

ON THE KATO PROBLEM FOR ELLIPTIC OPERATORS IN NON-DIVERGENCE FORM

LUIS ESCAURIAZA, PABLO HIDALGO-PALENCIA, AND STEVE HOFMANN

Dedicated to Prof. Carlos Kenig on the occasion of his 70th birthday, and to the memory of Luis Escauriaza.

ABSTRACT. We consider the Kato square root problem for non-divergence second order elliptic operators $L = -\sum_{i,j=1}^n a_{ij}D_iD_j$, and, especially, the normalized adjoints of such operators. In particular, our results are applicable to the case of real coefficients having sufficiently small BMO norm. We assume that the coefficients of the operator are smooth, but our quantitative estimates do not depend on the assumption of smoothness.

CONTENTS

1. Introduction and main result	2
1.1. Notation	7
2. Preliminaries	8
2.1. Gaussian bounds for kernels of semigroups	8
2.2. Weighted Littlewood-Paley theory	9
2.3. Uniform bounds for some operators	10
2.4. Off-diagonal estimates	10
2.5. Estimates for differences and gradients	13
3. The Kato problem for L	14
4. The Kato problem for $\tilde{L}$	17
4.1. Reduction to a quadratic estimate	18
4.2. Reduction to a Carleson measure estimate	19
4.3. The $T(b)$ argument	26
References	28

Date: February 26, 2024.

2020 Mathematics Subject Classification. 35J15, 42B25, 42B37, 47B44, 47D06.

Key words and phrases. Elliptic operators in non-divergence form, Kato square root problem, Muckenhoupt weights, Littlewood-Paley theory, Functional Calculus.

The second author is supported by the grant CEX2019-000904-S-20-3, funded by MCIN/AEI/ 10.13039/501100011033, and acknowledges financial support from MCIN/AEI/ 10.13039/501100011033 grants CEX2019-000904-S and PID2019-107914GB-I00. The third author was supported by NSF grant DMS-2000048. Part of this work was carried out while the first and third authors were visiting ICMAT in Madrid, and part of this work was carried out while the second author was visiting the third author at the University of Missouri - Columbia. The authors express their gratitude to these institutions.

1. INTRODUCTION AND MAIN RESULT

In [AHLMcT], the authors resolved in the affirmative the long-standing square root problem of Kato for divergence form complex-elliptic operators in $\mathbb{R}^n$. This was the culmination of a series of previous results: [CMcM] (treating the 1-d case); [CDM] and [FJK1, FJK2] (treating small perturbations of the constant coefficient case); [HMc] (the 2-d case); [AHLT] (small perturbations of the real-symmetric case, sometimes referred to as the restricted Kato problem); and [HLMc] (the case that the heat kernel satisfies a Gaussian upper bound and Nash-type Hölder continuity estimates). The solution of the Kato problem, and the circle of ideas involved in its proof, led to subsequent breakthroughs in the theory of elliptic boundary value problems, see, e.g., [AAAHK], [Hof] [HKMP, HKMP2], [AKMc], [AA], [AAH], [AAMc, AAMc2], [CUR], [HLeM], [EH], [HLMP1, HLMP2], [BHLMP]. See also the significant related ground-breaking work in the parabolic setting: [Ny], [AEN1, AEN2].

In this note, we initiate the study of the square root problem in the non-divergence setting in dimensions greater than 1, with the eventual goal of developing applications to the theory of boundary value problems, as has been done in the aforementioned divergence form case. Previously, the non-divergence problem had been treated only in the 1-dimensional setting, in [KM]; in fact, in that paper the authors treat the more general class of operators of the form $L = -aDbD$, where D denotes the ordinary differentiation operator on the line, and a, b are arbitrary bounded accretive complex-valued functions on the line.

At present, we are able to treat only the case of real coefficients. On the other hand, we point out that in the divergence form case, there is no known proof for real, non-symmetric coefficients that is fundamentally easier than the proof in the general case, owing to the non-selfadjointness of non-symmetric divergence form operators. Moreover, it is the real, non-symmetric case that underlies the breakthrough in the study of the Dirichlet problem obtained in [HKMP]. We observe that in the non-divergence setting, we may assume without loss of generality that the coefficient matrix is symmetric, but in contrast to the divergence form case, operators of non-divergence type are inherently non-selfadjoint, even with symmetric coefficients.

A fundamental difficulty that one encounters in the non-divergence setting, is that the Kato problem for non-divergence elliptic operators seems to be most naturally formulated in a weighted L^2 space, and in general, the weight need not belong to the Muckenhoupt A_2 class. Another difficulty inherent to the non-divergence setting is the lack of uniqueness (see the work of Nadirashvili [N]). On the other hand, working with real coefficients allows us to make use of the pioneering work of Krylov and Safonov [KS], as well as the important ideas of Baumann [B] and Escauriaza [E2]. We shall return to these matters in the sequel. First, let us set notation and definitions.

We will say that the operator L is a *second-order elliptic operator in non-divergence form* on $\mathbb{R}^n$ if

$$(1.1) \quad Lu = - \sum_{i,j=1}^n a_{ij} D_i D_j u,$$

where $A = (a_{ij}(\cdot))$ is a real and measurable coefficient matrix which (without loss of generality) we can take to be symmetric, and for which we also assume, for some $\lambda > 0$,

$$(1.2) \quad A(x)\xi \cdot \xi \geq \lambda |\xi|^2 \quad \text{and} \quad |A(x)\xi \cdot \zeta| \leq \lambda^{-1} |\xi| |\zeta| \quad \text{for all } \xi, \zeta \in \mathbb{R}^n \text{ and a.e. } x \in \mathbb{R}^n,$$

For such L , we have also its adjoint operator

$$(1.3) \quad L^*u = - \sum_{i,j=1}^n D_i D_j (a_{ij}u).$$

Following [E2], we say that the function $u \in L^1_{\text{loc}}(\mathbb{R}^n)$ is a solution of the adjoint equation $L^*u = 0$ if for every $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ we have

$$\int_{\mathbb{R}^n} u(x) L\varphi(x) dx = 0.$$

Let us recall also the definition of Muckenhoupt weights, that will be used throughout the text because of the properties of some particularly relevant adjoint solution, as we shall see shortly.

Definition 1.4 (Muckenhoupt weights). We say that the function w belongs to the Muckenhoupt class of weights A_p for some $1 < p < \infty$ if $w(x) > 0$ a.e. $x \in \mathbb{R}^n$ and

$$[w]_{A_p} := \sup_B \left(\frac{1}{|B|} \int_B w(x) dx \right) \left(\frac{1}{|B|} \int_B w(x)^{1-p'} dx \right)^{p-1} < \infty,$$

where the supremum is taken over all the balls $B \subset \mathbb{R}^n$, and also denote $A_\infty := \bigcup_{1 < p < \infty} A_p$.

We recall that it is well known that the A_∞ property is equivalent to the Reverse Hölder property, i.e., that there is an exponent $q > 1$, and a uniform constant C such that for every ball B ,

$$(RH_q) \quad \left(\frac{1}{|B|} \int_B w^q(x) dx \right)^{1/q} \leq C \frac{1}{|B|} \int_B w(x) dx$$

For details of the theory of Muckenhoupt weights, the reader may consult, e.g. [Duo, Chapter 7].

With this definition in mind, we recall a fundamentally important property of equations in non-divergence form.

Lemma 1.5 ([E2, Theorem 1.1]). *Let L be a second-order elliptic operator in non-divergence form, and L^* its adjoint. Then there exists a non-negative solution W of the adjoint equation $L^*W = 0$ in $\mathbb{R}^n$, satisfying $W(B_1(0)) = |B_1(0)|$, which we call the **global non-negative adjoint solution**. Furthermore, W satisfies a Reverse Hölder property with exponent $\frac{n}{n-1}$, so $W \in A_\infty$ (c.f. Definition 1.4). Moreover, the $RH_{n/(n-1)}$ constants depend only on dimension and ellipticity.*

If the coefficients of L are smooth, or even belong to VMO, then W (with the stated normalization) is unique. In general, it need not be unique. On the other hand, for any given L , any choice of such a W will enjoy the same quantitative estimates, with uniform control of all relevant constants. In the case that W is not unique, we may therefore simply fix an arbitrary choice of W .

It is a well known fact that A_∞ weights are doubling. In our case, this means that there exists a constant $C_D = C_D([W]_{A_\infty}) \geq 1$ such that $W(2B) \leq C_D W(B)$ for every ball B .

From now on we will work most of the time in the weighted Hilbert space

$$L_W^2 := L^2(\mathbb{R}^n, W(x)dx).$$

In the non-divergence setting, this space is more natural in many ways than unweighted L^2 . In particular, the following identity holds, as may be seen formally by using $L^*W = 0$ and integrating by parts:

$$(1.6) \quad \int_{\mathbb{R}^n} uLuWdx = \int_{\mathbb{R}^n} A\nabla u \cdot \nabla uWdx.$$

In fact, one may readily deduce that (1.6) holds when the coefficients are smooth, and more generally, it also holds at least when the coefficients have sufficiently small BMO norm (depending only on dimension and ellipticity), for all $u \in \mathcal{D}(L)$ (the domain of L), defined by

$$\mathcal{D}(L) := \{u \in L_W^2 : Lu \in L_W^2\}.$$

Indeed, if the coefficients have sufficiently small BMO norm, then $W \in A_2$ (see [E1, Theorem 1.2], and its proof¹). In turn, using this fact, one may prove² the regularity estimate

$$(1.7) \quad \|\nabla^2 u\|_{L_W^2} \lesssim \|f\|_{L_W^2}$$

for solutions of the Poisson problem $Lu = f \in L_W^2$, and hence, that

$$(1.8) \quad \mathcal{D}(L) = H_W^2(\mathbb{R}^n) := \{u \in L_W^2 : \nabla u, \nabla^2 u \in L_W^2\}.$$

The identity (1.6) then follows readily for all $u \in \mathcal{D}(L)$.

Remark 1.9. We observe that $H_W^2(\mathbb{R}^n)$ is dense in L_W^2 when $W \in A_2$ (indeed, even $\mathcal{C}_c^\infty$ is dense in that case), and therefore L is densely defined when the coefficients have sufficiently small BMO norm.

We shall also consider the *normalized adjoint* of L , which we denote by $\tilde{L}$, and which we define to be the adjoint of L with respect to the space L_W^2 . Thus, $\tilde{L}$ is given, at least for smooth coefficients, by the formula

$$(1.10) \quad \tilde{L}u := -\frac{1}{W} \sum_{i,j=1}^n D_i D_j (a_{ij}uW) = \frac{1}{W} L^*(uW).$$

If the coefficients are merely measurable, then we interpret $\tilde{L}$ in the weak sense: we say that $u \in L_W^2$ belongs to $\mathcal{D}(\tilde{L})$, the domain of $\tilde{L}$, provided that there is a function $g \in L_W^2$ such that for every $v \in \mathcal{D}(L)$,

$$\int_{\mathbb{R}^n} Lv u W dx = \int_{\mathbb{R}^n} v g W dx,$$

and in this case we set $\tilde{L}u = g$. Just as for (1.6), integrating by parts and using $L^*W = 0$, we obtain (at least in the case of smooth coefficients, and for $u \in H_W^2$)

$$(1.11) \quad \int_{\mathbb{R}^n} u \tilde{L}u W dx = \int_{\mathbb{R}^n} A\nabla u \cdot \nabla u W dx.$$

We shall henceforth make the *qualitative* assumption that the coefficients a_{ij} are smooth, with qualitative L^∞ bounds on ∇a_{ij} and $\nabla^2 a_{ij}$. Thus, (1.11) will be valid in the sequel, for $u \in H_W^2$. On the other hand, we emphasize that our *quantitative* bounds will never depend on smoothness, nor on estimates for the derivatives of a_{ij} .

¹In fact, [E1] precedes [E2] chronologically; the result in [E1] treats explicitly local versions of W , but the same arguments may be applied to the global W constructed in [E2].

²We caution the reader that the proof of (1.7) is a somewhat non-trivial matter, and is based on the boundedness of the commutator $[T, b]$, where T is a singular integral and $b \in \text{BMO}$ [CRW], along with a suitable expansion in terms of spherical harmonics; see [CFL1, CFL2] for related results in the unweighted case.

Remark 1.12. Using (1.6), (1.11), (1.7), and (1.8)³, one may then show that L , and $\tilde{L}$, viewed as unbounded operators on L_W^2 , are each closed, sectorial and m-accretive, and hence each has an m-accretive square root (see [Ka, Theorem 3.35, p. 281], or [H, Sections 3 and 7]). Moreover, L generates a heat semigroup $z \mapsto e^{-zL}$, which is well-defined and analytic in a sector containing the positive real axis (hence, the analogous statement is also true for $\tilde{L}$).

Let us now sketch the proofs of the functional analytic facts listed in Remark 1.12.

L and $\tilde{L}$ are closed operators. For $\tilde{L}$, this follows immediately from the fact that L is densely defined (see [Ka, p. 168]). Thus we consider L . Suppose that $\{u_n\}_n \subset \mathcal{D}(L)$, that $u_n \rightarrow u$ in L_W^2 , and that

$$f_n := Lu_n \rightarrow f \text{ in } L_W^2.$$

We need to verify that $u \in \mathcal{D}(L)$, and that $Lu = f$, i.e., that the graph $\{(u, Lu) : u \in \mathcal{D}(L)\}$ is a closed set in $L_W^2 \times L_W^2$. Applying (1.6) and (1.7) to $u_n - u_m$, we see that $\{\nabla u_n\}_n$ and $\{\nabla^2 u_n\}_n$ are each convergent in L_W^2 , thus, $\{u_n\}_n$ is convergent in H_W^2 , and since $u_n \rightarrow u$ in L_W^2 , we see that $u \in H_W^2 = \mathcal{D}(L)$, and that $u_n \rightarrow u$ in H_W^2 . In particular, $D_i D_j u_n \rightarrow D_i D_j u$ in L_W^2 for each $i, j = 1, 2, \dots, n$, hence $Lu_n \rightarrow Lu$, so that $Lu = f$, as desired.

L and $\tilde{L}$ are sectorial. It follows readily from (1.6) (respectively, (1.11)) that the numerical ranges

$$\Theta := \{ \langle Lu, u \rangle \in \mathbb{C} : \|u\|_{L_W^2} = 1 \}, \quad \tilde{\Theta} := \{ \langle \tilde{L}u, u \rangle \in \mathbb{C} : \|u\|_{L_W^2} = 1 \}$$

are each contained in a sector $S_\omega := \{z \in \mathbb{C} : |\arg z| \leq \omega\} \cup \{0\}$, with $0 < \omega < \pi/2$, depending only on ellipticity. We omit the standard argument.

L and $\tilde{L}$ are m-accretive. By [Ka, Problem 3.31, p 279], it suffices to show that L is m-accretive. To this end, set

$$\Delta := \mathbb{C} \setminus S_\omega,$$

and for $\zeta \in \Delta$, set $\delta = \delta(\zeta) := \text{dist}(\zeta, S_\omega)$. By symmetry, we also have $\delta = \text{dist}(\bar{\zeta}, S_\omega)$. We now claim that

$$(1.13) \quad \|(L - \zeta)u\|_{L_W^2} \geq \delta \|u\|_{L_W^2}, \quad u \in \mathcal{D}(L), \quad \zeta \in \Delta.$$

Indeed, to verify the claim, we may assume without loss of generality that $\|u\|_{L_W^2} = 1$, in which case

$$\|(L - \zeta)u\|_{L_W^2} \geq |\langle (L - \zeta)u, u \rangle| = |\langle Lu, u \rangle - \zeta| \geq \delta,$$

since $\langle Lu, u \rangle \in \Theta \subset S_\omega$.

Fix $\zeta \in \Delta$. Then $L - \zeta$ is 1-1 on $\mathcal{D}(L)$, and has closed range (since L is a closed operator). Similarly, $\tilde{L} - \bar{\zeta}$ is 1-1 on $\mathcal{D}(\tilde{L})$. Since L is densely defined,

$$\mathcal{N}(\tilde{L} - \bar{\zeta}) = \mathcal{R}(L - \zeta)^\perp,$$

i.e., the null space of $\tilde{L} - \bar{\zeta}$ is the orthogonal complement of the range of $L - \zeta$. Thus, $L - \zeta$ has dense range, since $\tilde{L} - \bar{\zeta}$ is 1-1. Hence, $L - \zeta$ is invertible as a mapping from $\mathcal{D}(L)$ onto L_W^2 . Combined with the estimate (1.13), this shows that L is m-accretive (see [Ka, p. 279]).

The heat semi-group. Given the preceding properties of L , we have existence, uniqueness, and analyticity of a contraction semigroup e^{-zL} , for z in the open sector $S_\alpha^0 := \{z \in \mathbb{C} : |\arg z| < \alpha\}$, provided $0 < \alpha < \pi/2 - \omega$. See, e.g., [Ka, pp. 480-493, especially Theorem 1.24, p. 492].

Our main result is the following:

³(1.8) is used to show that (1.6) holds for all $u \in \mathcal{D}(L)$.

Theorem 1.14. *Let L be a second-order elliptic operator in non-divergence form with smooth real coefficients satisfying (1.2), and let W be the associated global non-negative adjoint solution provided by Lemma 1.5. If $W \in A_2$ (see Definition 1.4), then we have*

$$\left\| \sqrt{L}f \right\|_{L^2_W} \approx \left\| \nabla f \right\|_{L^2_W} \approx \left\| \sqrt{\tilde{L}}f \right\|_{L^2_W},$$

where the implicit constants depend only on n, λ and $[W]_{A_2}$.

The main goal of this paper is to prove this Theorem. Hence, from here on we will always impose the extra assumption that $W \in A_2$, along with the qualitative assumption that the coefficients are smooth. The result will follow at once from Theorems 3.1 and 4.1.

Some additional remarks are in order.

Remark 1.15. As mentioned above, in general W belongs to the class A_∞ , and thus $W \in A_p$ for some p depending on dimension and ellipticity, but p may be strictly greater than 2, and in fact in the general case we have no precise upper bound on p . Thus, our result is only a partial one, and does not address the fundamental challenge of treating the non- A_2 case. On the other hand, as noted above, if the coefficients have sufficiently small BMO norm, then $W \in A_2$, and thus our result does apply in that setting.

Remark 1.16. In the case that the coefficient matrix has sufficiently small BMO norm, then as also noted above, we may identify the domain $\mathcal{D}(L)$ as the weighted Sobolev space $H^2_W(\mathbb{R}^n)$ (see (1.8)). Hence, by combining several known (or at least implicit) results, we may identify the domain of $\sqrt{L}$ as the Sobolev space $H^1_W(\mathbb{R}^n) := \{u \in L^2_W(\mathbb{R}^n) : \nabla u \in L^2_W(\mathbb{R}^n)\}$; this corresponds to the estimate $\left\| \sqrt{L}f \right\|_{L^2_W} \lesssim \left\| \nabla f \right\|_{L^2_W}$. Indeed, in [DY] it is shown that the operator L has a bounded holomorphic functional calculus in (unweighted) L^2 (even in L^p), provided that the BMO norm of the coefficients is sufficiently small. Under the same smallness assumption, the arguments of [DY] may be extended to the weighted case considered here, to deduce that L has a bounded holomorphic functional calculus in L^2_W . Combining the results of [Y] and [Mc], we find that $\mathcal{D}(\sqrt{L})$ is the complex interpolation space mid-way between L^2_W and $\mathcal{D}(L) = H^2_W$, i.e., $\mathcal{D}(\sqrt{L}) = H^1_W$.

The analogous strategy fails for $\tilde{L}$, as we have no idea how to identify $\mathcal{D}(\tilde{L})$ (similarly, the square root problem in the divergence form case entailed the same difficulty).

Remark 1.17. The assumption of smoothness of the coefficients is purely qualitative, and our quantitative estimates will not depend on smoothness, but only on the stated parameters n, λ and $[W]_{A_2}$. However, it is not clear at present how to make sense of the identity (1.11) for non-smooth coefficients, and as a consequence, in the absence of smoothness, we do not know how to prove certain estimates which rely on (1.11), such as Lemma 2.11 (viii) (in the case of measurable coefficients, we know how to give only a formal proof of the latter, assuming a priori finiteness of $\left\| \nabla e^{-t^2 \tilde{L}} f \right\|_{L^2_W}$).

On the other hand, as noted above, identity (1.6) holds without smoothness, in the case that the coefficients have sufficiently small BMO norm. Under the latter scenario, we require identity (1.11) and its consequences (and thus, the qualitative, a priori assumption of smoothness of the coefficients) in two places: 1) in the proof of Theorem 4.1 (the square root problem for $\tilde{L}$), where estimate (1.11) is heavily used, and 2) in the proof of the m-accretivity of L given above, where we used (1.11) to establish density of the range of $L - \zeta$. Otherwise, (1.11) is not used in the proof of Theorem 3.1 (the square root problem of L).

Although our operators L and $\tilde{L}$ are not of divergence form, there is a nice identity relating these two non-divergence operators with another one which is in divergence form, but degenerate elliptic⁴. Indeed, if we let $\widetilde{\operatorname{div}}$ denote the normalized divergence, defined for an $\mathbb{R}^n$ -valued function $\mathbf{v}$ by $\widetilde{\operatorname{div}} \mathbf{v} := \frac{1}{W} \operatorname{div}(W\mathbf{v})$, then $\widetilde{\operatorname{div}}$ is precisely the adjoint operator to $-\nabla$ inside L_W^2 , and we also have, using $L^*W = 0$,

$$(1.18) \quad Lu + \tilde{L}u = -2\widetilde{\operatorname{div}}(A\nabla u).$$

In the case of non-smooth coefficients, we interpret the latter identity in the weak sense described above: for $u, \varphi \in H_W^2$,

$$\int_{\mathbb{R}^n} (\varphi Lu + uL\varphi) W dx = 2 \int_{\mathbb{R}^n} A\nabla u \cdot \nabla \varphi W dx.$$

The identity (1.18) will be of great use to us in the sequel.

The paper is organized as follows:

- In Section 2 we give some definitions and estimates for some of the operators that which will appear repeatedly across the paper.
- In Section 3 we prove $\|\sqrt{L}f\|_{L_W^2} \lesssim \|\nabla f\|_{L_W^2}$, which turns out to be a relatively easy consequence of Littlewood-Paley theory because of the form of L (since L annihilates not only constants but also first degree monomials).
- In Section 4 we prove $\|\sqrt{\tilde{L}}f\|_{L_W^2} \lesssim \|\nabla f\|_{L_W^2}$, which is in fact the more difficult result in the paper. To treat $\tilde{L}$, we follow broadly the scheme provided by [AHLMcT], first reducing the problem to some square function estimates, which are handled using a $T1$ -like argument and then a local Tb argument. Of course, some significant modifications of the arguments in [AHLMcT] are needed; the identity (1.18) will be useful in this case.

We remark that the square root problem for $\tilde{L}$ is significantly more difficult than its analogue for L .

1.1. Notation.

- We use the notation $a \lesssim b$ to denote that there exists a positive harmless constant C (which can vary from line to line) such that $a \leq Cb$. We will also denote $a \approx b$ whenever $a \lesssim b$ and $b \lesssim a$.
- In the proofs of the results from now on, we will omit dependencies of constants on n , λ , $[W]_{A_\infty}$ and $[W]_{A_2}$ – treating them as harmless constants – although we will make these dependencies explicit in the statements.
- Euclidean balls are denoted by $B_t(x) := \{y \in \mathbb{R}^n : |y - x| < t\}$.
- If $B = B_t(x) \subset \mathbb{R}^n$ is a ball and $\kappa > 0$, by κB we denote the ball with same radius and scaled by a factor of κ , i.e., $B_{\kappa t}(x)$. The same applies to cubes.
- For $E \subset \mathbb{R}^n$, $|E|$ denotes the Lebesgue measure of E .
- If $E, F \subset \mathbb{R}^n$ are arbitrary subsets, we write $\operatorname{dist}(E, F) := \inf\{|x - y| : x \in E, y \in F\}$.
- For any subset $E \subset \mathbb{R}^n$, we denote $\mathbf{1}_E$ the characteristic function of E (i.e. $\mathbf{1}_E(x) = 1$ if $x \in E$ and 0 otherwise). Concretely, we write $\mathbf{1} := \mathbf{1}_{\mathbb{R}^n}$, the function constantly 1.
- We will denote vector-valued functions with boldface letters, e.g., $\mathbf{v} := (v_1, \dots, v_n)$.
- D_j denotes the differentiation operator in the direction of x_j , i.e., $D_j = \frac{\partial}{\partial x_j}$.

⁴Thus, our results here are somewhat related to those of Cruz-Uribe and Rios [CUR].

- We denote averages with respect to a measure ν by $f_E f d\nu := \nu(E)^{-1} \int_E f d\nu$. Often the measure with respect to which we take averages will be the weighted measure $W(x)dx$: it will be clear by the context. For the latter measure, we write $W(E) := \int_E W(x)dx$.
- We will frequently use cubes in our proofs: every time we cover $\mathbb{R}^n$ (or some portion of it) by cubes, we mean we are using a covering by cubes of the dyadic grid $\{2^j \mathbf{k} + [0, 2^j)^n : j \in \mathbb{Z}, \mathbf{k} \in \mathbb{Z}^n\}$. Anytime we use the letter Q , we will be referring to a dyadic cube. For such a cube Q , we let $\ell(Q)$ denote its sidelength.
- We let $\mathcal{M}$ and $\mathcal{M}_W$ denote, respectively, the classical Hardy-Littlewood maximal operator, and the Hardy-Littlewood maximal function with respect to the measure $W(x)dx$, that is,

$$\mathcal{M}f(x) := \sup_{B \ni x} \int_B |f(y)| dy = \sup_{B \ni x} \frac{1}{|B|} \int_B |f(y)| dy,$$

and

$$\mathcal{M}_W f(x) := \sup_{B \ni x} \int_B |f(y)| W(y) dy = \sup_{B \ni x} \frac{1}{W(B)} \int_B |f(y)| W(y) dy.$$

Since W is doubling (because $W \in A_\infty$), $\mathcal{M}_W$ is bounded on L_W^p for every $1 < p < \infty$. We will use this fact in the sequel.

- As explained before, we set $L_W^2 := L^2(\mathbb{R}^n, W(x)dx)$, and we define the weighted Sobolev space $H_W^2 := \{u \in L_W^2 : \nabla u \in L_W^2, \nabla^2 u \in L_W^2\}$. We will also write $L_W^2(E) := L^2(E, W(x)dx)$ for any subset $E \subset \mathbb{R}^n$.
- We denote the composition of two operators U and V by $UV(f) := U(V(f))$.
- For a function $f \in L^2(\mathbb{R}^n)$, we denote its Fourier transform by $\hat{f}$.
- $\mathcal{S}$ will denote the usual Schwarz class of smooth, rapidly decaying functions on $\mathbb{R}^n$.

Acknowledgements. The third author thanks Prof. X. T. Duong for an interesting conversation concerning the latter's joint work with L. Yan [DY], and in particular for pointing out to us the argument sketched in Remark 1.16.

2. PRELIMINARIES

2.1. Gaussian bounds for kernels of semigroups. From now on, we will use many times the parabolic semigroup (with elliptic homogeneity) $e^{-t^2 L}$, whose kernel is the fundamental solution $\Gamma_{t^2}(x, y)$; i.e. we have $e^{-t^2 L} f(x) = \int_{\mathbb{R}^n} \Gamma_{t^2}(x, y) f(y) dy$ for sufficiently regular f . The fundamental solution satisfies the following Gaussian estimate:

Lemma 2.1 ([E2, Theorem 1.2]). *The kernel $\Gamma_{t^2}(\cdot, \cdot)$ of $e^{-t^2 L}$ satisfies the Gaussian bounds*

$$(2.2) \quad \Gamma_{t^2}(x, y) \lesssim \min \left\{ \frac{1}{W(B_t(x))}, \frac{1}{W(B_t(y))} \right\} e^{-c \frac{|x-y|^2}{t^2}} W(y),$$

and

$$(2.3) \quad \max \left\{ \frac{1}{W(B_t(x))}, \frac{1}{W(B_t(y))} \right\} e^{-\frac{|x-y|^2}{ct^2}} W(y) \lesssim \Gamma_{t^2}(x, y),$$

where the implicit constants and c depend on n and λ .

Remark 2.4. The results stated above as Lemma 1.5 and Lemma 2.1, are stated in [E2] explicitly for smooth coefficients, but as the author points out, “the usual compactness arguments”, and the uniformity of the estimates depending only on n and λ , allow one to deduce the existence of (non-unique) W and Γ verifying the same bounds, in the general case of bounded measurable coefficients.

Remark 2.5. The doubling property of W , combined with the exponential decay factor, allow us to interchange “min” and “max” in (2.2) and (2.3), modulo an adjustment of the constants.

Remark 2.6. The (absolute values of) the kernels of the operators $t^2 L e^{-t^2 L}$ and $t^4 L^2 e^{-t^2 L}$ also satisfy the upper bound (2.2), by analyticity of the semigroup $z \mapsto e^{-zL}$ in a sector.

2.2. Weighted Littlewood-Paley theory. The following results are standard. We recall them here for the sake of convenience of exposition.

Lemma 2.7. *Let $W \in A_2$, and let $K_t f := k_t * f$, with k defined on $\mathbb{R}^n$ satisfying $|k(x)| \leq (1 + |x|)^{-n-1}$, where $k_t(x) := t^{-n} k(x/t)$. Then*

$$\| \sup_{t>0} |K_t f| \|_{L_W^2} \lesssim \| \mathcal{M} f \|_{L_W^2} \lesssim \| f \|_{L_W^2},$$

where $\mathcal{M}$ is the classical Hardy-Littlewood maximal operator, and the implicit constant depends on n and $[W]_{A_2}$.

Lemma 2.8. *Let $W \in A_2$, and let $Q_s f := \psi_s * f$, where $\psi \in \mathcal{S}$ and satisfies $\int_{\mathbb{R}^n} \psi = 0$. Then*

$$\int_0^\infty \|Q_s f\|_{L_W^2}^2 \frac{ds}{s} \lesssim \|f\|_{L_W^2}^2,$$

where the implicit constant depends on n, ψ , and $[W]_{A_2}$. Moreover, if in addition ψ is radial and non-trivial, then using a slight abuse of notation and then normalizing, we may assume that $\int_0^\infty \hat{\psi}(s)^2 \frac{ds}{s} = 1$, in which case we have the Calderón reproducing formula

$$\int_0^\infty Q_s^2 f \frac{ds}{s} = f \in L_W^2.$$

Remark 2.9. Regarding the last pair of lemmata:

- We will often denote by Q_t the operators satisfying the hypotheses in Lemma 2.8, while we use P_t for “nice” approximate identities (i.e. $P_t f := \varphi_t * f$, with $\varphi \in \mathcal{S}$ radial, and $\int \varphi = 1$).
- It is easy to check if P_t is a nice approximate identity, then $Q_t := t D_i P_t$ satisfies the hypotheses of the first part of Lemma 2.8 (where D_i denotes the partial derivative in any direction x_i).
- We will frequently further assume that the kernel k in Lemma 2.7 (in particular, φ and ψ as above) satisfies $\text{supp } k \subset B_1(0)$; in this case, we shall refer to K_t (in particular, P_t or Q_t) as having a “compactly supported kernel”.
- We will use repeatedly the fact that P_t and Q_s commute with derivatives, for they are convolution operators.

The following is an easy consequence of Lemma 2.8, by standard “almost-orthogonality” arguments. We omit the well-known proof.

Lemma 2.10. *Let $\{Q_s\}_{s>0}$ be a family operators satisfying the conditions in Lemma 2.8, and R_t be a family of operators, bounded on $L_W^2(\mathbb{R}^n)$ for each fixed $t > 0$, and satisfying, for some $\alpha > 0$, the almost-orthogonality condition*

$$\|R_t Q_s\|_{L_W^2 \rightarrow L_W^2} \leq C_1 \min \left\{ \frac{t}{s}, \frac{s}{t} \right\}^\alpha,$$

where C_1 is a uniform constant which does not depend on t, s . Then R_t satisfies the square function estimate

$$\int_0^\infty \|R_t f\|_{L_W^2}^2 \frac{dt}{t} \lesssim \|f\|_{L_W^2}^2,$$

where the implicit constant depends on n, C_1, α and $[W]_{A_2}$.

2.3. Uniform bounds for some operators. Let us show that some operators related with the semigroups are uniformly bounded, which we will use throughout the text.

Lemma 2.11. *The following operators are $L_W^2 \rightarrow L_W^2$ bounded, uniformly on t , with norm depending on n, λ and the doubling constant for W :*

$$\begin{aligned} (i) \quad & e^{-t^2 L}, & (ii) \quad & t^2 L e^{-t^2 L}, & (iii) \quad & t^4 L^2 e^{-t^2 L}, & (iv) \quad & t \nabla e^{-t^2 L}, \\ (v) \quad & e^{-t^2 \tilde{L}}, & (vi) \quad & t^2 \tilde{L} e^{-t^2 \tilde{L}}, & (vii) \quad & t^4 \tilde{L}^2 e^{-t^2 \tilde{L}}, & (viii) \quad & t \nabla e^{-t^2 \tilde{L}}, \\ (ix) \quad & t e^{-t^2 L} \widetilde{\text{div}}, & (x) \quad & t e^{-t^2 \tilde{L}} \widetilde{\text{div}}, & (xi) \quad & t^3 L e^{-t^2 L} \widetilde{\text{div}}, & (xii) \quad & t^3 \tilde{L} e^{-t^2 \tilde{L}} \widetilde{\text{div}}. \end{aligned}$$

Remark 2.12. Some of the operators are in fact defined for functions $\mathbf{f} \in (L_W^2)^n$. In this case, we obviously mean that each of their components are bounded.

Proof. (i) - (iii). Recall that $\mathcal{M}_W$ denotes the Hardy-Littlewood maximal operator with respect to the weighted measure $W(x)dx$. If we let T_t denote the operator under consideration in (i), (ii) or (iii), it suffices to observe that in each case we have the pointwise bound

$$\sup_t |T_t f| \lesssim \mathcal{M}_W f,$$

using only the Gaussian bounds in Lemma 2.1 (and Remark 2.6 in the case of (ii) and (iii)), and the doubling property of W . We omit the routine details.

To prove (iv), we set $u := e^{-t^2 L} f$. Note that by the identity (1.6), we can “interpolate” the estimates in (i) and (ii) using Cauchy-Schwarz as follows:

$$\begin{aligned} \int_{\mathbb{R}^n} |t \nabla u|^2 W dx &\lesssim \int_{\mathbb{R}^n} t^2 A \nabla u \cdot \nabla u W dx = \int_{\mathbb{R}^n} t^2 L u u W dx \\ &\leq \left(\int_{\mathbb{R}^n} |t^2 L u|^2 W dx \right)^{1/2} \left(\int_{\mathbb{R}^n} |u|^2 W dx \right)^{1/2} \lesssim \int_{\mathbb{R}^n} |f|^2 W dx, \end{aligned}$$

which shows (iv).

In turn, (v)-(vii) follow by duality from (i)-(iii), while (viii) follows “interpolating” (v) and (vi) in the same way that we did it for (iv), this time using (1.11) instead of (1.6).

Then, (ix) and (x) follow by duality from (viii), and (iv), respectively.

Lastly, (xi) (resp. (xii)) follows by first dualizing, and then “interpolating” (vi) and (vii) using (1.11) (resp. (ii) and (iii) using (1.6)). $\square$

2.4. Off-diagonal estimates.

Definition 2.13. We say that the operators S_t satisfy **off-diagonal estimates** (aka Gaffney estimates), if there exist $c, C > 0$ (independent of t) such that it holds for every $t > 0$, and every pair of measurable sets E and F ,

$$\int_F |S_t f(x)|^2 W(x) dx \leq C e^{-c \frac{\text{dist}(E, F)^2}{t^2}} \int_E |f(x)|^2 W(x) dx \quad \text{if } \text{supp } f \subset E.$$

Lemma 2.14. *The following operators satisfy off-diagonal estimates, with constants depending only on n, λ and the doubling constant of W :*

$$\begin{aligned}
(i) \quad & e^{-t^2 L}, & (ii) \quad & t^2 L e^{-t^2 L}, & (iii) \quad & t^4 L^2 e^{-t^2 L}, & (iv) \quad & t \nabla e^{-t^2 L}, \\
(v) \quad & e^{-t^2 \tilde{L}}, & (vi) \quad & t^2 \tilde{L} e^{-t^2 \tilde{L}}, & (vii) \quad & t^4 \tilde{L}^2 e^{-t^2 \tilde{L}}, & (viii) \quad & t e^{-t^2 \tilde{L}} \widetilde{\operatorname{div}}, \\
(ix) \quad & t \nabla e^{-t^2 \tilde{L}}, & (x) \quad & t e^{-t^2 L} \widetilde{\operatorname{div}}.
\end{aligned}$$

Proof. Fix sets $E, F \subset \mathbb{R}^n$. We may assume that $d := \operatorname{dist}(E, F) > 5t$, as otherwise we may invoke Lemma 2.11.

The bound for (i) is a straightforward consequence of the pointwise bounds in Lemma 2.1. We omit the routine details. The off-diagonal estimates for (ii)-(iii) follow in the same way as for (i), using Remark 2.6.

Let us now treat (iv). We argue as in the proof of Caccioppoli's inequality. Let $u := e^{-t^2 L} f$, with f supported in E . Choose $\psi \in \mathcal{C}^\infty(\mathbb{R}^n)$, where $0 \leq \psi \leq 1$ satisfies $\psi \equiv 1$ on F , $\operatorname{dist}(\operatorname{supp} \psi, E) \geq d/2$ (denoting $d := \operatorname{dist}(E, F)$), and $\|\psi\|_\infty + d \|\nabla \psi\|_\infty + d^2 \|\nabla^2 \psi\|_\infty \leq C$. For future reference, we note that

$$(2.15) \quad |\nabla \psi|^2 + |\nabla^2 \psi| \lesssim d^{-2} \mathbf{1}_{\operatorname{supp} \psi}.$$

Clearly, we have

$$(2.16) \quad \int_F |t \nabla u|^2 W dx \leq t^2 \int_{\mathbb{R}^n} |\nabla u|^2 \psi^2 W dx.$$

Compute, using (1.6) and the symmetry of A :

$$\begin{aligned}
(2.17) \quad & \int_{\mathbb{R}^n} A \nabla(u\psi) \cdot \nabla(u\psi) W dx = \int_{\mathbb{R}^n} L(u\psi) u \psi W = - \sum_{i,j=1}^n \int_{\mathbb{R}^n} a_{ij} D_i D_j(u\psi) u \psi W dx \\
& = \sum_{i,j=1}^n \left(- \int_{\mathbb{R}^n} a_{ij} D_i D_j u u \psi^2 W dx - 2 \int_{\mathbb{R}^n} a_{ij} D_i u D_j \psi u \psi W dx - \int_{\mathbb{R}^n} a_{ij} D_i D_j \psi u^2 \psi W dx \right) \\
& = \int_{\mathbb{R}^n} L u u \psi^2 W dx - 2 \int_{\mathbb{R}^n} A \nabla u \cdot \nabla \psi u \psi W dx + \int_{\mathbb{R}^n} L \psi u^2 \psi W dx.
\end{aligned}$$

Also,

$$\begin{aligned}
(2.18) \quad & \int_{\mathbb{R}^n} A \nabla(u\psi) \cdot \nabla(u\psi) W dx \geq \lambda \int_{\mathbb{R}^n} |\nabla(u\psi)|^2 W dx \\
& = \lambda \left(\int_{\mathbb{R}^n} |\nabla u|^2 \psi^2 W dx + 2 \int_{\mathbb{R}^n} \nabla u \cdot \nabla \psi u \psi W dx + \int_{\mathbb{R}^n} |\nabla \psi|^2 u^2 W dx \right) \\
& \geq \lambda \int_{\mathbb{R}^n} |\nabla u|^2 \psi^2 W dx + 2\lambda \int_{\mathbb{R}^n} \nabla u \cdot \nabla \psi u \psi W dx.
\end{aligned}$$

Combining (2.17) and (2.18), we may dominate the right hand side of (2.16) by

$$\begin{aligned}
(2.19) \quad & t^2 \int_{\mathbb{R}^n} |\nabla u|^2 \psi^2 W dx \\
& \lesssim t^2 \int_{\mathbb{R}^n} |\nabla u| |\nabla \psi| |u \psi| W dx + t^2 \int_{\mathbb{R}^n} |L u u \psi^2| W + t^2 \int_{\mathbb{R}^n} |L \psi u^2 \psi| W \\
& \quad =: I + II + III.
\end{aligned}$$

For $\varepsilon > 0$ to be chosen momentarily, we have

$$I \lesssim \varepsilon t^2 \int_{\mathbb{R}^n} |\nabla u|^2 \psi^2 W dx + \varepsilon^{-1} \left(\frac{t}{d} \right)^2 \int_{\text{supp } \psi} u^2 W dx,$$

where we have used “Cauchy’s inequality with ε ” in term I , and then (2.15). Choosing ε small enough, we may hide the first term, and use the off-diagonal bounds for (i) to obtain the desired bound for the second term, since $t \leq d$. We estimate term II using the Cauchy-Schwarz inequality, along with the off-diagonal bounds for (i) and (ii).

The operators (v)-(viii) are dual to the first four, and (x) is dual to (ix). It therefore remains to treat (ix). To this end, we now set $u := e^{-t^2 \tilde{L}} f$, with f supported in E , so with ψ as above, again using (2.16), we have

$$\begin{aligned} \int_F |t \nabla u|^2 W dx &\leq \int_{\mathbb{R}^n} |t \nabla u|^2 \psi^2 W dx \lesssim t^2 \int_{\mathbb{R}^n} a_{ij} D_j u D_i u \psi^2 W dx \\ &= t^2 \int_{\mathbb{R}^n} D_j u D_i (a_{ij} W u) \psi^2 dx - t^2 \int_{\mathbb{R}^n} u D_j u D_i (a_{ij} W) \psi^2 dx =: \tilde{I} + \tilde{II}, \end{aligned}$$

where summation over i, j is understood. We first note that

$$\begin{aligned} \tilde{I} &= -t^2 \int_{\mathbb{R}^n} u D_j D_i (a_{ij} W u) \psi^2 dx - 2t^2 \int_{\mathbb{R}^n} u D_i (a_{ij} W u) \psi D_j \psi dx \\ &= t^2 \int_{\text{supp } \psi} u \tilde{L} u \psi^2 W dx - 2t^2 \int_{\mathbb{R}^n} u D_i (a_{ij} W u) \psi D_j \psi dx =: \tilde{I}_1 + \tilde{I}_2. \end{aligned}$$

We obtain the desired estimate for $\tilde{I}_1$ by the off-diagonal bounds for (v) and (vi) like for term II in (2.19). Also, integrating by parts,

$$\tilde{I}_2 = 2t^2 \int_{\mathbb{R}^n} u D_i u a_{ij} \psi D_j \psi W dx + 2t^2 \int_{\mathbb{R}^n} u^2 a_{ij} (D_i \psi D_j \psi + \psi D_i D_j \psi) W dx =: \tilde{I}_{2,1} + \tilde{I}_{2,2}.$$

We will cancel term $\tilde{I}_{2,1}$ momentarily, but it may also be handled directly, exactly like term I in (2.19), using Cauchy’s inequality with ε , hiding the small term, and bounding the other term using the off-diagonal bounds for (v). We estimate term $\tilde{I}_{2,2}$ like terms I and III in (2.19), using (2.15), the fact that $t \leq d$, and the off-diagonal bounds for (v).

Since W is an adjoint solution,

$$\begin{aligned} \tilde{II} &= t^2 \int u^2 \psi D_j \psi D_i (a_{ij} W) dx \\ &= -2t^2 \int u D_i u \psi D_j \psi a_{ij} W dx - t^2 \int u^2 (D_i \psi D_j \psi + \psi D_i D_j \psi) a_{ij} W dx =: \tilde{II}_1 + \tilde{II}_2. \end{aligned}$$

Observe that $\tilde{II}_1 \equiv -\tilde{I}_{2,1}$, and $\tilde{II}_2 \equiv -\frac{1}{2} \tilde{I}_{2,2}$. $\square$

Lemma 2.20. *Let K_t be a convolution type operator as in Lemma 2.7, with a compactly supported kernel. If $h = h(x, t)$ is a function such that $f \mapsto h(\cdot, t) K_t f$ is bounded on L_W^2 , uniformly in t (i.e. $\|h(\cdot, t) K_t\|_{L_W^2 \rightarrow L_W^2} \lesssim 1$ uniformly in t), then $h(\cdot, t) K_t$ satisfies off-diagonal estimates.*

We omit the trivial proof.

Lemma 2.21. *Let $\{U_t\}_{t>0}$ and $\{U'_t\}_{t>0}$ be two families of operators, each satisfying off-diagonal estimates, then the composition $U_t U'_t$ also satisfies off-diagonal estimates for each t .*

We omit the routine proof.

2.5. Estimates for differences and gradients.

Lemma 2.22. *For $f \in \text{Lip}$ and $t \leq \ell(Q)$, we have*

$$\int_Q \left| e^{-t^2 L} f(x) - f(x) \right|^2 W(x) dx \lesssim t^2 \|\nabla f\|_\infty^2 W(Q)$$

and

$$\int_Q \left| e^{-t^2 \tilde{L}} f(x) - f(x) \right|^2 W(x) dx \lesssim t^2 \|\nabla f\|_\infty^2 W(Q),$$

where the implicit constants depend on n, λ and $[W]_{A_2}$.

Proof. Let us show the first estimate. Cover Q by non-overlapping cubes Q_k with sidelength $t/2 < \ell(Q_k) \leq t$. Note that $e^{-t^2 L} \mathbf{1} = \mathbf{1}$, since $L\mathbf{1} = 0$. Letting $[f]_E$ denote the average of f over the set E ,

$$(2.23) \quad \left\| e^{-t^2 L} f - f \right\|_{L_W^2(Q)}^2 = \sum_k \left\| e^{-t^2 L} (f - [f]_{2Q_k}) - (f - [f]_{2Q_k}) \right\|_{L_W^2(Q_k)}^2 =: \sum_k A_k^2$$

We then have

$$(2.24) \quad A_k \leq \left\| e^{-t^2 L} ((f - [f]_{2Q_k}) \mathbf{1}_{2Q_k}) \right\|_{L_W^2(Q_k)} + \|f - [f]_{2Q_k}\|_{L_W^2(Q_k)} \\ + \sum_{j=1}^{\infty} \left\| e^{-t^2 L} ((f - [f]_{2Q_k}) \mathbf{1}_{2^{j+1}Q_k \setminus 2^j Q_k}) \right\|_{L_W^2(Q_k)} =: I^{(k)} + \sum_{j=1}^{\infty} II_j^{(k)}.$$

Using the boundedness of $e^{-t^2 L}$ from Lemma 2.11 and the weighted version of Poincaré's inequality [HKM, 15.26], we deduce

$$I^{(k)} \lesssim \|f - [f]_{2Q_k}\|_{L_W^2(2Q_k)} \lesssim \ell(Q_k) \|\nabla f\|_{L_W^2(2Q_k)} \leq t \|\nabla f\|_{L_W^2(2Q_k)} \lesssim t \|\nabla f\|_\infty \sqrt{W(Q_k)}.$$

For convenience of notation in the rest of this argument, we replace the constant c by $4c$ in the off-diagonal estimates for $e^{-t^2 L}$ from Lemma 2.14. Thus,

$$II_j^{(k)} \lesssim e^{-c4^{j+1} \frac{\ell(Q_k)^2}{t^2}} \|f - [f]_{2Q_k}\|_{L_W^2(2^{j+1}Q_k)} \\ \leq e^{-c4^j} \|f - [f]_{2^{j+1}Q_k}\|_{L_W^2(2^{j+1}Q_k)} + \sum_{i=1}^j e^{-c4^j} \|[f]_{2^{i+1}Q_k} - [f]_{2^i Q_k}\|_{L_W^2(2^{j+1}Q_k)} \\ \lesssim e^{-c4^j} \|f - [f]_{2^{j+1}Q_k}\|_{L_W^2(2^{j+1}Q_k)} + \sum_{i=1}^j e^{-c4^j} C_D^{(j-i)/2} \|f - [f]_{2^{i+1}Q_k}\|_{L_W^2(2^{i+1}Q_k)}.$$

Hence, summing on j and interchanging the order of summation we obtain

$$\sum_{j=1}^{\infty} \sum_{i=1}^j e^{-c4^j} C_D^{(j-i)/2} \|f - [f]_{2^{i+1}Q_k}\|_{L_W^2(2^{i+1}Q_k)} \lesssim \sum_{i=1}^{\infty} e^{-\frac{c}{2}4^i} \|f - [f]_{2^{i+1}Q_k}\|_{L_W^2(2^{i+1}Q_k)},$$

and therefore we have, by Poincaré's inequality,

$$\begin{aligned}
\sum_{j=1}^{\infty} II_j^{(k)} &\lesssim \sum_{j=1}^{\infty} e^{-\frac{\varepsilon}{2}4^j} \|f - [f]_{2^{j+1}Q_k}\|_{L_W^2(2^{j+1}Q_k)} \lesssim \sum_{j=1}^{\infty} e^{-\frac{\varepsilon}{2}4^j} 2^j \ell(Q_k) \|\nabla f\|_{L_W^2(2^{j+1}Q_k)} \\
&\lesssim t \sum_{j=1}^{\infty} e^{-\frac{\varepsilon}{4}4^j} \|\nabla f\|_{\infty} \sqrt{W(2^{j+1}Q_k)} \lesssim t \|\nabla f\|_{\infty} \sum_{j=1}^{\infty} e^{-\frac{\varepsilon}{5}4^j} \sqrt{W(Q_k)} \lesssim t \|\nabla f\|_{\infty} \sqrt{W(Q_k)}.
\end{aligned}$$

Thus, going back to A_k (see (2.23)-(2.24)) we obtain

$$A_k \lesssim t \|\nabla f\|_{\infty} \sqrt{W(Q_k)},$$

and therefore, since the cubes Q_k are non-overlapping,

$$\left\| e^{-t^2 L} f - f \right\|_{L_W^2(Q)}^2 \lesssim t^2 \|\nabla f\|_{\infty}^2 \sum_k W(Q_k) = t^2 \|\nabla f\|_{\infty}^2 W(Q).$$

The corresponding estimate for $e^{-t^2 \tilde{L}}$ holds with exactly the same proof, for the operator is also bounded and satisfies off-diagonal estimates (see Lemmas 2.11 and 2.14), and $e^{-t^2 \tilde{L}} \mathbf{1} = \mathbf{1}$ since $\tilde{L} \mathbf{1} = \frac{1}{W} L^* W = 0$. $\square$

Lemma 2.25. *For $f \in \text{Lip}$ and $t \leq \ell(Q)$, we have*

$$\int_Q \left| \nabla e^{-t^2 L} f(x) \right|^2 W(x) dx \lesssim \|\nabla f\|_{\infty}^2 W(Q)$$

and

$$\int_Q \left| \nabla e^{-t^2 \tilde{L}} f(x) \right|^2 W(x) dx \lesssim \|\nabla f\|_{\infty}^2 W(Q),$$

where the implicit constant depends on n, λ and $[W]_{A_2}$.

Proof. The reader can check that the proof is the same as for Lemma 2.22, this time using the operators $t \nabla e^{-t^2 L}$ and $t \nabla e^{-t^2 \tilde{L}}$. Indeed, the proof of the preceeding Lemma used only the boundedness of the operator (Lemma 2.11), the off-diagonal estimates (Lemma 2.14), and the conservation property which allows us to subtract constants. In this case,

$$\left\| \nabla e^{-t^2 L} f \right\|_{L_W^2(Q)}^2 = \sum_k \left\| \nabla e^{-t^2 L} (f - [f]_{2Q_k}) \right\|_{L_W^2(Q_k)}^2,$$

and similarly for $\tilde{L}$, because $e^{-t^2 L} \mathbf{1} = \mathbf{1} = e^{-t^2 \tilde{L}} \mathbf{1}$, as before. After this, the proof is the same as that of Lemma 2.22. $\square$

3. THE KATO PROBLEM FOR L

In this section our goal is to prove the following result:

Theorem 3.1. *It holds*

$$\left\| \sqrt{L} f \right\|_{L_W^2} \lesssim \|\nabla f\|_{L_W^2},$$

where the hidden constant depends on n, λ and $[W]_{A_2}$.

As noted above (see Remark 1.16), in the case that the coefficient matrix has small enough BMO norm, we could deduce Theorem 3.1 as an easy consequence of certain known results. Instead, following the easier part of our proof of Theorem 4.1 below, we shall give a self-contained, direct argument, which does not rely on an explicit assumption of smallness in BMO, but only on the validity of (1.6), and the assumption that $W \in A_2$.

To prove the theorem, let us use the representation of the square root operator via the Functional Calculus formula

$$\sqrt{L}f = a \int_0^\infty t^3 L^2 e^{-2t^2 L} f \frac{dt}{t},$$

where $a = (\int_0^\infty t^3 e^{-2t^2} \frac{dt}{t})^{-1} = \sqrt{\frac{128}{\pi}}$ is just a normalizing constant, to estimate, using duality and later Cauchy-Schwarz,

$$\left| \langle \sqrt{L}f, g \rangle_{L_W^2} \right|^2 \leq a^2 \left(\int_0^\infty \left\| t L e^{-t^2 L} f(x) \right\|_{L_W^2}^2 \frac{dt}{t} \right) \left(\int_0^\infty \left\| t^2 \tilde{L} e^{-t^2 \tilde{L}} g(x) \right\|_{L_W^2}^2 \frac{dt}{t} \right).$$

With this decomposition, we will finish the proof of Theorem 3.1 by duality once we prove the following Lemma 3.2 and Lemma 3.6.

The desired bound for the second factor is the following.

Lemma 3.2. *It holds*

$$(3.3) \quad \int_0^\infty \int_{\mathbb{R}^n} \left| t^2 \tilde{L} e^{-t^2 \tilde{L}} g(x) \right|^2 W(x) dx \frac{dt}{t} \lesssim \|g\|_{L_W^2}^2,$$

where the implicit constant depends only on n, λ and the doubling constant for W .

We could obtain the conclusion of the lemma by invoking the abstract McIntosh and Yagi theorem [Y], [Mc], but instead, we will give a self-contained and more elementary proof, using quasi-orthogonality arguments.

Proof. We abbreviate $V_t := t^2 L e^{-t^2 L}$. Its adjoint within L_W^2 is $\tilde{V}_t := t^2 \tilde{L} e^{-t^2 \tilde{L}}$. To make the argument rigorous, given a small positive ε , we set $V_t \equiv 0 \equiv \tilde{V}_t$ whenever $t \leq \varepsilon$ or $t \geq 1/\varepsilon$, and we obtain quantitative bounds that are uniform in ε . We compute, using duality, Fubini and Cauchy-Schwarz,

$$(3.4) \quad \begin{aligned} \int_0^\infty \int_{\mathbb{R}^n} \left| \tilde{V}_t g(x) \right|^2 W(x) dx \frac{dt}{t} &= \int_0^\infty \int_{\mathbb{R}^n} \tilde{V}_t g(x) \tilde{V}_t g(x) W(x) dx \frac{dt}{t} \\ &= \int_0^\infty \int_{\mathbb{R}^n} g(x) V_t \tilde{V}_t g(x) W(x) dx \frac{dt}{t} = \int_{\mathbb{R}^n} \int_0^\infty V_t \tilde{V}_t g(x) \frac{dt}{t} g(x) W(x) dx \\ &\leq \|g\|_{L_W^2} \left\| \int_0^\infty V_t \tilde{V}_t g \frac{dt}{t} \right\|_{L_W^2}. \end{aligned}$$

To deal with the second term, let us first establish a useful fact:

Claim 3.5. *It holds, for any $t, s > 0$,*

$$\left\| \tilde{V}_t V_s \right\|_{L_W^2 \rightarrow L_W^2} \lesssim \min \left\{ \frac{s}{t}, \frac{t}{s} \right\}.$$

Proof of the claim. We may assume that $\varepsilon < s, t < 1/\varepsilon$. If $s \leq t$, we can compute using (1.18) and the uniform bounds from Lemma 2.11,

$$\begin{aligned} \left\| \tilde{V}_t V_s \right\| &= \left\| t^2 s^2 \tilde{L} e^{-t^2 \tilde{L}} L e^{-s^2 L} \right\| \lesssim \left\| t^2 s^2 \tilde{L} e^{-t^2 \tilde{L}} \tilde{L} e^{-s^2 L} \right\| + \left\| t^2 s^2 \tilde{L} e^{-t^2 \tilde{L}} \operatorname{div} A \nabla e^{-s^2 L} \right\| \\ &\leq \frac{s^2}{t^2} \left\| t^4 \tilde{L}^2 e^{-t^2 \tilde{L}} \right\| \left\| e^{-s^2 L} \right\| + \frac{s}{t} \left\| t^3 \tilde{L} e^{-t^2 \tilde{L}} \operatorname{div} \right\| \|A\| \left\| s \nabla e^{-s^2 L} \right\| \lesssim \frac{s^2}{t^2} + \frac{s}{t} \lesssim \frac{s}{t}. \end{aligned}$$

In a similar fashion we can compute, when $s > t$, using again (1.18) and Lemma 2.11,

$$\begin{aligned}
\|\tilde{V}_t V_s\| &= \|t^2 s^2 e^{-t^2 \tilde{L}} \tilde{L} L e^{-s^2 L}\| \lesssim \|t^2 s^2 e^{-t^2 \tilde{L}} L L e^{-s^2 L}\| + \|t^2 s^2 e^{-t^2 \tilde{L}} \operatorname{div} A \nabla L e^{-s^2 L}\| \\
&\leq \frac{t^2}{s^2} \|e^{-t^2 \tilde{L}}\| \|s^4 L^2 e^{-s^2 L}\| + \frac{t}{s} \|t e^{-t^2 \tilde{L}} \operatorname{div}\| \|A\| \|s^3 \nabla L e^{-s^2 L}\| \lesssim \frac{t^2}{s^2} + \frac{t}{s} \lesssim \frac{t}{s}.
\end{aligned}$$

□

With this almost-orthogonality result, we can estimate the last term in (3.4) as follows:

$$\begin{aligned}
\left\| \int_0^\infty V_t \tilde{V}_t g \frac{dt}{t} \right\|_{L_W^2}^2 &= \int_{\mathbb{R}^n} \left(\int_0^\infty V_t \tilde{V}_t g(x) \frac{dt}{t} \right) \left(\int_0^\infty V_s \tilde{V}_s g(x) \frac{ds}{s} \right) W(x) dx \\
&= \int_0^\infty \int_0^\infty \int_{\mathbb{R}^n} V_t \tilde{V}_t g(x) V_s \tilde{V}_s g(x) W(x) dx \frac{dt}{t} \frac{ds}{s} \\
&= \int_0^\infty \int_0^\infty \int_{\mathbb{R}^n} \tilde{V}_t g(x) \tilde{V}_t V_s \tilde{V}_s g(x) W(x) dx \frac{dt}{t} \frac{ds}{s} \\
&\leq \int_0^\infty \int_0^\infty \left(\int_{\mathbb{R}^n} |\tilde{V}_t g|^2 W dx \right)^{1/2} \left(\int_{\mathbb{R}^n} |\tilde{V}_t V_s \tilde{V}_s g|^2 W dx \right)^{1/2} \frac{dt}{t} \frac{ds}{s} \\
&\leq \int_0^\infty \int_0^\infty \min \left\{ \frac{s}{t}, \frac{t}{s} \right\} \left(\int_{\mathbb{R}^n} |\tilde{V}_t g|^2 W dx \right)^{1/2} \left(\int_{\mathbb{R}^n} |\tilde{V}_s g|^2 W dx \right)^{1/2} \frac{dt}{t} \frac{ds}{s} \\
&\lesssim \int_0^\infty \int_0^\infty \min \left\{ \frac{s}{t}, \frac{t}{s} \right\} \int_{\mathbb{R}^n} |\tilde{V}_t g(x)|^2 W(x) dx \frac{dt}{t} \frac{ds}{s} \\
&\lesssim \int_0^\infty \int_{\mathbb{R}^n} |\tilde{V}_t g(x)|^2 W(x) dx \frac{dt}{t},
\end{aligned}$$

Plugging this estimate into (3.4), we have

$$\int_0^\infty \int_{\mathbb{R}^n} |\tilde{V}_t g(x)|^2 W(x) dx \frac{dt}{t} \lesssim \|g\|_{L_W^2} \left(\int_0^\infty \int_{\mathbb{R}^n} |\tilde{V}_t g(x)|^2 W(x) dx \frac{dt}{t} \right)^{1/2},$$

from which the result readily follows (recall that we have effectively truncated so $\varepsilon < t < 1/\varepsilon$, hence the integrals are finite). □

Let us turn our attention to the other square function estimate.

Lemma 3.6. *It holds*

$$(3.7) \quad \int_0^\infty \int_{\mathbb{R}^n} |t L e^{-t^2 L} f(x)|^2 W(x) dx \frac{dt}{t} \lesssim \|\nabla f\|_{L_W^2}^2,$$

where the implicit constant depends on n, λ and $[W]_{A_2}$.

We will devote the rest of the section to the proof of Lemma 3.6. As above, set $V_t := t^2 L e^{-t^2 L}$ and decompose, with the help of an approximate identity P_t ,

$$(3.8) \quad t L e^{-t^2 L} = t^{-1} V_t (I - P_t) + t^{-1} V_t P_t =: R_t + T_t.$$

The proof of Lemma 3.6, and hence of Theorem 3.1, will come immediately from the next two lemmas.

Lemma 3.9. *With the notations of (3.8), we have*

$$\int_0^\infty \|R_t f\|_{L_W^2}^2 \frac{dt}{t} \lesssim \|\nabla f\|_{L_W^2}^2,$$

where the implicit constant depends on n, λ and $[W]_{A_2}$.

Proof. Choose the approximate identity $P_t := e^{-t^2(-\Delta)}$. With it, we can compute, using the Fundamental Theorem of Calculus,

$$\begin{aligned} R_t f &= t^{-1} V_t (I - P_t) f = V_t \left(\frac{1}{t} (I - P_t) f \right) = V_t \left(-\frac{1}{t} \int_0^t \frac{\partial}{\partial s} P_s f ds \right) \\ &= -2V_t \left(\frac{1}{t} \int_0^t s P_s \Delta f ds \right) = -2V_t \left(\frac{1}{t} \int_0^t s P_s \operatorname{div} \nabla f ds \right) =: 2V_t \left(\frac{1}{t} \int_0^t \vec{Q}_s \nabla f ds \right). \end{aligned}$$

Now, using the boundedness on L_W^2 of $V_t = t^2 L e^{-t^2 L}$ (see Lemma 2.11), Hardy's inequality and the fact that $\vec{Q}_s$ satisfies the square function estimate of Lemma 2.8 (see Remark 2.9), we obtain the desired estimate:

$$\begin{aligned} \int_0^\infty \|R_t f\|_{L_W^2}^2 \frac{dt}{t} &= \int_0^\infty \int_{\mathbb{R}^n} \left| 2V_t \left(\frac{1}{t} \int_0^t \vec{Q}_s \nabla f ds \right) (x) \right|^2 W(x) dx \frac{dt}{t} \\ &\lesssim \int_0^\infty \int_{\mathbb{R}^n} \left| \frac{1}{t} \int_0^t \vec{Q}_s \nabla f(x) ds \right|^2 W(x) dx \frac{dt}{t} \lesssim \int_0^\infty \int_{\mathbb{R}^n} \left| \vec{Q}_t \nabla f(x) \right|^2 W(x) dx \frac{dt}{t} \lesssim \|\nabla f\|_{L_W^2}^2. \end{aligned}$$

□

Lemma 3.10. *With the notations of (3.8), we have*

$$\int_0^\infty \|T_t f\|_{L_W^2}^2 \frac{dt}{t} \lesssim \|\nabla f\|_{L_W^2}^2,$$

where the hidden constant depends on n, λ and $[W]_{A_2}$.

Proof. Simply compute, using the boundedness of $e^{-t^2 L}$ from Lemma 2.11, and the square function bounds of Lemma 2.8 (see Remark 2.9),

$$\begin{aligned} \int_0^\infty \|T_t f\|_{L_W^2}^2 \frac{dt}{t} &\lesssim \int_0^\infty \|t L P_t f\|_{L_W^2}^2 \frac{dt}{t} \lesssim \sum_{i,j=1}^n \int_0^\infty \|t D_i D_j P_t f\|_{L_W^2}^2 \frac{dt}{t} \\ &= \sum_{i,j=1}^n \int_0^\infty \|t D_i P_t D_j f\|_{L_W^2}^2 \frac{dt}{t} =: \sum_{i,j=1}^n \int_0^\infty \|Q_t^{(i)} D_j f\|_{L_W^2}^2 \frac{dt}{t} \\ &\lesssim \sum_{j=1}^n \|D_j f\|_{L_W^2}^2 = \|\nabla f\|_{L_W^2}^2. \end{aligned}$$

□

4. THE KATO PROBLEM FOR $\tilde{L}$

In this section our goal is to prove the following result, which is really the main result in this paper:

Theorem 4.1. *It holds*

$$\left\| \sqrt{\tilde{L}} f \right\|_{L_W^2} \lesssim \|\nabla f\|_{L_W^2},$$

where the implicit constant depends on n, λ and $[W]_{A_2}$.

4.1. Reduction to a quadratic estimate. To prove Theorem 4.1, let us again use the representation of the square root operator via the formula

$$\sqrt{\tilde{L}}f = a \int_0^\infty t^3 \tilde{L}^2 e^{-2t^2 \tilde{L}} f \frac{dt}{t},$$

so that

$$\left| \langle \sqrt{\tilde{L}}f, g \rangle_{L_W^2} \right|^2 \leq a^2 \left(\int_0^\infty \left\| t \tilde{L} e^{-t^2 \tilde{L}} f(x) \right\|_{L_W^2}^2 \frac{dt}{t} \right) \left(\int_0^\infty \left\| t^2 L e^{-t^2 L} g(x) \right\|_{L_W^2}^2 \frac{dt}{t} \right).$$

Theorem 4.1 then follows immediately from Lemma 4.2 and Theorem 4.4 below.

We estimate the second square function via the following lemma.

Lemma 4.2. *It holds*

$$(4.3) \quad \int_0^\infty \int_{\mathbb{R}^n} \left| t^2 L e^{-t^2 L} g(x) \right|^2 W(x) dx \frac{dt}{t} \lesssim \|g\|_{L_W^2}^2,$$

where the implicit constant depends on n, λ and $[W]_{A_\infty}$.

Proof. The lemma can be proved either by invoking the McIntosh and Yagi theorem [Y], [Mc], or via a self-contained elementary proof using quasi-orthogonality. For the latter path, the proof follows that of Lemma 3.2 *mutatis mutandis*, simply reversing the roles of V_t and $\tilde{V}_t$. We omit the details. $\square$

Let us now turn the attention to the other square function estimate, which is in fact the core of this paper.

Theorem 4.4. *It holds*

$$(4.5) \quad \int_0^\infty \int_{\mathbb{R}^n} \left| t \tilde{L} e^{-t^2 \tilde{L}} f(x) \right|^2 W(x) dx \frac{dt}{t} \lesssim \|\nabla f\|_{L_W^2}^2,$$

where the implicit constant depends on n, λ and $[W]_{A_2}$.

The rest of this section is devoted to the proof of Theorem 4.4. We start by splitting

$$(4.6) \quad t \tilde{L} e^{-t^2 \tilde{L}} = t \tilde{L} e^{-t^2 \tilde{L}} (I - P_t^2) + t \tilde{L} e^{-t^2 \tilde{L}} P_t^2 =: \tilde{R}_t + \tilde{T}_t,$$

where P_t is a nice approximate identity with a smooth compactly supported convolution kernel $\varphi_t(x) = t^{-n} \varphi(x/t)$, which we take to be even. For future reference, let us record the following well-known observation:

$$t \partial_t (\widehat{\varphi}(t\xi))^2 = 2(\nabla \widehat{\varphi})(t\xi) \cdot t\xi \widehat{\varphi}(t\xi) = c(\widehat{x\varphi(x)})(t\xi) \cdot (\widehat{\nabla \varphi})(t\xi) =: c\widehat{\psi^{(1)}}(t\xi) \widehat{\psi^{(2)}}(t\xi),$$

where c is a harmless constant, and $\psi^{(1)}(x) := x\varphi(x)$ and $\psi^{(2)} := \nabla \varphi$ are both $\mathcal{C}_c^\infty$ functions with mean value zero (here we are using that φ is even, in the case of $\psi^{(1)}$). Hence,

$$(4.7) \quad t \partial_t P_t^2 = c Q_t^{(1)} Q_t^{(2)},$$

where $Q_t^{(k)}$ is the convolution kernel with kernel $\psi_t^{(k)}(x) := t^{-n} \psi^{(k)}(x/t)$, $k = 1, 2$, and therefore each of $Q_t^{(1)}$, $Q_t^{(2)}$ satisfies the square function bound of Lemma 2.8 (and each is bounded on L_W^2 , uniformly in t).

Lemma 4.8. *With the notations of (4.6), we have*

$$(4.9) \quad \int_0^\infty \left\| \tilde{R}_t f \right\|_{L_W^2}^2 \frac{dt}{t} \lesssim \|\nabla f\|_{L_W^2}^2,$$

where the implicit constant depends on n, λ and $[W]_{A_2}$.

Proof. Using the preceeding observations, we may follow the proof of Lemma 3.9, invoking Lemma 2.11 to obtain that $t^2 \tilde{L} e^{-t^2 \tilde{L}}$ is L_W^2 bounded, to obtain (4.9). $\square$

Applying now (1.18) to $u = P_t^2 f$ we obtain

$$(4.10) \quad \tilde{T}_t f = t e^{-t^2 \tilde{L}} \tilde{L} P_t^2 f = -t e^{-t^2 \tilde{L}} L P_t^2 f - 2t e^{-t^2 \tilde{L}} \widetilde{\operatorname{div}}(A \nabla (P_t^2 f)).$$

Lemma 4.11. *We have*

$$\int_0^\infty \left\| t e^{-t^2 \tilde{L}} L P_t^2 f \right\|_{L_W^2}^2 \frac{dt}{t} \lesssim \|\nabla f\|_{L_W^2}^2,$$

where the implicit constant depends on n, λ and $[W]_{A_2}$.

Proof. The proof is the same as that in Lemma 3.10 once we use that $e^{-t^2 \tilde{L}} : L_W^2 \rightarrow L_W^2$ is uniformly bounded by Lemma 2.11, and that P_t are uniformly bounded on L_W^2 by Lemma 2.7. $\square$

Therefore, to finish the proof of Theorem 4.4 (and hence of Theorem 4.1), it remains to show

$$(4.12) \quad \int_0^\infty \left\| t e^{-t^2 \tilde{L}} \widetilde{\operatorname{div}}(A \nabla (P_t^2 f)) \right\|_{L_W^2}^2 \frac{dt}{t} \lesssim \|\nabla f\|_{L_W^2}^2.$$

4.2. Reduction to a Carleson measure estimate. For $\mathbf{g} = (g_1, g_2, \dots, g_n)$, write

$$(4.13) \quad \theta_t \mathbf{g} := t e^{-t^2 \tilde{L}} \widetilde{\operatorname{div}}(A \mathbf{g}) \left(= t e^{-t^2 \tilde{L}} \frac{1}{W} \operatorname{div}(W A \mathbf{g}) \right).$$

With this notation, the remaining estimate (4.12) becomes

$$(4.14) \quad \int_0^\infty \left\| \theta_t \nabla (P_t^2 f) \right\|_{L_W^2}^2 \frac{dt}{t} \lesssim \|\nabla f\|_{L_W^2}^2.$$

Let us also define the operator

$$\tilde{\theta}_t \mathbf{g} := t e^{-t^2 \tilde{L}} \left(\widetilde{\operatorname{div}}(A \mathbf{g}) - \frac{1}{2} \sum_{i,j=1}^n a_{ij} D_i \mathbf{g}_j \right),$$

so that, taking $\mathbf{g} = \nabla u$, and using (1.18),

$$(4.15) \quad \tilde{\theta}_t \nabla u = -\frac{1}{2} t e^{-t^2 \tilde{L}} \tilde{L} u.$$

It will be convenient to use both operators at different stages of the proof. Note that trivially, $\tilde{\theta}_t \mathbf{e} = \theta_t \mathbf{e}$, for any constant vector $\mathbf{e}$. In particular, if $\mathbb{1}$ denotes the $n \times n$ identity matrix, then

$$(4.16) \quad \tilde{\theta}_t \mathbb{1} = \theta_t \mathbb{1},$$

where we naturally define $\tilde{\theta}_t \mathbb{1} = \theta_t \mathbb{1}$ as a vector-valued function whose k^{th} entry is $\tilde{\theta}_t \mathbf{e}^k = \theta_t \mathbf{e}^k$, with $\mathbf{e}^k$ equal to the standard unit basis vector in the x_k direction.

To prove (4.12), as in the divergence form case treated in [AHLMT], we begin with a “T1” reduction.

Lemma 4.17. *We have*

$$\int_0^\infty \left\| \theta_t P_t^2 \mathbf{g} - (\theta_t \mathbb{1}) \cdot (P_t^2 \mathbf{g}) \right\|_{L_W^2}^2 \frac{dt}{t} \lesssim \|\mathbf{g}\|_{L_W^2}^2,$$

where the implicit constant depends on n, λ and $[W]_{A_2}$.

Proof. Write $U_t \mathbf{g} := \theta_t P_t^2 \mathbf{g} - (\theta_t \mathbf{1}) \cdot (P_t^2 \mathbf{g})$. By Lemma 2.10, it suffices to show that

$$\|U_t\|_{L_W^2 \rightarrow L_W^2} \lesssim 1$$

uniformly on t , and for some $\alpha > 0$, and for any nice operator Q_s as in Lemma 2.8, with a compactly supported kernel,

$$\|U_t Q_s\|_{L_W^2 \rightarrow L_W^2} \lesssim \min \left\{ \frac{s}{t}, \frac{t}{s} \right\}^\alpha.$$

These two estimates, and hence the conclusion of Lemma 4.17, will follow at once from the next claims and Lemma 2.7.

Claim 4.18. *We have, uniformly on t ,*

$$\|\theta_t P_t\|_{L_W^2 \rightarrow L_W^2} \lesssim 1,$$

where the implicit constant depends on n, λ and $[W]_{A_2}$.

Proof of the claim. Just compute, with the aid of Lemmas 2.7 and 2.11,

$$\|\theta_t P_t\|_{L_W^2 \rightarrow L_W^2} \leq \left\| t e^{-t^2 \tilde{L}} \widetilde{\operatorname{div}} \right\|_{L_W^2 \rightarrow L_W^2} \|A\|_\infty \|P_t\|_{L_W^2 \rightarrow L_W^2} \lesssim 1.$$

□

Claim 4.19. *We have, uniformly on t ,*

$$\|(\theta_t \mathbf{1}) \cdot P_t\|_{L_W^2 \rightarrow L_W^2} \lesssim 1,$$

where the implicit constant depends on n, λ and $[W]_{A_2}$.

Proof of the claim. This proof will follow that of [CUR, (4.10)]. Let us cover $\mathbb{R}^n$ by cubes Q_k satisfying $t/2 < \ell(Q_k) \leq t$. In this way, we obtain

$$(4.20) \quad \int_{\mathbb{R}^n} |(\theta_t \mathbf{1})(x) \cdot P_t \mathbf{g}(x)|^2 W(x) dx = \sum_k \int_{Q_k} |(\theta_t \mathbf{1})(x) \cdot P_t \mathbf{g}(x)|^2 W(x) dx.$$

We first establish an L^∞ bound for $P_t \mathbf{g}(x)$, in the cube Q_k . Note that for $x \in Q_k$ we have that $P_t \mathbf{g}(x) = P_t(\mathbf{g} \mathbf{1}_{3Q_k})(x)$ because $t \leq 2\ell(Q_k)$ (and $\operatorname{supp} \varphi \subset B(0, 1)$). Hence,

$$(4.21) \quad \begin{aligned} |P_t(\mathbf{g} \mathbf{1}_{3Q_k})(x)|^2 &\leq \left(\|\varphi_t\|_\infty \int_{3Q_k} |\mathbf{g}(y)| dy \right)^2 \leq \left(t^{-n} \int_{3Q_k} |\mathbf{g}(y)| dy \right)^2 \\ &\approx \left(\int_{3Q_k} |\mathbf{g}(y)| dy \right)^2 \leq \frac{1}{|3Q_k|^2} \int_{3Q_k} |\mathbf{g}(y)|^2 W(y) dy \int_{3Q_k} W^{-1}(y) dy \\ &= \frac{W(3Q_k) W^{-1}(3Q_k)}{|3Q_k|^2} \int_{3Q_k} |\mathbf{g}(y)|^2 W(y) dy \lesssim \int_{3Q_k} |\mathbf{g}(y)|^2 W(y) dy \end{aligned}$$

where we used the definition of an A_2 weight in the last step (see Definition 1.4).

We next claim that

$$(4.22) \quad \frac{1}{W(Q_k)} \int_{Q_k} |(\theta_t \mathbf{b})(x)|^2 W(x) dx \lesssim \|\mathbf{b}\|_\infty^2.$$

Taking this claim for granted momentarily, we obtain

$$\int_{\mathbb{R}^n} |(\theta_t \mathbf{1})(x) \cdot P_t \mathbf{g}(x)|^2 W(x) dx \lesssim \sum_k \int_{Q_k} |(\theta_t \mathbf{1})(x)|^2 \int_{3Q_k} |\mathbf{g}(y)|^2 W(y) dy W(x) dx$$

$$\begin{aligned}
&\lesssim \sum_k \frac{1}{W(Q_k)} \int_{Q_k} |(\theta_t \mathbf{1})(x)|^2 W(x) dx \int_{3Q_k} |\mathbf{g}(y)|^2 W(y) dy \\
&\lesssim \sum_k \int_{3Q_k} |\mathbf{g}(y)|^2 W(y) dy \lesssim \int_{\mathbb{R}^n} |\mathbf{g}(y)|^2 W(y) dy,
\end{aligned}$$

where in the last two steps we used first (4.22), and then the bounded overlap property of the cubes $3Q_k$.

It remains to verify (4.22). We dualize: choose $\mathbf{h} = (h_1, h_2, \dots, h_n) \in L_W^2$, with $\text{supp } \mathbf{h} \subset Q_k$, and write

$$\begin{aligned}
(4.23) \quad &\int_{\mathbb{R}^n} \theta_t \mathbf{b}(x) \cdot \mathbf{h}(x) W(x) dx = \int_{\mathbb{R}^n} t e^{-t^2 \tilde{L}} \widetilde{\text{div}} A \mathbf{b}(x) \cdot \mathbf{h}(x) W(x) dx \\
&= \int_{\mathbb{R}^n} A \mathbf{b}(x) \cdot t \nabla e^{-t^2 L} \mathbf{h}(x) W(x) dx \lesssim \|\mathbf{b}\|_\infty \int_{\mathbb{R}^n} |t \nabla e^{-t^2 L} \mathbf{h}(x)| W(x) dx \\
&\leq \|\mathbf{b}\|_\infty \left(\int_{2Q_k} |t \nabla e^{-t^2 L} \mathbf{h}(x)| W(x) dx + \sum_{j=2}^{\infty} \int_{2^j Q_k \setminus 2^{j-1} Q_k} |t \nabla e^{-t^2 L} \mathbf{h}(x)| W(x) dx \right) \\
&=: \|\mathbf{b}\|_\infty \left(I^{(k)} + II^{(k)} \right).
\end{aligned}$$

For the first term we may simply compute, using Jensen's inequality and the boundedness of $t \nabla e^{-t^2 L}$ from Lemma 2.11,

$$I^{(k)} \lesssim \left(W(Q_k) \int_{2Q_k} |t \nabla e^{-t^2 L} \mathbf{h}(x)|^2 W(x) dx \right)^{1/2} \lesssim \sqrt{W(Q_k)} \|\mathbf{h}\|_{L_W^2}.$$

For the second term we use Jensen again, and later the off-diagonal estimates from Lemma 2.14 (taking advantage of $\ell(Q_k) \approx t$) to obtain

$$\begin{aligned}
II^{(k)} &\lesssim \sum_{j=2}^{\infty} \left(W(2^j Q_k) \int_{2^j Q_k \setminus 2^{j-1} Q_k} |t \nabla e^{-t^2 L} \mathbf{h}(x)|^2 W(x) dx \right)^{1/2} \\
&\lesssim \sum_{j=2}^{\infty} \left(C_D^j W(Q_k) e^{-c4^j} \int_{Q_k} |\mathbf{h}(x)|^2 W(x) dx \right)^{1/2} \lesssim \sqrt{W(Q_k)} \|\mathbf{h}\|_{L_W^2}.
\end{aligned}$$

With the estimates for $I^{(k)}$ and $II^{(k)}$, we can substitute back in (4.23) and obtain

$$\int_{\mathbb{R}^n} \theta_t \mathbf{b}(x) \cdot \mathbf{h}(x) W(x) dx \lesssim \sqrt{W(Q_k)} \|\mathbf{b}\|_\infty \|\mathbf{h}\|_{L_W^2},$$

which after squaring gives (4.22) by duality, as desired. This completes the proof of Claim 4.19. $\square$

Claim 4.24. *Suppose $s \leq t$. Then*

$$\|U_t Q_s\|_{L_W^2 \rightarrow L_W^2} \lesssim \frac{s}{t}.$$

Proof of the claim. Note that we have the pointwise estimate

$$(4.25) \quad |P_t Q_s f| \lesssim \frac{s}{t} \mathcal{M}^2 f,$$

where $\mathcal{M}^2 := \mathcal{M} \circ \mathcal{M}$ is the iterated Hardy-Littlewood maximal operator (with respect to Lebesgue measure). One may verify (4.25) by a standard argument using the size estimates and compact support of the kernels of P_t and Q_s , along with the smoothness of the former, and the cancellation property of the latter. We omit the well-known details. Since $\mathcal{M}$ is bounded on L_W^2 (recall that $W \in A_2$), we find, with the aid of Claims 4.18 and 4.19, and (4.25):

$$\begin{aligned} & \|U_t Q_s\|_{L_W^2 \rightarrow L_W^2} \\ & \leq \|\theta_t P_t\|_{L_W^2 \rightarrow L_W^2} \|P_t Q_s\|_{L_W^2 \rightarrow L_W^2} + \|(\theta_t \mathbf{1}) \cdot P_t\|_{L_W^2 \rightarrow L_W^2} \|P_t Q_s\|_{L_W^2 \rightarrow L_W^2} \\ & \lesssim \|P_t Q_s\|_{L_W^2 \rightarrow L_W^2} \lesssim \frac{s}{t}. \end{aligned}$$

□

Claim 4.26. *We have, uniformly on t ,*

$$\|U_t \mathbf{g}\|_{L_W^2} \lesssim t \|\nabla \mathbf{g}\|_{L_W^2}.$$

Proof. The proof is inspired by [AAHK, Lemma 3.5], and in fact is similar in spirit to that of Lemma 2.22, relying strongly in a decomposition in subcubes of the right size to use Poincaré's inequality, and some boundedness and off-diagonal estimates. Nevertheless, let us show it in detail, because some parts of it will be reused later. Cover $\mathbb{R}^n$ by a grid of non-overlapping dyadic cubes Q_k with sidelength $t/2 < \ell(Q_k) \leq t$. Using the easy fact that $U_t \mathbf{1} = 0$ we compute

$$\begin{aligned} (4.27) \quad & \|U_t \mathbf{g}\|_{L_W^2}^2 = \sum_k \|U_t (\mathbf{g} - [\mathbf{g}]_{2Q_k})\|_{L_W^2(Q_k)}^2 \\ & \lesssim \sum_k \|\theta_t P_t^2 (\mathbf{g} - [\mathbf{g}]_{2Q_k})\|_{L_W^2(Q_k)}^2 + \sum_k \|(\theta_t \mathbf{1}) \cdot P_t^2 (\mathbf{g} - [\mathbf{g}]_{2Q_k})\|_{L_W^2(Q_k)}^2 =: A + B. \end{aligned}$$

Let us first deal with A , denoting $S_t := \theta_t P_t^2$ because we intend to reuse some computations later on. For each term in the series, simply using linearity and the triangle inequality

$$\begin{aligned} (4.28) \quad & \|S_t (\mathbf{g} - [\mathbf{g}]_{2Q_k})\|_{L_W^2(Q_k)} \leq \|S_t ((\mathbf{g} - [\mathbf{g}]_{2Q_k}) \mathbf{1}_{2Q_k})\|_{L_W^2(Q_k)} \\ & + \sum_{j=1}^{\infty} \|S_t ((\mathbf{g} - [\mathbf{g}]_{2Q_k}) \mathbf{1}_{2^{j+1}Q_k \setminus 2^j Q_k})\|_{L_W^2(Q_k)} =: I^{(k)} + \sum_{j=1}^{\infty} II_j^{(k)}. \end{aligned}$$

Using the boundedness of S_t (in this case, this follows from Lemmas 2.7 and 2.11) and Poincaré's inequality, we deduce

$$(4.29) \quad I^{(k)} \lesssim \|\mathbf{g} - [\mathbf{g}]_{2Q_k}\|_{L_W^2(2Q_k)} \lesssim \ell(Q_k) \|\nabla \mathbf{g}\|_{L_W^2(2Q_k)} \leq t \|\nabla \mathbf{g}\|_{L_W^2(2Q_k)}.$$

And for the other terms, we can use the off-diagonal estimates for S_t (in this case, this follows from Lemmas 2.14 and 2.21), and taking advantage of $\ell(Q_k) \approx t$ and Poincaré, we obtain, similarly to the situation in Lemma 2.22,

$$\begin{aligned} & \sum_{j=1}^{\infty} II_j^{(k)} \lesssim e^{-c4^j} \|\mathbf{g} - [\mathbf{g}]_{2Q_k}\|_{L_W^2(2^{j+1}Q_k)} \\ & \lesssim \sum_{j=1}^{\infty} e^{-c4^j} \|\mathbf{g} - [\mathbf{g}]_{2^{j+1}Q_k}\|_{L_W^2(2^{j+1}Q_k)} + \sum_{j=1}^{\infty} \sum_{i=1}^j e^{-c4^j} C_D^{(j-i)/2} \|\mathbf{g} - [\mathbf{g}]_{2^{i+1}Q_k}\|_{L_W^2(2^{i+1}Q_k)} \end{aligned}$$

$$\lesssim \sum_{j=1}^{\infty} e^{-\frac{\varepsilon}{2}4^j} 2^j \ell(Q_k) \|\nabla \mathbf{g}\|_{L_W^2(2^{j+1}Q_k)} \lesssim t \sum_{j=1}^{\infty} e^{-\frac{\varepsilon}{4}4^j} \|\nabla \mathbf{g}\|_{L_W^2(2^{j+1}Q_k)}.$$

Thus, going back to (4.28) we obtain

$$\|S_t(\mathbf{g} - [\mathbf{g}]_{2Q_k})\|_{L_W^2(Q_k)} \lesssim t \|\nabla \mathbf{g}\|_{L_W^2(2Q_k)} + t \sum_{j=1}^{\infty} e^{-\frac{\varepsilon}{4}4^j} \|\nabla \mathbf{g}\|_{L_W^2(2^{j+1}Q_k)},$$

and hence

$$A \lesssim t^2 \sum_k \|\nabla \mathbf{g}\|_{L_W^2(2Q_k)}^2 + t^2 \sum_k \left(\sum_{j=1}^{\infty} e^{-\frac{\varepsilon}{4}4^j} \|\nabla \mathbf{g}\|_{L_W^2(2^{j+1}Q_k)} \right)^2 =: t^2(A_1 + A_2).$$

By bounded overlap of the cubes $2Q_k$ we easily get

$$A_1 \lesssim \|\nabla \mathbf{g}\|_{L_W^2}^2.$$

For the other term, we note that $|x - y| \lesssim 2^j \ell(Q_k) \approx 2^j t$, whenever $x \in Q_k$, and $y \in 2^{j+1}Q_k$. We further note that $W(Q_k) \approx W(B_t(x))$, for $x \in Q_k$, and that for all $x \in \mathbb{R}^n$,

$$e^{-\frac{\varepsilon}{4}4^j} \int_{|x-y| \lesssim 2^j t} W(B_t(x))^{-1} W(x) dx \lesssim e^{-\frac{\varepsilon}{8}4^j},$$

by the doubling property of W . We now use these observations, along with Cauchy-Schwarz, Fubini's theorem, and the fact that the cubes Q_k are non-overlapping, to obtain

$$\begin{aligned} A_2 &\leq \sum_k \left(\sum_{j=1}^{\infty} e^{-\frac{\varepsilon}{4}4^j} \right) \left(\sum_{j=1}^{\infty} e^{-\frac{\varepsilon}{4}4^j} \|\nabla \mathbf{g}\|_{L_W^2(2^{j+1}Q_k)}^2 \right) = \sum_k \sum_{j=1}^{\infty} e^{-\frac{\varepsilon}{4}4^j} \|\nabla \mathbf{g}\|_{L_W^2(2^{j+1}Q_k)}^2 \\ &\lesssim \sum_k \sum_{j=1}^{\infty} e^{-\frac{\varepsilon}{4}4^j} \int_{Q_k} W(B_t(x))^{-1} W(x) \int_{|x-y| \lesssim 2^j t} |\nabla \mathbf{g}(y)|^2 W(y) dy dx \\ &= \sum_{j=1}^{\infty} e^{-\frac{\varepsilon}{4}4^j} \int_{\mathbb{R}^n} W(B_t(x))^{-1} W(x) \int_{|x-y| \lesssim 2^j t} |\nabla \mathbf{g}(y)|^2 W(y) dy dx \\ &\lesssim \sum_{j=1}^{\infty} e^{-\frac{\varepsilon}{8}4^j} \int_{\mathbb{R}^n} |\nabla \mathbf{g}(y)|^2 W(y) dy \lesssim \|\nabla \mathbf{g}\|_{L_W^2}^2. \end{aligned}$$

Consequently, we have shown that

$$A \lesssim t^2 \|\nabla \mathbf{g}\|_{L_W^2}^2.$$

We can apply a similar, but simpler argument to handle term B in (4.27). We now set $S_t := (\theta_t \mathbf{1}) \cdot P_t^2$, and note that S_t is uniformly bounded on L_W^2 , by Claim 4.19 and Lemma 2.7. Moreover, the kernel of P_t^2 is compactly supported in the ball of radius $2t$, so the same is true for S_t . Hence, for the current version of S_t , we obtain a simplified variant of (4.28), in which only the term $I^{(k)}$ appears, enjoying the same bound as in (4.29). Thus,

$$B \lesssim t^2 \|\nabla \mathbf{g}\|_{L_W^2}^2.$$

The proof of Claim 4.26 is now complete. $\square$

Claim 4.30. *For $t \leq s$, we have*

$$\|U_t Q_s\|_{L_W^2 \rightarrow L_W^2} \lesssim \frac{t}{s}.$$

Proof of the claim. Using Claim 4.26 and Lemma 2.7, we have

$$\|U_t Q_s \mathbf{g}\|_{L_W^2} \lesssim t \|\nabla Q_s \mathbf{g}\|_{L_W^2 \rightarrow L_W^2} = \frac{t}{s} \|s \nabla Q_s \mathbf{g}\|_{L_W^2} \lesssim \frac{t}{s} \|\mathbf{g}\|_{L_W^2},$$

as desired. □

As noted above, the preceding claims conclude the proof of Lemma 4.17. □

We are now ready to reduce matters to a Carleson measure estimate. Recall that to prove Theorem 4.4 (and hence Theorem 4.1), it suffices to verify estimate (4.14) (equivalently, (4.12)).

Lemma 4.31. *Theorem 4.4 (and hence Theorem 4.1) follows from the Carleson measure estimate*

$$(4.32) \quad \sup_Q \frac{1}{W(Q)} \int_0^{\ell(Q)} \int_Q \left| \tilde{\theta}_t \mathbf{1}(x) \right|^2 W(x) \frac{dx dt}{t} < \infty.$$

Proof. Recalling that $\tilde{\theta}_t \mathbf{1} = \theta_t \mathbf{1}$, we see that by Lemma 4.17 and a weighted version of Carleson's embedding inequality (see [CUR, Lemma 2.2]), the estimate (4.32) implies (4.14). □

Our goal then, is to prove (4.32). To this end, let us first establish a few more estimates to be used in the sequel. We define the dyadic averaging operator by

$$A_t \mathbf{f}(x) := \int_{Q_{x,t}} \mathbf{f}(y) dy,$$

where $Q_{x,t}$ is the half-open dyadic cube containing x for which $t/2 < \ell(Q_{x,t}) \leq t$.

Lemma 4.33. *We have*

$$\int_0^\infty \|(\theta_t \mathbf{1}) \cdot (P_t^2 - A_t) \mathbf{g}\|_{L_W^2}^2 \frac{dt}{t} \lesssim \|\mathbf{g}\|_{L_W^2}^2,$$

where the implicit constant depends on n, λ and $[W]_{A_2}$.

Proof. The proof of this estimate will be very similar to that of Lemma 4.17. We set

$$\tilde{U}_t := (\theta_t \mathbf{1}) \cdot (P_t^2 - A_t),$$

and note that it is enough to show that $\tilde{U}_t$ satisfies the hypotheses of the weighted Littlewood-Paley almost-orthogonality result Lemma 2.10. The uniform boundedness of $\tilde{U}_t$ arises immediately from that of $(\theta_t \mathbf{1}) \cdot P_t^2$ (see Claim 4.19 and Lemma 2.7), along with the following result:

Claim 4.34. *We have, uniformly on t ,*

$$\|(\theta_t \mathbf{1}) \cdot A_t\|_{L_W^2 \rightarrow L_W^2} \lesssim 1.$$

Proof of the claim. The proof is the same as that of Claim 4.19, which treated $(\theta_t \mathbf{1}) \cdot P_t$. Indeed, the only properties of P_t that were used in that argument were the size and support condition of its kernel. The kernel of A_t enjoys similar properties, in fact

$$|A_t \mathbf{f}(x)|^2 \leq \left(\int_{Q_k} |\mathbf{f}(y)| dy \right)^2 \lesssim \int_{Q_k} |\mathbf{f}(y)|^2 W(y) dy,$$

hence, the same proof may be repeated. $\square$

To prove the quasi-orthogonality with the Q_s operators, the next result will be useful.

Claim 4.35. *We have, uniformly on t ,*

$$\left\| \tilde{U}_t \mathbf{g} \right\|_{L_W^2} \lesssim t \|\nabla \mathbf{g}\|_{L_W^2}.$$

Proof. The proof is similar to that of Claim 4.26, but simpler: now one has to deal only with terms like “ B ” associated to $S_t = (\theta_t \mathbf{1}) \cdot A_t$ in (4.27), so that there is no “tail” as in (4.28), but rather only a local term analogous to $I^{(k)}$. We omit the routine details. $\square$

The following two claims finish the proof of Lemma 4.33, and are analogous to those in the proof of Lemma 4.17.

Claim 4.36. *We have, uniformly for $t \leq s$,*

$$\left\| \tilde{U}_t Q_s \right\|_{L_W^2 \rightarrow L_W^2} \lesssim \frac{t}{s}.$$

Proof. In view of Claim 4.35, repeating the proof of Claim 4.30, we simply write

$$\left\| \tilde{U}_t Q_s \right\|_{L_W^2 \rightarrow L_W^2} \lesssim t \|\nabla Q_s\|_{L_W^2 \rightarrow L_W^2} \lesssim \frac{t}{s}.$$

$\square$

Claim 4.37. *We have, uniformly in $s \leq t$, and for some fixed $\alpha > 0$,*

$$\left\| \tilde{U}_t Q_s \right\|_{L_W^2 \rightarrow L_W^2} \lesssim \left(\frac{s}{t} \right)^\alpha.$$

Proof. On the one hand, as in Claim 4.24, and using the boundedness of Claim 4.19,

$$\left\| (\theta_t \mathbf{1}) \cdot P_t^2 Q_s \right\|_{L_W^2 \rightarrow L_W^2} \lesssim \|P_t Q_s\|_{L_W^2 \rightarrow L_W^2} \lesssim \frac{s}{t}.$$

On the other hand, by [AHLMT, Lemma 4.7 and its proof], we have the unweighted quasi-orthogonality estimate

$$\|A_t Q_s\|_{L^2 \rightarrow L^2} \lesssim \left(\frac{s}{t} \right)^\alpha,$$

for some exponent $\alpha > 0$, uniformly for $s \leq t$. Consequently, we may use the technique of Duoandikoetxea and Rubio de Francia [DR], in which one first self-improves the weight W , and then uses Stein-Weiss interpolation with change of measure [SW], to deduce the weighted quasi-orthogonality estimate

$$\|A_t Q_s\|_{L_W^2 \rightarrow L_W^2} \lesssim \left(\frac{s}{t} \right)^\beta,$$

for some positive $\beta < \alpha$ (see Lemma 2.5 in [CUR] for more details). Hence, by Claim 4.34,

$$\left\| (\theta_t \mathbf{1}) \cdot A_t Q_s \right\|_{L_W^2 \rightarrow L_W^2} = \left\| (\theta_t \mathbf{1}) \cdot A_t^2 Q_s \right\|_{L_W^2 \rightarrow L_W^2} \lesssim \|A_t Q_s\|_{L_W^2 \rightarrow L_W^2} \lesssim \left(\frac{s}{t} \right)^\beta.$$

$\square$

Collecting all the above claims, the proof of Lemma 4.33 is completed. $\square$

Corollary 4.38. *We have the square function bound*

$$\int_0^\infty \left\| \tilde{\theta}_t \nabla f - (\tilde{\theta}_t \mathbf{1}) \cdot A_t \nabla f \right\|_{L_W^2}^2 \frac{dt}{t} \lesssim \|\nabla f\|_{L_W^2}^2,$$

where the implicit constant depends on n, λ and $[W]_{A_2}$.

Proof. With Lemma 4.33 in hand, since $\tilde{\theta}_t \mathbf{1} = \theta_t \mathbf{1}$, it is enough to prove the following:

$$\int_0^\infty \left\| \tilde{\theta}_t \nabla f - (\theta_t \mathbf{1}) \cdot P_t^2 \nabla f \right\|_{L_W^2}^2 \frac{dt}{t} \lesssim \|\nabla f\|_{L_W^2}^2.$$

To this end, we write

$$\tilde{\theta}_t \nabla f - (\theta_t \mathbf{1}) \cdot P_t^2 \nabla f = \tilde{\theta}_t \nabla (I - P_t^2) f + \left[\tilde{\theta}_t \nabla P_t^2 f - (\theta_t \mathbf{1}) \cdot P_t^2 \nabla f \right] =: Y_t f + Z_t f.$$

By (4.15),

$$-2Y_t = t\tilde{L}e^{-t^2\tilde{L}}(I - P_t^2) =: \tilde{R}_t,$$

where $\tilde{R}_t$ is precisely the same operator defined in (4.6), enjoying the square function bound established in Lemma 4.8. In addition, again using (4.15),

$$-2\tilde{\theta}_t \nabla P_t^2 = t\tilde{L}e^{-t^2\tilde{L}}P_t^2 =: \tilde{T}_t,$$

where $\tilde{T}_t$ is precisely the same operator defined in (4.6). We now repeat the splitting of $\tilde{T}_t$, exactly as in (4.10):

$$\tilde{T}_t f = -te^{-t^2\tilde{L}}LP_t^2 f - 2te^{-t^2\tilde{L}}\widetilde{\text{div}}(A\nabla(P_t^2 f)).$$

Note that the the second term equals $-2\theta_t \nabla P_t^2 f$ (see (4.13)). Combining these observations, we see that

$$Z_t f = -\frac{1}{2}\tilde{T}_t f - (\theta_t \mathbf{1}) \cdot P_t^2 \nabla f = \frac{1}{2}te^{-t^2\tilde{L}}LP_t^2 f + [\theta_t P_t^2 \nabla f - (\theta_t \mathbf{1}) \cdot P_t^2 \nabla f] =: E_t f + U_t f,$$

where the term $E_t f$ (which is actually the error $(\tilde{\theta}_t - \theta_t) \nabla P_t^2 f$), satisfies the desired square function bound, by Lemma 4.11. The last term also enjoys the desired square function bound, by Lemma 4.17. This concludes the proof of the Corollary. $\square$

4.3. The $T(b)$ argument. Recall that our goal is to prove the Carleson measure estimate (4.32). We now turn to this task, which will finish the proof of Theorem 4.1 (and therefore also the proof of Theorem 1.14). Our arguments here will be an adaptation of the proof of the Kato conjecture in the divergence form setting, see [AHLMT], and in particular, the extension of that proof to the degenerate elliptic case in [CUR].

We note that by the doubling property of W , we may assume that the supremum in (4.32) is taken over dyadic cubes Q . Given any such cube Q , a sufficiently small number $\varepsilon \in (0, 1)$ to be chosen, and $v \in \mathbb{R}^n$ with $|v| = 1$, we define

$$(4.39) \quad f_{Q,v}^\varepsilon := e^{-(\varepsilon \ell(Q))^2 \tilde{L}}(\Phi_Q \chi_Q \cdot v),$$

where $\Phi_Q(x) = x - x_Q$, x_Q denotes the center of Q , and $\chi_Q \in \mathcal{C}_0^\infty$ is a cut-off function such that $\chi_Q \equiv 1$ in $2Q$, $\text{supp } \chi_Q \subset 4Q$ and $\|\chi_Q\|_\infty + \ell(Q) \|\nabla \chi_Q\|_\infty + \ell(Q)^2 \|\nabla^2 \chi_Q\|_\infty \lesssim 1$. Clearly,

$$(4.40) \quad \|\nabla(\Phi_Q \chi_Q \cdot v)\|_\infty \lesssim 1,$$

and also

$$(4.41) \quad \int_{\mathbb{R}^n} |\Phi_Q \chi_Q \cdot v|^2 W(x) dx \lesssim \ell(Q)^2 W(Q).$$

The following estimates hold for $f_{Q,v}^\varepsilon$, with constants that are uniform on Q, v and ε :

$$(4.42) \quad \int_{5Q} |f_{Q,v}^\varepsilon - \Phi_Q \chi_Q \cdot v|^2 W dx \lesssim \varepsilon^2 \ell(Q)^2 W(Q),$$

$$(4.43) \quad \int_{5Q} |\nabla f_{Q,v}^\varepsilon|^2 W dx + \int_{5Q} |\nabla (f_{Q,v}^\varepsilon - \Phi_Q \chi_Q \cdot v)|^2 W dx \lesssim W(Q),$$

These estimates follow at once from (4.40), Lemmas 2.22 and 2.25 (with $t = \varepsilon \ell(Q)$), and the doubling property of W .

The proof of (4.32) (and hence of Theorem 4.1 by Corollary 4.31) follows from the next two lemmas.

Lemma 4.44. *There exists $0 < \varepsilon = \varepsilon(\lambda, n, [W]_{A_2}) \ll 1$ and a finite set V of unit vectors in $\mathbb{R}^n$, whose cardinality depends only on ε and n , such that*

$$\begin{aligned} \sup_Q \frac{1}{W(Q)} \int_0^{\ell(Q)} \int_Q \left| (\tilde{\theta}_t \mathbf{1})(x) \right|^2 W(x) \frac{dx dt}{t} \\ \lesssim \sum_{v \in V} \sup_Q \frac{1}{W(Q)} \int_0^{\ell(Q)} \int_Q \left| (\tilde{\theta}_t \mathbf{1})(x) \cdot (A_t \nabla f_{Q,v}^\varepsilon)(x) \right|^2 W(x) \frac{dx dt}{t}, \end{aligned}$$

where the implicit constant depends on n, λ and $[W]_{A_2}$.

Proof. The reader may check that the proof of [CUR, Lemma 5.1] (which in turn is an adaptation to the weighted case of [AHLMcT, Lemma 5.4]) works perfectly well in our situation: as long as $W \in A_2$ and $f_{Q,v}^\varepsilon$ satisfies the estimates (4.42) and (4.43), the proof in [CUR] goes through⁵. $\square$

With Lemma 4.44 in hand, estimate (4.32) will follow immediately from the next lemma.

Lemma 4.45. *For every cube Q and unit vector v , we have*

$$\int_0^{\ell(Q)} \int_Q \left| (\tilde{\theta}_t \mathbf{1})(x) \cdot (A_t \nabla f_{Q,v}^\varepsilon)(x) \right|^2 W(x) \frac{dx dt}{t} \lesssim W(Q),$$

where the implicit constant depends on $n, \lambda, [W]_{A_2}$ and ε , but is uniform on Q and v .

Proof. Fix Q and v , and abbreviate $f := f_{Q,v}^\varepsilon$. By Corollary 4.38, we have

$$\begin{aligned} \int_0^{\ell(Q)} \int_Q \left| (\tilde{\theta}_t \mathbf{1})(x) \cdot (A_t \nabla f)(x) \right|^2 W(x) \frac{dx dt}{t} \\ \lesssim \|\nabla f\|_{L_W^2}^2 + \int_0^{\ell(Q)} \int_Q \left| (\tilde{\theta}_t \nabla f)(x) \right|^2 W(x) \frac{dx dt}{t} =: I + II \lesssim W(Q) + II, \end{aligned}$$

where in the last step, we have used (4.43) to obtain the desired bound for term I .

⁵To clarify a possible point of confusion, we mention that in [AHLMcT] and [CUR], the unit vectors were taken in $\mathbb{C}^n$, because in the divergence form setting of those papers, one treats the case of complex coefficients; at present, our results in the non-divergence form case treat only the case of real coefficients, so we need only consider real unit vectors.

Term II can be treated as follows, using (4.15), Lemma 2.11, and the definition of $f = f_{Q,v}^\varepsilon$,

$$\begin{aligned}
II &\approx \int_0^{\ell(Q)} \int_Q \left| t e^{-t^2 \tilde{L}} \tilde{L} f(x) \right|^2 W(x) \frac{dx dt}{t} \\
&\lesssim \int_0^{\ell(Q)} t dt \int_{\mathbb{R}^n} \left| \tilde{L} e^{-(\varepsilon \ell(Q))^2 \tilde{L}} (\Phi_Q \chi_Q \cdot v)(x) \right|^2 W(x) dx \\
&\approx \ell(Q)^2 \int_{\mathbb{R}^n} \left| \tilde{L} e^{-(\varepsilon \ell(Q))^2 \tilde{L}} (\Phi_Q \chi_Q \cdot v)(x) \right|^2 W(x) dx \\
&\lesssim \varepsilon^{-4} \ell(Q)^{-2} \int_{\mathbb{R}^n} |\Phi_Q \chi_Q \cdot v|^2 W(x) dx \lesssim \varepsilon^{-4} W(Q),
\end{aligned}$$

where in the last two steps we have first used Lemma 2.11 (vi) with $t = \varepsilon \ell(Q)$, and then (4.41). Since ε has been fixed depending only on allowable parameters, the dependence on ε is harmless.

Collecting all the preceeding estimates, we have finished the proof. $\square$

REFERENCES

- [AAAHK] M.A. Alfonseca, P. Auscher, A. Axelsson, S. Hofmann and S. Kim, Analyticity of layer potentials and L^2 solvability of boundary value problems for divergence form elliptic equations with complex L^∞ coefficients, *Adv. Math.* **226** (2011) 4533–4606.
- [AA] P. Auscher and A. Axelsson, Weighted maximal regularity estimates and solvability of non-smooth elliptic systems I, *Invent. Math.* **184** (2011), 47–115.
- [AAH] P. Auscher, A. Axelsson and S. Hofmann, Functional calculus of Dirac operators and complex perturbations of Neumann and regularity problems, *J. Functional Analysis* **255** (2008), 374–448.
- [AAMc] P. Auscher, A. Axelsson and McIntosh, Solvability of elliptic systems with square integrable boundary data, *Ark. Mat.* **48** (2010), 253–287.
- [AAMc2] P. Auscher, A. Axelsson, and A. McIntosh, On a quadratic estimate related to the Kato conjecture and boundary value problems, *Contemporary Mathematics, AMS*, **505** (2010), 105–129.
- [AEN1] P. Auscher, M. Egert, K. Nyström, The Dirichlet problem for second order parabolic operators in divergence form, *J. Ec. Polytech. Math.* **5** (2018), 407–441.
- [AEN2] P. Auscher, M. Egert, K. Nyström, L^2 well-posedness of boundary value problems and the Kato square root problem for parabolic systems with measurable coefficients, *Journal of the European Mathematical Society* **22** (9) (2020), 2943–3058.
- [AHLMcT] P. Auscher, S. Hofmann, M. Lacey, A. McIntosh, P. Tchamitchian, The solution of the Kato square root problem for second order elliptic operators on $\mathbb{R}^n$, *Ann. of Math.* (2) **156** (2002), no. 2, 633–654.
- [AHLT] P. Auscher, S. Hofmann, J.L. Lewis, P. Tchamitchian, Extrapolation of Carleson measures and the analyticity of Kato’s square root operators, *Acta Math.* **187** (2001), no. 2, 161–190.
- [AKMc] A. Axelsson, S. Keith and A. McIntosh, Quadratic estimates and functional calculi of perturbed Dirac operators, *Invent. Math.* **163** (2006), 455–497.
- [B] P. Bauman, Positive solutions of elliptic equations in nondivergence form and their adjoints, *Ark. Mat.* **22** (1984), 153–173.
- [BHLMP] S. Bortz, S. Hofmann, J. L. Luna Garcia, S. Mayboroda, and B. Poggi, Critical Perturbations for Second Order Elliptic Operators. Part I: Square function bounds for layer potentials, to appear, *Analysis and PDE*, accepted December 2020.
- [CFL1] F. Chiarenza, M. Frasca, and P. Longo, Interior $W^{2,p}$ estimates for nondivergence elliptic equations with discontinuous coefficients, *Ricerche Mat.* **40** (1991), 149–168.
- [CFL2] F. Chiarenza, M. Frasca, and P. Longo, $W^{2,p}$ -solvability of the Dirichlet problem for non-divergence elliptic equations with VMO coefficients, *Trans. Amer. Math. Soc.* **336** (1993), 841–853.
- [CDM] R. Coifman, D. Deng, and Y. Meyer, Domaine de la racine carrée de certains operateurs différentiels accréatifs, *Annales de l’Institut Fourier* **33** (1983), 123–134.
- [CMcM] R. Coifman, A. McIntosh, and Y. Meyer. L’intégrale de Cauchy définit un opérateur borné sur $L^2(\mathbb{R})$ pour les courbes lipschitziennes. *Ann. Math.* **116** (1982), 361–387.

- [CRW] R. Coifman, R. Rochberg, and G. Weiss, Factorization Theorems for Hardy Spaces in Several Variables, *Annals of Math.* **103** (1976), 611-635.
- [CUR] D. Cruz-Uribe, C. Ríos, The Kato problem for operators with weighted ellipticity, *Trans. Amer. Math. Soc.* **367** (2015), 4727-4756.
- [CUR2] D. Cruz-Uribe, C. Ríos, The solution of the Kato problem for degenerate elliptic operators with Gaussian bounds, *Trans. Amer. Math. Soc.* **364** (2012), 3449-3478.
- [Duo] J. Duoandikoetxea, *Fourier Analysis*, American Mathematical Soc. **29** (2001).
- [DR] J. Duoandikoetxea, J.L. Rubio de Francia, Maximal and singular integral operators via Fourier transform estimates, *Invent. Math.* **84** (1986), 541-561.
- [DY] X. T. Duong and L. Yan, Bounded holomorphic functional calculus for non-divergence form differential operators, *Differential and Integral Equations* **15** (2002), 709-730.
- [E1] L. Escauriaza, Weak-type (1,1) inequalities and regularity properties of adjoint and normalized adjoint solutions to linear nondivergence form operators with VMO coefficients, *Duke Math J.* **74** (1994), 177-201.
- [E2] L. Escauriaza, Bounds for the fundamental solution of elliptic and parabolic equations, *Communications in Partial Differential Equations* **25** (2000), 821-845.
- [EH] L. Escauriaza, S. Hofmann, Kato square root problem with unbounded leading coefficients, *Proc. Amer. Math. Soc.* **146** (2018), 5295-5310.
- [FJK1] E. Fabes, D. Jerison and C. Kenig, Multilinear square functions and partial differential equations, *Amer. J. of Math.* **107** (1985), 1325-1367.
- [FJK2] E. Fabes, D. Jerison and C. Kenig, Necessary and sufficient conditions for absolute continuity of elliptic-harmonic measure, *Annals of Math.* **119** (1984), 121-141.
- [H] M. Haase, *The functional calculus for sectorial operators*, Operator Theory: Advances and Applications, vol. 169, Birkhäuser Verlag, Basel (2006).
- [HKM] J. Heinonen, T. Kilpeläinen, O. Martio, *Nonlinear potential theory of degenerate elliptic equations*, Courier Dover Publ. (2018).
- [Hof] S. Hofmann, Local T_b Theorems and applications in PDE, *Proceedings of the ICM Madrid*, Vol. **II**, pp. 1375-1392, European Math. Soc., 2006.
- [HKMP] S. Hofmann, C. Kenig, S. Mayboroda, and J. Pipher, Square function/Non-tangential maximal function estimates and the Dirichlet problem for non-symmetric elliptic operators, *Jour. Amer. Math. Soc.* **28** (2015), pp 483-529.
- [HKMP2] S. Hofmann, C. Kenig, S. Mayboroda, and J. Pipher, The Regularity problem for second order elliptic operators with complex-valued bounded measurable coefficients, *Math. Ann.* **361** (2015), pp 863-907.
- [HLMc] S. Hofmann, M. Lacey and A. McIntosh. The solution of the Kato problem for divergence form elliptic operators with Gaussian heat kernel bounds. *Annals of Math.* **156** (2002), pp 623-631.
- [HLeM] S. Hofmann, P. Le and A. J. Morris, Carleson measure estimates and the Dirichlet problem for degenerate elliptic equations, *Analysis and PDE* **12** (2019), 2095-2146.
- [HLMP1] S. Hofmann, L. Li, S. Mayboroda and J. Pipher, L^p theory for the square roots and square functions of elliptic operators having a BMO anti-symmetric part, to appear, *Math. Z.* accepted Dec 2021.
- [HLMP2] S. Hofmann, L. Li, S. Mayboroda and J. Pipher, The Dirichlet problem for elliptic operators having a BMO anti-symmetric part, *Math. Ann.* **382** (2022), no. 1-2, 103-168.
- [HMc] S. Hofmann and A. McIntosh, The solution of the Kato problem in two dimensions, Proceedings of the Conference on Harmonic Analysis and PDE held in El Escorial, Spain in July 2000, *Publ. Mat.* Vol. extra, 2002 pp. 143-160.
- [Ka] T. Kato, *Perturbation Theory for Linear Operators*, Springer-Verlag, New York 1966.
- [KM] C. Kenig and Y. Meyer, The Cauchy integral on Lipschitz curves and the square root of second order accretive operators are the same, in *Recent Progress in Fourier Analysis*, I. Peral, ed., North Holland Math Studies **111** (1985), 123-145.
- [KS] N. Krylov and M. Safonov, A certain property of solutions of parabolic equations with measurable coefficients, *Math. USSR Izv.* **16** (1981), 151-164.
- [Mc] A. McIntosh, Operators which have an H^∞ functional calculus, *Miniconference on Operator Theory and Partial Differential Equations, 1986*, Proceedings of the Centre for Mathematical Analysis, ANU, Canberra, 1986, 210-231.
- [N] N.S. Nadirashvili, Non-uniqueness in the martingale problem and the Dirichlet problem for uniformly elliptic operators, *Ann. Sc. Norm. Super. Pisa Cl. Sci. (4)*, **24** (1997), pp. 537-550.
- [Ny] K. Nyström, L^2 Solvability of boundary value problems for divergence form parabolic equations with complex coefficients, *Journal of Differential Equations* **262** (2017), 2808-2939.

- [SW] E. M. Stein and G. Weiss, Interpolation of operators with change of measures, *Trans. Amer. Math. Soc.* **87** (1958), 159–172.
- [Y] A. Yagi, Coincidence entre des espaces d'interpolation et des domaines de puissances fractionnaires d'opérateurs, *C. R. Acad. Sc. Paris* **299** Série I (1984), 173–176.

LUIS ESCAURIAZA, UNIVERSIDAD DEL PAÍS VASCO, APARTADO 644, 48080 BILBAO, SPAIN.

PABLO HIDALGO-PALENCIA, INSTITUTO DE CIENCIAS MATEMÁTICAS CSIC-UAM-UC3M-UCM, CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS, E-28049 MADRID, SPAIN

AND

DEPARTAMENTO DE ANÁLISIS MATEMÁTICO Y MATEMÁTICA APLICADA, FACULTAD DE MATEMÁTICAS, UNIVERSIDAD COMPLUTENSE DE MADRID, E-28040 MADRID, SPAIN

Email address: `pablo.hidalgo@icmat.es`

STEVE HOFMANN, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MISSOURI, COLUMBIA, MO 65211, USA

Email address: `hofmanns@missouri.edu`