{\Delta}} |\gamma^\epsilon_{\lambda,\zeta}(x,t)|^2 \frac{\mathrm{d}\mu\,\mathrm{d}\lambda}{\lambda},$$

and where the supremum is taken over all dyadic parabolic subcubes $\tilde{\Delta} \subset \Delta$. By (8-18), we have

$$\mathrm{II} \le \frac{16}{\mu(\Delta)} \int_0^{\ell(\Delta)} \iint_\Delta |(\mathcal{U}_\lambda w^{-1} A)(x,t) \cdot (\mathcal{A}_\lambda \nabla_x f^\zeta_{\Delta,\epsilon})(x,t)|^2 \frac{\mathrm{d}\mu\,\mathrm{d}\lambda}{\lambda}.$$

Since these estimates hold for all dyadic parabolic cubes, in particular those which are subcubes of Δ, we conclude that

$$A^\epsilon_\zeta \le (1-\eta') A^\epsilon_\zeta + \sup_{\tilde{\Delta}} \frac{16}{\mu(\tilde{\Delta})} \int_0^{\ell(\tilde{\Delta})} \iint_{\tilde{\Delta}} |(\mathcal{U}_\lambda w^{-1} A)(x,t) \cdot (\mathcal{A}_\lambda \nabla_x f^\zeta_{\tilde{\Delta},\epsilon})(x,t)|^2 \frac{\mathrm{d}\mu\,\mathrm{d}\lambda}{\lambda}.$$

Summing with respect to $\zeta \in W$ completes the proof of Lemma 8.5 under the a priori assumption that A^ϵ_ζ is qualitatively finite, since it can then be absorbed into the left-hand side.

Step 4: Removing the a priori assumption. The a priori assumption that A_ζ^ϵ is qualitatively finite can be removed by setting $\gamma_{\lambda,\zeta}^\epsilon(x,t)$ to 0 for λ small and large, repeating the argument from (8-18) on and passing to the limit at the end. For the truncated $\gamma_{\lambda,\zeta}^\epsilon(x,t)$, we get $A_\zeta^\epsilon < \infty$ from (7-3). Indeed, for $0 < \delta < 1$ small, we have

$$A_\zeta^\epsilon \le \int_\delta^{\delta^{-1}} \left(\sup_{\tilde{\Delta}} \frac{1}{\mu(\tilde{\Delta})} \iint_{\tilde{\Delta}} |\mathcal{U}_\lambda w^{-1} A|^2 \, \mathrm{d}\mu \right) \frac{\mathrm{d}\lambda}{\lambda} \le \int_\delta^{\delta^{-1}} C \, \frac{\mathrm{d}\lambda}{\lambda} < \infty,$$

where C depends on $\ell(\Delta)$ and δ. This completes the argument. □

8.2. ***The Carleson measure estimate: proof of Lemma 8.3.*** Thanks to Lemma 8.5, it suffices to prove

$$\int_0^{\ell(\Delta)} \iint_\Delta |(\mathcal{U}_\lambda w^{-1} A) \cdot \mathcal{A}_\lambda \nabla_x f_{\Delta,\epsilon}^\zeta|^2 \, \frac{\mathrm{d}\mu \, \mathrm{d}\lambda}{\lambda} \lesssim \mu(\Delta). \tag{8-19}$$

The left-hand side in (8-19) is bounded by

$$|||(\lambda \mathcal{E}_\lambda \mathcal{H} + (\mathcal{U}_\lambda w^{-1} A) \cdot \mathcal{A}_\lambda \nabla_x) f_{\Delta,\epsilon}^\zeta|||_{2,\mu}^2 + \int_0^{\ell(\Delta)} \iint_\Delta |\lambda \mathcal{E}_\lambda \mathcal{H} f_{\Delta,\epsilon}^\zeta|^2 \, \frac{\mathrm{d}\mu \, \mathrm{d}\lambda}{\lambda} =: \mathrm{I} + \mathrm{II}.$$

By Proposition 8.1 and Lemma 8.4, we have

$$\mathrm{I} \lesssim \|\mathbb{D} f_{\Delta,\epsilon}^\zeta\|_{2,\mu}^2 \lesssim \mu(\Delta).$$

As for II, we obtain from (8-6) that

$$\mathcal{H} f_{\Delta,\epsilon}^\zeta = \frac{(L_\Delta^\zeta - f_{\Delta,\epsilon}^\zeta)}{(\epsilon \ell(\Delta))^2}.$$

Using the L_μ^2-boundedness of $\mathcal{E}_\lambda$, see Lemma 4.3, and then Lemma 8.4, we obtain

$$\begin{aligned} \mathrm{II} &\lesssim \int_0^{\ell(\Delta)} \|\lambda (\epsilon \ell(\Delta))^{-2} (L_\Delta^\zeta - f_{\Delta,\epsilon}^\zeta)\|_{2,\mu}^2 \frac{\mathrm{d}\lambda}{\lambda} \\ &= \frac{1}{2\epsilon^4 \ell(\Delta)^2} \|L_\Delta^\zeta - f_{\Delta,\epsilon}^\zeta\|_{2,\mu}^2 \lesssim \epsilon^{-2} \mu(\Delta). \end{aligned}$$

This completes the proof of (8-19), and hence the proof of Theorem 1.1.

References

[Alfonseca et al. 2011] M. A. Alfonseca, P. Auscher, A. Axelsson, S. Hofmann, and S. Kim, "Analyticity of layer potentials and L^2 solvability of boundary value problems for divergence form elliptic equations with complex L^∞ coefficients", *Adv. Math.* **226**:5 (2011), 4533–4606. MR Zbl

[Amenta and Auscher 2018] A. Amenta and P. Auscher, *Elliptic boundary value problems with fractional regularity data: the first order approach*, CRM Monogr. Ser. **37**, Amer. Math. Soc., Providence, RI, 2018. MR Zbl

[Auscher and Axelsson 2011] P. Auscher and A. Axelsson, "Weighted maximal regularity estimates and solvability of non-smooth elliptic systems, I", *Invent. Math.* **184**:1 (2011), 47–115. MR Zbl

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

[Auscher and Mourgoglou 2019] P. Auscher and M. Mourgoglou, "Representation and uniqueness for boundary value elliptic problems via first order systems", *Rev. Mat. Iberoam.* **35**:1 (2019), 241–315. MR Zbl

[Auscher and Rosén 2012] P. Auscher and A. Rosén, "Weighted maximal regularity estimates and solvability of nonsmooth elliptic systems, II", *Anal. PDE* **5**:5 (2012), 983–1061. MR Zbl

[Auscher and Tchamitchian 1998] P. Auscher and P. Tchamitchian, *Square root problem for divergence operators and related topics*, Astérisque **249**, Soc. Math. France, Paris, 1998. MR Zbl

[Auscher et al. 2002] P. Auscher, S. Hofmann, M. Lacey, A. McIntosh, and P. Tchamitchian, "The solution of the Kato square root problem for second order elliptic operators on $\mathbb{R}^n$", *Ann. of Math.* (2) **156**:2 (2002), 633–654. MR Zbl

[Auscher et al. 2015] P. Auscher, A. Rosén, and D. Rule, "Boundary value problems for degenerate elliptic equations and systems", *Ann. Sci. École Norm. Sup.* (4) **48**:4 (2015), 951–1000. MR Zbl

[Auscher et al. 2018] P. Auscher, M. Egert, and K. Nyström, "The Dirichlet problem for second order parabolic operators in divergence form", *J. Éc. Polytech. Math.* **5** (2018), 407–441. MR Zbl

[Auscher et al. 2020] P. Auscher, M. Egert, and K. Nyström, "L^2 well-posedness of boundary value problems for parabolic systems with measurable coefficients", *J. Eur. Math. Soc.* **22**:9 (2020), 2943–3058. MR Zbl

[Castro et al. 2016] A. J. Castro, K. Nyström, and O. Sande, "Boundedness of single layer potentials associated to divergence form parabolic equations with complex coefficients", *Calc. Var. Partial Differential Equations* **55**:5 (2016), art. id. 124. MR Zbl

[Chiarenza and Serapioni 1985] F. Chiarenza and R. Serapioni, "A remark on a Harnack inequality for degenerate parabolic equations", *Rend. Sem. Mat. Univ. Padova* **73** (1985), 179–190. MR Zbl

[Cruz-Uribe and Rios 2008] D. Cruz-Uribe and C. Rios, "Gaussian bounds for degenerate parabolic equations", *J. Funct. Anal.* **255**:2 (2008), 283–312. Correction in **267**:9 (2014), 3507–3513. MR Zbl

[Cruz-Uribe and Rios 2012] D. Cruz-Uribe and C. Rios, "The solution of the Kato problem for degenerate elliptic operators with Gaussian bounds", *Trans. Amer. Math. Soc.* **364**:7 (2012), 3449–3478. MR Zbl

[Cruz-Uribe and Rios 2015] D. Cruz-Uribe and C. Rios, "The Kato problem for operators with weighted ellipticity", *Trans. Amer. Math. Soc.* **367**:7 (2015), 4727–4756. MR Zbl

[Cruz-Uribe et al. 2018] D. Cruz-Uribe, J. M. Martell, and C. Rios, "On the Kato problem and extensions for degenerate elliptic operators", *Anal. PDE* **11**:3 (2018), 609–660. MR Zbl

[David et al. 1985] G. David, J.-L. Journé, and S. Semmes, "Opérateurs de Calderón–Zygmund, fonctions para-accrétives et interpolation", *Rev. Mat. Iberoam.* **1**:4 (1985), 1–56. MR Zbl

[Di Nezza et al. 2012] E. Di Nezza, G. Palatucci, and E. Valdinoci, "Hitchhiker's guide to the fractional Sobolev spaces", *Bull. Sci. Math.* **136**:5 (2012), 521–573. MR Zbl

[Dier and Zacher 2017] D. Dier and R. Zacher, "Non-autonomous maximal regularity in Hilbert spaces", *J. Evol. Equ.* **17**:3 (2017), 883–907. MR Zbl

[Escauriaza and Hofmann 2018] L. Escauriaza and S. Hofmann, "Kato square root problem with unbounded leading coefficients", *Proc. Amer. Math. Soc.* **146**:12 (2018), 5295–5310. MR Zbl

[García-Cuerva and Rubio de Francia 1985] J. García-Cuerva and J. L. Rubio de Francia, *Weighted norm inequalities and related topics*, North-Holland Math. Stud. **116**, North-Holland, Amsterdam, 1985. MR Zbl

[Haase 2006] M. Haase, *The functional calculus for sectorial operators*, Operator Theory: Adv. Appl. **169**, Birkhäuser, Basel, 2006. MR Zbl

[Heinonen et al. 1993] J. Heinonen, T. Kilpeläinen, and O. Martio, *Nonlinear potential theory of degenerate elliptic equations*, Oxford Univ. Press, 1993. MR Zbl

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

[Hofmann et al. 2015] S. Hofmann, C. Kenig, S. Mayboroda, and J. Pipher, "Square function/non-tangential maximal function estimates and the Dirichlet problem for non-symmetric elliptic operators", *J. Amer. Math. Soc.* **28**:2 (2015), 483–529. MR Zbl

[Hofmann et al. 2019] S. Hofmann, P. Le, and A. J. Morris, "Carleson measure estimates and the Dirichlet problem for degenerate elliptic equations", *Anal. PDE* **12**:8 (2019), 2095–2146. MR Zbl

[Hofmann et al. 2022] S. Hofmann, L. Li, S. Mayboroda, and J. Pipher, "The Dirichlet problem for elliptic operators having a BMO anti-symmetric part", *Math. Ann.* **382**:1-2 (2022), 103–168. MR Zbl

[Ishige et al. 2017] K. Ishige, Y. Kabeya, and E. M. Ouhabaz, "The heat kernel of a Schrödinger operator with inverse square potential", *Proc. Lond. Math. Soc.* (3) **115**:2 (2017), 381–410. MR Zbl

[Kaplan 1966] S. Kaplan, "Abstract boundary value problems for linear parabolic equations", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (3) **20**:2 (1966), 395–419. MR Zbl

[Kato 1966] T. Kato, *Perturbation theory for linear operators*, Grundl. Math. Wissen. **132**, Springer, 1966. MR Zbl

[Kilpeläinen 1994] T. Kilpeläinen, "Weighted Sobolev spaces and capacity", *Ann. Acad. Sci. Fenn. Ser. A I Math.* **19**:1 (1994), 95–113. MR Zbl

[Lions 1962] J.-L. Lions, "Espaces d'interpolation et domaines de puissances fractionnaires d'opérateurs", *J. Math. Soc. Japan* **14** (1962), 233–241. MR Zbl

[Litsgård and Nyström 2023] M. Litsgård and K. Nyström, "On local regularity estimates for fractional powers of parabolic operators with time-dependent measurable coefficients", *J. Evol. Equ.* **23**:1 (2023), art. id. 3. MR Zbl

[McIntosh 1986] A. McIntosh, "Operators which have an H_∞ functional calculus", pp. 210–231 in *Miniconference on operator theory and partial differential equations* (North Ryde, Australia, 1986), edited by B. Jefferies et al., Proc. Centre Math. Anal. Austral. Nat. Univ. **14**, Austral. Nat. Univ., Canberra, 1986. MR Zbl

[Morris 2012] A. J. Morris, "The Kato square root problem on submanifolds", *J. Lond. Math. Soc.* (2) **86**:3 (2012), 879–910. MR Zbl

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

[Nyström 2017] K. Nyström, "L^2 solvability of boundary value problems for divergence form parabolic equations with complex coefficients", *J. Differential Equations* **262**:3 (2017), 2808–2939. MR Zbl

[Ouhabaz 2021] E. M. Ouhabaz, "The square root of a parabolic operator", *J. Fourier Anal. Appl.* **27**:3 (2021), art. id. 59. MR Zbl

[Stein 1993] E. M. Stein, *Harmonic analysis*: *real-variable methods*, *orthogonality*, *and oscillatory integrals*, Princeton Math. Ser. **43**, Princeton Univ. Press, 1993. MR Zbl

Received 22 Sep 2022. Revised 30 Aug 2023. Accepted 19 Sep 2023.

ALIREZA ATAEI: alireza.ataei@math.uu.se
Department of Mathematics, Uppsala University, Uppsala, Sweden

MORITZ EGERT: egert@mathematik.tu-darmstadt.de
Fachbereich Mathematik, TU Darmstadt, Darmstadt, Germany

KAJ NYSTRÖM: kaj.nystrom@math.uu.se
Department of Mathematics, Uppsala University, Uppsala, Sweden

Analysis & PDE

msp.org/apde

EDITOR-IN-CHIEF

Clément Mouhot — Cambridge University, UK — c.mouhot@dpmms.cam.ac.uk

BOARD OF EDITORS

Massimiliano Berti — Scuola Intern. Sup. di Studi Avanzati, Italy — berti@sissa.it

Zbigniew Błocki — Uniwersytet Jagielloński, Poland — zbigniew.blocki@uj.edu.pl

Charles Fefferman — Princeton University, USA — cf@math.princeton.edu

David Gérard-Varet — Université de Paris, France — david.gerard-varet@imj-prg.fr

Colin Guillarmou — Université Paris-Saclay, France — colin.guillarmou@universite-paris-saclay.fr

Ursula Hamenstaedt — Universität Bonn, Germany — ursula@math.uni-bonn.de

Peter Hintz — ETH Zurich, Switzerland — peter.hintz@math.ethz.ch

Vadim Kaloshin — Institute of Science and Technology, Austria — vadim.kaloshin@gmail.com

Izabella Laba — University of British Columbia, Canada — ilaba@math.ubc.ca

Anna L. Mazzucato — Penn State University, USA — alm24@psu.edu

Richard B. Melrose — Massachussets Inst. of Tech., USA — rbm@math.mit.edu

Frank Merle — Université de Cergy-Pontoise, France — merle@ihes.fr

William Minicozzi II — Johns Hopkins University, USA — minicozz@math.jhu.edu

Werner Müller — Universität Bonn, Germany — mueller@math.uni-bonn.de

Igor Rodnianski — Princeton University, USA — irod@math.princeton.edu

Yum-Tong Siu — Harvard University, USA — siu@math.harvard.edu

Terence Tao — University of California, Los Angeles, USA — tao@math.ucla.edu

Michael E. Taylor — Univ. of North Carolina, Chapel Hill, USA — met@math.unc.edu

Gunther Uhlmann — University of Washington, USA — gunther@math.washington.edu

András Vasy — Stanford University, USA — andras@math.stanford.edu

Dan Virgil Voiculescu — University of California, Berkeley, USA — dvv@math.berkeley.edu

Jim Wright — University of Edinburgh, UK — j.r.wright@ed.ac.uk

Maciej Zworski — University of California, Berkeley, USA — zworski@math.berkeley.edu

PRODUCTION

production@msp.org

Silvio Levy, Scientific Editor

Cover image: Eric J. Heller: “Linear Ramp”

See inside back cover or msp.org/apde for submission instructions.

The subscription price for 2025 is US $475/year for the electronic version, and $735/year (+$70, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscriber address should be sent to MSP.

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online.

APDE peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

© 2025 Mathematical Sciences Publishers

ANALYSIS & PDE

Volume 18 No. 1 2025