Boundary value problems for degenerate elliptic equations and systems

Pascal Auscher, Andreas Rosén, David Rule

To cite this version:

BOUNDARY VALUE PROBLEMS FOR DEGENERATE ELLIPTIC EQUATIONS AND SYSTEMS

PASCAL AUSCHER, ANDREAS ROSÉN, AND DAVID RULE

ABSTRACT. We study boundary value problems for degenerate elliptic equations and systems with square integrable boundary data. We can allow for degeneracies in the form of an A_2 weight. We obtain representations and boundary traces for solutions in appropriate classes, perturbation results for solvability and solvability in some situations. The technology of earlier works of the first two authors can be adapted to the weighted setting once the needed quadratic estimate is established and we even improve some results in the unweighted setting. The proof of this quadratic estimate does not follow from earlier results on the topic and is the core of the article.

CONTENTS

1. Introduction

In the series of articles [21, 19, 20], degenerate elliptic equations in divergence form with real symmetric coefficients are studied. There, the degeneracy is given in terms of an A_2 weight or a power of the jacobian of a quasi-conformal map. The first article gives interior estimates, the second article deals with the Wiener test and the third one study boundary behavior and harmonic measure. Further work along these lines, for example [22, 13], has been done. However, little work on the fundamental L_p Dirichlet and Neumann problems in the degenerate setting seems to have been done. Here, we want to initiate such study of boundary value problems, with L_2 boundary data, for a large class of weights, in the case of domains which are Lipschitz diffeomorphic to the upper half space $\mathbb{R}^{1+n}_+ := \{(t,x) \in \mathbb{R} \times \mathbb{R}^n : t > 0\}$, $n \geq 1$. Thus, our work includes the case of special Lipschitz domains. Another difference to earlier work that we want to stress is that we consider general elliptic divergence form systems, and not only scalar equations, as this is the natural setting for the methods used. In this generality, interior pointwise regularity estimates fail in general, even in the uniform elliptic case. However, we emphasize that the methods used in this paper do not require such pointwise estimates.

We consider divergence form second order, real and complex, **degenerate** elliptic systems

$$\text{(1)} \qquad \sum_{i,j=0}^{n} \sum_{\beta=1}^{m} \partial_i \Big(A_{i,j}^{\alpha,\beta}(t,x) \partial_j u^\beta(t,x) \Big) = 0, \ \alpha = 1, \ldots, m$$

in $\mathbb{R}^{1+n}_+$, where $\partial_0 = \frac{\partial}{\partial t}$ and $\partial_i = \frac{\partial}{\partial x_i}$, $1 \leq i \leq n$, which we abbreviate as $\mathrm{div} A \nabla u = 0$, where

$$\text{(2)} \qquad A = (A_{i,j}^{\alpha,\beta}(t,x))_{i,j=0,\ldots,n}^{\alpha,\beta=1,\ldots,m}.$$

We assume A to be degenerate in the sense that for some $w \in A_2(\mathbb{R}^n)$ and $C < \infty$,

$$\text{(3)} \qquad |A(t,x)| \leq Cw(x), \text{ for a.e. } (t,x) \in \mathbb{R}^{1+n}_+$$

and elliptic degenerate in the sense that $w^{-1}A$ is accretive on a space $\mathcal{H}^0$ that we define below. This ellipticity condition means that there

exists $\kappa > 0$ such that

$$(4) \qquad \operatorname{Re} \int_{\mathbb{R}^n} (Af(x), f(x)) dx \geq \kappa \sum_{i=0}^{n} \sum_{\alpha=1}^{m} \int_{\mathbb{R}^n} |f_i^\alpha(x)|^2 w(x) \, dx,$$

for all $f \in \mathcal{H}^0$ and a.e. $t > 0$. We have set

$$(A\xi, \xi) = \sum_{i,j=0}^{n} \sum_{\alpha,\beta=1}^{m} A_{i,j}^{\alpha,\beta}(t,x) \xi_j^\beta \, \overline{\xi_i^\alpha}.$$

The space $\mathcal{H}^0$ is the closed subspace of $L^2(\mathbb{R}^n, w; \mathbb{C}^{m(1+n)})$ consisting of those functions with $\operatorname{curl}_x(f_i^\alpha)_{i=1,\dots,n} = 0$ for all α. The case of equations is when $m = 1$ or, equivalently, when $A_{i,j}^{\alpha,\beta} = A_{i,j}\delta_{\alpha,\beta}$. In this case, the accretivity condition becomes the usual pointwise accretivity

$$(5) \qquad \operatorname{Re} \sum_{i,j=0}^{n} A_{i,j}\xi_j\overline{\xi_i} \geq \kappa \sum_{i=0}^{n} |\xi_i|^2 w(x),$$

for all $\xi \in \mathbb{C}^{1+n}$ and a.e. $(t,x) \in \mathbb{R}_+^{1+n}$. Observe that the function $(t,x) \mapsto w(x)$ is an A_2 weight in $\mathbb{R}^{1+n}$ if w is an A_2 weight in $\mathbb{R}^n$. So, the degeneracy is a special case of that considered in the works mentioned above. However, for the boundary value problems we wish to consider, this seems a natural class. To our knowledge, this has not been considered before.

A natural question is whether weights could depend on both variables or only on the t-variable. Already in the non-degenerate case there are regularity conditions without which the Dirichlet problem is ill-posed. As for the degenerate case, the well known example from [14] when $A = t^{1-2s}I$, $0 < s < 1$, leading to representation of the fractional Laplacian $(-\Delta)^s$ as the Dirichlet to Neumann operator, is of a nature that our theory cannot cover. In fact, our weights need to have a trace at the boundary in some sense. Another option is to assume the weight depends on the x-variable only and put the t-perturbations in the coefficients. See below in this introduction and Section 4.

The equation (1) must be properly interpreted. Solutions are taken with u and ∇u locally in $L^2(\mathbb{R}_+^{1+n}, w(x)dxdt)$, and the equation is taken in the sense of distributions. Note that we allow complex coefficients and systems, so most of the theory developed for real symmetric equations does not apply, and even for real coefficients we want to develop methods regardless of regularity theory.

The boundary value problems can be formulated as follows: find weak solutions u with appropriate interior estimates of $\nabla_{t,x}u$ satisfying one of the following three natural boundary conditions.

- The Dirichlet condition $u = \varphi$ on $\mathbb{R}^n$, given the Dirichlet datum $\varphi \in L^2(\mathbb{R}^n, w; \mathbb{C}^m)$.

- The Dirichlet regularity condition $\nabla_x u = \varphi$ on $\mathbb{R}^n$, given the regularity datum $\varphi \in L^2(\mathbb{R}^n, w; \mathbb{C}^{mn})$ satisfying $\mathrm{curl}_x \varphi = 0$. Alternately, φ is the tangential gradient of the Dirichlet datum.
- The Neumann condition $\partial_{\nu_A} u = (A\nabla_{t,x} u, e_0) = \varphi$ on $\mathbb{R}^n$, given $\varphi \in L^2(\mathbb{R}^n, w^{-1}; \mathbb{C}^m)$. Here e_0 is the upward unit vector in the t-direction.

Observe that the natural space for the gradient at the boundary is $L^2(\mathbb{R}^n, w; \mathbb{C}^{m(1+n)})$, and since A is of the size w, multiplication by A maps into $L^2(w^{-1})$. Thus the weight w^{-1} is natural for the conormal derivative. In order to work with the same weighted space for all three problems, we shall consider the w-normalized conormal derivative $\partial_{\nu_{w^{-1}A}} u|_{t=0} = w^{-1} \partial_{\nu_A} u|_{t=0} \in L^2(\mathbb{R}^n, w; \mathbb{C}^m)$.

Boundary value problems can be formulated for L^p data with $p \neq 2$. This is for later work and here we restrict our attention to $p = 2$. We mention that there are also "intermediate" boundary value problems for regularity/Neumann data in some negative Sobolev spaces based on $L^2(w)$ using fractional powers of the Laplacian $-\Delta_w$ (defined later) that can be treated by the same methods. One important case is the treatment of variational solutions in this context, *i.e.* those with $\iint_{\mathbb{R}^{1+n}_+} |\nabla u|^2 \, dw(x)dt < \infty$, in which these problems are always well-posed. In the case $w = 1$, the methods have been worked out completely in [34].

Let us also comment on the corresponding degenerate inhomogeneous problem $\mathrm{div} A \nabla u = f$, with $\mathrm{div} A \nabla u$ denoting the left hand side in (1). A study of such equations would be of interest in its own right, but could also prove useful in the study of boundary value problems of the type we study here. Such applications were implemented in [17, 16] in the non-degenerate setting where the coefficients of the operator were assumed to satisfy a Carleson measure condition in place of t-independence. In [17] a duality argument reduced the desired estimate for solutions to the homogeneous equation to an estimate on a solution to an inhomogeneous equation, which could subsequently be proved. While such investigations in the degenerate setting would certainly make the theory more complete, studying the inhomogeneous equation is not our goal here.

We shall obtain *a priori* representations of solutions and existence of boundary traces (in the almost everywhere sense) for appropriate solution spaces together with various estimates involving non-tangential maximal functions, a characterization of well-posedness, a duality principle between regularity and Dirichlet problems, perturbation results for well-posedness on perturbing the coefficients, and well-posedness for hermitian systems and some block-triangular coefficients.

To this end, we follow the two step strategy developed in [3] and [4] and we assume the reader has these references handy. The first step is to obtain a priori representations and boundary traces (in weighted L^2

and almost everywhere) for solutions in appropriate classes. Actually, we shall even improve upon known results on almost everywhere convergence when $w = 1$. The starting point is to transform the second order equation to a first order system in the w-normalized conormal gradient

$$\nabla_{w^{-1}A} u = \begin{bmatrix} \partial_{\nu_{w^{-1}A}} u \\ \nabla_x u \end{bmatrix},$$

and then prove a weighted quadratic estimate and a non-tangential maximal estimate for functions of a bisectorial operator DB when the coefficients are t-independent. Then one can develop a semigroup representation of those w-normalized conormal gradients. This semigroup method for t-independent A and their t-dependent perturbations is presented in such an abstract way in [3] that it applies in the weighted space without any change once the initial quadratic estimate is established. We shall basically define the necessary objects and prove only what is not *mutatis mutandi* the same. Here, the discrepancy function should be renormalized by the weight, so it reads $w^{-1}(x)(A(t,x) - A(0,x))$, and it is measured by a weighted Carleson-Dahlberg condition. Also some estimates such as weighted Carleson embedding require the fact that dw is doubling, which follows from the A_2 condition.

The quadratic estimate for bisectorial DB is also a fundamental ingredient in the proof of non-tangential maximal estimates and almost everywhere convergence. As in [4], the almost everywhere convergence is not in a pointwise sense but in some averaged sense on Whitney regions approaching the boundary. We obtain convergence results for solutions in spaces relative to the Dirichlet problem, and for solutions, their gradients and w-normalized conormal derivatives for solutions in spaces relative to the regularity and Neumann problems. We stress that even when $w = 1$, this convergence at the level of gradients is new in this generality.

Once representations and boundary traces of solutions are obtained, the second step is to formulate the boundary value problems and show that solvability amounts to inverting boundary operators. As for positive results, we shall obtain well-posedness results for t-independent A which, in addition, are hermitian, block-diagonal, block-triangular or merely satisfy a divergence free condition on the triangular block. Let us remark that in the non-degenerate case $w = 1$, all these BVPs are well posed with constant matrices. Here, constant would mean that A is a constant multiple of w, or that DB comes with constant B. However, we do not know well-posedness for such B, in particular for the example below with $\varepsilon \sim 2$. As usual, all positive well-posedness results are stable under perturbation, that is, when $w^{-1}(A' - A)$ is small in L^∞. We also obtain perturbation results for well-posedness with

t-dependent coefficients A under smallness of the discrepancy function $w^{-1}(A(t,x) - A(0,x))$ in a weighted Carleson-Dahlberg sense.

The key result is therefore the quadratic estimates for bisectorial operators DB in $L_2(\mathbb{R}^n, w)$, stated in Theorem 3.3. The differential operator D, although not explicit in our notation, depends on the weight w as well. The structure of the operator dictates a different reduction to a Carleson measure estimate than in the uniformly elliptic case. In particular there appears a dyadic averaging/expectation operator E_t which acts on vector valued functions by unweighted averages of some components and weighted averages of some other components. It seems therefore from the proof that the A_2 condition on the weight is sharp for this quadratic estimate, since the $L_2(\mathbb{R}^n, w)$ bound of E_t requires the A_2 condition. In particular, our method does not to apply to the quasiconformal Jacobians mentioned above.

The quadratic estimate from Theorem 3.3 is the central estimate without which the method breaks down. We summarize the harvest of results using in an essential way this quadratic estimate:

1. The estimates of the functional calculus in Proposition 3.19;

2. The non-tangential maximal estimates in L^2 and almost everywhere convergence in Theorem 4.9;

3. The estimates, through operational calculus, of the singular integral with operator-valued kernel $S_{\mathcal{E}}$ in Theorems 4.11 and 4.12;

4. The perturbation results for boundary value problems in Theorems 5.4 and 5.5.

The proof of the quadratic estimate is not an easy matter and will be the hard part of the article. Let us comment on this. First, it does not follow from available extensions of the proof of the Kato square root problem such as the ones in [10], [6] or [12]. This is because D will no longer be a constant coefficient operator. Let us give an example to illustrate the differences. When $w = 1$, the basic example is

$$D = \begin{bmatrix} 0 & -\frac{d}{dx} \\ \frac{d}{dx} & 0 \end{bmatrix}.$$

On $\mathbb{R}$, if B is any constant elliptic matrix, the domain of DB is the Sobolev space $H^1(\mathbb{R}; \mathbb{C}^2)$. This can be seen using a Fourier transform argument. If $w \neq 1$, then D becomes

$$D = \begin{bmatrix} 0 & -\frac{1}{w}\frac{d}{dx}w \\ \frac{d}{dx} & 0 \end{bmatrix}$$

and (the column vector) (f, g) is in the domain of D if and only if $f \in H^1(\mathbb{R}, w)$ and $wg \in H^1(\mathbb{R}, w^{-1})$ (the weighted Sobolev spaces). Note that elements in these spaces are absolutely continuous functions when $w \in A_2$. Let B be the constant matrix

$$B = \begin{bmatrix} 1 & 0 \\ \varepsilon & 1 \end{bmatrix}.$$

Then for any $\varepsilon \neq 0$, the domain of DB cannot be the same as the domain of D. Indeed, the conditions to be in the domain of DB are $w(g + \varepsilon f) \in H^1(\mathbb{R}, w^{-1})$ and $f \in H^1(\mathbb{R}, w)$. If wg were in $H^1(\mathbb{R}, w^{-1})$, then so would wf, and both f and wf would be continuous on $\mathbb{R}$, which is a restriction on f.

Moreover, D cannot be coercive on its range. In the previous example, that would lead to an estimate like g' controlled by $\frac{1}{w}(gw)'$, which is absurd.

Nevertheless, Cruz-Uribe and Rios [15] recently proved the Kato conjecture for square roots of degenerate elliptic operators $-\text{div}_{\mathbb{R}^n} A \nabla_{\mathbb{R}^n}$ on $\mathbb{R}^n$, which corresponds here to the special case of block diagonal, bounded, measurable and accretive $B = \begin{bmatrix} 1 & 0 \\ 0 & w^{-1}A \end{bmatrix}$. The preceding example indicates that non-block terms introduce a difficulty that was not in the above work; indeed, it makes the argument harder. Let us explain where this difficulty appears in the proof and how we overcome it. As is customary in this area, we look for a Tb argument to prove a Carleson estimate which proceeds via a stopping time argument on the averages on dyadic cubes of some test functions. In our situation, averages are taken with respect to both the unweighted and weighted measure dx and $w(x)dx$. But if the weight varies too much on a given cube Q, the stopping time does not give useful information. We are forced to organize the collection of all subcubes of Q in subclasses on which the dx-averages of w do not vary too much from a parent cube. The ideal situation is that all cubes fall into one of these subclasses. This is not the case, but the next best thing happens, namely that the those subcubes which are left out satisfy a packing condition with uniform, possibly large, constant. This organization of cubes is intrinsic to the given weight and is in fact a weighted version of a result found in Garnett's book [24]. Here, it allows us to run the stopping time on the test functions for the Carleson estimate. In the situation of [15], one can split things so as to run two separate stopping time arguments each concerned with one measure so this step is not necessary.

To conclude this introduction, we discuss two different boundary geometries. First, all of the results here are invariant under bilipschitz changes of variables. Thus, whenever one can pull back a boundary value problem on a special Lipschitz domain to one that fits our hypotheses, one obtains a result for the initial problem. As an example, this shows that if $w \in A_2(\mathbb{R}^n)$, the three boundary value problems for $-\text{div}_{x,t}(w(x)\nabla_{x,t}u)(x,t) = 0$ with appropriate interior estimates (see Section 5) are well-posed on any special Lipschitz domain $\Omega = \{t > \varphi(x)\}$ when the corresponding datum is in $L^2(\partial\Omega; \tilde{w}d\sigma)$ where σ is the surface measure and $\tilde{w}(\varphi(x)) := w(x)$. Applying the standard pullback $(x,t) \mapsto (y, t - \varphi(y)) = (y, s)$, we obtain an equation of the form $\text{div}_{y,s}(A(y)\nabla_{y,s}v)(y,s) = 0$ that is degenerate elliptic.

Secondly, it is natural to expect results on bounded domains. Bilipschitz invariance implies that one can look at the case of the unit ball as the non-smoothness is carried by the coefficients. There, the setup of [4] applies to radially independent weights and degenerate coefficients, and perturbations of the latter. It will be clear from the present article that it all depends on the weighted quadratic estimate on the boundary. This requires a proof that is left to further work.

2. Preliminaries on weights

2.1. **Muckenhoupt weights.** Recall that for a weight w on $\mathbb{R}^n$ and $p > 1$, the A_p condition reads

$$\left(\fint_Q w \, dx \right) \left(\fint_Q w^{1-p'} \, dx \right)^{p/p'} \leq C,$$

for all cubes Q with p' the conjugate exponent to p. The smallest possible C is denoted by $[w]_{A_p}$. The notation $\fint_E$ means the average with respect to the indicated measure on E.

We identify w with the measure $dw = w(x) \, dx$ and write $w(E)$ for $\int_E dw$ while $|E| = \int_E dx$. Recall that $w \in A_p$ implies $w \in A_q$ for all $q > p - \varepsilon$ where $\varepsilon > 0$ depends on $[w]_{A_p}$. Every $w \in A_p$ is an A_∞ weight: there exist constants $0 < \sigma \leq 1 \leq \tau < \infty$ such that

$$(6) \qquad \left(\frac{|E|}{|Q|} \right)^\tau \lesssim \frac{w(E)}{w(Q)} \lesssim \left(\frac{|E|}{|Q|} \right)^\sigma$$

for all cubes Q and measurable subsets E of Q (actually, $\tau = p$ if $w \in A_p$ and $(1/\sigma)'$ is the reverse Hölder exponent of w, which can be arbitrary in $(1, \infty]$). In particular, dw is a doubling measure. There also exists a constant $c_0 > 0$ such that

$$(7) \qquad \fint_Q \ln w(x) dx \leq \ln \left(\fint_Q w(x) dx \right) \leq \fint_Q \ln w(x) dx + c_0,$$

the first inequality being Jensen's inequality and the second being a reverse form of it.

Denote $L^p(w; \mathbb{C}^d) = L^p(\mathbb{R}^n, w; \mathbb{C}^d)$ for $d \geq 1$, and $L^p(w) = L^p(w; \mathbb{C})$. If $w \in A_p$,

$$(8) \qquad \fint_Q |f(y)| \, dy \leq [w]_{A_p}^{1/p} \left(\fint_Q |f(y)|^p \, dw(y) \right)^{1/p}.$$

In particular, for $w \in A_2$, since $w \in A_{2/p}$ for some $p > 1$, this implies Muckenhoupt's theorem: the Hardy-Littlewood maximal operator M with respect to dx is bounded on $L^2(w)$.

2.2. A corona decomposition for A_2 weights.

We use the following dyadic decomposition of $\mathbb{R}^n$. Let $\triangle = \bigcup_{j=-\infty}^{\infty} \triangle_{2^j}$ where $\triangle_{2^j} := \{2^j(k + (0,1]^n) : k \in \mathbb{Z}^n\}$. For a dyadic cube $Q \in \triangle_{2^j}$, denote by $\ell(Q) = 2^j$ its *sidelength* and by $|Q| = 2^{nj}$ its *Lebesgue volume*. We set $\triangle_t = \triangle_{2^j}$ if $2^{j-1} < t \le 2^j$.

Here we describe a decomposition of a dyadic cube $Q \in \triangle$ with respect to a Muckenhoupt weight $w \in A_2$.

We follow a construction of Garnett [24]. For a fixed $\sigma_w > 0$, to be chosen later, we consider $B^w(Q)$ the collection of those ("bad for $\ln w$") maximal sub-cubes of Q for which

$$|(\ln w)_R - (\ln w)_Q| > \sigma_w.$$

In this section, we use the notation $f_Q = (f)_Q := \fint_Q f \, dx$. We can then define $B_j^w(Q)$ inductively as $B_1^w(Q) = B^w(Q)$ and for $j = 2, 3, \ldots$ by

$$B_j^w(Q) = \bigcup_{R \in B_{j-1}^w(Q)} B^w(R), \quad \text{and set} \quad B_*^w(Q) = \bigcup_{j=1}^{\infty} B_j^w(R).$$

The following proposition shows that the "number" of cubes on which the oscillation of $\ln w$ differs too much from the one on some ancestor can be controlled.

Proposition 2.1. *For a cube $R \in \triangle$, if $R \subseteq Q$ is not contained in a cube of $B^w(Q)$ then $|(\ln w)_R - (\ln w)_Q| \le \sigma_w$. We also have that*

$$(9) \qquad \sum_{R \in B_*^w(Q)} w(R) \le \frac{C}{\sigma_w^2} w(Q)$$

for some $C < \infty$ which depends only on $[w]_{A_2}$.

Proof. The first statement of the Proposition is clear from the definition of $B^w(Q)$, so we only need to prove (9).

For each $R \in B_*^w(Q)$, there is a unique $R' \in B_*^w(Q) \cup \{Q\}$ such that $R \in B_1^w(R')$, that is, R' is the stopping-time parent of R. Thus

$$w(R) \le \int_R \frac{|(\ln w)_R - (\ln w)_{R'}|^2}{\sigma_w^2} \, dw$$

and so

$$\sum_{R \in B_*^w(Q)} w(R) \lesssim \frac{1}{\sigma_w^2} \sum_{R \in B_*^w(Q)} \int_R |(\ln w)_R - (\ln w)_{R'}|^2 \, dw$$

$$= \frac{1}{\sigma_w^2} \int_Q \left(\sum_{R \in B_*^w(Q)} 1_R |(\ln w)_R - (\ln w)_{R'}|^2 \right) dw,$$

where 1_R is the characteristic function of the set R. For any $b \in L^1_{loc}(\mathbb{R}^n, dx)$ consider the square function

$$S(b)(x) := \left(\sum_{R \in B^w_*(Q)} 1_R(x) |b_R - b_{R'}|^2 \right)^{1/2}.$$

For any $w \in A_\infty$ we have (see Lemma 6.4 in [24] for the unweighted estimate and [30] for the weighted one) the estimate

$$\int_{\mathbb{R}^n} |S(b)|^2 dw \lesssim \int_{\mathbb{R}^n} |M(b)|^2 dw,$$

where M is the Hardy-Littlewood maximal operator with respect to dx. Applying this to $b(x) := (\ln w(x) - (\ln w)_Q) 1_Q(x)$ and using Muckenhoupt's theorem, it suffices to show that

(10) $$\int_Q |\ln w - (\ln w)_Q|^2 \, dw \lesssim w(Q).$$

This inequality follows from the fact that $\ln w \in BMO(\mathbb{R}^n, dx)$ when $w \in A_2$ and the fact that $BMO(\mathbb{R}^n, dx) = BMO(\mathbb{R}^n, dw)$ with equivalent norms for $w \in A_\infty$. The latter statement is a consequence of the John-Nirenberg inequality and reverse Hölder inequality for w. However, we present a direct proof. It follows from Jensen's inequality and the A_2 condition on w that

$$\int_Q e^{|\ln w(x) - (\ln w)_Q|} \, dx \lesssim |Q|.$$

Setting $M_\lambda = \{x \in Q \,;\, |\ln w(x) - (\ln w)_Q| > \lambda\}$, we can write

$$\int_0^\infty e^\lambda |M_\lambda| d\lambda = \int_Q \int_0^{|\ln w(x) - (\ln w)_Q|} e^\lambda d\lambda dx \le \int_Q e^{|\ln w(x) - (\ln w)_Q|} \, dx,$$

so

$$\int_0^\infty e^\lambda |M_\lambda| d\lambda \lesssim |Q|.$$

Similarly,

$$\int_Q |\ln w - (\ln w)_Q|^2 \, dw = \int_0^\infty 2\lambda w(M_\lambda) d\lambda.$$

Condition (6), in which we take $\sigma < 1$ as one can always do, ensures that

$$\int_0^\infty \lambda w(M_\lambda) d\lambda \lesssim \int_0^\infty \lambda w(Q) \left(\frac{|M_\lambda|}{|Q|} \right)^\sigma d\lambda$$

$$= w(Q) \int_0^\infty \left(\frac{|M_\lambda|}{|Q|} e^\lambda \right)^\sigma (\lambda e^{-\lambda\sigma}) d\lambda$$

$$\lesssim w(Q) \left(\int_0^\infty \frac{|M_\lambda|}{|Q|} e^\lambda d\lambda \right)^\sigma \left(\int_0^\infty (\lambda e^{-\lambda\sigma})^{1/(1-\sigma)} d\lambda \right)^{1-\sigma}$$

$$\lesssim w(Q),$$

which proves (10) and with it the proof of the Proposition. □

2.3. **Review of weighted Littlewood-Paley inequalities.** We recall here a few facts of Littlewood-Paley theory. The treatment in Wilson's book used in [15] is not completely adapted for our needs. In particular, we use Calderón's reproducing formula in a different way. Here w denotes a weight in A_2. We begin with approximation issues.

Lemma 2.2. *Let $\varphi \in L^1(\mathbb{R}^n)$ with a radially decreasing integrable majorant ϕ. Let $\varphi_\varepsilon(x) = \varepsilon^{-n} \varphi(x/\varepsilon)$ for $\varepsilon > 0$.*

- *Convolution with φ_ε is a bounded operator on $L^2(w)$, uniformly with respect to ε.*
- *For every $f \in L^2(w)$, $\varphi_\varepsilon \star f \to cf$ and $\varphi_{1/\varepsilon} \star f \to 0$ in $L^2(w)$ as $\varepsilon \to 0$, where $c = \int_{\mathbb{R}^n} \varphi(x) \, dx$.*

Proof. For the first point, we observe that $\sup_{\varepsilon > 0} |\varphi_\varepsilon \star f| \leq \|\phi\|_1 Mf$ almost everywhere (See [25, Corollary 2.1.12]) and recall that M is bounded on $L^2(w)$. For the second point, it is easy to see that $L_c^\infty(\mathbb{R}^n)$, the set of bounded functions with bounded support, is dense in $L^2(w)$. Thus, it is enough to assume $f \in L_c^\infty(\mathbb{R}^n)$. Then $\varphi_\varepsilon \star f - cf$ and $\varphi_{1/\varepsilon} \star f$ converge almost everywhere (for dx, thus for dw) to 0 as $\varepsilon \to 0$ (See [25, Corollary 2.1.19]). The conclusion follows using the dominated convergence theorem in $L^2(w)$ as $Mf \in L^2(w)$. □

Corollary 2.3. *$C_0^\infty(\mathbb{R}^n)$ is dense in $L^2(w)$. Also the space $\mathcal{E}$ of functions of the form $\int_\varepsilon^R \psi_t \star f \frac{dt}{t}$ where $f \in C_0^\infty(\mathbb{R}^n)$, $\psi_t(x) = t^{-n} \psi(x/t)$ with $\psi \in S(\mathbb{R}^n)$ with Fourier transform supported away from 0 and ∞, $\varepsilon > 0, R < \infty$ and $\int_0^\infty \hat{\psi}(t\xi) \frac{dt}{t} = 1$ for all $\xi \neq 0$, form a dense subspace of $L^2(w)$. [Here, $\hat{g}$ is the Fourier transform of g in $\mathbb{R}^n$.]*

Proof. As $L_c^\infty(\mathbb{R}^n)$ is dense in $L^2(w)$, choosing $\varphi \in C_0^\infty(\mathbb{R}^n)$ in the previous lemma proves the first density. Next, with ψ as in the statement, it is easy to see that $\hat{\varphi}(\xi) := \int_1^\infty \hat{\psi}(t\xi) \frac{dt}{t} = 1 - \int_0^1 \hat{\psi}(t\xi) \frac{dt}{t}$ for $\xi \neq 0$ and $\hat{\varphi}(0) = 1$ is a Schwartz class function and that $\int_\varepsilon^R \psi_t \star f \frac{dt}{t} = \varphi_\varepsilon \star f - \varphi_R \star f$. So the lemma follows. □

Corollary 2.4. *Assume ψ and φ are integrable functions, that φ is as in Lemma 2.2 with $\int \varphi = 1$ and that $\int_\varepsilon^R \hat{\psi}(t\xi) \frac{dt}{t} = \hat{\varphi}(\varepsilon\xi) - \hat{\varphi}(R\xi)$ for all $\xi \in \mathbb{R}^n, 0 < \varepsilon < R$. Set $Q_t f = \psi_t \star f$. Then the Calderón representation formula $\int_0^\infty Q_t f \frac{dt}{t} = f$ holds for all $f \in L^2(w)$ in the sense that $\int_\varepsilon^R Q_t f \frac{dt}{t}$ converges to f in $L^2(w)$ as $\varepsilon \to 0$ and $R \to \infty$.*

Proof. The lemma follows immediately from the formula $\int_\varepsilon^R \psi_t \star f \frac{dt}{t} = \varphi_\varepsilon \star f - \varphi_R \star f$ and Lemma 2.2. □

Next, we recall that any Calderón-Zygmund operator in the sense of Meyer's book [33] is bounded on $L^2(w)$ (the sharp dependance of the norm with respect to $[w]_{A_2}$ has been obtained by T. Hytönen [28]

but we do not need such a refined estimate). This result extends to vector-valued Calderón-Zygmund operators, and consequently we have the following useful corollary.

Lemma 2.5. *Let ψ be an integrable function in $\mathbb{R}^n$ with $\int_{\mathbb{R}^n} \psi(x)\,dx = 0$ and such that $(\psi_t(x-y))_{t>0}$ is an $L^2(\mathbb{R}_+, \frac{dt}{t})$-valued Calderón-Zygmund kernel and set $Q_t f = \psi_t \star f$. Then for all $f \in L^2(w)$,*

$$\int_0^\infty \|Q_t f\|_{L^2(w)}^2 \frac{dt}{t} \le C_{w,\psi} \|f\|_{L^2(w)}^2.$$

Proof. Note that $T : f \mapsto (Q_t f)_{t>0}$ is a Calderón-Zygmund operator, bounded from $L^2(\mathbb{R}^n)$ to $L^2(\mathbb{R}^n, L^2(\mathbb{R}_+, \frac{dt}{t}))$. Thus weighted L^2 theory for A_2 weights applies (see [23], Chapter V). $\square$

To conclude this section, we recall a trick proved by Cruz-Uribe and Rios [15] originating from the proof of [18, Corollary 4.2].

Lemma 2.6. *If T is a linear operator that is bounded on $L^2(w)$ for any $w \in A_2$ with norm depending only on $[w]_{A_2}$, then for any fixed $w \in A_2$, there exists $\theta \in (0,1)$ and $C > 0$ such that $\|T\|_{\mathcal{L}(L^2(w))} \le C\|T\|_{\mathcal{L}(L^2(dx))}^\theta$.*

2.4. Weighted Sobolev spaces and the gradient operator.

For $1 < p < \infty$, $W^{1,p}(w)$ is the closure of $C_0^\infty(\mathbb{R}^n)$ in the norm

$$\left(\int_{\mathbb{R}^n} \left(|f|^p + |\nabla f|^p \right) dw \right)^{1/p}.$$

For $p = 2$, we set $H^1(w) = W^{1,2}(w)$.

It follows from [21, Theorem 1.5] that $w \in A_p$ satisfies a Poincaré inequality:

$$(11) \qquad \left(\fint_Q |\psi - c|^p \, dw \right)^{1/p} \lesssim \ell(Q) \left(\fint_Q |\nabla \psi|^p \, dw \right)^{1/p}$$

where $\psi \in C_0^\infty(\mathbb{R}^n)$, c is the mean of ψ with respect to either dw or dx and the implicit constant depends only on $[w]_{A_p}$, n and p. In particular, this implies that $W^{1,p}(w)$ can be identified with a space of measurable functions imbedded in $L^p(w)$. The gradient (taken in the sense of distributions) extends to an operator on $W^{1,p}(w)$, still denoted by ∇, and the Poincaré inequality extends to all $W^{1,p}(w)$ functions. The space of Lemma 2.3 is also dense in $W^{1,2}(w)$ when $w \in A_2$.

We shall need the following result.

Proposition 2.7. *If $w \in A_2$, the spaces $W^{1,p}(w)$ interpolate by the real or complex methods for $r(w) < p < \infty$ where $r(w) < 2$ is the infimum of exponents p for which $w \in A_p$.*

Proof. This follows from N. Badr's interpolation theorem [11, Theorem 8.4], using the Poincaré inequality and doubling of dw. $\square$

We denote the Riesz transforms by R_j, $j = 1, \ldots, n$. Set $\mathcal{R}f = (R_1 f, \ldots, R_n f)$ and $\mathcal{R}^* g = R_j^* g_1 + \cdots + R_n^* g_n$ where $R_j^* = -R_j$, the adjoint being for the $L^2(dx)$ duality. The Riesz transforms are bounded on $L^2(w)$.

In this paper, if A is an unbounded linear operator on a Banach space, then $\mathsf{D}(A), \mathsf{N}(A), \mathsf{R}(A)$ denote respectively, its domain, null space and range. Closure of the range is denoted by $\overline{\mathsf{R}(A)}$.

Lemma 2.8. *Let $w \in A_2$. The following statements hold.*
  (1) *The operator ∇ on $L^2(w)$ into $L^2(w; \mathbb{C}^n)$ with domain $\mathsf{D}(\nabla) = H^1(w)$ is densely defined, closed and injective.*
  (2) *Its adjoint $\nabla^* := -\mathrm{div}_w := -\frac{1}{w}\mathrm{div}w$ is also closed, densely defined and $\overline{\mathsf{R}(\nabla^*)} = L^2(w)$. Moreover, $\frac{1}{w}C_0^\infty(\mathbb{R}^n; \mathbb{C}^n)$ is a dense subspace of $\mathsf{D}(\mathrm{div}_w)$.*
  (3) *The operator $\Delta_w = -\nabla^*\nabla = \mathrm{div}_w\nabla$ constructed by the method of sesquilinear forms from $\int_{\mathbb{R}^n} \nabla u \cdot \overline{\nabla v}\, dw$ on $H^1(w)$ is a negative self-adjoint, injective operator on $L^2(w)$.*
  (4) *The operator $\mathcal{R}_w := \nabla(-\Delta_w)^{-1/2}$ is bounded from $L^2(w)$ into $L^2(w; \mathbb{C}^n)$ and its range is $\overline{\mathsf{R}(\nabla)}$.*
  (5) $\overline{\mathsf{R}(\nabla)} = \overline{\mathcal{R}(L^2(w))}$.
  (6) *Moreover $\overline{\mathsf{R}(\nabla)} = \overline{\nabla(C_0^\infty(\mathbb{R}^n))} = \{g \in L^2(w; \mathbb{C}^n)\,;\, \mathrm{curl}\,g = 0\}$.*

As we can see, the closure of the range of ∇ has several characterizations: the one in (4) comes from spectral theory for the weighted Laplacian, the one in (5) comes from weighted theory for the usual Laplacian. Both will be useful.

Proof. To prove (1), observe that ∇ is densely defined by construction and the fact that it is closed is proved in [21]. Injectivity follows from Poincaré inequalities: indeed if $u \in H^1(w)$ such that $\nabla u = 0$ implies u constant. As w is doubling and $\mathbb{R}^n$ is unbounded, we have that $w(\mathbb{R}^n) = +\infty$, so that the constant must be 0.

The first part of (2) follows upon standard functional analysis from (1). To show the density of $\frac{1}{w}C_0^\infty(\mathbb{R}^n; \mathbb{C}^n)$ in $\mathsf{D}(\mathrm{div}_w)$, we note that this is equivalent to the density of $C_0^\infty(\mathbb{R}^n; \mathbb{C}^n)$ in $\mathsf{D}(\mathrm{div}) := \{f \in L^2(w^{-1}; \mathbb{C}^n)\,;\, \mathrm{div}f \in L^2(w^{-1})\}$ (equipped with the graph norm). This equivalent statement follows using regularisation by convolution as in Lemma 2.2 and smooth truncations.

The construction in (3) is classical (see for example, Kato's book [31]). The injectivity follows from that of the gradient. Next, the boundedness of $\mathcal{R}_w$ in (4) follows from the equality $\|\nabla f\| = \|(-\Delta_w)^{1/2}f\|$, which itself is obtained using self-adjointness and the form. The identification of its range follows immediately.

Let us turn to (5). As the Riesz transforms are bounded on $L^2(w)$ when $w \in A_2$ and $\mathcal{R}^*\mathcal{R} = I$, $\mathcal{R}(L^2(w))$ is a closed subspace of $L^2(w; \mathbb{C}^n)$. For f in the dense subspace $\mathcal{E}$ of Corollary 2.3, it is easy to see that

$\mathcal{R}f = \nabla g$ with $g = (-\Delta)^{-1/2}f \in \mathcal{S}(\mathbb{R}^n)$ using the properties of the function ψ. Thus, $\mathcal{R}(\mathcal{E}) \subseteq \mathsf{R}(\nabla)$ so that $\mathcal{R}(L^2(w)) \subseteq \overline{\mathsf{R}(\nabla)}$. Conversely, if g belongs to this dense subspace $\mathcal{E}$ then $\nabla g = \mathcal{R}f$ for $f = (-\Delta)^{1/2}g \in \mathcal{S}(\mathbb{R}^n)$. Thus $\nabla(\mathcal{E}) \subseteq \mathcal{R}(L^2(w))$ and, by density, $\overline{\mathsf{R}(\nabla)} \subseteq \mathcal{R}(L^2(w))$.

We now prove (6). We introduce $\dot{H}^1(w)$ as the closure of $C_0^\infty(\mathbb{R}^n)$ for the semi norm $\|\nabla f\|_{L^2(w;\mathbb{C}^n)}$. We claim that this space coincides with the space of $f \in L^2_{\mathrm{loc}}(w)$ such that $\nabla f \in L^2(w;\mathbb{C}^n)$. In one direction, if one has a Cauchy sequence ∇g_k in $L^2(w)$ with $g_k \in C_0^\infty(\mathbb{R}^n)$, then using the Poincaré inequality, if c_k is the mean of g_k on the unit ball, one can see that $g_k - c_k$ converges in $L^2_{\mathrm{loc}}(w)$ to some g and thus ∇g_k converges to the distributional gradient of g in $L^2(w;\mathbb{C}^n)$. Conversely, if $g \in L^2_{loc}(w)$ and $\nabla g \in L^2(w;\mathbb{C}^n)$, then set $g_k = (g-c_k)\varphi_k$ where c_k is the mean of g on the ball $B(0,2^{k+1})$ and $\varphi_k(x) = \varphi(2^{-k}x)$ with φ is a smooth function that vanishes outside $B(0,2)$ and that is 1 on $B(0,1)$. It follows from Poincaré's inequality that ∇g_k converges weakly to ∇g in $L^2(w)$. By Mazur's theorem, this implies that a sequence of convex combinations converges strongly to ∇g. Noticing that $g_k \in H^1(w)$ in which $C_0^\infty(\mathbb{R}^n)$ is dense concludes the proof of the claim. Clearly, $\overline{\mathsf{R}(\nabla)} = \nabla(\dot{H}^1(w))$, so this proves the first equality.

Next, we have $\nabla(C_0^\infty(\mathbb{R}^n)) \subseteq \{g \in L^2(w;\mathbb{C}^n)\,;\, \mathrm{curl}g = 0\}$. This inclusion is preserved by taking the closure in $L^2(w;\mathbb{C}^n)$ as the second set is clearly closed. Conversely, using the boundedness of the Riesz transforms, the idendity $g = (I - \mathcal{R}\mathcal{R}^*)g + \mathcal{R}\mathcal{R}^*g$ holds for $g \in L^2(w;\mathbb{C}^n)$. Moreover $\mathrm{div}((I - \mathcal{R}\mathcal{R}^*)g) = 0$ and $\mathrm{curl}(\mathcal{R}\mathcal{R}^*g) = 0$ in the distributions sense. Indeed, it is certainly true for $g \in C_0^\infty(\mathbb{R}^n;\mathbb{C}^n)$ and one can use a limiting argument. If one assumes $\mathrm{curl}g = 0$ then one also has $\mathrm{curl}((I - \mathcal{R}\mathcal{R}^*)g) = 0$. Hence $(I - \mathcal{R}\mathcal{R}^*)g$ is constant as a distribution, and this equals to 0 since it is in $L^2(w;\mathbb{C}^n)$. Thus $g = \mathcal{R}\mathcal{R}^*g \in \mathcal{R}(L^2(w))$ and we conclude using (5). $\qquad\square$

3. The main quadratic estimate

3.1. **Review of the DB formalism.** Let $\mathcal{H}$ be a separable complex Hilbert space. One uses $(\,,\,)$ and $\|\,\|$ for the hermitian product and norm on $\mathcal{H}$. Let D be an unbounded self-adjoint operator on $\mathcal{H}$. There is an orthogonal splitting

(12) $$\mathcal{H} = \mathsf{N}(D) \oplus \overline{\mathsf{R}(D)}.$$

Define closed double sectors in the complex plane by

$$S_\mu := \{z \in \mathbb{C} : |\arg(\pm z)| \le \mu\} \cup \{0\}.$$

Consider $B \in \mathcal{L}(\mathcal{H})$ and assume that B is (strictly) accretive on $\overline{\mathsf{R}(D)}$, that is, there exists $\kappa > 0$ such that

(13) $$\mathrm{Re}(Bv, v) \ge \kappa\|v\|^2, \ \forall v \in \overline{\mathsf{R}(D)}.$$

In this case, let

$$\mu(B) := \sup_{v \in \mathsf{R}(D), v \neq 0} |\arg(Bv, v)| < \pi/2$$

denote the *angle of accretivity* of B on $\mathsf{R}(D)$. Note that B may not be invertible on $\mathcal{H}$. Still for X a subspace of $\mathcal{H}$, we set $B^{-1}X = \{u \in \mathcal{H} \, ; \, Bu \in X\}$.

Proposition 3.1. *With the above assumptions, we have the following facts.*

   (i) *The operator DB, with domain $B^{-1}\mathsf{D}(D)$, is $\mu(B)$-bisectorial, i.e. $\sigma(DB) \subseteq S_{\mu(B)}$ and there are resolvent bounds $\|(\lambda I - DB)^{-1}\| \lesssim 1/\mathrm{dist}\,(\lambda, S_\mu)$ when $\lambda \notin S_\mu$, $\mu(B) < \mu < \pi/2$.*
   (ii) *The operator DB has range $\mathsf{R}(DB) = \mathsf{R}(D)$ and null space $\mathsf{N}(DB) = B^{-1}\mathsf{N}(D)$ such that topologically (but not necessarily orthogonally) one has*

$$\mathcal{H} = \overline{\mathsf{R}(DB)} \oplus \mathsf{N}(DB).$$

   (iii) *The restriction of DB to $\overline{\mathsf{R}(D)}$ is a closed, injective operator with dense range in $\overline{\mathsf{R}(D)}$. Moreover, the same statements on spectrum and resolvents as in (i) hold.*
   (iv) *Statements similar to (i) and (ii) hold for BD with $\mathsf{D}(BD) = \mathsf{D}(D)$, defined as the adjoint of DB^* or equivalently by $BD = B(DB)B^{-1}$ on $\mathsf{R}(BD) \cap \mathsf{D}(D)$, with $\mathsf{R}(BD) := B\mathsf{R}(D)$ and $BD = 0$ on the null space $\mathsf{N}(BD) := \mathsf{N}(D)$.*

For a proof, see [1]. Note that the accretivity is only needed on $\mathsf{R}(D)$. For $t \in \mathbb{R}$, we shall

$$R_t^B = (I + itDB)^{-1},$$

$$Q_t^B = \frac{1}{2i}(R_{-t}^B - R_t^B) = tDB(I + t^2 DBDB)^{-1},$$

$$P_t^B = \frac{1}{2}(R_{-t}^B + R_t^B) = (I + t^2 DBDB)^{-1}.$$

It follows from the previous result that R_t^B, P_t^B and Q_t^B are uniformly bounded operators on $\mathcal{H}$.

3.2. **The main theorem.** Let us specify in this section the operator D we consider throughout. In $\mathbb{C}^{m(n+1)} = \mathbb{C}^m \oplus \mathbb{C}^{mn} = \mathbb{C}^m \oplus (\mathbb{C}^m \otimes \mathbb{C}^n)$, we use the notation $v = \begin{bmatrix} v_\perp \\ v_\parallel \end{bmatrix}$, with $v_\perp \in \mathbb{C}^m$ and $v_\parallel \in \mathbb{C}^{mn} = \mathbb{C}^m \otimes \mathbb{C}^n$ which we call respectively the normal (or scalar) and tangential parts of v. To simplify the exposition, we carry out the detailed proof only in the "scalar" case $m = 1$ as it carries in the vector-valued case $m > 1$ without change. See Section 3.9 for the notation.

Proposition 3.2. *Let $w \in A_2(\mathbb{R}^n)$ and set $\mathcal{H} = L^2(w; \mathbb{C}^{n+1})$ with norm $\|f\| = (\int_{\mathbb{R}^n} |f|^2 \, dw)^{1/2}$.*

(1) *The operator* $D := \begin{bmatrix} 0 & \mathrm{div}_w \\ -\nabla & 0 \end{bmatrix}$ *with domain* $\begin{bmatrix} \mathsf{D}(\nabla) \\ \mathsf{D}(\mathrm{div}_w) \end{bmatrix}$ *is self-adjoint.*

(2) $\begin{bmatrix} C_0^\infty(\mathbb{R}^n) \\ \frac{1}{w}C_0^\infty(\mathbb{R}^n;\mathbb{C}^n) \end{bmatrix}$ *is a dense subspace of* $\mathsf{D}(D)$.

(3) $\overline{R(D)} = \begin{bmatrix} L^2(w) \\ \mathcal{R}_w(L^2(w)) \end{bmatrix}$.

(4) $\overline{R(D)} = \begin{bmatrix} L^2(w) \\ \mathcal{R}(L^2(w)) \end{bmatrix} = \{g \in L^2(\mathbb{R}^n, w; \mathbb{C}^{1+n}); \mathrm{curl}_x(g_\parallel) = 0\}$,

where the closure is taken in $\mathcal{H}$ *and* $\overline{R(D)} \cap C_0^\infty(\mathbb{R}^n; \mathbb{C}^{n+1})$ *is dense in* $\overline{R(D)}$.

The proof is a direct consequence of Lemma 2.8. We omit the details.

Let us specify the required assumption on B: this is the operator of multiplication by an $(n+1) \times (n+1)$ matrix $B(x)$ which has bounded entries and is accretive on $\overline{R(D)}$. Associated to B are the constants $\|B\|_\infty$ and (the best) $\kappa > 0$ in (13).

In this section, $\|f\|$ systematically designates the weighted $L^2(w; \mathbb{C}^d)$ norm with $d = 1, n$ or $n + 1$ depending on the context.

Theorem 3.3. *With the preceding assumptions, one has*

$$(14) \qquad \int_0^\infty \|Q_t^B v\|^2 \frac{dt}{t} \lesssim \|v\|^2, \qquad \text{for all } v \in \mathcal{H}.$$

Theorem 3.3 differs from previous results in the field. First, D still being a first order differential operator, no longer has constant coefficients. Secondly, the coercivity assumption $\|\nabla u\| \lesssim \|Du\|$ for $u \in R(D) \cap D(D)$ which is of utmost importance in other proofs cannot be true here, thus we cannot apply the metric measure space (here $\mathbb{R}^n$ with Euclidean distance and measure dw) generalisation made by Bandara [12] of the Euclidean results in [10] or [6].

Note that the block-matrix case $B = \begin{bmatrix} 1 & 0 \\ 0 & d \end{bmatrix}$ in the splitting $\mathbb{C}^{n+1} = \mathbb{C} \oplus \mathbb{C}^n$ gives a proof of the Kato conjecture for degenerate operators, first proved by Cruz-Uribe and Rios [15]. See Section 3.10.

Our strategy to prove Theorem 3.3 is to follow the line of argument in [6] but with some necessary twists. In particular, we will use different spectral decompositions on the scalar and tangential parts. Also the stopping-time argument requires the use of the Corona decomposition for w (Proposition 2.1).

3.3. **Reduction to a Carleson measure estimate.**

Lemma 3.4 (Off-diagonal decay). *For every integer N there exists $C_N > 0$ such that*

$$(15) \qquad \|1_E R_t^B u\| + \|1_E Q_t^B u\| \le C_N \langle \mathrm{dist}\,(E, F)/|t| \rangle^{-N} \|u\|$$

for all $t \neq 0$, whenever $E, F \subseteq \mathbb{R}^n$ are closed sets, $u \in \mathcal{H}$ is such that $\operatorname{supp} u \subseteq F$. We have set $\langle x \rangle := 1 + |x|$ and $\operatorname{dist}(E, F) := \inf\{|x - y| ; x \in E, y \in F\}$.

Proof. Repeat the argument in [6, Proposition 5.1], which only uses that D is of order 1 and that the commutator $[\chi, D]$ of D with a Lipschitz function χ is bounded on the space $\mathcal{H}$. More precisely, it is the pointwise multiplication by

$$(16) \qquad \begin{bmatrix} 0 & -(\nabla \chi)^t \\ \nabla \chi & 0 \end{bmatrix}.$$

We omit further details. $\qquad \square$

Let A_t be the dyadic averaging operator with respect to Lebesgue measure and A_t^w the one with respect to dw. Set

$$E_t u(x) := \begin{bmatrix} A_t^w u_\perp(x) \\ A_t u_\parallel(x) \end{bmatrix} = \begin{bmatrix} \fint_Q u_\perp(y) \, dw(y) \\ \fint_Q u_\parallel(y) \, dy \end{bmatrix},$$

where Q is the unique dyadic cube $Q \in \triangle_t$ that contains x. The notation for dyadic cubes is the same as in Section 2.2. Recall that $\fint_Q$ means the average on Q with respect to the indicated measure. We also set

$$E_Q u = \begin{bmatrix} \fint_Q u_\perp(y) \, dw(y) \\ \fint_Q u_\parallel(y) \, dy \end{bmatrix}.$$

Observe that E_t acts as a linear operator componentwise. Using the inequality (8) with $p = 2$, we have that E_t is a bounded operator on $\mathcal{H}$, uniformly in t. We also have the pointwise estimate

$$\sup_{t > 0} |E_t u| \leq M_{d,w} |u_\perp| + M_d |u_\parallel|$$

where $M_{d,w}$ is the dyadic maximal function with respect to dw and M_d the dyadic maximal function with respect to dx. Both are bounded on $L^2(w)$ by the Hardy-Littlewood theorem for the doubling measure dw and by Muckenhoupt's theorem since $w \in A_2$. Thus, $u \mapsto \sup_{t>0} |E_t u|$ is bounded from $\mathcal{H}$ into $L^2(w)$.

Definition 3.5. By the *principal part* of $(Q_t^B)_{t>0}$, we mean the function $(x, t) \mapsto \gamma_t(x)$ defined from $\mathbb{R}_+^{n+1}$ to $\mathcal{L}(\mathbb{C}^{n+1})$ by

$$\gamma_t(x) z := (Q_t^B z)(x)$$

for every $z \in \mathbb{C}^{n+1}$. We view z on the right-hand side of the above equation as the constant function valued in $\mathbb{C}^{n+1}$ defined on $\mathbb{R}^n$ by $z(x) := z$. We denote by $|\gamma_t(x)|$ its norm in $\mathcal{L}(\mathbb{C}^{n+1})$ subordinated to the hermitian structure on $\mathbb{C}^{n+1}$. We identify $\gamma_t(x)$ with the (possibly unbounded) multiplication operator $\gamma_t : u(x) \mapsto \gamma_t(x) u(x)$, $u \in \mathcal{H}$.

Lemma 3.6. *The operator Q_t^B extends to a bounded operator from $L^\infty(w; \mathbb{C}^{n+1})$ into $\mathcal{H}_{loc} = L^2_{loc}(w; \mathbb{C}^{n+1})$. In particular we have $\gamma_t \in L^2_{loc}(w; \mathcal{L}(\mathbb{C}^{n+1}))$ with bounds*

$$\fint_Q |\gamma_t(y)|^2 \, dw(y) \lesssim 1$$

for all $Q \in \triangle_t$. Moreover, $\gamma_t E_t$ are bounded on $\mathcal{H}$ with $\|\gamma_t E_t u\| \lesssim \|E_t u\|$ uniformly for all $t > 0$ and $u \in \mathcal{H}$.

Proof. Fix a cube $Q \in \triangle_t$ and $u \in L^\infty(w; \mathbb{C}^{n+1})$ with $\|u\|_\infty = 1$. Then write $u = u_0 + u_1 + u_2 + \ldots$ where $u_0 = u$ on $2Q$ and 0 elsewhere and if $j \geq 1$, $u_j = u$ on $2^{j+1}Q \setminus 2^j Q$ and 0 elsewhere. Then apply Q_t^B and use Lemma 3.4 for each term $Q_t^B u_j$ with N large enough and sum to obtain

$$\fint_Q |(Q_t^B u)(y)|^2 \, dw(y) \leq C.$$

If we do this for the constant functions with values describing an orthonormal basis of $\mathbb{C}^{n+1}$ and sum, we obtain an upper bound for the desired average of γ_t. Next, for a function $u \in \mathcal{H}$ and $Q \in \triangle_t$, as $E_t u$ is constant on Q,

$$\int_Q |\gamma_t(y) E_t u(y)|^2 \, dw(y) \leq \int_Q |\gamma_t(y)|^2 \, dw(y) \times \fint_Q |E_t u(y)|^2 \, dw(y)$$

$$\lesssim \int_Q |E_t u(y)|^2 \, dw(y).$$

Thus

$$\|\gamma_t E_t u\|^2 \lesssim \sum_{Q \in \triangle_t} \int_Q |E_t u(y)|^2 \, dw(y) = \|E_t u\|^2 \lesssim \|u\|^2.$$

$\square$

As Q_t^B vanishes on $\mathsf{N}(DB)$, it is enough to prove the quadratic estimate (14) for $v \in \overline{\mathsf{R}(DB)} = \overline{\mathsf{R}(D)}$. Our principal part approximation reads as follows.

Proposition 3.7. *We have*

$$(17) \qquad \int_0^\infty \|Q_t^B v - \gamma_t E_t v\|^2 \frac{dt}{t} \lesssim \|v\|^2, \quad v \in \overline{\mathsf{R}(D)}.$$

The function from $\mathbb{R}_+^{n+1}$ to $\mathbb{R}_+$ defined by $(x, t) \mapsto |\gamma_t(x)|^2$, is a *weighted dyadic Carleson function* if there exists $C < \infty$ such that

$$\iint_{\widehat{Q}} |\gamma_t(x)|^2 \frac{dw(x) dt}{t} \leq C^2 w(Q)$$

for all dyadic cubes $Q \subseteq \mathbb{R}^n$. Here $\widehat{Q} := Q \times (0, l(Q)]$ is the Carleson box over Q. We define the dyadic Carleson norm $\|\gamma_t\|_C$ to be the smallest constant C for which this inequality holds. The form of Carleson's

lemma that we need and will apply componentwise is the following (see [9], p.168 and references therein).

Lemma 3.8. *For all $u \in \mathcal{H}$,*

$$\int_0^\infty \|\gamma_t E_t u\|^2 \frac{dt}{t} \lesssim \|\gamma_t\|_C^2 \| \sup_{t>0} |E_t u|\|^2 \lesssim \|\gamma_t\|_C^2 \|u\|^2.$$

Corollary 3.9. *If $|\gamma_t(x)|^2$ is a weighted dyadic Carleson function, then Theorem 3.3 holds.*

This corollary clearly follows from the above lemma and (17).

3.4. **Proof of the principal part approximation.** We begin the proof of the principal part approximation (17) with some further notation. Define

$$P_t = \begin{bmatrix} (I - t^2 \Delta_w)^{-1} & 0 \\ 0 & (I - t^2 \Delta)^{-1} I_{\mathbb{C}^n} \end{bmatrix}.$$

Here $\Delta_w = \mathrm{div}_w \nabla$ is the negative self-adjoint operator on $L^2(w)$ defined in Lemma 2.8 while Δ is the usual negative Laplacian on $L^2(dx)$. From Lemma 2.2, the convolution operator $(I - t^2 \Delta)^{-1}$ is bounded on $L^2(w)$ uniformly with respect to $t > 0$: It is indeed classical that $(I - t^2 \Delta)^{-1}$ is the convolution with $t^{-n} G_2(x/t)$ where the Bessel potential G_2 is integrable and radially decreasing (see [25, Chapter 6]). Thus, P_t is uniformly bounded on $\mathcal{H}$.

Lemma 3.10. *For all $v \in \overline{R(D)}$, one has*

$$(18) \qquad \int_0^\infty \|Q_t^B (I - P_t) v\|^2 \frac{dt}{t} \lesssim \|v\|^2.$$

Proof. The key point is that $\overline{R(D)}$ is preserved by P_t. This follows from the characterization of $\overline{R(D)}$ in Proposition 3.2 as $\begin{bmatrix} L^2(w) \\ \mathcal{R}(L^2(w)) \end{bmatrix}$, since the Riesz transforms and $(I - t^2 \Delta)^{-1}$ commute and $\mathcal{R}(-\Delta)^{1/2} = \nabla$. Write $v = \begin{bmatrix} f \\ \mathcal{R}g \end{bmatrix}$, with $f, g \in L^2(w)$. Then,

$$(I - P_t) v = \begin{bmatrix} -t^2 \Delta_w (I - t^2 \Delta_w)^{-1} f \\ -t^2 \Delta (I - t^2 \Delta)^{-1} \mathcal{R} g \end{bmatrix}$$

$$= \begin{bmatrix} -t^2 \mathrm{div}_w \nabla (I - t^2 \Delta_w)^{-1} f \\ t^2 \nabla (-\Delta)^{1/2} (I - t^2 \Delta)^{-1} g \end{bmatrix}$$

$$= -tD \begin{bmatrix} t(-\Delta)^{1/2} (I - t^2 \Delta)^{-1} g \\ t \nabla (I - t^2 \Delta_w)^{-1} f \end{bmatrix}.$$

Thus using that $tQ_t^B D$ is uniformly bounded on $\mathcal{H}$, it suffices to prove

$$\int_0^\infty \|t \nabla (I - t^2 \Delta_w)^{-1} f\|^2 \frac{dt}{t} \lesssim \|f\|^2$$

and

$$\int_0^\infty \|t(-\Delta)^{1/2}(I - t^2\Delta)^{-1}g\|^2 \frac{dt}{t} \lesssim \|g\|^2.$$

The first estimate is a simple consequence of the construction of $-\Delta_w$ as the self-adjoint operator $\nabla^*\nabla$ on $L^2(w)$. Indeed, $\|\nabla(I - t^2\Delta_w)^{-1}f\| = \|(-\Delta_w)^{1/2}(I - t^2\Delta_w)^{-1}f\|$ and one concludes using the spectral theorem for the self-adjoint operator Δ_w that $\int_0^\infty \|t\nabla(I - t^2\Delta_w)^{-1}f\|^2 \frac{dt}{t} = c\|f\|^2$ with $c = \int_0^\infty t(1 + t^2)^{-1}\frac{dt}{t}$. The second estimate is a consequence of Lemma 2.5 as soon as the conditions for the function ψ defined on the Fourier transform side by $\hat{\psi}(\xi) = |\xi|(1 + |\xi|^2)^{-1}$ have been checked. $\square$

Lemma 3.11. *For all $v \in \overline{R(D)}$, one has*

$$\int_0^\infty \|Q_t^B P_t v - \gamma_t E_t P_t v\|^2 \frac{dt}{t} \lesssim \|v\|^2.$$

Proof. We remark that for $t > 0$ fixed and $x \in \mathbb{R}^n$, we have

$$(Q_t^B P_t v - \gamma_t E_t P_t v)(x) = Q_t^B(u - u_Q)(x),$$

where $u = P_t v$, Q is the unique dyadic cube in $\triangle_t$ containing x and u_Q is the value of $E_t u$ on Q. Define $C_0(Q) = 2Q$ and $C_j(Q) = 2^{j+1}Q \setminus 2^j Q$ if $j \in \mathbb{N}^*$. Then,

$$\|Q_t^B P_t v - \gamma_t E_t P_t v\|^2 = \sum_{Q \in \triangle_t} \int_Q |Q_t^B(u - u_Q)|^2 \, dw$$

$$\le \sum_{Q \in \triangle_t} \left(\sum_{j \ge 0} \left(\int_Q |Q_t^B(\mathbf{1}_{C_j(Q)}(u - u_Q))|^2 \, dw \right)^{1/2} \right)^2$$

$$\lesssim \sum_{Q \in \triangle_t} \left(\sum_{j \ge 0} 2^{-jN} \left(\int_{C_j(Q)} |u - u_Q|^2 \, dw \right)^{1/2} \right)^2$$

$$\lesssim \sum_{Q \in \triangle_t} \sum_{j \ge 0} 2^{-jN} \int_{C_j(Q)} |u - u_Q|^2 \, dw$$

$$\lesssim \sum_{Q \in \triangle_t} \sum_{j \ge 0} 2^{-jN} 2^{2j} \ell(Q)^2 2^{jd} \int_{2^{j+1}Q} |\nabla u|^2 \, dw$$

$$\lesssim t^2 \sum_{j \ge 0} 2^{-jN} 2^{2j} 2^{jd} 2^{jn} \int_{\mathbb{R}^n} |\nabla u|^2 \, dw.$$

$$\lesssim t^2 \|\nabla u\|^2.$$

We used the Minkowski inequality on the second line, off-diagonal decay on the third, Cauchy–Schwarz inequality on the fourth, Poincaré inequality on the fifth (recalling that one can take the average with respect to either dx or dw), and a telescoping argument which produces the doubling exponent d of w, the covering inequality $\sum_{Q \in \triangle_t} \mathbf{1}_{2^{j+1}Q} \lesssim$

2^{jn} and $\ell(Q) \sim t$ on the sixth. Finally, we choose $N > n+d+2$ in the last. Hence

$$\int_0^\infty \|Q_t^B P_t v - \gamma_t E_t P_t v\|^2 \, \frac{dt}{t} \lesssim \int_0^\infty \|t\nabla P_t v\|^2 \, \frac{dt}{t}.$$

For the scalar part of v, we have to control the weighted quadratic estimate for $t\nabla(I - t^2\Delta_w)^{-1} v_\perp$ which we have seen already. Using $v \in \overline{\mathsf{R}(D)}$, the tangential part $v_\parallel$ is of the form $\mathcal{R}g$ for some $g \in L^2(w)$. Hence we have to control the quadratic estimate of $t\nabla(I - t^2\Delta)^{-1} R_j g = R_j \mathcal{R} t(-\Delta)^{1/2}(I - t^2\Delta)^{-1}g$ for $j = 1, \ldots, n$. We can eliminate R_j and $\mathcal{R}$ as the Riesz transforms are bounded on $L^2(w)$ and the weighted Littlewood-Paley estimate for $t(-\Delta)^{1/2}(I - t^2\Delta)^{-1}$ has been already seen. □

Lemma 3.12. *There are constants $C < \infty$ and $\tau_1 \in (0,1)$ such that for all $f \in \mathsf{D}(\mathrm{div}_w)$ and all dyadic cubes Q,*

$$\left| \fint_Q \mathrm{div}_w f \, dw \right| \leq \frac{C}{\ell(Q)^{\tau_1}} \left(\fint_Q |\mathrm{div}_w f|^2 \, dw \right)^{\frac{1-\tau_1}{2}} \left(\fint_Q |f|^2 \, dw \right)^{\frac{\tau_1}{2}}.$$

Proof. Observe that if f has support contained in Q, then $\int_Q \mathrm{div}_w f \, dw = 0$. Thus this lemma follow from [12]. Here is a simple proof in our situation. Let $A = \left(\fint_Q |\mathrm{div}_w f|^2 \, dw \right)^{1/2}$ and $B = \left(\fint_Q |f|^2 \, dw \right)^{1/2}$. If $B \geq A\ell$ a simple application of the Cauchy-Schwarz inequality gives the result.

Assume next that $B < A\ell$. Let $\varphi \colon \mathbb{R}^n \to [0,1]$ be a smooth function which equals to 1 on $(1-t)Q$, 0 on Q^c with $\|\nabla\varphi\|_\infty \leq C(t\ell)^{-1}$ for some $t \in (0,1)$ to be chosen. We can write

$$\fint_Q \mathrm{div}_w f \, dw = \fint_Q \varphi \, \mathrm{div}_w f \, dw + \fint_Q (1-\varphi)\mathrm{div}_w f \, dw = I + II.$$

Using that $-\mathrm{div}_w$ is the adjoint of ∇ (writing the integral as an integral on $\mathbb{R}^n$ thanks to the support of φ),

$$|I| = \left| \fint_Q \nabla\varphi f \, dw \right| \leq C(t\ell)^{-1}B.$$

For II, using Cauchy-Schwarz inequality and the right hand side of (6) with E being the support of $1-\varphi$ in Q,

$$|II| \leq A\frac{w(Q \setminus (1-t)Q)^{1/2}}{w(Q)^{1/2}} \leq C_w A t^{\sigma/2}.$$

Hence, choosing $t^{1+\sigma/2} = B/A\ell$, we obtain the inequality with $\tau_1 = \frac{\sigma/2}{1+\sigma/2}$. □

Lemma 3.13. *There are constants $C < \infty$ and $\tau_2 \in (0, 1)$ such that for all $g \in D(\nabla)$ and all dyadic cubes Q,*

$$\left| \fint_Q \nabla g \, dx \right| \leq \frac{C}{\ell(Q)^{\tau_2}} \left(\fint_Q |\nabla g|^2 \, dw \right)^{\frac{1-\tau_2}{2}} \left(\fint_Q |g|^2 \, dw \right)^{\frac{\tau_2}{2}}.$$

Proof. The proof follows a similar pattern to that of Lemma 3.12. Let $A = \left(\fint_Q |\nabla g|^2 \, dw \right)^{1/2}$ and $B = \left(\fint_Q |g|^2 \, dw \right)^{1/2}$. If $B \geq A\ell$ a simple application of (8) with $p = 2$ gives the result.

Assume next that $B < A\ell$. Let $\varphi \colon \mathbb{R}^n \to [0, 1]$ be a smooth function which equals to 1 on $(1-t)Q$, 0 on Q^c with $\|\nabla \varphi\|_\infty \leq C(t\ell)^{-1}$ for some $t \in (0, 1)$ to be chosen. We can write

$$\fint_Q \nabla g \, dx = \fint_Q \varphi \nabla g \, dx + \fint_Q (1 - \varphi) \nabla g \, dx = I + II.$$

Interpreting I as a distribution-test function pairing, we may integrate by parts coordinatewise, and using that $w \in A_2$ we find

$$|I| = \left| \fint_Q g \nabla \varphi \, dx \right| \leq C(t\ell)^{-1} \fint_Q |g| \, dx \leq C[w]_{A_2}(t\ell)^{-1}B.$$

For II, pick $p > 1$ such that $w \in A_{2/p}$ and use Hölder's inequality with conjugate exponents p and p', the support properties of $1 - \varphi$, and the fact that $w \in A_{2/p}$, to conclude that

$$|II| \lesssim \left(\fint_Q |\nabla g|^p \, dx \right)^{1/p} t^{1/p'} \leq [w]_{A_{2/p}}^{1/p} A t^{1/p'}.$$

Hence, choosing $t^{1+1/p'} = B/A\ell$, we obtain the inequality with $\tau_2 = \frac{1/p'}{1+1/p'}$. $\qquad\square$

Lemma 3.14. *For all $u \in \mathcal{H} = L^2(w; \mathbb{C}^{n+1})$,*

$$\int_0^\infty \|\gamma_t E_t (P_t - I) u\|^2 \, \frac{dt}{t} \lesssim \|u\|^2.$$

Proof. It follows from Lemma 3.6 that $\|\gamma_t E_t(P_t - I)u\| \lesssim \|E_t(P_t - I)u\|$. Given the definitions of E_t and P_t, the lemma reduces to the scalar inequalities

$$\int_0^\infty \|A_t^w((I - t^2\Delta_w)^{-1} - I)f\|^2 \, \frac{dt}{t} \lesssim \|f\|^2$$

and

$$\int_0^\infty \|A_t((I - t^2\Delta)^{-1} - I)f\|^2 \, \frac{dt}{t} \lesssim \|f\|^2.$$

For the first one, we follow [10] with a minor simplification. Let $Q_s^w = s^2\Delta_w e^{s^2\Delta_w}$. Then for $f \in L^2(w)$ we have by the spectral theorem, $f = 8 \int_0^\infty (Q_s^w)^2 f \, \frac{ds}{s}$ and also $\|f\|^2 = 8 \int_0^\infty \|Q_s^w f\|^2 \, \frac{ds}{s}$. By Schur's lemma, it is enough to show that the operator norm of $A_t^w((I - t^2\Delta_w)^{-1} - I)Q_s^w$

in $L^2(w)$ is bounded by $h(s/t)$ with $h \geq 0$ and $\int_0^\infty h(u) \frac{du}{u} < \infty$. We shall find $h(u) = C \inf(u^{\tau_1}, u^{-2})$.

If $t < s$,

$$\|A_t^w((I - t^2\Delta_w)^{-1} - I)Q_s^w f\| \lesssim \|((I - t^2\Delta_w)^{-1} - I)s^2\Delta_w e^{s^2\Delta_w} f\|$$
$$\lesssim C(t/s)^2 \|f\|.$$

If $s < t$,

$$\|A_t^w(I - t^2\Delta_w)^{-1}Q_s^w f\| \lesssim \|(I - t^2\Delta_w)^{-1}s^2\Delta_w e^{s^2\Delta_w} f\|$$
$$\lesssim C(s/t)^2 \|f\|.$$

Thus, it remains to study the operator norm of $A_t^w Q_s^w$ for $s < t$. For this, we remark that $A_t^w Q_s^w f(x)$ is an average with respect to dw of $\mathrm{div}_w(sg)$ with $g = -s\nabla e^{s^2\Delta_w} f$ so that we can use Lemma 3.12. Thus,

$$\|A_t^w Q_s^w f\|^2 = \sum_{Q \in \Delta_t} w(Q) \left| \fint_Q \mathrm{div}_w(sg)\, dw \right|^2$$

$$\lesssim \frac{1}{t^{2\tau_1}} \sum_{Q \in \Delta_t} w(Q) \left(\fint_Q |\mathrm{div}_w(sg)|^2\, dw \right)^{1-\tau_1} \left(\fint_Q |sg|^2\, dw \right)^{\tau_1}$$

$$= \frac{s^{2\tau_1}}{t^{2\tau_1}} \sum_{Q \in \Delta_t} \left(\int_Q |Q_s^w f|^2\, dw \right)^{1-\tau_1} \left(\int_Q |s\nabla e^{s^2\Delta_w} f|^2\, dw \right)^{\tau_1}$$

$$\leq \frac{s^{2\tau_1}}{t^{2\tau_1}} \|Q_s^w f\|^{2(1-\tau_1)} \|s\nabla e^{s^2\Delta_w} f\|^{2\tau_1}$$

$$\lesssim \frac{s^{2\tau_1}}{t^{2\tau_1}} \|f\|^2$$

where we used Hölder's inequality for the sum.

The proof of the second inequality is as follows. Setting $Q_s = s^2\Delta e^{s^2\Delta}$, as before we have that the operator norm of $A_t((I - t^2\Delta)^{-1} - I)Q_s$ on $L^2(dx)$ is bounded by $h(s/t)$ with h similar to above. As this operator is also bounded on $L^2(w)$ uniformly in s and t for all $w \in A_2$ (for the convolution operators, it follows from Lemma 2.2 and for A_t this has been seen before), its operator norm on $L^2(w)$ is bounded by $h(s/t)^\theta$ for some power $\theta > 0$ depending only on w using Lemma 2.6. Thus, we can use the fact that the integral $f = 8 \int_0^\infty (Q_s)^2 f \frac{ds}{s}$ converges in $L^2(w)$ from Corollary 2.4, Schur's Lemma and that $\int_0^\infty \|Q_s f\|^2 \frac{ds}{s} \lesssim \|f\|^2$ from Lemma 2.5. $\square$

Proof of Proposition 3.7. It is enough to write

$$Q_t^B v - \gamma_t E_t v = Q_t^B(I - P_t)v + (Q_t^B P_t v - \gamma_t E_t P_t v) + \gamma_t E_t(P_t - I)v$$

and to use successively Lemma 3.10, Lemma 3.11 and Lemma 3.14. $\square$

3.5. **Preamble to the Carleson measure estimate.** We are now ready to prove that $|\gamma_t(x)|^2$ is a weighted dyadic Carleson function and so, via Proposition 3.7, complete the proof of Theorem 3.3. The first step towards this is a compactness argument. As was seen in the solution of the Kato square root problem ([7]), the application of a stopping-time argument was made possible by restricting $\gamma_t(x)$ so that, once normalised, it is close to a fixed element in the unit sphere of $\mathcal{L}(\mathbb{C}^{1+n})$, the set of bounded linear transformations on $\mathbb{C}^{1+n}$. We will make use of the same stopping-time argument, but also require a second stopping-time related to the oscillation of the weight, and with it comes a second compactness argument which restricts our attention to Whitney boxes on which the average of the weight is close to that of the top cube.

A convenient way to define the Whitney box W_Q associated to a given dyadic cube Q is $W_Q = \{(x,t)\,;\, x \in Q, Q \in \Delta_t\}$. With our definition $\widehat{Q}$ is the union all $W_{Q'}$ for which Q' is a dyadic subcube of Q.

Consider the compact unit sphere in $\mathcal{L}(\mathbb{C}^{1+n})$ and the compact interval $[0, c_0]$, where c_0 is as in (7). For each $\nu \in \mathcal{L}(\mathbb{C}^{1+n})$ such that $|\nu| = 1$, $\tau \in [0, c_0]$ and $\sigma_1, \sigma_2 > 0$, define G_{τ,σ_2} as the union of those Whitney boxes W_Q for which $|\ln(w_Q) - (\ln w)_Q - \tau| < \sigma_2$, and

(19)
$$
\widetilde{\gamma}_t(x) = \begin{cases} \gamma_t(x) & \text{if } \gamma_t(x) \neq 0,\ \left|\frac{\gamma_t(x)}{|\gamma_t(x)|} - \nu\right| \leq \sigma_1 \text{ and } (x,t) \in G_{\tau,\sigma_2}, \\ 0 & \text{otherwise.} \end{cases}
$$

We recall the notation $B^w(Q)$ from Section 2.2 and set

$$
\Omega^w(Q) = \widehat{Q} \setminus \cup_{R \in B^w(Q)} \widehat{R}.
$$

Lemma 3.15. *Suppose that we can show*

$$
K = \sup_{\nu,\tau} \sup_{Q \in \triangle} \frac{1}{w(Q)} \iint_{\Omega^w(Q)} |\widetilde{\gamma}_t(x)|^2 \frac{dw(x)dt}{t} < \infty
$$

for some choice of parameters σ_1 and σ_2 depending only on $\|B\|_\infty$, κ, $[w]_{A_2}$ and n. Then $|\gamma_t(x)|^2$ is a weighted dyadic Carleson function.

Proof. Fix σ_1 and σ_2 so that the hypothesis applies. Let $Q \in \Delta$. Observe that the sets $\Omega^w(R)$ form a partition of $\widehat{Q}$ when R runs over elements of $B_*^w(Q) \cup \{Q\}$. Thus, by the hypothesis and Proposition 2.1,

$$
\iint_{\widehat{Q}} |\widetilde{\gamma}_t(x)|^2 \frac{dw(x)dt}{t} = \sum_{R \in B_*^w(Q) \cup \{Q\}} \iint_{\Omega^w(R)} |\widetilde{\gamma}_t(x)|^2 \frac{dw(x)dt}{t}
$$

$$
\leq K \sum_{R \in B_*^w(Q) \cup \{Q\}} w(R) \leq \frac{KC}{\sigma_w^2} w(Q) + Kw(Q).
$$

By the compactness of the unit sphere in $\mathcal{L}(\mathbb{C}^{1+n})$ and the interval $[0, c_0]$, there exist a finite index set $A \subseteq \mathbb{N}$ and, for each $j \in A$, choices of ν_j and τ_j such that $|\gamma_t(x)|^2 \leq \sum_{j \in A} |\widetilde{\gamma}_t^j(x)|^2$, where $\widetilde{\gamma}_t^j(x) = \widetilde{\gamma}_t(x)$ with the choice $\nu = \nu_j$ and $\tau = \tau_j$. This completes the proof. $\square$

3.6. **Stopping-time arguments for test functions.** We fix an arbitrary vector ξ in the unit sphere of $\mathbb{C}^{1+n}$. For any $Q_1 \in \triangle$ and $\sigma_3 > 0$ to be chosen, define a test function

$$(20) \qquad f_{Q_1}^\xi := \left(I + \left(\sigma_3 \ell(Q_1) DB \right)^2 \right)^{-1} (1_{Q_1}\xi) = P_{\sigma_3\ell(Q_1)}^B (1_{Q_1}\xi),$$

where 1_{Q_1} is the indicator of Q_1. Note that $\|f_{Q_1}^\xi\|^2 \lesssim w(Q_1)$ and $\|\sigma_3\ell(Q_1)DB f_{Q_1}^\xi\|^2 \lesssim w(Q_1)$ with uniform implicit constants with respect to $|\xi| = 1$, $\sigma_3 > 0$ and Q_1, as can be seen using the uniform boundedness in $t > 0$ of Q_t^B and P_t^B.

Lemma 3.16. *There exist a constant c depending only on $\|B\|_\infty$, κ, $[w]_{A_2}$ and n, and a constant $\delta > 0$ depending only on $[w]_{A_2}$, such that for all such ξ, Q_1 and σ_3,*

$$|E_{Q_1}(f_{Q_1}^\xi) - \xi| \leq c\sigma_3^\delta.$$

Proof. We have $E_{Q_1}(f_{Q_1}^\xi) - \xi = E_{Q_1} Du$ with

$$u := -(\sigma_3\ell(Q_1))^2 BDB \left(I + \left(\sigma_3\ell(Q_1)DB \right)^2 \right)^{-1} (1_{Q_1}\xi)$$

and notice that $E_{Q_1} D$ acts on $u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ componentwise by averaging $\mathrm{div}_w u_2$ with respect to dw and ∇u_1 with respect to dx. Lemma 3.12 says

$$\left| \fint_{Q_1} \mathrm{div}_w u_2 \, dw \right| \leq \frac{C}{\ell(Q_1)^{\tau_1}} \left(\fint_{Q_1} |\mathrm{div}_w u_2|^2 \, dw \right)^{\frac{1-\tau_1}{2}} \left(\fint_{Q_1} |u_2|^2 \, dw \right)^{\frac{\tau_1}{2}}$$

$$\leq C\sigma_3^{\tau_1} \left(\fint_{Q_1} |\xi - f_{Q_1}^\xi|^2 \, dw \right)^{\frac{1-\tau_1}{2}} \left(\fint_{Q_1} |\sigma_3\ell(Q_1)DB f_{Q_1}^\xi|^2 \, dw \right)^{\frac{\tau_1}{2}}$$

$$\leq C\sigma_3^{\tau_1}$$

and Lemma 3.13 says

$$\left| \fint_Q \nabla u_1 \, dx \right| \leq \frac{C}{\ell(Q)^{\tau_2}} \left(\fint_Q |\nabla u_1|^2 \, dw \right)^{\frac{1-\tau_2}{2}} \left(\fint_Q |u_1|^2 \, dw \right)^{\frac{\tau_2}{2}}$$

$$\leq C\sigma_3^{\tau_2} \left(\fint_{Q_1} |\xi - f_{Q_1}^\xi|^2 \, dw \right)^{\frac{1-\tau_2}{2}} \left(\fint_{Q_1} |\sigma_3\ell(Q_1)DB f_{Q_1}^\xi|^2 \, dw \right)^{\frac{\tau_2}{2}}$$

$$\leq C\sigma_3^{\tau_2}.$$

So taking $\delta = \min(\tau_1, \tau_2)$ completes the proof. $\square$

Recall that $w_Q = \fint_Q w \, dx$ and similarly for $(\ln w)_Q$.

Lemma 3.17. *Fix $\tau \in [0, c_0]$ and ξ in the unit sphere of $\mathbb{C}^{1+n}$. Let*

$$S^{\tau}_{Q_1} = \begin{bmatrix} w_{Q_1} e^{-\tau - (\ln w)_{Q_1}} & 0 \\ 0 & I \end{bmatrix},$$

and define the collection of 'bad' cubes $B^{\tau,\xi}(Q_1)$ to be the set of maximal $Q' \in \triangle$ such that $Q' \subseteq Q_1$ and

either $|E_{Q'}(f^{\xi}_{Q_1})| > \frac{1}{\sigma_4}$ or $\mathrm{Re}\left(S^{\tau}_{Q_1}(\xi), \begin{bmatrix} w_{Q'}/w_{Q_1} & 0 \\ 0 & I \end{bmatrix} E_{Q'}(f^{\xi}_{Q_1}) \right) < \sigma_5.$

We can then choose positive σ_3, σ_4 and σ_5 depending only on $\|B\|_{\infty}$, κ, $[w]_{A_2}$ and n, in particular independently on τ, ξ, Q_1, so that

$$(21) \qquad \sum_{R \in B^{\tau,\xi}(Q_1)} w(R) \leq (1 - \sigma_6) w(Q_1),$$

with $0 < \sigma_6 \leq 1$.

Proof. There are two sets of cubes to consider. The first is the set of those maximal Q' for which

$$(22) \qquad \mathrm{Re}\left(S^{\tau}_{Q_1}(\xi), \begin{bmatrix} w_{Q'}/w_{Q_1} & 0 \\ 0 & I \end{bmatrix} E_{Q'}(f^{\xi}_{Q_1}) \right) < \sigma_5.$$

By (7), we know that

$$-c_0 \leq \ln(w_{Q_1}) - \tau - (\ln w)_{Q_1} \leq c_0$$

so $S^{\tau}_{Q_1}$ is a constant self-adjoint matrix with $e^{-c_0} I \leq S^{\tau}_{Q_1} \leq e^{c_0} I$. Applying Lemma 3.16,

$$\mathrm{Re}\left(S^{\tau}_{Q_1}(\xi), E_{Q_1}(f^{\xi}_{Q_1}) \right) = \left(S^{\tau}_{Q_1}(\xi), \xi \right) + \mathrm{Re}\left(S^{\tau}_{Q_1}(\xi), E_{Q_1}(f^{\xi}_{Q_1}) - \xi \right)$$

$$(23) \qquad\qquad \geq e^{-c_0} - c e^{c_0} \sigma_3^{\delta} \geq \frac{1}{2} e^{-c_0},$$

on choosing σ_3 so that $2c\sigma_3^{\delta} \leq e^{-2c_0}$. Consequently, setting $G = Q_1 \setminus (\cup Q')$ and $f^{\xi}_{Q_1} = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}$, we have

$$E_{Q_1}(f^{\xi}_{Q_1}) = \begin{bmatrix} \frac{1}{w(Q_1)} \int_{Q_1} f_1 \, dw \\ \frac{1}{|Q_1|} \int_{Q_1} f_2 \, dx \end{bmatrix}$$

$$= \sum_{Q'} \frac{|Q'|}{|Q_1|} \begin{bmatrix} w_{Q'}/w_{Q_1} & 0 \\ 0 & I \end{bmatrix} E_{Q'}(f^{\xi}_{Q_1}) + \begin{bmatrix} \frac{1}{w(Q_1)} \int_G f_1 \, dw \\ \frac{1}{|Q_1|} \int_G f_2 \, dx \end{bmatrix},$$

where the subcubes Q' are those of (22). Using (22) and (23), we obtain

$$
\frac{1}{2} e^{-c_0} \leq \operatorname{Re} \left(S_{Q_1}^\tau(\xi), E_{Q_1}(f_{Q_1}^\xi) \right)
$$

$$
\leq \sigma_5 \sum_{Q'} \frac{|Q'|}{|Q_1|} + \operatorname{Re} \left(S_{Q_1}^\tau(\xi), \begin{bmatrix} \frac{1}{w(Q_1)} \int_G f_1 \, dw \\ \frac{1}{|Q_1|} \int_G f_2 \, dx \end{bmatrix} \right)
$$

$$
\leq \sigma_5 + \operatorname{Re} \left(S_{Q_1}^\tau(\xi), \begin{bmatrix} \frac{1}{w(Q_1)} \int_G f_1 \, dw \\ \frac{1}{|Q_1|} \int_G f_2 \, dx \end{bmatrix} \right)
$$

and, using the estimate $\int_{Q_1} |f_{Q_1}^\xi|^2 \, dw \lesssim w(Q_1)$ and again the A_2 condition for w,

$$
\left| \left(S_{Q_1}^\tau(\xi), \begin{bmatrix} \frac{1}{w(Q_1)} \int_G f_1 \, dw \\ \frac{1}{|Q_1|} \int_G f_2 \, dx \end{bmatrix} \right) \right|
$$

$$
\leq e^{c_0} \left(\frac{w(G)}{w(Q_1)^2} \left(\int_G |f_1|^2 dw \right) + \frac{w^{-1}(G)}{|Q_1|^2} \left(\int_G |f_2|^2 dw \right) \right)^{1/2}
$$

$$
\lesssim \left(\frac{w(G)}{w(Q_1)} \right)^{1/2} + \frac{(w^{-1}(G))^{1/2} w(Q_1)^{1/2} (w^{-1}(Q_1))^{1/2}}{(w^{-1}(Q_1))^{1/2} |Q_1|}
$$

$$
\lesssim \left(\frac{w(G)}{w(Q_1)} \right)^{1/2} + \left(\frac{w^{-1}(G)}{w^{-1}(Q_1)} \right)^{1/2} \lesssim \left(\frac{w(G)}{w(Q_1)} \right)^\theta.
$$

for some $\theta > 0$ by (6) applied to w^{-1} and w. Therefore, for a small enough choice of σ_5, we have that

$$
\left(\frac{w(G)}{w(Q_1)} \right)^\theta \gtrsim 1
$$

which implies that $w(G) \geq 2\sigma_6 w(Q_1)$ for some small $\sigma_6 > 0$ and so

$$
(24) \qquad \sum_{Q'} w(Q') \leq (1 - 2\sigma_6) w(Q_1),
$$

where the sum is taken over those cubes Q' which satisfy (22).

Now we consider the set of maximal dyadic subcubes Q' of Q_1 for which

$$
(25) \qquad |E_{Q'}(f_{Q_1}^\xi)| > \frac{1}{\sigma_4}.
$$

Then $w(Q') \leq \sigma_4^2 \int_{Q'} |f_{Q_1}^\xi|^2 dw$ and

$$
\sum_{Q'} w(Q') \leq \sum_{Q'} \sigma_4^2 \int_{Q'} |f_{Q_1}^\xi|^2 dw \leq \sigma_4^2 \int_{Q_1} |f_{Q_1}^\xi|^2 dw \lesssim \sigma_4^2 \, w(Q_1)
$$

where the sum is now over those cubes Q' which satisfy (25). So we can choose σ_4 so small that

$$(26) \qquad \sum_{Q'} w(Q') \leq \sigma_6\, w(Q_1).$$

Combining (24) and (26) proves the lemma. $\qquad\square$

3.7. Conclusion of the Carleson measure estimate.

Consider $\tilde{\gamma}_t(x)$ depending on ν and τ as defined in (19). Associate to ν a vector $\xi \in \mathbb{C}^{1+n}$ such that $|\nu(\xi)| = 1$ and $|\xi| = 1$. Such a ξ may not be uniquely defined but we pick one. For a cube $Q_1 \in \triangle$, consider the test function $f_{Q_1}^{\xi}$ of (20) and set $\Omega^{\tau,\xi}(Q_1) = \widehat{Q_1} \setminus \cup_{R \in B^{\tau,\xi}(Q_1)} \widehat{R}$ with $B^{\tau,\xi}(Q_1)$ defined in Lemma 3.17.

Lemma 3.18. *Suppose that σ_3, σ_4 and σ_5 are chosen as in Lemma 3.17 so that (21) holds. Then there exists a choice of σ_w, σ_1 and σ_2 so that for all $Q_0, Q_1 \in \triangle$ with $Q_1 \subseteq Q_0$, and all ν and τ,*

$$(27) \qquad |\tilde{\gamma}_t(x)| \leq C|\gamma_t E_t(f_{Q_1}^{\xi})(x)|, \quad \text{for } (x,t) \in \Omega^w(Q_0) \cap \Omega^{\tau,\xi}(Q_1),$$

where $C > 0$ depends only on the choice of σ_w, σ_1, σ_2, σ_3, σ_4 and σ_5.

Proof. We assume that $\Omega^w(Q_0) \cap \Omega^{\tau,\xi}(Q_1)$ is non-empty, otherwise, there is nothing to prove. Recall that it is a union of Whitney boxes $W_{Q'}$ and (27) follows from $\left| \frac{\tilde{\gamma}_t(x)}{|\tilde{\gamma}_t(x)|} \left(E_{Q'}(f_{Q_1}^{\xi}) \right) \right| \gtrsim 1$ for $(x,t) \in W_{Q'}$ with $\tilde{\gamma}_t(x) \neq 0$. For a Whitney box $W_{Q'} \subseteq \Omega^w(Q_0) \cap \Omega^{\tau,\xi}(Q_1)$ we have that

$$|(\ln w)_{Q'} - (\ln w)_{Q_0}| \leq \sigma_w,$$

$$|E_{Q'}(f_{Q_1}^{\xi})| \leq \frac{1}{\sigma_4} \text{ and}$$

$$\text{Re}\left(S_{Q_1}^{\tau}(\xi), \begin{bmatrix} w_{Q'}/w_{Q_1} & 0 \\ 0 & I \end{bmatrix} E_{Q'}(f_{Q_1}^{\xi}) \right) \geq \sigma_5.$$

The last two inequalities are the definition of $\Omega^{\tau,\xi}(Q_1)$. The first comes from the fact Q' is not contained in a cube of $B^w(Q_0)$ by Proposition 2.1. As $Q' \subseteq Q_1 \subseteq Q_0$, Q_1 is also not contained in a cube of $B^w(Q_0)$ and we also have

$$|(\ln w)_{Q_1} - (\ln w)_{Q_0}| \leq \sigma_w.$$

Moreover, recall that if $\tilde{\gamma}_t(x) \neq 0$, then

$$\left| \frac{\tilde{\gamma}_t(x)}{|\tilde{\gamma}_t(x)|} - \nu \right| \leq \sigma_1.$$

Finally, if $\tilde{\gamma}_t(x) \neq 0$ and $(x,t) \in W_{Q'}$ then

$$|\ln(w_{Q'}) - (\ln w)_{Q'} - \tau| < \sigma_2.$$

Clearly then, we may assume the six inequalities above.

We begin by observing that

$$\begin{bmatrix} w_{Q'}/w_{Q_1} & 0 \\ 0 & I \end{bmatrix} S_{Q_1}^\tau = \begin{bmatrix} e^{\ln(w_{Q'})-\tau-(\ln w)_{Q_1}} & 0 \\ 0 & I \end{bmatrix}$$

and

$$
\begin{aligned}
|\ln(w_{Q'}) - \tau - (\ln w)_{Q_1}| &\leq |\ln(w_{Q'}) - (\ln w)_{Q'} - \tau| \\
&\quad + |(\ln w)_{Q'} - (\ln w)_{Q_0}| + |(\ln w)_{Q_1} - (\ln w)_{Q_0}| \\
&\leq \sigma_2 + \sigma_w + \sigma_w.
\end{aligned}
$$
(28)

Recall that we chose σ_3 in (23) so that $0 < c\sigma_3^\delta \leq e^{-2c_0}/2 \leq 1/2$. Therefore, using Lemma 3.16,

$$\left| \nu\left(E_{Q'}(f_{Q_1}^\xi)\right) \right| \geq |\nu(\xi)| - \left| \nu\left(E_{Q'}(f_{Q_1}^\xi) - \xi\right) \right| \geq 1 - c\sigma_3^\delta \geq 1/2$$

and

$$\operatorname{Re}\left(\xi, E_{Q'}(f_{Q_1}^\xi)\right) = (\xi, \xi) + \operatorname{Re}\left(\xi, E_{Q'}(f_{Q_1}^\xi) - \xi\right) \leq 1 + c\sigma_3^\delta \leq 2,$$

so

$$\left| \nu\left(E_{Q'}(f_{Q_1}^\xi)\right) \right| \geq \frac{1}{4}\operatorname{Re}\left(\xi, E_{Q'}(f_{Q_1}^\xi)\right).$$

It then follows that for $(x,t) \in W_{Q'}$ with $\tilde{\gamma}_t(x) \neq 0$,

$$
\begin{aligned}
&\left| \frac{\tilde{\gamma}_t(x)}{|\tilde{\gamma}_t(x)|}\left(E_{Q'}(f_{Q_1}^\xi)\right) \right| \\
&\geq \left| \nu\left(E_{Q'}(f_{Q_1}^\xi)\right) \right| - \left| \left(\frac{\tilde{\gamma}_t(x)}{|\tilde{\gamma}_t(x)|} - \nu\right)\left(E_{Q'}(f_{Q_1}^\xi)\right) \right| \\
&\geq \frac{1}{4}\operatorname{Re}\left(\xi, E_{Q'}(f_{Q_1}^\xi)\right) - \frac{\sigma_1}{\sigma_4} \\
&= \frac{1}{4}\operatorname{Re}\left(\begin{bmatrix} e^{\ln(w_{Q'})-\tau-(\ln w)_{Q_1}} & 0 \\ 0 & I \end{bmatrix}\xi, E_{Q'}(f_{Q_1}^\xi)\right) \\
&\quad + \frac{1}{4}\operatorname{Re}\left(\begin{bmatrix} 1 - e^{\ln(w_{Q'})-\tau-(\ln w)_{Q_1}} & 0 \\ 0 & 0 \end{bmatrix}\xi, E_{Q'}(f_{Q_1}^\xi)\right) - \frac{\sigma_1}{\sigma_4} \\
&\geq \frac{1}{4}\operatorname{Re}\left(S_{Q_1}^\tau(\xi), \begin{bmatrix} w_{Q'}/w_{Q_1} & 0 \\ 0 & I \end{bmatrix}E_{Q'}(f_{Q_1}^\xi)\right) - \frac{e}{\sigma_4}(\sigma_2 + 2\sigma_w) - \frac{\sigma_1}{\sigma_4} \\
&\geq \frac{1}{4}\sigma_5 - \frac{e}{\sigma_4}(\sigma_2 + 2\sigma_w + \sigma_1).
\end{aligned}
$$

We have used (28) and $|1 - e^u| \leq eu$ if u is a real with $|u| \leq 1$, assuming $\sigma_2 + 2\sigma_w \leq 1$. We have already chosen σ_4 and σ_5, but we are still free to choose σ_w, σ_1 and σ_2 small so that (27) holds. $\qquad\square$

Proof of Theorem 3.3. By Corollary 3.9 and Lemma 3.15, we know it is enough to show, for fixed ν and τ, and for any cube Q_0, that

$$\iint_{\Omega^w(Q_0)} |\tilde{\gamma}_t(x)|^2 \frac{dw(x)dt}{t} \leq K w(Q_0).$$
(29)

Fix $Q_1 \in \Delta$ with $Q_1 \subseteq Q_0$. Having fixed the parameters in Lemma 3.18, we apply (27) in the first inequality to obtain

$$
\iint_{\Omega^w(Q_0) \cap \Omega^{\tau,\xi}(Q_1)} |\widetilde{\gamma}_t(x)|^2 \frac{dw(x)dt}{t}
$$

$$
\leq \iint_{\widehat{Q_1}} |\gamma_t E_t(f_{Q_1}^\xi)(x)|^2 \frac{dw(x)dt}{t}
$$

$$
\lesssim \iint_{\widehat{Q_1}} |(Q_t^B f_{Q_1}^\xi)(x)|^2 \frac{dw(x)dt}{t}
$$

$$
+ \iint_{\widehat{Q_1}} |(Q_t^B - \gamma_t E_t)(f_{Q_1}^\xi - 1_{Q_1}\xi)(x)|^2 \frac{dw(x)dt}{t}
$$

$$
+ \iint_{\widehat{Q_1}} |(Q_t^B - \gamma_t E_t)(1_{Q_1}\xi)(x)|^2 \frac{dw(x)dt}{t}.
$$

Since $Q_t^B f_{Q_1}^\xi = \frac{t}{\sigma_3 \ell(Q_1)} P_t^B(\sigma_3 \ell(Q_1) DB f_{Q_1}^\xi)$, one has that

$$
\iint_{\widehat{Q_1}} |(Q_t^B f_{Q_1}^\xi)(x)|^2 \frac{dw(x)dt}{t} \lesssim \int_0^{\ell(Q_1)} \frac{t^2 \|\sigma_3 \ell(Q_1) DB f_{Q_1}^\xi\|^2}{(\sigma_3 \ell(Q_1))^2} \frac{dt}{t} \lesssim w(Q_1),
$$

and, by Proposition 3.7 because $f_{Q_1}^\xi - 1_{Q_1}\xi \in \mathsf{R}(D)$,

$$
\iint_{\widehat{Q_1}} |(Q_t^B - \gamma_t E_t)(f_{Q_1}^\xi - 1_{Q_1}\xi)(x)|^2 \frac{dw(x)dt}{t} \lesssim \|f_{Q_1}^\xi - 1_{Q_1}\xi\|^2 \lesssim w(Q_1).
$$

For the last term, using that by definition $Q_t^B - \gamma_t E_t$ annihilates constants and $E_t((1_{Q_1} - 1)\xi)(x) = 0$ when $(x, t) \in \widehat{Q_1}$, we can rewrite

$$
(Q_t^B - \gamma_t E_t)(1_{Q_1}\xi)(x) = (Q_t^B - \gamma_t E_t)((1_{Q_1} - 1)\xi)(x) = Q_t^B((1_{Q_1} - 1)\xi)(x).
$$

Using off-diagonal estimates for Q_t^B as in Lemma 3.6, one can easily show $\iint_{\widehat{Q_1}} |Q_t^B((1_{2Q_1} - 1)\xi)(x)|^2 \frac{dw(x)dt}{t} \lesssim w(Q_1)$. Next, decompose $2Q_1 \setminus Q_1 = \partial_{a(t)} \cup (2Q_1 \setminus \partial_{a(t)})$ where $a(t) = \sqrt{t/\ell(Q_1)}\ell(Q_1)$ and $\partial_a = \{y \notin Q_1 ; d(y, Q_1) \leq a\}$. Again, using the off-diagonal estimates for each t, the function $1_{2Q_1 \setminus \partial_{a(t)}}$ contributes $w(Q_1)$. It remains to control the integral corresponding to $1_{\partial_{a(t)}}$. From the uniform boundedness of Q_t^B, one has

$$
\iint_{\widehat{Q_1}} |Q_t^B(1_{\partial_{a(t)}}\xi)(x)|^2 \frac{dw(x)dt}{t} \lesssim \int_0^{\ell(Q_1)} w(\partial_{a(t)}) \frac{dt}{t}
$$

$$
\lesssim \int_0^{\ell(Q_1)} \left(\frac{|\partial_{a(t)}|}{|Q_1|}\right)^\sigma \frac{dt}{t} w(Q_1)
$$

$$
\lesssim w(Q_1)
$$

using (6) and $\frac{|\partial_{a(t)}|}{|Q_1|} \lesssim \left(\frac{t}{\ell(Q_1)}\right)^{1/2}$ obtained from elementary observations. Summarizing the estimates above, we have proved that

$$\iint_{\Omega^w(Q_0) \cap \Omega^{\tau,\xi}(Q_1)} |\widetilde{\gamma}_t(x)|^2 \frac{dw(x)dt}{t} \lesssim w(Q_1).$$

We can now prove (29). Define

$$B_0^{\tau,\xi}(Q_0) = \{Q_0\}, \quad B_1^{\tau,\xi}(Q_0) = B^{\tau,\xi}(Q_0),$$
$$B_{j+1}^{\tau,\xi}(Q_0) = \cup_{R \in B_j^{\tau,\xi}(Q_0)} B^{\tau,\xi}(R) \text{ for } j = 1, 2, \ldots,$$
$$\text{and } B_*^{\tau,\xi}(Q_0) = \cup_{j=0}^\infty B_j^{\tau,\xi}(Q_0).$$

Using

$$\iint_{\Omega^w(Q_0)} |\widetilde{\gamma}_t(x)|^2 \frac{dw(x)dt}{t} = \sum_{Q_1 \in B_*^{\tau,\xi}(Q_0)} \iint_{\Omega^w(Q_0) \cap \Omega^{\tau,\xi}(Q_1)} |\widetilde{\gamma}_t(x)|^2 \frac{dw(x)dt}{t},$$

and summing the estimate above together with an iteration of Lemma 3.17 imply

$$\iint_{\Omega^w(Q_0)} |\widetilde{\gamma}_t(x)|^2 \frac{dw(x)dt}{t} \lesssim \sum_{Q_1 \in B_*^{\tau,\xi}(Q_0)} w(Q_1)$$
$$\leq \sum_{j=0}^\infty (1 - \sigma_6)^j w(Q_0) \lesssim w(Q_0),$$

which proves (29) and with it Theorem 3.3. $\qquad\square$

3.8. **The case of block matrices.** We show how to simplify the argument in this case. Recall that B is a $(n+1) \times (n+1)$ matrix. Assume here that it is block diagonal, namely

$$B(x) = \begin{bmatrix} a(x) & 0 \\ 0 & d(x) \end{bmatrix}$$

with $a(x)$ scalar-valued and $d(x)$ $n \times n$ matrix-valued. Define the normal and tangential spaces

$$\mathcal{H}_\perp = \begin{bmatrix} L^2(\mathbb{R}^n, w; \mathbb{C}) \\ 0 \end{bmatrix} \quad \text{and} \quad \mathcal{H}_\parallel = \begin{bmatrix} 0 \\ L^2(\mathbb{R}^n, w; \mathbb{C}^n) \end{bmatrix}.$$

In this case, both operators BD and DB swap the normal and tangential spaces. So do Q_t^B and multiplication by γ_t. This means that

$$\gamma_t(x) = \begin{bmatrix} 0 & \alpha_t(x) \\ \beta_t(x) & 0 \end{bmatrix}$$

so that the Carleson function norms for $|\alpha_t(x)|^2$ and $|\beta_t(x)|^2$ can be estimated separately. The normal and tangential parts of our test functions can be used in two separate stopping-time arguments, which do not require the Corona decomposition (Proposition 2.1), following

the usual proof in the unweighted case, since for each stopping-time we use the average against *one* measure: either dx or dw.

3.9. **A vector-valued extension.** The proof of the main quadratic estimate carries straightforwardly to the case of systems where

- $\mathcal{H} = L_2(\mathbb{R}^n, w; \mathbb{C}^{m(1+n)})$, where $\mathbb{C}^{m(1+n)} = (\mathbb{C}^m)^{1+n}$,
- D acts componentwise on $\mathcal{H}$ by $(Du)^\alpha = D(u^\alpha)$ for $\alpha = 1, \ldots, m$ (in other words, the new D is $D \otimes I_{\mathbb{C}^m}$ but we shall not use this notation), and
- $B(x)$ is an $(n + 1) \times (n + 1)$ matrix whose entries are $m \times m$ matrices and the multiplication by $B(x)$ is assumed to be bounded on $\mathcal{H}$ and accretive on $\overline{R(D)}$.

3.10. **Consequences.** We gather here some consequences on the functional calculus for the convenience of the reader.

Proposition 3.19. *Let w, D and B be as above. If $T = DB$ or $T = BD$, then one has the equivalence*

$$(30) \qquad \int_0^\infty \|tT(1 + t^2T^2)^{-1}u\|^2 \frac{dt}{t} \sim \|u\|^2, \qquad for \ all \ \ u \in \overline{R(T)}.$$

Proof. First, the square function estimate for BD follows from that for DB. Indeed, on $\mathsf{N}(BD)$, $tBD(1 + t^2(BD)^2)^{-1}u = 0$. On $\overline{R(BD)}$, BD is similar to DB so the square function inequality for BD follows. We conclude using the splitting $\mathcal{H} = \mathsf{N}(BD) \oplus \overline{R(BD)}$. Now, if one changes B to B^*, this means we have proved (14) for both $T = DB$ (resp. $T = BD$) and its adjoint. It is classical ([1]) that this implies the equivalence on the range. $\qquad \square$

The next result summarizes consequences of quadratic estimates that are needed.

Proposition 3.20. *Let T be an ω-bisectorial operator on a separable Hilbert space $\mathcal{H}$ with $0 \leq \omega < \pi/2$. Assume that the quadratic estimate*

$$(31) \qquad \int_0^\infty \|tT(1 + t^2T^2)^{-1}u\|^2 \frac{dt}{t} \sim \|u\|^2 \ holds \ for \ all \ \ u \in \overline{R(T)}.$$

Then, the following statements hold.

- *T has a bounded holomorphic functional calculus on $\overline{R(T)}$ on any bisector $|\arg(\pm z)| < \mu$ for any $\omega < \mu < \pi/2$, which can be extended to all $\mathcal{H}$ by setting $f(T) = f(0)I$ on $\mathsf{N}(T)$ whenever f is also defined at 0.*
- *The comparison*

$$(32) \qquad \int_0^\infty \|\psi(tT)u\|^2 \frac{dt}{t} \sim \|u\|^2 \ holds \ for \ all \ \ u \in \overline{R(T)}.$$

for any $\omega < \mu < \pi/2$ and for any holomorphic function ψ in the bisector $|\arg(\pm z)| < \mu$, which is not identically zero on

each connected component of the bisector and which satisfies $|\psi(z)| \leq C \inf(|z|^\alpha, |z|^{-\alpha})$ for some $C < \infty$ and $\alpha > 0$.

- The operator $\mathrm{sgn}(T)$ is a bounded involution on $\overline{R(T)}$.
- $\overline{R(T)}$ splits topologically into two spectral subspaces

(33) $$\overline{R(T)} = \mathcal{H}_T^+ \oplus \mathcal{H}_T^-$$

with $\mathcal{H}_T^\pm = E_T^\pm(\overline{R(T)})$ and $E_T^+ = \chi^\pm(T)$ are projections with $\chi^\pm(z) = 1$ if $\pm\,\mathrm{Re}\,z > 0$ and $\chi^\pm(z) = 0$ otherwise.

- The operator $|T| = sgn(T)T = \sqrt{T^2}$ with $D(|T|) = D(T)$ is a ω-sectorial operator and $-|T|$ generates an analytic semigroup of operators $(e^{-z|T|})_{|\arg z| < \pi/2 - \omega}$.
- For $h \in D(T)$, $h \in \mathcal{H}_T^\pm$ if and only if $|T|h = \pm Th$. As a consequence $e^{\mp zT}$ are well-defined operators on $\mathcal{H}_T^\pm$ respectively, and $e^{-zT}E_T^+$ and $e^{+zT}E_T^-$ are well-defined operators on $\mathcal{H}$ for $|\arg z| < \pi/2 - \omega$.

As announced in Section 3.2, we recall here why this implies the Kato conjecture for block diagonal

$$B = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$$

identifying the functions a and d with the corresponding multiplication operators. We have

$$BD = \begin{bmatrix} 0 & a\,\mathrm{div}_w \\ -d\nabla & 0 \end{bmatrix}, \quad (BD)^2 = \begin{bmatrix} -a\,\mathrm{div}_w d\nabla & 0 \\ 0 & -d\nabla a\,\mathrm{div}_w \end{bmatrix},$$

so that for $u \in H^1(\mathbb{R}^n, w; \mathbb{C}^m)$, $v = \begin{bmatrix} u \\ 0 \end{bmatrix} \in \mathsf{D}(BD) = \mathsf{D}(|BD|)$ and

$$\|\sqrt{-a\,\mathrm{div}_w d\nabla}\,u\| \sim \||BD|v\| \sim \|BDv\| \sim \|d\nabla u\| \sim \|\nabla u\|.$$

4. Representations for solutions of degenerate elliptic systems

From now on, we write points in the upper half-space $\mathbb{R}_+^{1+n}$ as $\mathbf{x} = (t, x)$, $t > 0$, $x \in \mathbb{R}^n$.

4.1. **From second order to first order.** We shall now follow closely [3], and its extension [34], but in the weighted setting. It is necessary to have these references handy. The estimates of these two articles obtained in abstract Hilbert spaces evidently apply here. Some other estimates use harmonic analysis (tent spaces, maximal functions). Thus we shall try to extract the relevant information and give proofs only when the argument uses a particular feature of the weighted situation.

We recall the notation $\mathcal{H} = L^2(\mathbb{R}^n, w; \mathbb{C}^{m(1+n)})$ and use $\mathcal{H}^0 = \overline{R(D)}$ where D was defined in Section 3.9. Beware that in [3], $\mathcal{H}$ was taken as

$\overline{R(D)}$. We continue to use $\| \; \|$ to denote the norm in $\mathcal{H}$, and occasionally use other notation when needed.

We construct solutions u to the divergence form system (1), by solving the equivalent vector-valued ODE (35) below for the w-normalized conormal gradient

$$f = \nabla_{w^{-1}A} u = \begin{bmatrix} \partial_{\nu_{w^{-1}A}} u \\ \nabla_x u \end{bmatrix},$$

and $\partial_{\nu_{w^{-1}A}} u$ denotes the upward (hence inward for $\mathbb{R}^{1+n}_+$) w-normalized conormal derivative of u.

Using the normal/tangential decomposition for $\mathbb{C}^{m(1+n)} = \mathbb{C}^m \oplus \mathbb{C}^{mn} = \mathbb{C}^m \oplus (\mathbb{C}^m \otimes \mathbb{C}^n)$ (see Section 3.2), we write matrices acting on $\mathbb{C}^{m(1+n)}$ as

$$M = \begin{bmatrix} M_{\perp\perp} & M_{\perp\|} \\ M_{\|\perp} & M_{\|\|} \end{bmatrix},$$

the entries being matrices acting from and into the various spaces in the splitting.

Proposition 4.1. *The transformation*

$$C \mapsto \widehat{C} := \begin{bmatrix} I & 0 \\ C_{\|\perp} & C_{\|\|} \end{bmatrix} \begin{bmatrix} C_{\perp\perp} & C_{\perp\|} \\ 0 & I \end{bmatrix}^{-1} = \begin{bmatrix} C_{\perp\perp}^{-1} & -C_{\perp\perp}^{-1}C_{\perp\|} \\ C_{\|\perp}C_{\perp\perp}^{-1} & C_{\|\|} - C_{\|\perp}C_{\perp\perp}^{-1}C_{\perp\|} \end{bmatrix}$$

is a self-inverse bijective transformation of the set of operator-valued matrices which are bounded on $\mathcal{H}$ and accretive on $\overline{R(D)}$.

The proof is analogous to that of [3].
We set

$$\widehat{w^{-1}A} = B$$

in what follows. Our assumption is that as a multiplication operator, $w^{-1}A(t, \cdot)$ is bounded on $\mathcal{H}$ and accretive on $\overline{R(D)}$ for a.e. $t > 0$ with uniform bounds with respect to t. In particular, the matrix $w^{-1}A_{\perp\perp}(t, \cdot)$ is invertible as an operator acting on $L^2(\mathbb{R}^n; w; \mathbb{C}^m)$, hence it is also invertible in $L^\infty(\mathbb{R}^n, \mathcal{L}(\mathbb{C}^m))$, with uniform bounds a.e. in $t > 0$. Thus, $B(t, \cdot)$ is also a multiplication operator.

We now introduce some notation. Let

$$\mathcal{D}_w = \begin{bmatrix} C_0^\infty(\mathbb{R}^{1+n}_+; \mathbb{C}^m) \\ w^{-1}C_0^\infty(\mathbb{R}^{1+n}_+; \mathbb{C}^{mn}) \end{bmatrix}.$$

Let $\mathcal{C}url_{\|,0} = \{f \in \mathcal{D}'(\mathbb{R}^n; \mathbb{C}^{m(1+n)}) \, ; \, \mathrm{curl}_x f_\| = 0\}$, where the curl operator is computed componentwise. Let $\mathcal{H}_{\mathrm{loc}} = L^2_{\mathrm{loc}}(\mathbb{R}^n, w; \mathbb{C}^{m(1+n)})$.

Proposition 4.2. *For a pair of coefficient matrices A and B related by $A = w\widehat{B}$, or equivalently $B = \widehat{w^{-1}A}$, the pointwise map $g \mapsto f = \begin{bmatrix} (w^{-1}Ag)_\perp \\ g_\| \end{bmatrix}$ gives a one-to-one correspondence, with inverse $g =$*

$\begin{bmatrix} (Bf)_\perp \\ f_\parallel \end{bmatrix}$, *between solutions g to the equations*

(34) $\qquad \begin{cases} g \in L^2_{\mathrm{loc}}(\mathbb{R}_+; \mathcal{H}_{\mathrm{loc}}), \\ \mathrm{div}_{t,x}(Ag) = 0, \\ \mathrm{curl}_{t,x} g = 0, \end{cases}$

in the sense of distributions on $\mathbb{R}_+^{1+n}$ *and solutions f to the generalized Cauchy–Riemann equations*

(35) $\qquad \begin{cases} f \in L^2_{\mathrm{loc}}(\mathbb{R}_+; \mathcal{H}_{\mathrm{loc}} \cap \mathcal{C}url_{\parallel,0}), \\ \partial_t f + DBf = 0, \end{cases}$

in the weak sense

(36) $\qquad \displaystyle\int_0^\infty -(f, \partial_t \varphi) + (Bf, D\varphi)\, dt = 0 \quad \forall \varphi \in \mathcal{D}_w,$

where (,) is the complex inner product with respect to dw.

The proof is almost completely identical to the one in [3].

Proof. The transformation $g \mapsto f = \begin{bmatrix} (w^{-1}Ag)_\perp \\ g_\parallel \end{bmatrix}$ is easily seen to be invertible on $L^2_{\mathrm{loc}}(\mathbb{R}_+; \mathcal{H}^{\mathrm{loc}})$. Consider a pair of functions g and f in $L^2_{\mathrm{loc}}(\mathbb{R}_+; \mathcal{H}^{\mathrm{loc}})$ related in this fashion. Equations (34) for g are equivalent to

(37) $\qquad \begin{cases} \partial_t(Ag)_\perp + \mathrm{div}_x(A_{\parallel\perp}g_\perp + A_{\parallel\parallel}g_\parallel) = 0, \\ \partial_t g_\parallel - \nabla_x g_\perp = 0, \\ \mathrm{curl}_x g_\parallel = 0, \end{cases}$

each in the sense of distributions on $\mathbb{R}_+^{1+n}$. The last equation is equivalent to $f_t = f(t, \cdot) \in \mathcal{C}url_{\parallel,0}$. Moreover, using that $(w^{-1}Ag)_\perp = f_\perp$, $g_\parallel = f_\parallel$ and $g_\perp = (Bf)_\perp = A_{\perp\perp}^{-1}(wf_\perp - A_{\perp\parallel}f_\parallel)$, the first two equations are seen to be equivalent to the equation $\partial_t f + DBf = 0$ in the prescribed sense. $\qquad \square$

Next, the strategy in [3] is to integrate the weak differential equation (36) to obtain an equivalent formulation in the Duhamel sense. Again, this can be followed almost line by line, once we have the following density lemma.

Lemma 4.3. *The space $\mathcal{D}_w$ is dense in $H^1_{\mathrm{c}}(\mathbb{R}_+; \mathcal{H}) \cap L^2_{\mathrm{c}}(\mathbb{R}_+; \mathsf{D}(D))$, where the subscript* c *means that elements have compact support in* $\mathbb{R}_+$. *Thus, if $f \in L^2_{\mathrm{loc}}(\mathbb{R}_+; \mathcal{H}^0)$, (36) holds for any $\varphi \in H^1_{\mathrm{c}}(\mathbb{R}_+; \mathcal{H}) \cap L^2_{\mathrm{c}}(\mathbb{R}_+; \mathsf{D}(D))$ if it does for any $\varphi \in \mathcal{D}_w$.*

Proof. The density of $\mathcal{D}_w$ in $H^1_{\mathrm{c}}(\mathbb{R}_+; \mathcal{H}) \cap L^2_{\mathrm{c}}(\mathbb{R}_+; \mathsf{D}(D))$ can be easily established using (2) in Proposition 3.2 and standard truncation and regularization in the t-variable. If $f \in L^2_{\mathrm{loc}}(\mathbb{R}_+; \mathcal{H}^0)$ and $\varphi \in$

$H^1_c(\mathbb{R}_+; \mathcal{H}) \cap L^2_c(\mathbb{R}_+; \mathsf{D}(D))$ then the integral in (36) makes sense and vanishes by approximating φ by elements in $\mathcal{D}_w$. $\qquad\square$

In the above proposition, we are mostly interested in having $g \in L^2_{\mathrm{loc}}(\mathbb{R}_+; \mathcal{H})$. Recall that

$$\mathcal{H}^0 = \overline{\mathsf{R}(D)} = L^2(\mathbb{R}^n, w; \mathbb{C}^{m(1+n)}) \cap \mathcal{C}url_{\|,0} = \mathcal{H} \cap \mathcal{C}url_{\|,0}.$$

In particular, $g \in L^2_{\mathrm{loc}}(\mathbb{R}_+; \mathcal{H})$ if and only if $f \in L^2_{\mathrm{loc}}(\mathbb{R}_+; \mathcal{H}^0)$ and we can apply the above lemma. Formally writing $\partial_t f + DB_0 f = D(\mathcal{E}f)$, $\mathcal{E} = B_0 - B$, where B_0 is now multiplication by a t-independent matrix, the integration of (36) leads to the following equation

$$(38) \qquad\qquad f_t = e^{-tDB_0} E_0^+ h^+ + (S_A f)_t,$$

for a unique $h^+ \in \mathcal{H}^+_{DB_0}$ and where S_A is the vector-valued singular integral operator given

$$(39) \quad (S_A f)_t := \int_0^t e^{-(t-s)DB_0} E_0^+ D\mathcal{E}_s f_s \, ds - \int_t^\infty e^{(s-t)DB_0} E_0^- D\mathcal{E}_s f_s \, ds.$$

Here $E_0^\pm = \chi^\pm(DB_0)$ are the projections defined in Section 3.10 and $\mathcal{H}^+_{DB_0} := E_0^\pm \mathcal{H}$ are the ranges of the respective projections. We also use the notation $g_t = g(t, \cdot)$. This operator can be rigorously defined using the maximal regularity operator for $|DB_0|$ viewed from the operational calculus point of view as $F(|DB_0|)$ with $F(z)$ being the operator-valued analytic function given by

$$(F(z)g)_t := \int_0^t z e^{-(t-s)z} g_s \, ds, \ \ \mathrm{Re}\, z > 0,$$

so (39) becomes

$$(40) \qquad\qquad S_A = F(|DB_0|)\widehat{E}_0^+ \mathcal{E} + F^*(|DB_0|)\widehat{E}_0^- \mathcal{E},$$

where $\widehat{E}_0^\pm$ are bounded operators on $\mathcal{H}$ such that $E_0^\pm D = (DB_0)\widehat{E}_0^\pm$. These two representations and Proposition 3.19 allow us to prove most relevant boundedness results concerning the regularity and Neumann problems. For the Dirichlet problem, they can be used as well in an appropriate sense (see [34]) but there is another useful representation using the operator

$$(41) \quad (\widetilde{S}_A f)_t := \int_0^t e^{-(t-s)B_0 D} \widetilde{E}_0^+ \mathcal{E}_s f_s \, ds - \int_t^\infty e^{(s-t)B_0 D} \widetilde{E}_0^- \mathcal{E}_s f_s \, ds,$$

where $\widetilde{E}_0^\pm = \chi^\pm(B_0 D)$, and the vector field defined by

$$v_t := e^{-tB_0 D} \widetilde{E}_0^+ \tilde{h}^+ + (\widetilde{S}_A f)_t,$$

for some $\tilde{h}^+$. From the intertwining property $b(DB_0)D = Db(B_0 D)$ of the functional calculi of DB_0 and $B_0 D$, one has $D\widetilde{S}_A = S_A$, so that the relation $D\tilde{h}^+ = h^+$, which uniquely determines a choice $\tilde{h}^+ \in \mathcal{H}^+_{B_0 D}$, shows that $Dv = f$. Solutions u to the second order equations are related to v in the sense that there exists a constant $c \in \mathbb{C}^m$ such that

$u = c - v_\perp$. This means that the tangential part $-v_\parallel$ encodes a conjugate to the solution u. This notion of conjugate was further developed in [4].

These representations are justified provided one has the operator bounds in the next section.

4.2. **Functions spaces and operator estimates.** Here, we give the definition of the functions spaces associated to the BVPs. What changes compared to [3] is that the Lebesgue measure dx on $\mathbb{R}^n$ is replaced by dw and Lebesgue measure $d\mathbf{x} = dt\,dx$ on $\mathbb{R}^{1+n}$ by $d\underline{w} = dt\,dw$ where $\underline{w}$ is the A_2 weight on $\mathbb{R}^{1+n}$ defined by $\underline{w}(t, x) = w(x)$. The only property required for w in this section is the doubling property of dw, except when we use the quadratic estimate which uses the A_2 property. Also we incorporate the posterior duality and multiplier results of [29] (for dx), which can be extended to the weighted setting too. See below.

Definition 4.4. For an L^q_{loc} function f, $1 \leq q \leq \infty$, define $W_q f(t, x) = \left(\fint_{W(t,x)} |f|^q \, d\underline{w} \right)^{1/q}$ with the usual essential supremum definition if $q = \infty$ and where $W(t, x) := (c_0^{-1}t, c_0 t) \times B(x; c_1 t)$ is a Whitney region, for some fixed constants $c_0 > 1$, $c_1 > 0$. The *weighted non-tangential maximal function* of an L^2_{loc} function f in $\mathbb{R}^{1+n}_+$ is

$$\widetilde{N}_*(f)(x) := \sup_{t>0} W_2 f(t, x), \qquad x \in \mathbb{R}^n,$$

The *weighted Carleson functional* of an L^1_{loc} function f is

$$Cf(x) := \sup_{Q \ni x} \frac{1}{w(Q)} \int_{(0, \ell(Q)) \times Q} |f(t, y)| \, d\underline{w}(t, y), \qquad x \in \mathbb{R}^n,$$

where the supremum is taken over all cubes Q in $\mathbb{R}^n$ containing x, with $\ell(Q)$ denoting their sidelengths. The *modified weighted Carleson norm* of a measurable function g in $\mathbb{R}^{1+n}_+$ is

$$\|g\|_* := \|C(W_\infty(|g|^2/t))\|_\infty^{1/2}.$$

We will use the modified Carleson norm to measure the size of perturbations of t-independent coefficients A_0. The proof of Lemma 2.2 in [3] adapts to show that if there exists $A_0(x)$ with $\|w^{-1}(A - A_0)\|_* < \infty$, then it is unique and $w^{-1}A_0$ is bounded, and accretive on $\overline{\mathsf{R}(D)}$, so that we may call A_0 the trace of A.

Definition 4.5. Define the Banach/Hilbert spaces

$$\mathcal{X} := \{f \in L^2_{loc}(\mathbb{R}^{1+n}_+; \mathbb{C}^{m(1+n)}) : \|\widetilde{N}_*(f)\| < \infty\},$$

$$\mathcal{C} := \{f \in L^2_{loc}(\mathbb{R}^{1+n}_+; \mathbb{C}^{m(1+n)}) : \|C(W_2 f)\| < \infty\},$$

$$\mathcal{Y} := \{f \in L^2_{loc}(\mathbb{R}^{1+n}_+; \mathbb{C}^{m(1+n)}) : \int_0^\infty \|f_t\|^2 \, t\,dt < \infty\},$$

$$\mathcal{Y}^* := \{f \in L^2_{loc}(\mathbb{R}^{1+n}_+; \mathbb{C}^{m(1+n)}) : \int_0^\infty \|f_t\|^2 \, \frac{dt}{t} < \infty\},$$

with the obvious norms.

We use the same notation as in [3], but of course, here all norms are weighted. Note that $\mathcal{Y}^*$ is the dual space of $\mathcal{Y}$ with respect to the inner product $\langle \, , \, \rangle_{\underline{w}}$ of $\mathcal{H} = L^2(\mathbb{R}^{1+n}_+, d\underline{w}; \mathbb{C}^{m(1+n)})$.

Lemma 4.6. *There are estimates*

$$\sup_{t>0} \frac{1}{t} \int_t^{2t} \|f_s\|^2 \, ds \lesssim \|\widetilde{N}_*(f)\|^2 \lesssim \int_0^\infty \|f_s\|^2 \, \frac{ds}{s}.$$

In particular $\mathcal{Y}^* \subseteq \mathcal{X}$.

A fundamental quantity is the norm of multiplication operators mapping $\mathcal{X}$ into $\mathcal{Y}^*$ or $\mathcal{Y}$ into $\mathcal{C}$.

Lemma 4.7. *The dual of $\mathcal{X}$ with respect to the pairing $\langle \, , \, \rangle_{\underline{w}}$ is $\mathcal{C}$. For functions $\mathcal{E} : \mathbb{R}^{1+n}_+ \to \mathcal{L}(\mathbb{C}^{m(1+n)})$, we have estimates*

$$\|\mathcal{E}\|_\infty \lesssim \|\mathcal{E}\|_* \sim \sup_{\|f\|_{\mathcal{X}}=1} \|\mathcal{E}f\|_{\mathcal{Y}^*} \sim \sup_{\|f\|_{\mathcal{Y}}=1} \|\mathcal{E}f\|_{\mathcal{C}}.$$

Proof. When $w = 1$, the duality was established in [29] and recently another more direct proof was given in [27]. This second proof passes to the doubling measure setting (personal communication of Amenta and Huang). Next, the first inequality is proved in a similar way than in [3]. The equivalences for the pointwise multiplier operator norms were also established in [29], and reproved in [27] when $w = 1$, and the latter proof extends in a doubling measure context as well (personal communication of Amenta and Huang). $\square$

Proposition 4.8. *Let $u \in W^{1,2}_{loc}(\mathbb{R}^{1+n}_+, w)$ be such that $\|\widetilde{N}_*(\nabla_{t,x}u)\| < \infty$. Then there exists $u_0 \in \dot{H}^1(\mathbb{R}^n, w)$ (as defined in the proof of Lemma 2.8) such that $\|u_t - u_0\| \lesssim t$, $\|\nabla_x u_0\| \lesssim \|\widetilde{N}_*(\nabla_{t,x}u)\|$, and for dw almost every $x_0 \in \mathbb{R}^n$,*

$$(42) \qquad \fint_{W(t,x_0)} |u - u_0(x_0)|^2 \, d\underline{w} \le t^2 g(x_0).$$

with $g \in L^1(\mathbb{R}^n, w)$. Conversely, if $u_0 \in \dot{H}^1(\mathbb{R}^n, w)$, then $u = e^{t^2 \Delta_w} u_0$ satisfies $\|\widetilde{N}_*(\nabla_{t,x}u)\| \lesssim \|\nabla_x u_0\|$.

Proof. The first part is the weighted version of a result in [32]. First, it is easy to show $\|u_t - u_{t'}\| \lesssim |t - t'|$ by using $u_t - u_{t'} = \int_{t'}^t \partial_s u_s \, ds$, Cauchy-Schwarz inequality and the left hand inequality in Lemma 4.6. This gives the existence of $u_0 \in L^2_{\text{loc}}(\mathbb{R}^n, w)$ with $\|u_t - u_0\| \lesssim t$ (observe that only the difference is in $L^2(\mathbb{R}^n, w)$). Next, the Poincaré inequality and a telescopic sum argument implies

$$\left| \fint_{W(t,x_0)} u \, d\underline{w} - \fint_{W(t',x_0)} u \, d\underline{w} \right| \leq C\tau \widetilde{N}^1_*(\nabla_{t,x} u)(x_0)$$

whenever $t, t' \leq \tau$ up to using a non-tangential maximal function with appropriately large Whitney regions and $\widetilde{N}^1_*$ is the analogue of $\widetilde{N}_*$ with L^1-averages. Thus for every x_0 where $u_0(x_0)$ exists

$$\left| \fint_{W(t,x_0)} u \, d\underline{w} - u_0(x_0) \right| \lesssim t \widetilde{N}^1_*(\nabla_{t,x} u)(x_0)$$

and $\widetilde{N}^1_*(\nabla_{t,x} u) \in L^2(\mathbb{R}^n, w)$ by equivalences of norms if we change the Whitney regions. By the Poincaré inequality again,

$$\fint_{W(t,x_0)} \left| u - \fint_{W(t,x_0)} u \right|^2 d\underline{w} \leq C t^2 \widetilde{N}_*(\nabla_{t,x} u)(x_0)$$

and we deduce (42) on combining the last two inequalities. Finally, note that if x_0, y_0 are different points and $t = 10(c_0 + c_1)|x_0 - y_0|$, then

$$\left| \fint_{W(t,x_0)} u \, d\underline{w} - \fint_{W(t',y_0)} u \, d\underline{w} \right| \leq C|x_0 - y_0| \widetilde{N}^1_*(\nabla_{t,x} u)(x_0)$$

again with slightly larger Whitney regions in the definition of $\widetilde{N}^1_*$, so that combining with inequalities above, we obtain

$$|u_0(x_0) - u_0(y_0)| \leq C|x_0 - y_0|(\widetilde{N}^1_*(\nabla_{t,x} u)(x_0) + \widetilde{N}^1_*(\nabla_{t,x} u)(y_0)).$$

Using the theory of Sobolev spaces on the complete doubling metric-measure space $(\mathbb{R}^n, |\ |, w)$, it follows that $u_0 \in \dot{H}^1(\mathbb{R}^n, w)$ (identified with the Hajłash space), see [26].

The converse will be proved after Theorem 4.9. $\qquad\square$

At this stage, we do not know if $\nabla_{t,x} u$ has almost everywhere limits or even strong $L^2(w)$ limits in the above averaged sense (although weak $L^2(w)$ convergence can be shown as in [32]). This will be the case, however, when u is a solution of our systems. We remark that in comparison, the space defined by $\int_0^\infty \|\nabla_{t,x} u_t\|^2 \, t dt < \infty$ does not have a trace on $\mathbb{R}^n$.

With the above notation, we can state our main theorem for t-independent B_0, thus for semigroups only.

Theorem 4.9. *Let $T = DB_0$ or $B_0 D$. Then one has the estimate*

$$(43) \qquad \|e^{-t|T|} h\|_{\mathcal{X}} \sim \|h\| \sim \|\partial_t e^{-t|T|} h\|_{\mathcal{Y}}, \ \forall h \in \overline{R(T)}.$$

Furthermore, for any $h \in \mathcal{H}$ (not just $\overline{R(T)}$), we have that the Whitney averages of $e^{-t|T|}h$ converge to h in L^2 sense, that is for dw almost every $x_0 \in \mathbb{R}^n$,

$$\text{(44)} \qquad \lim_{t \to 0} \fint_{W(t,x_0)} |e^{-s|T|}h - h(x_0)|^2 \, d\underline{w} = 0.$$

In particular, this implies the dw almost everywhere convergence of Whitney averages

$$\text{(45)} \qquad \lim_{t \to 0} \fint_{W(t,x_0)} e^{-s|T|}h \, d\underline{w} = h(x_0).$$

More generally, one can replace $e^{-s|T|}$ by any $\varphi(sT)$ where φ is holomophic and bounded in some bisector containing $\sigma(T)$ and satisfies $|\varphi(z)| \lesssim |z|^{-\alpha}$ and $|\varphi(z)-a| \lesssim |z|^\alpha$ for some $\alpha > 0$, $a \in \mathbb{C}$. In this case, convergence is towards ah, and only the upper bound $\|\varphi(tT)h\|_{\mathcal{X}} \lesssim \|h\|$ holds if $a = 0$.

The proof of this theorem will be given in Section 6. The last equivalence in (43) is nothing but (32), we put it here for completeness.

If $B_0 = I$, then $T^2 = D^2 = \begin{bmatrix} -\Delta_w & 0 \\ 0 & -\nabla \mathrm{div}_w \end{bmatrix}$, so that

$$\nabla_x e^{t^2 \Delta_w} u_0 = -\left(D e^{-t^2 D^2} \begin{bmatrix} u_0 \\ 0 \end{bmatrix} \right)_{\parallel} = \left(e^{-t^2 D^2} \begin{bmatrix} 0 \\ \nabla_x u_0 \end{bmatrix} \right)_{\parallel}$$

and

$$\nabla_t e^{t^2 \Delta_w} u_0 = \left(2t D e^{-t^2 D^2} \begin{bmatrix} 0 \\ \nabla_x u_0 \end{bmatrix} \right)_{\perp}$$

Thus, $\|\widetilde{N}_*(\nabla_{t,x} e^{t^2 \Delta_w} u_0)\| \lesssim \|\nabla_x u_0\|$ follows from this result, proving the converse statement in Proposition 4.8.

We observe that only the weak type bound $\|\widetilde{N}_*(e^{-t|T|}h)\|_{L^{2,\infty}(w)} \lesssim \|h\|$ holds if $h \in \mathsf{N}(T)$. Concerning the convergence (44), this is new even when $w = 1$ for $T = DB_0$ in this generality. What was proved in [4] is (44) for $|B_0 e^{-t|DB_0|}h - (B_0 h)(x_0)|^2$ (which is also true in this situation), and the removal of B_0 was done only when B_0^{-1} is given by pointwise multiplication. It turns out this is not necessary. This will yield the almost everywhere limits in full generality in Theorem 4.13 as compared to [4].

Remark 4.10. the almost everywhere limit (45) is stated with respect to $d\underline{w}$, which is natural. However, as they are derived from the weighted $L^2(\underline{w})$ limits (44), using that $\underline{w} \in A_2$, we also have unweighted L^1 averages that converge to 0 almost everywhere. This means that (45) holds also with Lebesgue measure replacing $d\underline{w}$.

The next two theorems are for the t-dependent S_A and $\widetilde{S}_A$. Note that we may rewrite $S_A f = S(\mathcal{E}f) := S_{\mathcal{E}}f$ and $\widetilde{S}_A f = \widetilde{S}(\mathcal{E}f) := \widetilde{S}_{\mathcal{E}}f$ where $\mathcal{E}$ may not be related to A. We use this notation in what follows.

Theorem 4.11. *Assume that $\|\mathcal{E}\|_* < \infty$. Then we have the following estimates for arbitrary $f \in \mathcal{X}$.*

$$\text{(46)} \qquad \|S_\mathcal{E} f\|_\mathcal{X} \lesssim \|\mathcal{E}\|_* \|f\|_\mathcal{X}.$$

The function $h^- := -\int_0^\infty e^{sDB_0} E_0^- D\mathcal{E}_s f_s ds$ belongs to $E_0^- \mathcal{H} = \mathcal{H}_{DB_0}^-$ and

$$\text{(47)} \qquad \|h^-\| \lesssim \|\mathcal{E}\|_* \|f\|_\mathcal{X},$$

$$\text{(48)} \qquad \|S_\mathcal{E} f - e^{tDB_0} E_0^- h^-\|_{\mathcal{Y}^*} \lesssim \|\mathcal{E}\|_* \|f\|_\mathcal{X},$$

$$\text{(49)} \quad \lim_{t \to 0} t^{-1} \int_t^{2t} \|(S_\mathcal{E} f)_s - h^-\|^2 \, ds = 0 = \lim_{t \to \infty} t^{-1} \int_t^{2t} \|(S_\mathcal{E} f)_s\|^2 ds,$$

and

$$\text{(50)} \qquad \lim_{t \to 0} \fint_{W(t,x_0)} |S_\mathcal{E} f - h^-(x_0)|^2 \, d\underline{w} = 0, \text{ for a.e. } x_0 \in \mathbb{R}^n.$$

Moreover, $\tilde{h}^- := -\int_0^\infty e^{sB_0 D} \widetilde{E}_0^- \mathcal{E}_s f_s \, ds$ satisfies $D\tilde{h}^- = h^- \in E_0^- \mathcal{H}$,

$$\text{(51)} \qquad \|(\widetilde{S}_\mathcal{E} f)_t - \tilde{h}^-\| \lesssim t\|\mathcal{E}\|_* \|f\|_\mathcal{X}.$$

In addition, if $\|\mathcal{E}\|_*$ is sufficiently small and $\mathcal{E}$ satisfies the t-regularity condition $\|t\partial_t \mathcal{E}\|_* < \infty$, then

$$\text{(52)} \qquad \|\partial_t(S_\mathcal{E} f)\|_\mathcal{Y} \lesssim (\|\mathcal{E}\|_* + \|t\partial_t \mathcal{E}\|_*)\|f\|_\mathcal{X} + \|\mathcal{E}\|_\infty \|\partial_t f\|_\mathcal{Y},$$

and one has $t \mapsto (S_\mathcal{E} f)_t$ is continuous into $\mathcal{H}$ if $\|f\|_\mathcal{X} + \|\partial_t f\|_\mathcal{Y} < \infty$ with improved limits

$$\text{(53)} \qquad \lim_{t \to 0} \|(S_\mathcal{E} f)_t - h^-\| = 0 = \lim_{t \to \infty} \|(S_\mathcal{E} f)_t\|.$$

Theorem 4.12. *Assume that $\|\mathcal{E}\|_* < \infty$. Then we have the following estimates for arbitrary $f \in \mathcal{Y}$.*

$$\text{(54)} \qquad \|S_\mathcal{E} f\|_\mathcal{Y} \lesssim \|\mathcal{E}\|_* \|f\|_\mathcal{Y}.$$

The operator $\widetilde{S}_\mathcal{E}$ maps $\mathcal{Y}$ into $C([0,\infty); \mathcal{H})$ with

$$\text{(55)} \qquad \sup_{t \geq 0} \|(\widetilde{S}_\mathcal{E} f)_t\| \lesssim \|\mathcal{E}\|_* \|f\|_\mathcal{Y}.$$

Moreover, $\tilde{h}^- := -\int_0^\infty e^{sB_0 D} \widetilde{E}_0^- \mathcal{E}_s f_s \, ds \in \widetilde{E}_0^- \mathcal{H} = \mathcal{H}_{B_0 D}^-$,

$$\text{(56)} \qquad \lim_{t \to 0} \|(\widetilde{S}_\mathcal{E} f)_t - \tilde{h}^-\| = 0 = \lim_{t \to \infty} \|(\widetilde{S}_\mathcal{E} f)_t\|.$$

Furthermore, if $p < 2$ and $\widetilde{N}_*^p$ is the p-modified version of $\widetilde{N}_*$ by taking W_p functionals on Whitney regions,

$$\text{(57)} \qquad \|\widetilde{N}_*^p(\widetilde{S}_\mathcal{E} f)\| \lesssim \|\mathcal{E}\|_* \|f\|_\mathcal{Y}$$

and

$$\text{(58)} \qquad \fint_{W(t,x_0)} |\widetilde{S}_\mathcal{E} f - \tilde{h}^-(x_0)|^p \, d\underline{w} = 0, \text{ for a.e. } x_0 \in \mathbb{R}^n.$$

These two theorems can be proved following the corresponding results in Sections 6, 7, 8, 9 and 10 of [3] (some of the arguments were simplified in [34]) and, concerning the almost everywhere convergence limits (50) and (58), in Section 15 of [4] : they hold in a doubling weighted context when coming to use maximal functions and Carleson estimates. The inequality (51) is not proved in [3] and merely sketched in [4, Section 13], but is easy to prove following the same decompositions as there. We shall not give details. We just mention that $\tilde{h}^-$ in Theorem 4.11 is not an element of $\mathcal{H}$: it is only defined as a limit of the integrals truncated away from 0 and ∞ in the sense that $-\int_\varepsilon^R De^{sB_0D}\widetilde{E}_0^-\mathcal{E}_s f_s\,ds$ converges to h^- in $\mathcal{H}$. Thus, only the difference $(\widetilde{S}_\mathcal{E}f)_t - \tilde{h}^-$ in (51) makes sense in $\mathcal{H}$. The scalar part of $\tilde{h}^-$ belongs to the homogeneous Sobolev space $\dot{H}^1(\mathbb{R}^n, w; \mathbb{C}^m)$ (as defined in the proof of Lemma 2.8) and as such is also an $L_{loc}^2(w)$ function.

4.3. **A priori estimates.** In this subsection, we derive a priori estimates for solutions of $\mathrm{div}A\nabla u = 0$ with $\nabla u \in \mathcal{X}$ or $\mathcal{Y}$. Again, these are obtained as in [3], together with [4] for the almost everywhere statements and the improvements noticed in Theorem 4.9.

Theorem 4.13. *Consider coefficients $w^{-1}A \in L^\infty(\mathbb{R}_+^{1+n}; \mathcal{L}(\mathbb{C}^{m(1+n)}))$ such that $w^{-1}A$ is accretive on $\mathcal{H}^0$ and assume there exists t-independent measurable coefficients A_0 such that $\|w^{-1}(A - A_0)\|_* < \infty$ or equivalently that $\|\mathcal{E}\|_* \sim \|w^{-1}(A - A_0)\|_* < \infty$ where $\mathcal{E} = B_0 - B$ and $B = \widehat{w^{-1}A}$, $B_0 = \widehat{w^{-1}A_0}$.*

Let u be a weak solution of $\mathrm{div}A\nabla u = 0$ in $\mathbb{R}_+^{1+n}$ with $\|\widetilde{N}_*(\nabla_{t,x}u)\| < \infty$. Then

$$\lim_{t\to 0} t^{-1}\int_t^{2t} \|\nabla_{s,x}u_s - g_0\|^2 ds = 0 = \lim_{t\to\infty} t^{-1}\int_t^{2t} \|\nabla_{s,x}u_s\|^2 ds,$$

for some $g_0 \in L^2(\mathbb{R}^n, w; \mathbb{C}^{m(1+n)})$, with estimate $\|g_0\| \lesssim \|\widetilde{N}_*(\nabla_{t,x}u)\|$, which we call the gradient of u at the boundary and we set $\nabla_{t,x}u|_{t=0} := g_0$. Furthermore, one has that for dw almost every $x_0 \in \mathbb{R}^n$,

$$(59) \qquad \lim_{t\to 0} \fint_{W(t,x_0)} \nabla_{s,x}u\,d\underline{w} = g_0(x_0).$$

All three limits hold with $\nabla u, g_0$ replaced by the w-normalized conormal gradient $f = \nabla_{w^{-1}A}u$ and $f_0 = \begin{bmatrix} (w^{-1}A_0 g_0)_\perp \\ (g_0)_\parallel \end{bmatrix} := \nabla_{w^{-1}A}u|_{t=0}$ (in particular, they hold for the w-normalised conormal derivative $\partial_{\nu_{w^{-1}A}}u$). Moroever, one has the representation

$$(60) \qquad \nabla_{w^{-1}A}u = e^{-tDB_0}h^+ + S_A(\nabla_{w^{-1}A}u).$$

for a unique $h^+ \in \mathcal{H}_{DB_0}^+$ and

$$(61)\quad \nabla_{w^{-1}A}u|_{t=0} = h^+ + h^-, \qquad h^- = -\int_0^\infty e^{sDB_0}E_0^- D\mathcal{E}_s(\nabla_{w^{-1}A}u)_s ds.$$

Finally, there exists $u_0 \in \dot{H}^1(w)$ (as defined in Lemma 2.8) such that $\nabla_x u_0 = (g_0)_{\parallel}$ and one has $\|u_t - u_0\| \lesssim t$ and for dw almost every $x_0 \in \mathbb{R}^n$

$$(62) \qquad \lim_{t \to 0} \fint_{W(t,x_0)} u \, d\underline{w} = u_0(x_0).$$

Remark 4.10 about replacing $d\underline{w}$ by Lebesgue measure in the almost everywhere limit applies here too.

Theorem 4.14. *Consider coefficients $w^{-1}A \in L^\infty(\mathbb{R}^{1+n}_+; \mathcal{L}(\mathbb{C}^{m(1+n)}))$ such that $w^{-1}A$ is accretive on $\mathcal{H}^0$ and assume there exists t-independent measurable coefficients A_0 such that $\|w^{-1}(A - A_0)\|_* < \infty$, or equivalently that $\|\mathcal{E}\|_* \sim \|w^{-1}(A - A_0)\|_* < \infty$ where $\mathcal{E} = B_0 - B$ and $B = \widehat{w^{-1}A}, B_0 = \widehat{w^{-1}A_0}$.*

Let u be a weak solution of $\mathrm{div}A\nabla u = 0$ in $\mathbb{R}^{1+n}_+$ and assume that $\int_0^\infty \|\nabla_{t,x}u\|^2 \, t\,dt < \infty$. Then $u = \hat{u} + c$ almost everywhere, for a unique constant $c \in \mathbb{C}^m$ and $\hat{u} \in C([0, \infty); L^2(\mathbb{R}^n, w; \mathbb{C}^m))$ given by $\hat{u} = -v_\perp$ with

$$(63) \qquad v = e^{-tB_0 D}\tilde{h}^+ + \widetilde{S}_A(\nabla_{w^{-1}A}u),$$

for a unique $\tilde{h}^+ \in \widetilde{E}_0^+\mathcal{H}$. Moreover,

$$(64) \qquad v_0 = \tilde{h}^+ + \tilde{h}^- \text{ with } \tilde{h}^- = -\int_0^\infty e^{sB_0 D}\widetilde{E}_0^-\mathcal{E}_s(\nabla_{w^{-1}A}u)_s \, ds,$$

and we call $-v$ the conjugate system associated to u. In addition, we have $Dv = \nabla_{w^{-1}A}u$.

Identifying the functions u and $\hat{u} + c$, we have limits

$$\lim_{t \to 0} \|u_t - \hat{u}_0 - c\| = 0 = \lim_{t \to \infty} \|u_t - c\|,$$

for $\hat{u}_0 = -(\tilde{h}^+)_\perp \in L^2(\mathbb{R}^n, w; \mathbb{C}^m)$, and we have the estimates

$$\|\hat{u}_0\| \lesssim \max(\|\widetilde{N}_*(\hat{u})\|, \sup_{t>0} \|\hat{u}_t\|) \lesssim \left(\int_0^\infty \|\nabla_{t,x}u\|^2 \, t\,dt\right)^{1/2}.$$

Finally, for dw almost every $x_0 \in \mathbb{R}^n$,

$$(65) \qquad \lim_{t \to 0} \fint_{W(t,x_0)} u \, d\underline{w} = u_0(x_0).$$

Remark 4.10 about replacing $d\underline{w}$ by Lebesgue measure in the almost everywhere limit applies here too.

The representation formula suggests a possible construction of solution given $\tilde{h}^+$ provided $\|\mathcal{E}\|_*$ is sufficiently small. This is what leads to well-posedness results.

Remark 4.15. We also have the representation (60) with both $e^{-t|DB_0|}$ and h^+ interpreted in a suitable sense with Sobolev spaces of order $s = -1$. This point of view is developed more systematically in [34]

when $w = 1$ and with Sobolev regularity $-1 \leq s \leq 0$. We refer the reader there to make the straightforward adaptation, as it is again an abstract argument. We just warn the reader that the $\mathcal{E}$ in [34] is not exactly the same as ours because the author assumed pointwise accretivity to simplify matters. One should use representation (39) for S_A instead.

Corollary 4.16. *Assume that A satisfies $\|w^{-1}(A - A_0)\|_* < \infty$ for some t-independent A_0 and is such that all weak solutions u to the system $\mathrm{div} A \nabla u = 0$ in a ball $B \subseteq \mathbb{R}_+^{1+n}$ satisfy the local boundedness property*

$$\sup_{\alpha B} |u| \leq C \left(\fint_{\beta B} |u|^2 d\underline{w} \right)^{1/2},$$

for any fixed constant $\alpha < \beta < 1$, with C independent of u and B. Then any weak solution u with $\int_0^\infty \|\nabla_{t,x} u\|^2 \, t dt < \infty$ or $\|\widetilde{N}_*(\nabla_{t,x} u)\| < \infty$ converges non-tangentially almost everywhere to its boundary trace.

The proof is a straightforward consequence of the more precise almost everywhere convergences we stated in the previous section. We skip the details. This result applies in particular to real equations as a consequence of [21].

5. WELL-POSEDNESS

5.1. **Formulation and general results.**

Definition 5.1. Fix $w \in A_2(\mathbb{R}^n)$. Consider degenerate coefficients A with $w^{-1}A \in L^\infty(\mathbb{R}_+^{1+n}; \mathcal{L}(\mathbb{C}^{m(1+n)}))$ such that $w^{-1}A$ is accretive on $\mathcal{H}^0$.

- By the Dirichlet problem with coefficients A being well-posed, we mean that given $\varphi \in L^2(\mathbb{R}^n, w; \mathbb{C}^m)$, there is a unique weak solution u solving (1), with $\int_0^\infty \|\nabla_{t,x} u\|^2 \, t dt < \infty$ and trace $u_0 = \varphi$.
- By the regularity problem with coefficients A being well-posed, we mean that given $\varphi \in L^2(\mathbb{R}^n, w; \mathbb{C}^{mn})$, where φ satisfies $\mathrm{curl}_x \varphi = 0$, there is a weak solution u, unique modulo constants, solving (1), with $\|\widetilde{N}_*(\nabla_{t,x} u)\| < \infty$ and such that $\nabla_x u|_{t=0} = \varphi$.
- By the Neumann problem with coefficients A being well-posed, we mean that given $\varphi \in L^2(\mathbb{R}^n, w; \mathbb{C}^m)$, there is a weak solution u, unique modulo constants, solving (1), with $\|\widetilde{N}_*(\nabla_{t,x} u)\| < \infty$ and such that $\partial_{\nu_{w^{-1}A}} u|_{t=0} = \varphi$.

We write $A \in \mathrm{WP(BVP)}$, if the corresponding boundary value problem (BVP) is well-posed with coefficients A.

We remark that the definition does not include almost everywhere requirements. For the regularity and Neumann problems, one can make sense of the trace in a weak sense, but for the Dirichlet problem,

the trace may not even make sense. However, as soon as we assume $\|w^{-1}(A - A_0)\|_* < \infty$, which will be the case here, we know exactly the meaning of the trace from the results in Section 4.3.

The most important observation following the *a priori* estimates in Theorems 4.13 and 4.14 is the fact that in the t-independent coefficient case, we completely identify the trace spaces: $\mathcal{H}_{DB}^+$ is the trace space of w-normalized conormal gradients for solutions with $\|\widetilde{N}_*(\nabla_{t,x}u)\| < \infty$; $\mathcal{H}_{BD}^+$ is the trace space of conjugate systems v for solutions with $\int_0^\infty \|\nabla_{t,x}u\|^2 \, t\,dt < \infty$. In each case this is an isomorphism.

This leads to the following characterisation of well-posedness.

Theorem 5.2. *Consider coefficients $w^{-1}A \in L^\infty(\mathbb{R}_+^{1+n}; \mathcal{L}(\mathbb{C}^{m(1+n)}))$ such that $w^{-1}A$ is accretive on $\mathcal{H}_0$. Assume that A has t-independent coefficients. Let $B = \widehat{w^{-1}A}$. Then $A \in WP(Reg)/WP(Neu)/WP(Dir)$ if and only if*

$$\mathcal{H}_{DB}^+ \longrightarrow \{g \in L^2(\mathbb{R}^n, w; \mathbb{C}^{mn}) \colon \operatorname{curl}_x g = 0\} : f \longmapsto f_\parallel,$$
$$\mathcal{H}_{DB}^+ \longrightarrow L^2(\mathbb{R}^n, w; \mathbb{C}^m) : f \longmapsto f_\perp,$$
$$\mathcal{H}_{BD}^+ \longrightarrow L^2(\mathbb{R}^n, w; \mathbb{C}^m) : f \longmapsto f_\perp,$$

are isomorphisms respectively.

Observe the change of space in the third line.

Let us mention a connection to so-called Rellich estimates. The isomorphisms imply the Rellich estimates

$$\|f_\perp\| \lesssim \|f_\parallel\|, \quad \forall f \in \mathcal{H}_{DB}^+,$$
$$\|f_\parallel\| \lesssim \|f_\perp\|, \quad \forall f \in \mathcal{H}_{DB}^+,$$
$$\|f_\parallel\| \lesssim \|f_\perp\|, \quad \forall f \in \mathcal{H}_{BD}^+,$$

respectively. Assuming the Rellich estimates is not enough to conclude well-posedness because this only gives injectivity with closed range. The surjectivity usually follows from a continuity argument starting with a situation where one knows surjectivity. Thus, if, in a connected component of validity of a Rellich estimate, there is one B for which surjectivity holds, then surjectivity holds for all B in this connected component and the corresponding BVP is well-posed for all B in this connected component. Usually one considers $B = I$ so that, here, $A = wI$. See [5] for a discussion which applies *in extenso*. We remark that this depends on the continuous dependence on $B \in L^\infty$ of the projections E_0^+ and $\widetilde{E}_0^+$, which follows from Theorem 3.3. Let us mention also the duality principle between Dirichlet and Regularity, whose proof is the same as that of Proposition 17.6 in [4].

Theorem 5.3. *Let $w^{-1}A \in L^\infty(\mathbb{R}_+^{1+n}; \mathcal{L}(\mathbb{C}^{m(1+n)}))$ such that $w^{-1}A$ is accretive on $\mathcal{H}^0$. Assume there exists t-independent measurable coefficients A_0 such that $\|w^{-1}(A - A_0)\|_* < \varepsilon$. If ε is small enough, then $A \in WP(Dir)$ if and only if $A^* \in WP(Reg)$.*

We now turn to perturbation results for both t-dependent and t-independent coefficients. Adapting [2, 5], see especially Lemma 4.3 in [5], one obtains that each WP(BVP) is open under perturbation of t-independent coefficients in wL^∞. We refer to [3] for the proofs of the t-dependent perturbations, which carry over without change to our setting. We gather these observations together in the following statement.

Theorem 5.4. *Assume the Neumann problem with t-independent A_0 is well-posed. Then there exist $\varepsilon_0 > 0$ and $\varepsilon_1 > 0$ such that if $\|w^{-1}(A - A_1)\|_* < \varepsilon_1$ and A_1 has t-independent coefficients with $\|w^{-1}(A_1 - A_0)\|_\infty < \varepsilon_0$, then the Neumann problem with coefficients A is well-posed.*

The corresponding result holds when the Neumann problem is replaced by the regularity problem.

Moreover, for all such A, the solutions u of the BVP satisfy

$$\|\widetilde{N}_*(\nabla_{t,x} u)\| \approx \|g_0\| \approx \|\varphi\|,$$

with φ the w-normalized Neumann data or the regularity data, and one has the limits and regularity estimates as described in Theorem 4.13.

With the duality principle above, we obtain.

Theorem 5.5. *Assume the Dirichlet problem with t-independent A_0 is well-posed. Then there exist $\varepsilon_0 > 0$ and $\varepsilon_1 > 0$ such that if $\|w^{-1}(A - A_1)\|_* < \varepsilon_1$ and A_1 has t-independent coefficients with $\|w^{-1}(A_1 - A_0)\|_\infty < \varepsilon_0$, then the Dirichlet problem with coefficients A is well-posed. Moreover, one has*

$$\|\widetilde{N}_*(u)\| \approx \sup_{t>0} \|u_t\| \approx \left(\int_0^\infty \|\nabla_{t,x} u\|^2 \, t dt \right)^{1/2} \approx \|\varphi\|,$$

if φ is the Dirichlet data, and one has the limits and regularity estimates described in Theorem 4.14.

5.2. **Well-posedness for t-independent hermitian coefficients.**

Proposition 5.6. *Assume that $A = A^*$ and that A is t-independent and satisfies the usual degenerate boundedness and accretivity on $\mathcal{H}^0 = \overline{R(D)}$ conditions. Then the regularity, Neumann and Dirichlet problems with coefficients A are well-posed.*

Proof. Let $B = \widehat{w^{-1}A}$ and $f \in E_0^+ \mathcal{H} = \mathcal{H}_{DB}^+$. Theorem 4.13 in the case of t-independent coefficients implies that the vector field $F_t = e^{-tDB} f$ in $\mathbb{R}_+^{1+n}$ is such that $\partial_t F_t = -DB F_t$, $\lim_{t\to\infty} F_t = 0$ and $\lim_{t\to 0} F_t = f$. Let $N := \begin{bmatrix} -I & 0 \\ 0 & I \end{bmatrix}$ and note that $DN + ND = 0$. Now, the definition of $B = \widehat{w^{-1}A}$ and the Hermitian condition $A^* = A$ imply $B^*N = NB$.

Using the hermitian inner product $(\ ,\ )$ for dw, we have

$$\partial_t(NF_t, BF_t) = (NDBF_t, BF_t) + (NF_t, BDBF_t)$$
$$= (NDBF_t, BF_t) + (DB^*NF_t, BF_t) = ((ND + DN)BF_t, BF_t) = 0.$$

Hence, integrating in t and taking into account the limit at ∞ gives us $(Nf, Bf) = 0$. Thus, separating scalar and tangential parts, we obtain the Rellich equality:

$$(66) \qquad (f, Bf) = 2(f_\perp, (Bf)_\perp) = 2(f_\|, (Bf)_\|).$$

Consider first the Neumann problem. It follows from (66) and the accretivity of B on $\mathcal{H}^0$ that

$$\kappa\|f\|^2 \le \operatorname{Re}(f, Bf) = 2\operatorname{Re}(f_\perp, (Bf)_\perp) \lesssim 2\|B\|_\infty\|f_\perp\|\|f\|.$$

This shows that $\|f\| \lesssim \|f_\perp\|$ for the Neumann map for any hermitian A, which implies that this map is injective with closed range. The continuity argument explained above implies that $A \in \mathrm{WP(Neu)}$ provided that $I \in \mathrm{WP(Neu)}$. That $I \in \mathrm{WP(Neu)}$ can be seen from the equality $\|\nabla(-\Delta_w)^{-1/2}u\| = \|u\|$ and

$$\mathrm{sgn}(D) = \begin{bmatrix} 0 & (-\Delta_w)^{-1/2}\mathrm{div}_w \\ -\nabla(-\Delta_w)^{-1/2} & 0 \end{bmatrix}.$$

Thus, for $f \in \mathcal{H}^0$, $f \in \mathcal{H}_D^+$ if and only if $f_\| = -\nabla(-\Delta_w)^{-1/2}f_\perp$, which in turn holds if and only if $f_\perp = (-\Delta_w)^{-1/2}\mathrm{div}_w f_\|$. This implies that the map used for solving the Neumann problem are invertible.

That $A \in \mathrm{WP(Reg)}$ is proved in the similar way. Then, by Theorem 5.3, it follows that $A = A^* \in \mathrm{WP(Dir)}$. $\qquad\square$

5.3. Well-posedness with algebraic structure and t-independent coefficients.

Recall that we write our coefficients A as a 2×2 block matrix. We say that it is *block lower-triangular* if the upper off-diagonal block $A_{\perp\|}$ is 0, and *block upper-triangular* if the lower block $A_{\|\perp}$ is 0.

Theorem 5.7. *We assume that A is t-independent and satisfies the usual degenerate boundedness and accretivity conditions.*

- *The Neumann problem with block lower-triangular coefficients A is well-posed.*
- *The regularity problem with block upper-triangular coefficients A is well-posed. More generally, it suffices for the off-diagonal lower block of A to be divergence free and have real entries.*
- *The Dirichlet problem with block lower-triangular coefficients A is well-posed. More generally, it suffices for the off-diagonal upper block to be divergence free and have real entries.*

Let us clarify the statements above. The off-diagonal lower block is $A_{\|\perp} = (A_{i,0}^{\alpha,\beta})_{i=1,\dots,n}^{\alpha,\beta=1,\dots,m}$. Real entries means that all these coefficients

are real: it guarantees that A and

$$A' = A - \begin{bmatrix} 0 & -A_{\|\perp}^t \\ A_{\|\perp} & 0 \end{bmatrix} = \begin{bmatrix} A_{\perp\perp} & A_{\perp\|} + A_{\|\perp}^t \\ 0 & A_{\|\|} \end{bmatrix}$$

have the same accretivity bounds. The divergence free condition is $\sum_{i=1}^n \partial_i A_{i,0}^{\alpha,\beta} = 0$ for all α, β. It implies that weak solutions with coefficients A or A' are the same as can be seen by integrating by parts. In other words, we can reduce matters the special case where the off-diagonal lower block is zero This possibility does not appear to be available for the Neumann problem because the conormal derivative depends on the coefficients.

The proof of this theorem is obtained by a line by line adaptation of [8] to the weighted setting using well-posedness of the three problems (modulo constants) in the class of energy solutions, that is, having finite energy $\int_{\mathbb{R}_+^{1+n}} |\nabla u|^2 d\underline{w} < \infty$. We mention that to carry out the algebra there, one should replace the standard Riesz transforms by the Riesz transforms $\mathcal{R}_w$ defined in Lemma 2.8. We leave details to the interested reader.

6. NON-TANGENTIAL MAXIMAL ESTIMATES AND FATOU TYPE RESULTS

Recall that $\underline{w}$ is the A_2 weight on $\mathbb{R}^{1+n}$ defined by $\underline{w}(t, x) = w(x)$ and that $\underline{w}$ and w have identical A_2 constants. Writing equations $\mathrm{div}A\nabla u = 0$ as $\mathrm{div}_{\underline{w}}(w^{-1}A\nabla u) = 0$ allows one to carry some proofs to the degenerate case without much change from the non-degenerate case. We quote two results we will be using. The first is the usual Caccioppoli inequality with a completely analogous proof: all weak solutions u in a ball $B = B(\mathbf{x}, r) \subseteq \mathbb{R}_+^{1+n}$ of $\mathrm{div}A\nabla u = 0$ enjoy the Caccioppoli inequality

$$(67) \qquad \int_{\alpha B} |\nabla u|^2 \, d\underline{w} \leq Cr^{-2} \int_{\beta B} |u|^2 \, d\underline{w}$$

for any $0 < \alpha < \beta < 1$, the implicit constants depending only on $\|w^{-1}A\|_\infty$ the accretivity constant of $w^{-1}A$, n, m, α and β.

The second one is a corollary of this and Poincaré inequalities: there exists $1 < p_1 < 2$ such that for any $p_1 < p < 2$, all weak solutions u in a ball $B = B(\mathbf{x}, r) \subseteq \mathbb{R}_+^{1+n}$ of $\mathrm{div}A\nabla u = 0$ enjoy the reverse Hölder inequality

$$(68) \qquad \left(\fint_{\alpha B} |\nabla u|^2 \, d\underline{w} \right)^{1/2} \lesssim \left(\fint_{\beta B} |\nabla u|^p \, d\underline{w} \right)^{1/p}$$

for any $0 < \alpha < \beta < 1$, the implicit constants depending only on $\|w^{-1}A\|_\infty$ the accretivity constant of $w^{-1}A$, n, m, p, α, β and the A_2 constant of w. The usual proof follows from Caccioppoli's inequality applied to $u - u_{\beta B}$ and using the Poincaré inequality with the gradient

in L^p. Here, to be able to do that we need $\underline{w} \in A_{2-\varepsilon}$ for $\varepsilon > 0$ but this can be done for some $\varepsilon > 0$ depending only in the size of the A_2 constant of $\underline{w}$. Then, one can use [21, Theorem 1.5] which asserts that the gain of exponent in the Poincaré inequality for an A_p weight on $\mathbb{R}^{1+n}$ is at least $p\frac{1+n}{n}$. So for $p > \frac{2n}{1+n}$ (and $p > 2 - \varepsilon$) we are done.

We continue with the analogue of Lemma 10.3 in [3].

Lemma 6.1. *Let B be t-independent, bounded on $\mathcal{H}$ and accretive on $\mathcal{H}^0 = \overline{R(D)}$. Let $T = DB$ or BD. Then there exists $1 < p_0 < 2$, such that for $p_0 < q < p_0'$,*

$$\|(I + itT)^{-1}f\|_q \leq C_N \left(1 + \frac{\operatorname{dist}(E,F)}{t}\right)^{-N} \|f\|_q$$

for all integer N, $t > 0$ and sets $E, F \subseteq \mathbb{R}^n$ such that $\operatorname{supp} f \subseteq F$, with C_N independent of f, t, E, F. Here $\operatorname{dist}(E,F) := \inf\{|x - y| : x \in E, y \in F\}$ and $\| \ \|_q$ are the weighted L^q norms.

Proof. It suffices to prove the lemma for $T = DB$ as then it holds for $T^* = B^*D$, and hence for BD upon changing B^* to B.

For $q = 2$, this is contained in Lemma 3.4. By interpolation, it suffices to estimate the operator norm of $(I+itDB)^{-1}$ on $L^q(\mathbb{R}^n, w; \mathbb{C}^{m(1+n)})$, uniformly for t.

To this end, assume that $(I + itDB)\tilde{f} = f$. As in Proposition 4.2, but replacing ∂_t by $(it)^{-1}$, this equation is equivalent to

$$\begin{cases} (w^{-1}A\tilde{g})_\perp + it\operatorname{div}_w(w^{-1}A\tilde{g})_\parallel = (w^{-1}Ag)_\perp, \\ \tilde{g}_\parallel - it\nabla_x\tilde{g}_\perp = g_\parallel, \end{cases}$$

where A, g and $\tilde{g}$ are related to B, f and $\tilde{f}$ respectively, as in Proposition 4.2. Using the second equation to eliminate $\tilde{g}_\parallel$ in the first, shows that $\tilde{g}_\perp$ satisfies the divergence form equation

$$L_t\tilde{g}_\perp := \begin{bmatrix} 1 & it\operatorname{div}_w \end{bmatrix} (w^{-1}A) \begin{bmatrix} 1 \\ it\nabla_x \end{bmatrix} \tilde{g}_\perp = \begin{bmatrix} 1 & it\operatorname{div}_w \end{bmatrix} \begin{bmatrix} w^{-1}A_{\perp\perp}g_\perp \\ -w^{-1}A_{\parallel\parallel}g_\parallel \end{bmatrix}.$$

Let $r(w) < 2$ be the infimum of those exponents q for which $w \in A_q$. For $r(w) < q < r(w)'$, L_t is bounded from $W^{1,q}(\mathbb{R}^n, w; \mathbb{C}^m)$ equipped with the norm $\|u\|_q + t\|\nabla u\|_q$ into $W^{-1,q}(\mathbb{R}^n, w; \mathbb{C}^m) = (W^{1,q'}(\mathbb{R}^n, w; \mathbb{C}^m))'$ for the duality $\langle \ , \ \rangle_w$, uniformly in t. (Observe that the A_2 constant is unchanged by scaling, so we can change variable to reduce to $t = 1$ up to changing $w(x)$ to $w(tx)$.) The accretivity condition on $\mathcal{H}^0$ tells us that

$$\operatorname{Re}\langle L_t u, u\rangle_w \geq \kappa(\|u\|_2^2 + t^2\|\nabla u\|_2^2).$$

Hence L_t is an isomorphism for $p = 2$. Using Proposition 2.7 and the stability result of Šneĭberg [35], it follows that L_t is an isomorphism for $p_0 < q < p_0'$ for some $r(w) < p_0 < 2$. This gives us the desired estimate

in the weighted $L^q(w)$ norms,

$$\|\tilde{f}\|_q \approx \|\tilde{g}\|_q \lesssim \|\tilde{g}_\perp\|_q + t\|\nabla_x \tilde{g}_\perp\|_q + \|g_\parallel\|_q \lesssim \|g\|_q \approx \|f\|_q.$$

$\square$

We state a simple but useful lemma.

Lemma 6.2. *Let f be a function satisfying the square function estimate $\int_0^\infty \|f_t\|^2 \frac{dt}{t} < \infty$. Then the Whitney averages of f converge dw almost everywhere in L^2 sense to 0.*

Proof. Recall that c_0, c_1 are the parameters for the Whitney box $W(t, x)$. Let

$$h(t, x) := \int_{|x-y| \le c_1 c_0 s,\, s < c_0 t} |f(s, y)|^2 \frac{d\underline{w}(s, y)}{sw(B(y, c_1 s))}.$$

It is an increasing function of t with

$$\int_{\mathbb{R}^n} h(t, x)\, dw(x) \lesssim \int_0^{c_0 t} \|f_s\|^2 \frac{ds}{s} \to 0$$

as $t \to 0$. Thus $h(t, x)$ converges to 0 as $t \to 0$ for dw almost every x. We conclude observing that $W_2 f(t, x)^2 \lesssim h(t, x)$. $\square$

We now turn to the proof of Theorem 4.9.

Proof of (43). Here it remains to show $\|\widetilde{N}_*(e^{-t|T|}h)\| \sim \|h\|$ whenever $h \in \overline{\mathsf{R}(T)}$ and $T = DB$ or BD with B being t-independent (we drop the subscript 0). The bound from below is easy:

$$\|h\|^2 = \lim_{t \to 0} \frac{1}{t} \int_t^{2t} \|e^{-s|T|}h\|^2\, ds \lesssim \|\widetilde{N}_*(e^{-s|T|}h)\|^2$$

from Lemma 4.6.

Let us see the converse when $T = DB$. Split $h \in \overline{\mathsf{R}(DB)} = \mathcal{H}^0$ as $h = h^+ + h^-$ according to $h^\pm = \chi^\pm(DB)h$. Then $e^{-t|DB|}h^\pm$ is the w-normalized conormal gradient of a weak solution $u^\pm$ in the upper-(resp. lower-)half space. Thus one can use (68) and estimate $\widetilde{N}_*(e^{-t|DB|}h^\pm)$ by $\widetilde{N}_*^p(e^{-t|DB|}h^\pm)$ for some $p < 2$ which we can take larger than p_0 of the Lemma 6.1. From here follow the proof of Lemma 2.52 in [2]: decompose $e^{-t|DB|} = \psi(tDB) + (I + itDB)^{-1}$ and use the quadratic estimate (32) and Lemma 4.6 to obtain

$$\|\widetilde{N}_*^p(\psi(tDB)h^\pm)\|^2 \lesssim \int_0^\infty \|\psi(tDB)h^\pm\|^2 \frac{dt}{t} \lesssim \|h^\pm\|^2$$

and Lemma 6.1 to obtain the pointwise bound

$$\widetilde{N}_*^p((I + itDB)^{-1}h^\pm) \le C M_w(|h^\pm|^p)^{1/p},$$

where M_w is the Hardy-Littlewood maximal operator with respect to dw.

The result for $T = BD$ follows from that of DB: If $g \in \overline{\mathsf{R}(BD)}$, then $B^{-1}g = h \in \overline{\mathsf{R}(DB)}$ with $\|h\| \sim \|g\|$ and $e^{-t|BD|}g = Be^{-t|DB|}h$. Thus

$$\|\widetilde{N}_*(e^{-t|BD|}g)\| = \|\widetilde{N}_*(Be^{-t|DB|}h)\| \leq \|B\|_\infty \|\widetilde{N}_*(e^{-t|DB|}h)\| \sim \|h\|.$$

$\square$

Proof of (44). This time we begin with $T = BD$. We can split $h = h_N + h_R$ where $h_N \in \mathsf{N}(BD)$ and $h_R \in \overline{\mathsf{R}(BD)}$. As $e^{-s|BD|}h_N = h_N$, the almost everywhere limit $\lim_{t \to 0} \fint_{W(t,x_0)} |e^{-s|BD|}h_N - h_N(x_0)|^2 \, d\underline{w} = 0$ follows from the Lebesgue differentiation theorem. We can thus assume that $h \in \overline{\mathsf{R}(BD)}$.

Pick a Lebesgue point x_0 for the conditions

$$(69) \quad \lim_{t \to 0} \fint_{B(x_0,t)} |h - h(x_0)|^2 \, dw = 0 = \lim_{t \to 0} \fint_{B(x_0,t)} |w^{-1} - w^{-1}(x_0)|^2 \, dw.$$

The second equality is possible since $w^{-1} \in L^2_{\mathrm{loc}}(w)$ as $w \in A_2$. As a consequence of (69) and (8) with $p = 2$, we have

$$(70) \quad \lim_{t \to 0} \fint_{B(x_0,t)} w \, dx = w(x_0), \qquad \lim_{t \to 0} \fint_{B(x_0,t)} w^{-1} \, dx = w^{-1}(x_0)$$

and

$$\lim_{t \to 0} \fint_{B(x_0,t)} |h - (hw)(x_0)w^{-1}|^2 \, dw = 0.$$

Write as above, $e^{-s|BD|}h = \psi(sBD)h + (I + isBD)^{-1}h$. The quadratic estimate (32) and Lemma 6.2 imply that

$$\lim_{t \to 0} \fint_{W(t,x_0)} |\psi(sBD)h|^2 \, d\underline{w} = 0$$

for dw almost every $x_0 \in \mathbb{R}^n$. Now the key point is that $D \begin{bmatrix} c_\perp \\ c_\parallel w^{-1} \end{bmatrix} = 0$ if c is a constant, thus $(I + isBD)^{-1} \begin{bmatrix} h_\perp(x_0) \\ (h_\parallel w)(x_0)w^{-1} \end{bmatrix} = \begin{bmatrix} h_\perp(x_0) \\ (h_\parallel w)(x_0)w^{-1} \end{bmatrix}$. It follows that

$$(I + isBD)^{-1}h - h(x_0)$$
$$= (I + isBD)^{-1}h - \begin{bmatrix} h_\perp(x_0) \\ (h_\parallel w)(x_0)w^{-1} \end{bmatrix} + \begin{bmatrix} 0 \\ (h_\parallel w)(x_0)w^{-1} - h_\parallel(x_0) \end{bmatrix}$$
$$= (I + isBD)^{-1} \begin{bmatrix} h_\perp - h_\perp(x_0) \\ h_\parallel - (h_\parallel w)(x_0)w^{-1} \end{bmatrix} + (h_\parallel w)(x_0) \begin{bmatrix} 0 \\ w^{-1} - w^{-1}(x_0) \end{bmatrix}.$$

By the assumption on x_0,

$$\lim_{t \to 0} \fint_{W(t,x_0)} \left| (h_\parallel w)(x_0) \begin{bmatrix} 0 \\ w^{-1} - w^{-1}(x_0) \end{bmatrix} \right|^2 \, d\underline{w} = 0.$$

Decomposing the inner function of the other term using annuli centred around $B(x_0, t)$ and using Lemma 3.4, we have that

$$\fint_{W(t,x_0)} \left| (I + isBD)^{-1} \begin{bmatrix} h_\perp - h_\perp(x_0) \\ h_\parallel - (h_\parallel w)(x_0) w^{-1} \end{bmatrix} \right|^2 d\underline{w}$$

is bounded by

$$(71) \quad \sum_{j \geq 1} 2^{-jN} (w(B(x_0, t)))^{-1} \int_{B(x_0, 2^j t)} \left| \begin{bmatrix} h_\perp - h_\perp(x_0) \\ h_\parallel - (h_\parallel w)(x_0) w^{-1} \end{bmatrix} \right|^2 dw.$$

The scalar term is bounded by

$$\sum_{j \geq 1} 2^{-j(N-d_w)} \fint_{B(x_0, 2^j t)} |h_\perp - h_\perp(x_0)|^2 \, dw$$

$$\lesssim \sup_{\tau \leq \sqrt{t}} \fint_{B(x_0, \tau)} |h_\perp - h_\perp(x_0)|^2 \, dw + \sqrt{t} M_w(|h|^2)(x_0)$$

as can be seen using the doubling condition on w (d_w being its homogeneous dimension) and breaking the sum up at j_0 with $2^{j_0} \sim 1/\sqrt{t}$ if $N \geq d_w + 1$, where M_w is the Hardy-Littlewood maximal operator with respect to dw. The weak type $(1,1)$ estimate of M_w implies that $M_w(|h|^2)(x_0) < \infty$ for almost every $x_0 \in \mathbb{R}^n$. Hence, the scalar term goes to 0 as $t \to 0$ at those x_0 which meet all these requirements. The tangential term in (71) is bounded by

$$\sum_{j \geq 1} 2^{-j(N-d_w-1)} \fint_{B(x_0, 2^j t)} |h_\parallel - h_\parallel(x_0)|^2 \, dw$$

$$+ \sum_{j \geq 1} 2^{-jN} |h_\parallel(x_0)|^2 (w(B(x_0, t)))^{-1} \int_{B(x_0, 2^j t)} |w^{-1}(x_0) - w^{-1}|^2 \, dw.$$

The first sum can be treated as above. For the second when $j \leq j_0$, we also do as above. For $j \geq j_0$, we write $|w^{-1}(x_0) - w^{-1}|^2 \leq 2w^{-2}(x_0) + 2w^{-2}$. The first term leads to a bound $\sqrt{t} |h_\parallel(x_0)|^2 w^{-2}(x_0)$ if $N \geq d_w + 1$. For the second term, observe that

$$\int_{B(x_0, 2^j t)} w^{-2} dw = w^{-1}(B(x_0, 2^j t)) \lesssim 2^{jd_{w^{-1}}} w^{-1}(B(x_0, t))$$

and that by (70)

$$(w(B(x_0, t)))^{-1} w^{-1}(B(x_0, t)) \to 1/w^2(x_0), \quad t \to 0.$$

Thus, if also $N \geq d_{w^{-1}} + 1$, then the tangential term in (71) tends to 0.

We now turn to the proof for $T = DB$. If $g \in \mathsf{N}(DB)$, (44) is a consequence of the Lebesgue differentiation theorem for the measure dw on $\mathbb{R}^n$ as $e^{-s|DB|} g = g$ for all s. Now consider $g \in \overline{\mathsf{R}(DB)}$. As

$$\lim_{t \to 0} \fint_{W(t,x_0)} |g - g(x_0)|^2 \, d\underline{w} = \lim_{t \to 0} \fint_{B(x_0, c_1 t_0)} |g - g(x_0)|^2 \, dw = 0$$

for almost every $x_0 \in \mathbb{R}^n$, it is enough to show the almost everywhere limit

$$\lim_{t \to 0} \fint_{W(t,x_0)} |e^{-s|DB|}g - g|^2 \, d\underline{w} = 0.$$

Write $e^{-s|DB|}g - g = \psi(sDB)g + (I + isDB)^{-1}g - g$. The quadratic estimate (32) and Lemma 6.2 imply that

$$\lim_{t \to 0} \fint_{W(t,x_0)} |\psi(sDB)g|^2 \, d\underline{w} = 0$$

for almost every $x_0 \in \mathbb{R}^n$. Now $(I + isDB)^{-1}g - g = -isDh_s$ with $h_s = B(I + isDB)^{-1}g = (I + isBD)^{-1}(Bg)$ and $Bg \in \mathcal{H}$. Let

$$\tilde{h}_s := (I + isBD)^{-1}(Bg) - \begin{bmatrix} (Bg)_\perp(x_0) \\ ((Bg)_\parallel w)(x_0)w^{-1} \end{bmatrix}.$$

As above we have that $Dh_s = D\tilde{h}_s$. We now apply the following local coercivity inequality on $\mathbb{R}^n$: for any $u \in \mathcal{H}_{\text{loc}}$ with $Du \in \mathcal{H}_{loc}$ and any ball $B(x,r)$ in $\mathbb{R}^n$,

$$(72) \qquad \int_{B(x,r)} |Du|^2 \, dw \lesssim \int_{B(x,2r)} |BDu|^2 \, dw + r^{-2} \int_{B(x,2r)} |u|^2 \, dw,$$

where the implicit constant depends only on the ellipticity constants of B, the dimension and m. (Of course, if B^{-1} is a multiplication operator, this inequality improves considerably.)

Applying this inequality to $u = \tilde{h}_s$, using $Dh_s = D\tilde{h}_s$ and integrating with respect to s implies

$$\fint_{W(t,x_0)} |isDh_s|^2 \, d\underline{w} \lesssim \fint_{\widetilde{W}(t,x_0)} |isBDh_s|^2 \, d\underline{w} + \fint_{\widetilde{W}(t,x_0)} |\tilde{h}_s|^2 \, d\underline{w}$$

$$\lesssim \fint_{\widetilde{W}(t,x_0)} |(I + isBD)^{-1}(Bg) - Bg|^2 \, d\underline{w}$$

$$+ \fint_{\widetilde{W}(t,x_0)} \left| Bg - \begin{bmatrix} (Bg)_\perp(x_0) \\ ((Bg)_\parallel w)(x_0)w^{-1} \end{bmatrix} \right|^2 \, d\underline{w},$$

where $\widetilde{W}(t,x_0)$ is a slightly expanded version of $W(t,x_0)$ and, in the last inequality, we have written

$$\tilde{h}_s = (I + isBD)^{-1}(Bg) - Bg + Bg - \begin{bmatrix} (Bg)_\perp(x_0) \\ ((Bg)_\parallel w)(x_0)w^{-1} \end{bmatrix}.$$

The last two integrals have been shown to converge to 0 as $t \to 0$ for almost every $x_0 \in \mathbb{R}^n$ in the argument for BD. It only remains to prove (72).

For this inequality, we let χ be a smooth scalar-valued function with $\chi = 1$ on $B(x,r)$, χ supported in $B(x,2r)$ and $|\nabla\chi| \lesssim r^{-1}$. Using the

commutator identity (16) between χ and D, we have

$$\int_{B(x,r)} |Du|^2 \, dw \leq \int |\chi Du|^2 \, dw \lesssim \int |D(\chi u)|^2 \, dw + \int |\nabla \chi|^2 |u|^2 \, dw.$$

Since B is accretive on $\mathsf{R}(D)$ and $\chi u \in \mathsf{D}(D)$, we have $\int |D(\chi u)|^2 \, dw \lesssim \int |BD(\chi u)|^2 \, dw$. Now, we use again the commutation between χ and D together with $\|B\|_\infty$. This proves (72). $\qquad\square$

References

[1] ALBRECHT, D., DUONG, X., AND MCINTOSH, A. Operator theory and harmonic analysis. In *Instructional Workshop on Analysis and Geometry, Part III (Canberra, 1995)*, vol. 34 of *Proc. Centre Math. Appl. Austral. Nat. Univ.* Austral. Nat. Univ., Canberra, 1996, pp. 77–136. 15, 32

[2] AUSCHER, P., AXELSSON, A., AND HOFMANN, S. Functional calculus of Dirac operators and complex perturbations of Neumann and Dirichlet problems. *Journal of Functional Analysis, 255,* 2 (2008), 374-448. 46, 50

[3] AUSCHER, P., AND AXELSSON, A. Weighted maximal regularity estimates and solvability of non-smooth elliptic systems, I. *Invent. Math. 184* (2011) 47–115. 4, 5, 33, 34, 35, 37, 38, 42, 46, 49

[4] AUSCHER, P., AND ROSÉN, A. Weighted maximal regularity estimates and solvability of non-smooth elliptic systems, II. *Analysis and PDE, Vol. 5* (2012), No. 5, 983–1061. 4, 5, 8, 37, 40, 42, 45

[5] AUSCHER, P., AXELSSON, A., AND MCINTOSH, A. Solvability of elliptic systems with square integrable boundary data, in Harmonic Analysis and Partial Differential Equations, Contemporary Mathematics, vol. 505, Amer. Math. Soc., Providence, RI, 2010, pp. 105-129. 45, 46

[6] AUSCHER, P., AXELSSON, A., AND MCINTOSH, A. On a quadratic estimate related to the Kato conjecture and boundary value problems in Harmonic Analysis and Partial Differential Equations, Contemporary Mathematics, vol. 505, Amer. Math. Soc., Providence, RI, 2010, pp. 105-129. 6, 16, 17

[7] AUSCHER, P., HOFMANN, S., LACEY, M., MCINTOSH, A., AND TCHAMITCHIAN, P. The solution of the Kato square root problem for second order elliptic operators on $\mathbf{R}^n$. *Ann. of Math. (2) 156*, 2 (2002), 633–654. 24

[8] AUSCHER, P., MCINTOSH, A., AND MOURGOGLOU, M., On L^2 solvability of BVPs of elliptic systems, *J. Fourier Anal. Appl.* 2013, Volume 19, Issue 3, 478-494. 48

[9] AUSCHER, P., AND TCHAMITCHIAN, P. *Square root problem for divergence operators and related topics*, vol. 249 of *Astérisque*. Soc. Math. France, 1998. 19

[10] AXELSSON, A., KEITH, S., AND MCINTOSH, A. Quadratic estimates and functional calculi of perturbed Dirac operators. *Invent. Math. 163*, 3 (2006), 455–497. 6, 16, 22

[11] BADR N. Real interpolation of Sobolev spaces, *Math. Scand., vol. 31,* no. 4, (2009), 235–264. 12

[12] BANDARA, L. Quadratic estimates for perturbed Dirac type operators on doubling measure metric spaces. To appear in Proceedings of the AMSI International Conference on Harmonic Analysis and Applications, Proc. Centre Math. Appl. Austral. Nat. Univ. ArXiv e-prints: 1107.3905. 6, 16, 21

[13] BIROLI, M.; MOSCO, U. Wiener criterion and potential estimates for obstacle problems relative to degenerate elliptic operators. *Ann. Mat. Pura Appl.* 159 (1991), no. 4, 255281. 2

[14] CAFFARELLI, L., AND SILVESTRE, L. An extension problem related to the fractional Laplacian, *Comm. Partial Differential Equations 32* (2007), no. 7-9, 1245–1260. 3

[15] CRUZ-URIBE S.F.O., D. AND RIOS, C. The Kato Problem for operators with weighted ellipticity, Preprint 2012. 7, 11, 12, 16

[16] DINDOŠ, M., PIPHER, J. AND RULE, D. Boundary value problems for second order elliptic operators satisfying a Carleson condition. ArXiv e-prints: 1301.0426. 4

[17] DINDOŠ, M. AND RULE, D. Elliptic equations in the plane satisfying a Carleson measure condition. *Rev. Mat. Iberoam. 26,* 3 (2010), 1013-1034. 4

[18] DUOANDIKOETXEA, J., RUBIO DE FRANCIA, J.-L. Maximal and singular integral operators via Fourier transform estimates *Invent. Math. 84* (1986), 541–561. 12

[19] FABES, E., JERISON, D., AND KENIG, C. The Wiener test for degenerate elliptic equations. *Ann. Inst. Fourier (Grenoble) 32* (1982), no. 3, vi, 151–182. 2

[20] FABES, E., JERISON, D., AND KENIG, C. Boundary behavior of solutions to degenerate elliptic equations. Conference on harmonic analysis in honor of Antoni Zygmund, Vol. I, II (Chicago, Ill., 1981), 577–589, Wadsworth Math. Ser., Wadsworth, Belmont, CA, 1983. 2

[21] FABES, E.; KENIG, C.; SERAPIONI, R. The local regularity of solutions of degenerate elliptic equations. *Comm. Partial Differential Equations 7* (1982), no. 1, 77–116. 2, 12, 13, 44, 49

[22] FRANCHI, B.; SERAPIONI, R. Pointwise estimates for a class of strongly degenerate elliptic operators: a geometrical approach. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* 14 (1987), no. 4, 527-568. 2

[23] GARCA-CUERVA, J. RUBIO DE FRANCIA, J.L. *Weighted Norm Inequalities and Related Topics*, North-Holland Math. Studies 116, North-Holland, 1985. 12

[24] GARNETT, J. B. *Bounded analytic functions,* Academic Press, 1981. 7, 9, 10

[25] GRAFAKOS, L. *Classical and Modern Fourier Analysis*, Pearson Education, Inc. Prentice Hall, 2004. 11, 19

[26] HAJŁASH, P. & KOSKELA, P. *Sobolev met Poincaré*, Mem. Amer. Math. Soc., 145 (2000), x+101 p. 39

[27] HUANG, Y. Weighted tent spaces with Whitney averages: factorization, interpolation and duality arXiv:1303.5982. 38

[28] HYTÖNEN, T. The sharp weighted bound for general Calderón-Zygmund operators. *Ann. of Math. (2) 175* (2012), no. 3, 1473–1506. 11

[29] HYTÖNEN, T., AND ROSÉN. A. On the Carleson duality, *Ark. Mat.* (2012), 1–21. 37, 38

[30] HYTÖNEN, T., AND ROSÉN. A. Approximate and exact extensions of Lebesgue functions. In preparation. 10

[31] KATO, T. *Perturbation theory for linear operators.* Springer Verlag, New York, 1966. 13

[32] KENIG, C., AND PIPHER, J. The Neumann problem for elliptic equations with nonsmooth coefficients. *Invent. Math. 113*, 3 (1993), 447–509. 39

[33] MEYER, Y. *Ondelettes et opérateurs.* Hermann, volumes 1 & 2, 1990. 11

[34] ROSÉN, A. Cauchy non-integral formula. in Harmonic Analysis and Partial Differential Equations, Contemporary Mathematics, vol. 612, Amer. Math. Soc., Providence, RI, 2014, pp. 163-178. 4, 33, 36, 42, 43, 44

[35] ŠNEĬBERG, I. J. Spectral properties of linear operators in interpolation families of Banach spaces. *Mat. Issled. 9*, 2(32) (1974), 214–229, 254–255. 49

PASCAL AUSCHER, UNIVERSITÉ DE PARIS-SUD, UMR DU CNRS 8628, 91405 ORSAY CEDEX, FRANCE
  E-mail address: pascal.auscher@math.u-psud.fr

ANDREAS ROSÉN, MATEMATISKA VETENSKAPER, GÖTEBORGS UNIVERSITET, 41296 GÖTEBORG, SWEDEN
  E-mail address: andreas.rosen@chalmers.se

DAVID RULE, MATEMATISKA INSTITUTIONEN, LINKÖPINGS UNIVERSITET, 581 83 LINKÖPING, SWEDEN
  E-mail address: david.rule@liu.se