# The Kato square root problem for higher order elliptic operators and systems on $\mathbb{R}^{n}$ 

Pascal Auscher<br>Université de Picardie-Jules Verne 33, rue Saint Leu F-80039 Amiens Cedex 1<br>email: auscher@mathinfo.u-picardie.fr<br>Steve Hofmann*<br>Department of Mathematics<br>University of Missouri-Columbia<br>Columbia, MO 65211<br>email: hofmann@math.missouri.edu

Alan M ${ }^{\text {c }}$ Intosh
Centre for Mathematics and its Applications
Australian National University
Canberra, ACT 0200
email: alan@maths.anu.edu.au
Philippe Tchamitchian
Faculté des Sciences et Techniques de St.-Jérome
Université d'Aix-Marseille III
Avenue Escadrille Normandie-Niemen
F-13397 Marseille Cedex 20
email: tchamphi@math.u-3mrs.fr
May 2, 2001

## Dedicated to the memory of T. Kato


#### Abstract

We prove the Kato conjecture for elliptic operators and $N \times N$ systems in divergence form of arbitrary order $2 m$ on $\mathbb{R}^{n}$. More precisely, we assume the coefficients to be bounded measurable and the ellipticity is taken in the sense of a Gårding inequality. We identify the domain of their square roots as the natural Sobolev space $H^{m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$. We also make some remarks on the relation between various ellipticity conditions and Gårding inequality.


## 1 Introduction, statement of the main results and strategy

In [2] the Kato conjecture for scalar second order operators on $\mathbb{R}^{n}$ is established. We refer to this reference for an historical account on this conjecture. The main purpose of this paper is to complete the whole program on $\mathbb{R}^{n}$, $n \geq 1$, and to prove the Kato conjecture for elliptic systems of any order of the form

$$
\begin{equation*}
(L f)_{i}=\sum_{\substack{|\alpha|,|\beta| \leq m \\ 1 \leq j \leq N}}(-1)^{|\alpha|} \partial^{\alpha}\left(a_{\alpha \beta}^{i j} \partial^{\beta} f_{j}\right), \quad 1 \leq i \leq N \tag{1.1}
\end{equation*}
$$

where $N, m \in \mathbb{N}^{*}, f=\left(f_{1}, \ldots, f_{N}\right)$ and the coefficients $a_{\alpha \beta}^{i j}$ are complexvalued $L^{\infty}$ functions on $\mathbb{R}^{n}$. The ellipticity is in the sense of a Gårding inequality. In the sequel, we write (1.1) in a vector form by introducing the $N \times N$-matrix valued coefficient $a_{\alpha \beta}=\left(a_{\alpha \beta}^{i j}\right)$. We use the terminology "operator" for $L$ and, sometimes, we write "scalar operator" to stress the case $N=1$. We use the standard notations of differential calculus in $\mathbb{R}^{n}$ : multiindices, partials...

Consider first the homogeneous case. Then $L$ has a representation of the form

$$
\begin{equation*}
L f=(-1)^{m} \sum_{|\alpha|=|\beta|=m} \partial^{\alpha}\left(a_{\alpha \beta} \partial^{\beta} f\right), \tag{1.2}
\end{equation*}
$$

[^0]and we assume
\[

$$
\begin{equation*}
\left|\sum_{|\alpha|=|\beta|=m} \int_{\mathbb{R}^{n}} a_{\alpha \beta}(x) \partial^{\beta} f(x) \partial^{\alpha} \bar{g}(x) d x\right| \leq \Lambda\left\|\nabla^{m} f\right\|_{2}\left\|\nabla^{m} g\right\|_{2} \tag{1.3}
\end{equation*}
$$

\]

and the Gårding inequality

$$
\begin{equation*}
\operatorname{Re} \sum_{|\alpha|=|\beta|=m} \int_{\mathbb{R}^{n}} a_{\alpha \beta}(x) \partial^{\beta} f(x) \partial^{\alpha} \bar{f}(x) d x \geq \lambda\left\|\nabla^{m} f\right\|_{2}^{2} . \tag{1.4}
\end{equation*}
$$

for some $\lambda>0$ and $\Lambda<+\infty$ independent of $f, g \in H^{m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$. Here, $\nabla^{m} f=\left(\partial^{\alpha} f\right)_{|\alpha|=m}$ and $\left\|\nabla^{m} f\right\|_{2}=\left(\sum_{|\alpha|=m} \int_{\mathbb{R}^{n}}\left|\partial^{\alpha} f\right|^{2}\right)^{1 / 2}$. We remark that the ellipticity constants, that is the largest $\lambda$ and smallest $\Lambda$ for which the above inequalities hold, are uniquely determined given a sesquilinear form as in (1.3), but the representation (1.2) is not unique. The inequality (1.4) is the strict Gårding inequality. The operator $L$ is defined from $H^{m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$ into $H^{-m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$. Defining $\mathcal{D}(L)$ as the space of $f \in H^{m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$ such that $L f \in L^{2}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$, the restriction of $L$ to $\mathcal{D}(L)$ can be shown to be a maximal-accretive operator and $\mathcal{D}(L)$ is dense in $H^{m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$ [13]. By abuse, we do not distinguish in the notation $L$ from its restriction. By holomorphic functional calculus, $L$ has a unique maximal-accretive square root, $\sqrt{L}$, and the identification of its domain is known as the Kato conjecture.

Theorem 1.5. Under the above hypotheses, the square root of $L$ has domain equal to the Sobolev space $H^{m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$ and we have

$$
\begin{equation*}
\|\sqrt{L} f\|_{2} \sim\left\|\nabla^{m} f\right\|_{2} \tag{1.6}
\end{equation*}
$$

with constants depending only on $n, m, N, \lambda$ and $\Lambda$.
Here, $A \sim B$ means a two-sided inequality $C A \leq B \leq C^{\prime} A$.
The main result in [2] is for scalar second order operators. We note that for self-adjoint operators with bounded coefficients, (1.4) is equivalent to (1.6), so that there is nothing to prove. For non self-adjoint operators, the strict Gårding inequality gives a lower bound on the self-adjoint part of the form associated to $L$ and there is no control on the skew-adjoint part besides an upper bound; the gap from (1.4) to (1.6) explains in part the depth of Kato's conjecture. Another remark is that this conjecture cannot be solved
by purely abstract methods as the counterexample in [14] shows (see [6], Preliminaries, for a short proof).

Let us mention an immediate corollary obtained by interpolation and straightforward functional calculus.
Corollary 1.7. For all $s \in[0, m]$,

$$
\|\sqrt{L} f\|_{\dot{H}^{s-m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)} \sim\|f\|_{\dot{H}^{s}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)}
$$

Here, $\dot{H}^{s}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$ denotes the homogeneous Sobolev space of order $s$.
We stress that the novelty is in the endpoints $s=0$ and $s=m$ as the cases $0<s<m$ were known before from Kato's work [13]. Note also that the constants of the equivalence do not blow up when $s \rightarrow 0$ or $m$, which cannot be obtained from abstract methods.

Our argument for Theorem 1.5 contains two parts (Section 2). In the first part we assume that the semigroup kernel of $L$ satisfies a pointwise upper bound: we then follow [6] in reducing matters to establishing a Carleson measure estimate and next proceed as in [12] or [2] through a " $\mathrm{T}(\mathrm{b})$ " argument. In a second part, as it is always enjoyed by operators with high enough order, the assumption is removed by an argument based on an interpolation result in [4] consisting in raising the order of $L$. This argument, sketched in [6], is given in detail here. We point out that it is possible to replace pointwise bounds by bounds in an averaged sense as in [2] (see [9] for the self-adjoint case, but the proof goes through the non self-adjoint case), hence to obtain a direct proof without the interpolation argument. We choose the indirect approach for two reasons; first the direct proof contains long technical developments and, second, we feel that the interpolation result is an interesting tool. It also provides us with a different proof of the main result in [2].

Historically, this interpolation result had been known long before it was written in [4] and can be derived from Kato's work [13]. The first argument to prove the boundedness of the Cauchy integral on Lipschitz curves, which is different from the one in $[8]$ and was not published, went via the onedimensional solution of Kato's problem and this interpolation result. The latter was used again in the first "T(b)" theorem [15] but the proof was not given.

Let us next consider the version of Theorem 1.5 for inhomogeneous operators. Now, $L$ has a representation of the form

$$
\begin{equation*}
L f=\sum_{|\alpha|,|\beta| \leq m}(-1)^{|\alpha|} \partial^{\alpha}\left(a_{\alpha \beta} \partial^{\beta} f\right) \tag{1.8}
\end{equation*}
$$

and we assume

$$
\begin{equation*}
\left|\sum_{|\alpha|,|\beta| \leq m} \int_{\mathbb{R}^{n}} a_{\alpha \beta}(x) \partial^{\beta} f(x) \partial^{\alpha} \bar{g}(x) d x\right| \leq \widetilde{\Lambda}\|f\|_{H^{m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)}\|g\|_{H^{m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)} \tag{1.9}
\end{equation*}
$$ and the Gårding inequality

$$
\begin{equation*}
\operatorname{Re} \sum_{|\alpha|,|\beta| \leq m} \int_{\mathbb{R}^{n}} a_{\alpha \beta}(x) \partial^{\beta} f(x) \partial^{\alpha} \bar{f}(x) d x \geq \tilde{\lambda}\|f\|_{H^{m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)}^{2} \tag{1.10}
\end{equation*}
$$

for some $\tilde{\lambda}>0$ and $\tilde{\Lambda}<\infty$ independent of $f, g \in H^{m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$. The norm on the Sobolev space $H^{m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$ is $\left(\sum_{|\alpha| \leq m} \int_{\mathbb{R}^{n}}\left|\partial^{\alpha} f\right|^{2}\right)^{1 / 2}$. Again the best constants $\tilde{\lambda}$ and $\widetilde{\Lambda}$ in the above inequalities are uniquely determined given a sesquilinear form as in (1.9), but the representation (1.8) is not unique.

Theorem 1.11. Under the above hypotheses, the square root of $L$ has domain equal to the Sobolev space $H^{m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$ and we have

$$
\|\sqrt{L} f\|_{2} \sim\|f\|_{H^{m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)}
$$

with constants depending only on $n, m, \tilde{\lambda}$ and $\widetilde{\Lambda}$.
This implies the inhomogeneous version of Corollary 1.7,
Corollary 1.12. For all $s \in[0, m]$,

$$
\|\sqrt{L} f\|_{H^{s-m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)} \sim\|f\|_{H^{s}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)}
$$

where $H^{s}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$ denotes the inhomogeneous Sobolev space of order $s$.
Let us come back to homogeneous operators. Let $L$ be given with a representation (1.2) and satisfy (1.3). Replace the strict Gårding inequality (1.4) by the more often encountered weak Gårding inequality

$$
\begin{equation*}
\operatorname{Re} \sum_{|\alpha|=|\beta|=m} \int_{\mathbb{R}^{n}} a_{\alpha \beta}(x) \partial^{\beta} f(x) \partial^{\alpha} \bar{f}(x) d x \geq \lambda\left\|\nabla^{m} f\right\|_{2}^{2}-\kappa\|f\|_{2}^{2}, \tag{1.13}
\end{equation*}
$$

for some $\kappa>0$.

Proposition 1.14. Under the above hypotheses, the square root of $L+\kappa$ has domain equal to the Sobolev space $H^{m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$ and we have

$$
\|\sqrt{L+\kappa} f\|_{2} \leq C\left(\left\|\nabla^{m} f\right\|_{2}^{2}+\kappa\|f\|_{2}^{2}\right)^{1 / 2}
$$

where $C$ depends only on $n, m, \lambda$ and $\Lambda$ but not on $\kappa$.
One interest of this result is in the precise behavior with respect to the parameter $\kappa$. Indeed, formally let $\kappa$ tend to 0 and compare with Theorem 1.5. When $\kappa>0$, we only obtain a one-sided inequality. However, if we replace $L+\kappa$ by $L+s \kappa$ for some fixed $s>1$, we have a two-sided inequality and the constants depend also on $s$. Both Theorem 1.11 and Proposition 1.14 are proved in Section 4.

Applicability of our results when $m \geq 2$ or $N \geq 2$ depends on the validity of the Gårding inequality. Let us discuss on this now. Assume that $L$ is given with a representation (1.2) and satisfies (1.3), that is the coefficients $a_{\alpha \beta}$ are bounded. One does not know a necessary and sufficient condition on the coefficients for (1.4) or (1.13). But, as mentioned earlier, these inequalities depend only the self-adjoint part of the form which has coefficients $b_{\alpha \beta}=$ $\frac{1}{2}\left(a_{\alpha \beta}+a_{\beta \alpha}^{*}\right)$.

The strong ellipticity condition is

$$
\begin{equation*}
\operatorname{Re} \sum_{\substack{|\alpha|=|,|=m \\ 1 \leq i, j \leq N}} b_{\alpha \beta}^{i j}(x) \xi_{\beta, j} \overline{\xi_{\alpha, i}} \geq \lambda \sum_{\substack{|\alpha|=m \\ 1 \leq i \leq N}}\left|\xi_{\alpha, i}\right|^{2}, \quad \text { a.e. } \forall \xi_{\alpha, i} \in \mathbb{C} \tag{1.15}
\end{equation*}
$$

for some $\lambda>0$. This inequality implies immediately (1.4) with the same $\lambda$. Hence, we obtain

Theorem 1.16. Assume $m \geq 2$. If (1.3) and (1.15) hold, then the square root of $L$ has domain equal to the Sobolev space $H^{m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$ and we have

$$
\|\sqrt{L} f\|_{2} \sim\left\|\nabla^{m} f\right\|_{2}
$$

where the constants depend only on $n, m, N, \lambda$ and $\Lambda$.
Consider next the weaker ellipticity condition

$$
\begin{equation*}
\operatorname{Re} \sum_{|\alpha|=|\beta|=m} b_{\alpha \beta}(x) \xi^{\beta} \xi^{\alpha} \geq \lambda \sum_{|\alpha|=m} \xi^{2 \alpha}, \quad \text { a.e. } \forall \xi \in \mathbb{R}^{n} \tag{1.17}
\end{equation*}
$$

where the inequality holds in the sense of self-adjoint $N \times N$-matrices. When $m \geq 2$ and $N=1$ this is sometimes called the "Nirenberg" ellipticity condition. When $m=1$ and $N \geq 2$ this is the Legendre-Hadamard ellipticity condition.

A standard argument shows that (1.17) is necessary for (1.13) to hold for some $\kappa \geq 0$.

However, this is not sufficient. A counterexample is in [18]. Another one is (indirectly) in [11]: it is a second order system with bounded measurable coefficients and (1.17) which does not satisfy a Caccioppoli inequality, whereas it is shown in [5] that the Gårding inequality (1.13) governs Caccioppoli inequality.

It is classical that (1.17) implies (1.13) when $b_{\alpha \beta}$ are uniformly continuous coefficients (see, e.g., [10]). Actually, this smoothness condition can be relaxed. This was observed in [1] for second order operators. It is enough to assume that the distance in $B M O$ of the $b_{\alpha \beta}$ 's to $V M O$ (the closure in $B M O$ of the uniformly continuous functions) is small. More precisely, we prove in Section 5 the following result.

Proposition 1.18. Assume $m \geq 2$, (1.3) and (1.17). Consider the nonincreasing function

$$
w(r)=\sup _{\alpha, \beta} \sup _{x \in \mathbb{R}^{n}, 0<\rho \leq r}\left(\frac{1}{\rho^{n}} \int_{B(x, \rho)}\left|b_{\alpha \beta}-m_{B(x, \rho)} b_{\alpha \beta}\right|^{2}\right)^{1 / 2}, \quad r>0 .
$$

There exists $\varepsilon=\varepsilon(n, m, N, \lambda)>0$ such that if

$$
\begin{equation*}
\lim _{r \rightarrow 0} w(r)<\varepsilon, \tag{1.19}
\end{equation*}
$$

then (1.13) holds with constants $\lambda / 2$ and $\kappa$ depends on $\Lambda$ and $\sup w^{-1}([0, \varepsilon[)$.
Here $B(x, \rho)$ is a ball centered at $x$ with radius $\rho$ and $m_{B(x, \rho)} f$ is the mean of $f$ over that ball. For $N=n=2$ and $m=1$, we mention a nice result by K. Zhang in this spirit [19]. This gives us the following result.

Theorem 1.20. Assume $m \geq 2$. If (1.3), (1.17) and (1.19) hold, then for $\kappa$ as above, the square root $L+\kappa$ has domain equal to the Sobolev space $H^{m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$ and we have

$$
\|\sqrt{L+\kappa} f\|_{2} \leq C\left(\left\|\nabla^{m} f\right\|_{2}^{2}+\kappa\|f\|_{2}^{2}\right)^{1 / 2}
$$

where $C$ depends only on $n, m, \lambda$ and $\Lambda$ but not on $\kappa$.

This theorem improves over prior results in two ways. First, L. Escauriaza obtained the conclusion when $m=1, N=1$ and $a_{\alpha \beta} \in V M O$ (unpublished) and this was extended in [6] to $m \geq 2$ and the $B M O$-distance of the $a_{\alpha \beta}$ to $V M O$ small. Here, we relax the smoothness condition on the skew-adjoint parts of the coefficients. Secondly, in those works, the behavior of $C$ was not made as precise.

We leave to the interested reader the care of stating the correponding statements for inhomogeneous operators with principal parts as in the last two results.

## 2 Proof of Theorem 1.5

Let $m \geq 1$ and $N \geq 1$, and $L$ be an operator having a representation (1.2), with ellipticity constants $\lambda$ and $\Lambda$. As in [2], it is enough to prove a priori that

$$
\begin{equation*}
\|\sqrt{L} f\|_{2} \leq C\left\|\nabla^{m} f\right\|_{2}, \tag{K}
\end{equation*}
$$

for $f$ in the domain of $L$ with $C$ depending only on $n, m, N, \lambda$ and $\Lambda$.
Let us also set the subsequent algebra.
If $v=\left(v_{i}\right) \in X, X$ being a finite dimensional complex vector space, then $\bar{v}=\left(\overline{v_{i}}\right)$ denotes the complex conjugate of $v$.

We let $\mathbb{C}^{N}$ be equipped with an hermitian structure and we use implicitely the canonical basis in $\mathbb{C}^{N}$. For $u=\left(u_{i}\right) \in \mathbb{C}^{N}, v=\left(v_{i}\right) \in \mathbb{C}^{N}$, we set $u v=\sum_{i} u_{i} v_{i}$ so that the inner product writes $u \bar{v}$ and the norm is $|u|^{2}=u \bar{u}$.

Let $\mathcal{M}_{N}(\mathbb{C})$ be the space of $N \times N$ complex matrices equipped with the induced norm, denoted by $|M|$ on $\mathcal{L}\left(\mathbb{C}^{N}\right)$, the space of linear maps on $\mathbb{C}^{N}$.

If $X, Y$ and $Z$ are finite dimensional complex spaces for which we have $\|u v\|_{Z} \leq\|u\|_{X}\|v\|_{Y}$ and $p \in \mathbb{N}^{*}$, then for $u=\left(u_{\beta}\right) \in X^{p}$ and $v=\left(v_{\beta}\right) \in Y^{p}$, we set

$$
u \cdot v=\sum_{\beta} u_{\beta} v_{\beta} \in Z
$$

For example, we may take $X=\mathcal{M}_{N}(\mathbb{C})$ and $Y=\mathbb{C}^{N}$ or $X=Y=\mathbb{C}^{N}$ or $X=\mathbb{C}$ and $Y=\mathbb{C}^{N}$.

Hence, the inner product on $\left(\mathbb{C}^{N}\right)^{p}$ is $(u \mid v)=u \cdot \bar{v}$ and the norm is $|u|^{2}=u \cdot \bar{u}=\sum_{\beta}\left|u_{\beta}\right|^{2}$ if $u=\left(u_{\beta}\right), u_{\beta} \in \mathbb{C}^{N}$.

For $M=\left(M_{\beta}\right) \in\left(\mathcal{M}_{N}(\mathbb{C})\right)^{p}$, we set $|M|^{2}=\sum_{\beta}\left|M_{\beta}\right|^{2}$.
As the reader may observe, we are using the same single bar notation for the norms on different spaces : $\mathbb{C}, \mathbb{C}^{N} \ldots$ This will not create any problems as we shall give enough details. We reserve the double bar norms for function spaces.

### 2.1 Assuming a pointwise upper bound

We say that $L$ satisfies an order $2 m$ pointwise upper bound if

$$
\begin{equation*}
\left|W_{t^{2 m}}(x, y)\right| \leq C t^{-n} e^{-\left(\frac{|x-y|}{c t}\right)^{\nu}}, \tag{2.1}
\end{equation*}
$$

for almost every $(x, y) \in \mathbb{R}^{2 n}$ and all $t>0$ where $W_{t}(x, y)$ is the $\mathcal{M}_{N}(\mathbb{C})$ valued Schwartz kernel of $e^{-t L}$. Here and in the sequel $\nu=\frac{2 m}{2 m-1}$. For example, constant coefficient elliptic operators do satisfy such an estimate as easily seen from Fourier analysis.

We, henceforth, assume in this section that L satisfies this technical assumption without repeating it.

We now begin the proof of $(\mathrm{K})$. Introduce the $\left(\mathcal{M}_{N}(\mathbb{C})\right)^{p}$-valued function $\gamma_{t}$ defined by

$$
\gamma_{t}(x)=\left(\left((-1)^{m} e^{-t^{2 m} L} t^{m} \partial^{\alpha} a_{\alpha \beta}\right)(x)\right)_{|\beta|=m} .
$$

Here, $p$ is the number of multiindices $\alpha \in \mathbb{N}^{n}$ having length $m$ and we are using the summation convention on repeated indices.

We recall that Theorem 24, Chapter 2 of [6] reduces matters to a Carleson measure estimate. Actually, this theorem was stated in the scalar case with a further regularity hypothesis that is not needed as it was remarked in [6]. For completeness we include an argument.

Lemma 2.2. The inequality ( $K$ ) follows from the Carleson measure estimate

$$
\begin{equation*}
\sup \frac{1}{|Q|} \int_{Q} \int_{0}^{\ell(Q)}\left|\gamma_{t}(x)\right|^{2} \frac{d x d t}{t} \leq C \tag{2.3}
\end{equation*}
$$

where $C$ depends only on $n, m, N, \lambda, \Lambda$ and the constants in (2.1).
The supremum runs over all cubes in $\mathbb{R}^{n}$ with sides parallel to the axes. If $Q$ is a cube, $|Q|$ and $\ell(Q)$ denote respectively its Lebesgue measure and sidelength.

Proof. Recall that $a_{\alpha \beta}$ is $\mathcal{M}_{N}(\mathbb{C})$-valued so that for $\left(\mathbb{C}^{N}\right)^{p}$-valued or $\mathcal{M}_{N}(\mathbb{C})$ valued functions $F=\left(F_{\beta}\right)_{|\beta|=m}$, on defines an operator $\theta_{t}$ by

$$
\theta_{t} F=(-1)^{m} e^{-t^{2 m} L} t^{m} \partial^{\alpha}\left(a_{\alpha \beta} F_{\beta}\right)
$$

This operator $\theta_{t}$ is bounded from $L^{2}\left(\mathbb{R}^{n},\left(\mathbb{C}^{N}\right)^{p}\right)$ into $L^{2}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$ and $\theta_{t} \nabla^{m} f=$ $e^{-t^{2 m} L} t^{m} L f$ for $f \in H^{m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$.

Remark that $\gamma_{t}(x)=\theta_{t} \mathbf{1}(x)$ with $\mathbf{1}=\left(\delta_{\beta \beta^{\prime}} I\right) \in\left(\mathcal{M}_{N}(\mathbb{C})\right)^{p \times p}$ where $I$ is the identity matrix in $\mathcal{M}_{N}(\mathbb{C})$ and $\delta_{\beta \beta^{\prime}}$ the Kronecker symbol (apply $\theta_{t}$ to each of the $p\left(\mathcal{M}_{N}(\mathbb{C})\right)^{p}$-valued column $)$.

First, (2.3) and Carleson's inequality imply

$$
\int_{\mathbb{R}^{n}} \int_{0}^{+\infty}\left|\gamma_{t}(x) \cdot\left(P_{t} \nabla^{m} g\right)(x)\right|^{2} \frac{d x d t}{t} \leq C \int_{\mathbb{R}^{n}}\left|\nabla^{m} g\right|^{2}, \quad g \in H^{m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)
$$

Here, $P_{t}$ denotes the convolution operator with $\frac{1}{t^{n}} \varphi\left(\frac{x}{t}\right)$ where $\varphi$ is a smooth real-valued function supported in the unit ball of $\mathbb{R}^{n}$ with $\int \varphi=1$ and the moment condition $\int x^{\alpha} \varphi(x) d x=0$ for $1 \leq|\alpha| \leq m$. Using (2.4) below, we deduce that

$$
\int_{\mathbb{R}^{n}} \int_{0}^{+\infty}\left|\left(\theta_{t} \nabla^{m} g\right)(x)\right|^{2} \frac{d x d t}{t} \leq C \int_{\mathbb{R}^{n}}\left|\nabla^{m} g\right|^{2}
$$

which rewrites

$$
\int_{0}^{+\infty}\left\|e^{-t^{2 m} L} t^{m} L g\right\|_{2}^{2} \frac{d t}{t} \leq C\left\|\nabla^{m} g\right\|_{2}^{2}
$$

As the latter is equivalent to (K) by $H^{\infty}$-functional calculus for $L$ and a theorem by $\mathrm{M}^{c}$ Intosh and Yagi [16], we are done.

The keystone of our analysis is therefore

$$
\begin{equation*}
\int_{\mathbb{R}^{n}} \int_{0}^{+\infty}\left|\gamma_{t}(x) \cdot\left(P_{t} \nabla^{m} g\right)(x)-\left(\theta_{t} \nabla^{m} g\right)(x)\right|^{2} \frac{d x d t}{t} \leq C \int_{\mathbb{R}^{n}}\left|\nabla^{m} g\right|^{2}, \tag{2.4}
\end{equation*}
$$

where $C$ depends only on $n, m, N, \lambda, \Lambda$ and the constants in (2.1). This inequality is proved in [6] (Lemma 29 of Chapter 2). Here is a quicker argument. Write

$$
\begin{gathered}
\gamma_{t}(x) \cdot\left(P_{t} \nabla^{m} g\right)(x)-\left(\theta_{t} \nabla^{m} g\right)(x)=\gamma_{t}(x) \cdot\left(P_{t} \nabla^{m} g\right)(x)-\left(\theta_{t} P_{t} \nabla^{m} g\right)(x) \\
+\theta_{t}\left(P_{t}-I\right)\left(\nabla^{m} g\right)(x)
\end{gathered}
$$

For the last term, using that $P_{t}$ commutes with partial derivatives and that $\theta_{t} \nabla^{m}$ is bounded on $L^{2}$ with bound $C t^{-m}$, we have

$$
\begin{aligned}
\int_{\mathbb{R}^{n}} \int_{0}^{+\infty}\left|\theta_{t}\left(P_{t}-I\right)\left(\nabla^{m} g\right)(x)\right|^{2} \frac{d x d t}{t} & \leq C \int_{\mathbb{R}^{n}} \int_{0}^{+\infty}\left|\left(P_{t}-I\right)(g)(x)\right|^{2} \frac{d x d t}{t^{1+2 m}} \\
& \leq C\left\|\nabla^{m} g\right\|_{2}^{2} .
\end{aligned}
$$

The latter inequality easily follows from the Plancherel theorem and this is where we use the moment conditions on $\varphi$.

Next, $G(x, t)=\gamma_{t}(x) \cdot\left(P_{t} \nabla^{m} g\right)(x)-\left(\theta_{t} P_{t} \nabla^{m} g\right)(x)$ has a kernel representation

$$
\begin{aligned}
G(x, t) & =\iint \theta_{t}(x, y) \cdot t^{-n}\left(\varphi\left(\frac{x-z}{t}\right)-\varphi\left(\frac{y-z}{t}\right)\right) \nabla^{m} g(z) d z d y \\
& \equiv \int K_{t}(x, z) \cdot \nabla^{m} g(z) d z
\end{aligned}
$$

where $\theta_{t}(x, y)$ is the $\left(\mathcal{M}_{N}(\mathbb{C})\right)^{p}$-valued kernel of $\theta_{t}$ whose components are $t^{m} \partial_{y}^{\alpha} W_{t^{2 m}}(x, y) a_{\alpha \beta}(y) \in \mathcal{M}_{N}(\mathbb{C}),|\beta|=m$. By a mere repetition of the proof of [6, Chapter 1, Theorem 29] when $N=1$, we deduce from the order 2 m pointwise upper bound the weighted $L^{2}$ inequalities

$$
\begin{equation*}
\int_{\mathbb{R}^{n}}\left|\partial_{y}^{\alpha} W_{t^{2 m}}(x, y)\right|^{2} e^{\left(\frac{|x-y|}{c t}\right)^{\nu}} d y \leq C t^{-n-2|\alpha|} \tag{2.5}
\end{equation*}
$$

for all $t>0, x \in \mathbb{R}^{n}$ and $|\alpha| \leq m$, for some positive constants $C$ and $c$ depending only on $n, m, N, \lambda, \Lambda$ and the constants in (2.1). [The proof is quite technical and relies on the analyticity of the semigroup, a Caccioppoli inequality and integration by parts.] It easily follows from (2.5) and the properties of $\varphi$ that $\gamma_{t}(x)$ is bounded uniformly in $(x, t)$ and, next, that $K_{t}(x, y)$ satisfies the pointwise bounds

$$
\left|K_{t}(x, z)\right| \leq C t^{-n} e^{-\left(\frac{|x-z|}{c t}\right)^{\nu}}
$$

and

$$
\left|K_{t}(x, z)-K_{t}\left(x, z^{\prime}\right)\right| \leq C t^{-n-1}\left|z-z^{\prime}\right| .
$$

Moreover, one has $\int_{\mathbb{R}^{n}} K_{t}(x, z) d z=0$. Hence, we deduce from the usual almost orthogonality arguments (see [7] or the exposition in Chapter 2 of [6]) that

$$
\int_{\mathbb{R}^{n}} \int_{0}^{+\infty}|G(x, t)|^{2} \frac{d x d t}{t} \leq C\left\|\nabla^{m} g\right\|_{2}^{2}
$$

This concludes the proof of (2.4) and that of Lemma 2.2.

The next step is to use a " $\mathrm{T}(\mathrm{b})$ " argument as in [2], [12]. To this end, fix a cube $Q, \varepsilon \in(0,1)$ and a unit vector $w$ in $\left(\mathbb{C}^{N}\right)^{p}$, and define the $\mathbb{C}^{N}$-valued functions

$$
\begin{equation*}
f_{Q, w}^{\varepsilon}(x)=\left(e^{-(\varepsilon \ell(Q))^{2 m} L}\left(\Phi_{Q} \cdot w\right)\right)(x) \tag{2.6}
\end{equation*}
$$

where

$$
\Phi_{Q}(x)=\left(\frac{(x-x(Q))^{\beta}}{\beta!}\right)_{|\beta|=m}
$$

and where $x(Q)$ is the center of $Q$.
Lemma 2.7. There exist $\varepsilon>0$ and $C<\infty$ depending on $n, m, N, \lambda, \Lambda$ and the constants in (2.1), and a finite collection $W$ of unit vectors in $\left(\mathbb{C}^{N}\right)^{p}$ whose cardinality depends on $\varepsilon, n, m$ and $N$ such that

$$
\begin{aligned}
& \sup \frac{1}{|Q|} \int_{Q} \int_{0}^{\ell(Q)}\left|\gamma_{t}(x)\right|^{2} \frac{d x d t}{t} \leq \\
& \\
& \quad C \sum_{w \in W} \sup \frac{1}{|Q|} \int_{Q} \int_{0}^{\ell(Q)}\left|\gamma_{t}(x) \cdot S_{t}^{Q} \nabla^{m} f_{Q, w}^{\varepsilon}(x)\right|^{2} \frac{d x d t}{t}
\end{aligned}
$$

the suprema running over all cubes $Q$.
Hereafter, if $Q$ is a cube, $S_{t}^{Q}$ is the dyadic averaging operator defined on $\left(\mathbb{C}^{N}\right)^{p}$-valued functions by

$$
S_{t}^{Q} f(x)=\frac{1}{\left|Q^{\prime}\right|} \int_{Q^{\prime}} f(y) d y
$$

for $x \in Q^{\prime}$ and $\frac{1}{2} \ell\left(Q^{\prime}\right)<t \leq \ell\left(Q^{\prime}\right)$ where $Q^{\prime}$ describes a collection of dyadic cubes of $\mathbb{R}^{n}$ that includes $Q$. Note that $\gamma_{t}(x) \cdot S_{t}^{Q} \nabla^{m} f_{Q, w}^{\varepsilon}(x)$ takes its values in $\mathbb{C}^{N}$.

Proof. We follow closely the strategy of proof given in [2] for scalar second order operators. We begin with the following inequality: If $Q$ is a cube, $\varepsilon>0$ and $w=\left(w_{\beta}\right)_{|\beta|=m}, w_{\beta} \in \mathbb{C}^{N}$, is a unit vector in $\left(\mathbb{C}^{N}\right)^{p}$ then

$$
\begin{equation*}
\left|\int_{Q} 1-\operatorname{Re}\left(\nabla^{m} f_{Q, w}^{\varepsilon}(x) \mid w\right) d x\right| \leq C \varepsilon^{1 / 2}|Q| \tag{2.8}
\end{equation*}
$$

where $C$ depends only on $n, m, N, \lambda, \Lambda$ and the constants in (2.1), but not on $\varepsilon, Q$ and $w$. Indeed, by definition of the dot and inner products on $\left(\mathbb{C}^{N}\right)^{p}$,

$$
\left(\nabla^{m}\left(\Phi_{Q} \cdot w\right)(x) \mid w\right)=\partial^{\alpha}\left\{\frac{(x-x(Q))^{\beta} w_{\beta}}{\beta!}\right\} \overline{w_{\alpha}}=w_{\alpha} \overline{w_{\alpha}}=|w|^{2}=1 .
$$

Setting

$$
\begin{equation*}
g(x)=g_{Q, w}^{\varepsilon}(x)=\Phi_{Q}(x) \cdot w-f_{Q, w}^{\varepsilon}(x) \tag{2.9}
\end{equation*}
$$

we have

$$
1-\operatorname{Re}\left(\nabla^{m} f_{Q, w}^{\varepsilon}(x) \mid w\right)=\operatorname{Re}\left(\nabla^{m} g(x) \mid w\right)=\operatorname{Re}\left(\partial^{\alpha} g \overline{w_{\alpha}}\right)(x) .
$$

Next, for a fixed $|\alpha|=m$, write $\partial^{\alpha}=\partial \partial^{\beta}$ where $\partial$ denotes one first order partial and $|\beta|=m-1$, and invoke the following inequality proved in [2]: for some $C=C(n)$

$$
\left|\int_{Q} \partial^{\alpha} g\right| \leq C \ell(Q)^{\frac{n-1}{2}}\left(\int_{Q}\left|\partial^{\beta} g\right|^{2}\right)^{1 / 4}\left(\int_{Q}\left|\partial^{\alpha} g\right|^{2}\right)^{1 / 4}
$$

The technical estimates obtained in Section 3 imply that

$$
\begin{equation*}
\int_{Q}\left|\partial^{\alpha} g\right|^{2} \leq C|Q| \tag{2.10}
\end{equation*}
$$

and

$$
\begin{equation*}
\int_{Q}\left|\partial^{\beta} g\right|^{2} \leq C(\varepsilon \ell(Q))^{2}|Q| \tag{2.11}
\end{equation*}
$$

where $C$ depends only on $n, m, N, \lambda, \Lambda$ and the constants in (2.1), but not on $\varepsilon, Q$ and $w$. This proves (2.8).

Repeating the stopping-time argument in [2], this allows us to obtain
Proposition 2.12. There exists a small $\varepsilon>0$ depending on $n, m, N, \lambda$ and $\Lambda$, and $\eta=\eta(\varepsilon)>0$ such that for each unit vector $w$ in $\left(\mathbb{C}^{N}\right)^{p}$ and cube $Q$, one can find a collection $\mathcal{S}_{w}^{\prime}=\left\{Q^{\prime}\right\}$ of non-overlapping dyadic sub-cubes of $Q$ with the following properties
(i) The union of the cubes in $\mathcal{S}_{w}^{\prime}$ has measure not exceeding $(1-\eta)|Q|$
(ii) If $Q^{\prime \prime} \in \mathcal{S}_{w}^{\prime \prime}$, the collection of all dyadic sub-cubes of $Q$ not contained in any $Q^{\prime} \in \mathcal{S}_{w}^{\prime}$, then

$$
\begin{equation*}
\frac{1}{\left|Q^{\prime \prime}\right|} \int_{Q^{\prime \prime}} \operatorname{Re}\left(\nabla^{m} f_{Q, w}^{\varepsilon}(y) \mid w\right) d y \geq \frac{3}{4} \tag{2.13}
\end{equation*}
$$

and

$$
\begin{equation*}
\left(\frac{1}{\left|Q^{\prime \prime}\right|} \int_{Q^{\prime \prime}}\left|\nabla^{m} f_{Q, w}^{\varepsilon}(y)\right|^{2} d y\right)^{1 / 2} \leq(4 \varepsilon)^{-1} \tag{2.14}
\end{equation*}
$$

At this stage, we want to select a finite collection of unit vectors $w \in$ $\left(\mathbb{C}^{N}\right)^{p}$ to make use of this.

When $m=1$ and $N=1$, which is the case of $[2],\left(\mathbb{C}^{N}\right)^{p}=\mathbb{C}^{n}$ and this selection is achieved via a sectorial decomposition of $\mathbb{C}^{n}$ in which $\gamma_{t}(x)$ also takes its values. Here, the sectorial decomposition is performed on the $\gamma_{t}(x)$ 's and this induces the choice of the $w$ 's.

Identify $\gamma_{t}(x)$ with the element $\Gamma_{t}(x)$ in $\mathcal{L}\left(\left(\mathbb{C}^{N}\right)^{p}, \mathbb{C}^{N}\right)=\mathcal{L}$ by

$$
\Gamma_{t}(x)(v)=\gamma_{t}(x) \cdot v, \quad v \in\left(\mathbb{C}^{N}\right)^{p}
$$

Remark that this identification is isometric. We use as well a single bar norm for the norm induced on $\mathcal{L}$. For $\sigma \in \mathcal{L}$ with $|\sigma|=1$, let $C_{\sigma}$ be the cone of elements $\tau \in \mathcal{L}$ such that

$$
|\tau-|\tau| \sigma| \leq \varepsilon|\tau| .
$$

Choose a finite collection $\Sigma$ of such $\sigma$ 's so as to cover $\mathcal{L}$ with a finite number of cones $C_{\sigma}$. It suffices to argue for each $\sigma \in \Sigma$ fixed and to obtain a Carleson measure estimate for $\Gamma_{t, \sigma}(x) \equiv \mathbf{1}_{\mathcal{C}_{\sigma}}\left(\Gamma_{t}(x)\right) \Gamma_{t}(x)$, where $\mathbf{1}_{\mathcal{C}_{\sigma}}$ denotes the indicator function of $\mathcal{C}_{\sigma}$.

Fix $\sigma \in \Sigma$ and choose $z \in \mathbb{C}^{N}$ such that $|z|=1$ and $\left|\sigma^{*}(z)\right|=1$ where $\sigma^{*}$ is the adjoint of $\sigma$. Remark that $\Gamma_{t, \sigma}(x)^{*}(z) \in\left(\mathbb{C}^{N}\right)^{p}$ satisfies

$$
\left|\Gamma_{t, \sigma}(x)^{*}(z)-\left|\Gamma_{t, \sigma}(x)\right| \sigma^{*}(z)\right| \leq \varepsilon\left|\Gamma_{t, \sigma}(x)\right| .
$$

Set $w=\sigma^{*}(z) \in\left(\mathbb{C}^{N}\right)^{p}$ and apply Proposition 2.12 with that $w$. If $Q^{\prime \prime} \in \mathcal{S}_{w}^{\prime \prime}$ then

$$
v=\frac{1}{\left|Q^{\prime \prime}\right|} \int_{Q^{\prime \prime}} \nabla^{m} f_{Q, w}^{\varepsilon}(y) d y \in\left(\mathbb{C}^{N}\right)^{p}
$$

satisfies

$$
\begin{equation*}
\operatorname{Re}(v \mid w) \geq \frac{3}{4} \quad \text { and } \quad|v| \leq(4 \varepsilon)^{-1} \tag{2.15}
\end{equation*}
$$

Writing

$$
\left(v \mid \Gamma_{t, \sigma}(x)^{*}(z)\right)=\left(v\left|\Gamma_{t, \sigma}(x)^{*}(z)-\left|\Gamma_{t, \sigma}(x)\right| \sigma^{*}(z)\right)+\left|\Gamma_{t, \sigma}(x)\right|(v \mid w)\right.
$$

we deduce from $|z|=1$ and (2.15)

$$
\left|\Gamma_{t, \sigma}(x)(v)\right| \geq\left|\left(v \mid \Gamma_{t, \sigma}(x)^{*}(z)\right)\right| \geq \frac{1}{2}\left|\Gamma_{t, \sigma}(x)\right| .
$$

Next, if $x \in Q^{\prime \prime}$ and $\frac{1}{2} \ell\left(Q^{\prime \prime}\right)<t \leq \ell\left(Q^{\prime \prime}\right)$ then $v=\left(S_{t}^{Q} \nabla f_{Q, w}^{\varepsilon}\right)(x)$ and we have obtained using the identification of $\Gamma_{t, \sigma}(x)$ with $\gamma_{t, \sigma}(x)=\mathbf{1}_{\mathcal{C}_{\sigma}}\left(\Gamma_{t}(x)\right) \gamma_{t}(x)$ that

$$
\begin{equation*}
\left|\gamma_{t, \sigma}(x)\right| \leq 2\left|\gamma_{t, \sigma}(x) \cdot\left(S_{t}^{Q} \nabla f_{Q, w}^{\varepsilon}\right)(x)\right| \leq 2\left|\gamma_{t}(x) \cdot\left(S_{t}^{Q} \nabla f_{Q, w}^{\varepsilon}\right)(x)\right| \tag{2.16}
\end{equation*}
$$

The conclusion to the proof of Lemma 2.7 is now exactly as in [2] and $W$ is here the collection of vectors $w$ just constructed from the $\sigma$ 's in $\Sigma$.

The next lemma is the last step. Compared to [6], Chapter 3, the argument is slightly simplified due to the fact that functions are defined on all of $\mathbb{R}^{n}$.

Lemma 2.17. For some constant $C$ depending on $n, m, N, \lambda, \Lambda$, the constants in (2.1) and $\varepsilon>0$, but not on $Q$ and $w$, we have

$$
\int_{Q} \int_{0}^{\ell(Q)}\left|\gamma_{t}(x) \cdot S_{t}^{Q} \nabla^{m} f_{Q, w}^{\varepsilon}(x)\right|^{2} \frac{d x d t}{t} \leq C|Q|
$$

Proof. Fix $Q, w$ and $\varepsilon$, and set $f \equiv f_{Q, w}^{\varepsilon}$. Let $\mathcal{X}$ be a smooth function supported in $4 Q$, which is identically 1 on $2 Q$ and such that $\left\|\nabla^{j} \mathcal{X}\right\|_{\infty} \leq$ $C \ell(Q)^{-j}$ for all $j=0, \ldots, m$. For $x \in Q$ and $0<t \leq \ell(Q)$, we have

$$
\left(S_{t}^{Q} \nabla^{m} f\right)(x)=\left(S_{t}^{Q} \nabla^{m}(\mathcal{X} f)\right)(x)
$$

Using that

$$
\int_{\mathbb{R}^{n}} \int_{0}^{+\infty}\left|\left(S_{t}^{Q}-P_{t}\right) h(x)\right|^{2} \frac{d x d t}{t} \leq C \int_{\mathbb{R}^{n}}|h|^{2}
$$

(See [6], Appendix C for a proof) and (2.4), we obtain

$$
\begin{aligned}
\int_{Q} \int_{0}^{\ell(Q)}\left|\gamma_{t}(x) \cdot\left(S_{t}^{Q} \nabla^{m} f\right)(x)\right|^{2} \frac{d x d t}{t} & \leq C \int_{\mathbb{R}^{n}}\left|\nabla^{m}(\mathcal{X} f)\right|^{2} \\
& +4 \int_{Q} \int_{0}^{\ell(Q)}\left|\left(\theta_{t} \nabla^{m}(\mathcal{X} f)\right)(x)\right|^{2} \frac{d x d t}{t}
\end{aligned}
$$

We shall prove in Section 3 that

$$
\begin{equation*}
\int_{\mathbb{R}^{n}}\left|\nabla^{m}(\mathcal{X} f)\right|^{2} \leq C|Q| \tag{2.18}
\end{equation*}
$$

where $C$ is independent of $Q$ and $w$ (it may depend on $\varepsilon$ which we allow).
It remains to estimate $\int_{Q} \int_{0}^{\ell(Q)}\left|\left(\theta_{t} \nabla^{m}(\mathcal{X} f)\right)(x)\right|^{2} \frac{d x d t}{t}$. To this end, write

$$
\theta_{t} \nabla^{m}(\mathcal{X} f)=e^{-t^{2 m} L} t^{m} L f+e^{-t^{2 m} L} t^{m} L((\mathcal{X}-1) f) .
$$

and treat each term by separate arguments.
To handle the first term, observe that (see Section 3 for a proof)

$$
\begin{equation*}
|L f(x)| \leq C(\varepsilon \ell(Q))^{-m} \tag{2.19}
\end{equation*}
$$

with $C$ independent of $Q, w, \varepsilon$ and $x$. Using (2.1) which implies $L^{\infty}$ boundedness of $e^{-t^{2 m} L}$, we obtain

$$
\left|\left(e^{-t^{2 m} L} t^{m} L f\right)(x)\right| \leq C t^{m}(\varepsilon \ell(Q))^{-m}
$$

from which we deduce

$$
\int_{Q} \int_{0}^{\ell(Q)}\left|\left(e^{-t^{2 m} L} t^{m} L f\right)(x)\right|^{2} \frac{d x d t}{t} \leq C \varepsilon^{-2 m}|Q|
$$

To handle the second term, observe that the kernel of $e^{-t^{2 m}} t^{2 m} L$ satisfies an upper bound similar to that of $e^{-t^{2 m} L}$ with different constants $C, c$ (it is a classical consequence of the analyticity of the semigroup). Hence for $x \in Q$ and $t \leq \ell(Q)$,

$$
\begin{aligned}
\left|\left(e^{-t^{2 m} L} t^{m} L((\mathcal{X}-1) f)\right)(x)\right| & \leq C t^{-m-n} \int_{y \notin 2 Q} e^{-\left(\frac{|x-y|}{c t}\right)^{\nu}}|f(y)| d y \\
& \leq C t^{-m-n} e^{-\frac{1}{2}\left(\frac{\ell(Q)}{2 c t}\right)^{\nu}} \int_{y \notin 2 Q} e^{-\frac{1}{2}\left(\frac{\mu|x(Q)-y|}{c \ell(Q)}\right)^{\nu}}|f(y)| d y
\end{aligned}
$$

where $\mu$ is the largest number (depending only on $n$ ) so that $|x-y| \geq$ $\mu|x(Q)-y|$ for all $x \in Q$ and $y \notin 2 Q$. We used that $|x-y| \geq \ell(Q) / 2$ for $x \in Q$ and $y \notin 2 Q$. Recall also that $\nu=\frac{2 m}{2 m-1}$. We now invoke the following estimate, to be proved in Section 3,

$$
\begin{equation*}
\int_{\mathbb{R}^{n}} e^{-\left(\frac{|x(Q)-y|}{b \ell(Q)}\right)^{\nu}}|f(y)| d y \leq C(\ell(Q))^{n+m} \tag{2.20}
\end{equation*}
$$

for all $b>0$ with $C$ depending only $n, m, N, \lambda, \Lambda, \varepsilon, b$ and the constants in (2.1). Therefore, we obtain a pointwise bound

$$
\left|\left(e^{-t^{2 m} L} t^{m} L((\mathcal{X}-1) f)\right)(x)\right| \leq C t^{-m-n}(\ell(Q))^{n+m} e^{-\frac{1}{2}\left(\frac{\ell(Q)}{2 c t}\right)^{\nu}}
$$

and this yields straightforwardly

$$
\int_{Q} \int_{0}^{\ell(Q)}\left|\left(e^{-t^{2 m} L} t^{m} L((\mathcal{X}-1) f)\right)(x)\right|^{2} \frac{d x d t}{t} \leq C|Q|
$$

the latter $C$ depending only on $n, m, N, \lambda, \Lambda$, the constants in (2.1) and $\varepsilon$.
We have proved Lemma 2.17 modulo the technical estimates (2.18), (2.19) and (2.20) that we explain in the Section 3.

### 2.2 Removing the pointwise upper bound

In this section, we are still given an operator $L$ of order $2 m$ with a representation (1.2) and ellipticity constants $\lambda$ and $\Lambda$. We begin with recalling the following result.

Proposition 2.21. If, in addition, $2 m \geq n$, then $L$ satisfies the order $2 m$ pointwise upper bound and the constants in (2.1) depend only on $n, m, N, \lambda$ and $\Lambda$.

Proof. See [9] for the self-adjoint case with $2 m>n$ and $N=1$, and [6], Chapter 1, Proposition 28, for the general case when $N=1$. The proof is identical when $N \geq 2$. We do stress that the estimates are global in time (i.e., the constant $C$ does not depend on $t$ ).

To take advantage of this, we shall increase the order of $L$. Consider the positive self-adjoint unbounded operator on $L^{2}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$ by

$$
S g=\left((-1)^{m} \sum_{|\alpha|=m} \partial^{2 \alpha}\right)^{1 / 2} g, \quad g \in H^{m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)
$$

Then $S^{2}$ is an elliptic homogeneous operator of order $2 m$ and $\|S g\|_{2}=$ $\left\|\nabla^{m} g\right\|_{2}$. For $k \in 2 \mathbb{N}$, define $L_{k}$ as the maximal accretive operator associated with the form $\left\langle a_{\alpha \beta} \partial^{\alpha} S^{k} f, \partial^{\beta} S^{k} g\right\rangle$ on $H^{m(1+k)}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$. Formally, $L_{k}=S^{k} L S^{k}$. Then $L_{k}$ is an homogeneous operator of order $2 m(1+k)$ and it has a representation (1.2) with ellipticity constants $\lambda_{k}$ and $\Lambda_{k}$ in the inequalities corresponding to (1.3) and (1.4) depending only on $n, m, N, k, \lambda$ and $\Lambda$. These observations are easy consequences of $\left\|S^{k+1} f\right\|_{2} \sim\left\|\nabla^{m(1+k)} f\right\|_{2}$ deduced from the Plancherel theorem.

Proposition 2.22. Under the above hypotheses, the inequalities (K) for $L$ and $L_{k}$ are equivalent. More precisely, the inequalities $\|\sqrt{L} f\|_{2} \leq C\|S f\|_{2}$ and $\left\|\sqrt{L_{k}} f\right\|_{2} \leq C_{k}\left\|S^{k+1} f\right\|_{2}$ hold simultaneously. Furthermore, $C$ depends on $C_{k}, k, \lambda$ and $\Lambda$.

Proof. This was proved for scalar second order operators in [6], Chapter 0, Proposition 10. The argument is similar. Write $L=S B S$ where $B=$ $S^{-1} L S^{-1}$. It follows from its definition and (1.4) that $B$ is a bounded, invertible, and $\omega$-accretive operator on $L^{2}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$, with $\|B\|,\left\|B^{-1}\right\|$ and $\omega \in[0, \pi / 2)$ depending only on $\lambda$ and $\Lambda$. As $L_{k}=S^{k+1} B S^{k+1}$, it remains to apply the general interpolation result in Hilbert spaces from [4] as quoted in [6], Chapter 0, Proposition 9. The constant $C$ depends only on $k, C_{k},\|B\|$, $\left\|B^{-1}\right\|$ and $\omega$, and symetrically $C_{k}$ depends only on $k, C,\|B\|,\left\|B^{-1}\right\|$ and $\omega$.

We may finish the proof of Theorem 1.5. If $L$ satisfies the order $2 m$ pointwise upper bound with constants depending only on $n, m, N, \lambda$ and $\Lambda$ we are done. Otherwise choose the smallest integer $k$ so that $2 m(1+k) \geq$ $n$. By Proposition 2.21, the operator $L_{k}$ defined above satisfies the order $2 m(1+k)$ pointwise upper bound with constants depending only on $n, m$, $\lambda$ and $\Lambda$. Thus $\left\|\sqrt{L_{k}} f\right\|_{2} \leq C_{k}\left\|\nabla^{m(1+k)} f\right\|_{2}$ holds by what we just did in the previous section and $C_{k}$ depends only on $n, m, N, \lambda_{k}$ and $\Lambda_{k}$. By Proposition 2.22, $\|\sqrt{L} f\|_{2} \leq C\left\|\nabla^{m} f\right\|_{2}$ holds with $C$ depending on $n, m, N$ $\lambda$ and $\Lambda$. Hence (K) is completely proved.

## 3 Proof of technical estimates

Here, we prove the technical estimates left aside, namely (2.10), (2.11), (2.18), (2.19) and (2.20).

We begin with the appropriate extension to higher order operators of the classical conservation property for scalar second order operators, that is the reproduction of polynomials by the semigroup.

Lemma 3.1. If $L$ satisfies the order $2 m$ pointwise upper bound, then

$$
\begin{equation*}
e^{-t L} P=P \tag{3.2}
\end{equation*}
$$

holds pointwise for any $P=\left(P_{1}, \ldots, P_{N}\right)$ and $t>0$ where the $P_{i}$ 's are polynomials with degrees not exceeding $m-1$.

Remark. The upper bound assumption can be dropped for the equality to hold say in $L_{\text {loc }}^{2}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$ but the argument is much more involved.

Proof. To simplify the exposition, we assume that $N=1$. The proof in the general case is the same.

Let $\mathcal{X}$ be a smooth function with $\mathcal{X}(x)=1$ if $|x| \leq 1$ and $\mathcal{X}(x)=0$ if $|x| \geq 2$. Let $\mathcal{X}_{R}(x)=\mathcal{X}(x / R)$ for $R>0$. If $\phi$ is a smooth compactly supported function, then for $R>0$ and $t>0$

$$
\begin{equation*}
\left\langle P, e^{-t L^{*}} \phi\right\rangle=\left\langle P \mathcal{X}_{R}, e^{-t L^{*}} \phi\right\rangle+\left\langle P\left(1-\mathcal{X}_{R}\right), e^{-t L^{*}} \phi\right\rangle \tag{3.3}
\end{equation*}
$$

where the brackets make sense as Lebesgue integrals by the decay of the kernel of $e^{-t L^{*}}$ and the support of $\phi$. We use this representation twice, first to show that the left hand side does not depend on $t>0$ and, second, to find $\langle P, \phi\rangle$ as its value. This, indeed, shows that $e^{-t L} P=P$ in the sense of distributions. The equality also holds in the pointwise sense using again kernel decay.

Let us begin with differentiating (3.3) with respect to $t$. Since the kernel of $\frac{d}{d t} e^{-t L^{*}}$ has also Gaussian decay, we have

$$
\frac{d}{d t}\left\langle P\left(1-\mathcal{X}_{R}\right), e^{-t L^{*}} \phi\right\rangle=\left\langle P\left(1-\mathcal{X}_{R}\right), \frac{d}{d t} e^{-t L^{*}} \phi\right\rangle
$$

and by Lebesgue dominated convergence, this quantity tends to 0 as $R$ tends to $+\infty$.

Next, since $P \mathcal{X}_{R} \in L^{2}$ we have that

$$
\frac{d}{d t}\left\langle P \mathcal{X}_{R}, e^{-t L^{*}} \phi\right\rangle=\sum_{|\alpha|=|\beta|=m}\left\langle a_{\alpha \beta} \partial^{\beta}\left(P \mathcal{X}_{R}\right), \partial^{\alpha} e^{-t L^{*}} \phi\right\rangle
$$

Since the degree $d$ of $P$ does not exceed $m-1$, observe from Leibniz formula that $\partial^{\beta}\left(P \mathcal{X}_{R}\right)$ is supported in the annulus $R \leq|x| \leq 2 R$ and is dominated by $C R^{d-m}$. Using this remark, together with a weighted $L^{1}$ estimate (which follows from the weighted $L^{2}$ estimate (2.5) by CauchySchwarz) and the fact that the support of $\phi$ is compact, it is easy to show that $\left\langle a_{\alpha \beta} \partial^{\beta}\left(P \mathcal{X}_{R}\right), \partial^{\alpha} e^{-t L^{*}} \phi\right\rangle$ tends to 0 as $R$ tends to $+\infty(t$ is fixed).

This shows that the left hand side of (3.3) is independent of $t>0$. In the right hand side, choose and fix $R$ large enough so that the supports of $\phi$ and $P\left(1-\mathcal{X}_{R}\right)$ are far apart. The decay of the kernel of $e^{-t L^{*}}$ yields that $\left\langle P\left(1-\mathcal{X}_{R}\right), e^{-t L^{*}} \phi\right\rangle$ tends to 0 as $t$ tends to 0 . Eventually, since $e^{-t L^{*}}$ is a continuous semigroup on $L^{2}$ at $t=0$, we obtain that $\left\langle P \mathcal{X}_{R}, e^{-t L^{*}} \phi\right\rangle$ tends to $\left\langle P \mathcal{X}_{R}, \phi\right\rangle=\langle P, \phi\rangle$ as $t$ tends to 0 . This proves (3.2).

We continue with
Lemma 3.4. Assume that $L$ satisfies the order $2 m$ pointwise upper bound. If $P=\left(P_{1}, \ldots, P_{N}\right)$ with the $P_{i}$ 's polynomials of degrees not exceeding $m, Q$ is a cube and $0<t \leq \ell(Q)$ then

$$
\begin{equation*}
\int_{5 Q}\left|\nabla^{j}\left(e^{-t^{2 m} L} P-P\right)(x)\right|^{2} d x \leq C\left(\sup _{\substack{|\alpha|=m \\ 1 \leq i \leq N}}\left|P_{i}^{(\alpha)}\right|\right)^{2} t^{2(m-j)}|Q|, \tag{3.5}
\end{equation*}
$$

where $C$ depends on $n, m, N, \lambda, \Lambda$ and the constants in (2.1), but not on $P, Q, t$ and $0 \leq j \leq m$.

Remark. This lemma holds without the upper bound assumption as well.
Proof. Again we assume $N=1$ to simplify matters. Write

$$
\left(e^{-t^{2 m} L} P\right)(x)=\left(e^{-t^{2 m} L}\left(P-P_{z}\right)\right)(x)+P_{z}(x)
$$

where $P_{z}$ is the Taylor polynomial of $P$ at $z$ with degree $m-1$. Letting $z=x$, we obtain

$$
\left(e^{-t^{2 m} L} P\right)(x)=\int_{\mathbb{R}^{n}} W_{t^{2 m}}(x, y)\left(P-P_{x}\right)(y) d y+P(x)
$$

Since

$$
\left(P-P_{x}\right)(y)=\sum_{|\alpha|=m} \frac{p_{\alpha}}{\alpha!}(y-x)^{\alpha}
$$

with $p_{\alpha} \in \mathbb{C}$ the value of the constant polynomial $P^{(\alpha)}$, we arrive at

$$
\begin{equation*}
\left(e^{-t^{2 m} L} P-P\right)(x)=\sum_{|\alpha|=m} \frac{p_{\alpha}}{\alpha!} \int_{\mathbb{R}^{n}} W_{t^{2 m}}(x, y)(y-x)^{\alpha} d y \tag{3.6}
\end{equation*}
$$

Hence, by the Leibniz formula we have

$$
\left|\nabla^{j}\left(e^{-t^{2 m} L} P-P\right)(x)\right| \leq C\left(\sup \left|p_{\alpha}\right|\right) \sum_{k=0}^{j} f_{j, k}(x)
$$

with

$$
f_{j, k}(x)=\int_{\mathbb{R}^{n}}\left|\nabla_{x}^{k} W_{t^{2 m}}(x, y) \| x-y\right|^{m-j+k} d y
$$

Choose $\mu>0$ depending on $n$ such that if $x \in 5 Q$ and $y \notin 10 Q$ we have

$$
\mu|x(Q)-y| \leq \frac{1}{2}|x-y|
$$

One easily checks that for $x \in 5 Q$ and $t \in(0, \ell(Q)]$,

$$
\int_{\mathbb{R}^{n}}|x-y|^{2 \ell} e^{-\left(\frac{|x-y|}{c t}\right)^{\nu}} e^{\nu\left(\frac{\mu x(Q)-y \mid}{c \ell(Q)}\right)^{\nu}} d y \leq C t^{2 \ell+n}
$$

where $c$ is the constant in (2.5), in which the roles of $x$ and $y$ are switched, and $C$ is independent of $x, Q, t \in(0, \ell(Q)]$ and $\ell \in\{0, \ldots, m\}$. Hence by Cauchy-Schwarz inequality and (2.5),

$$
\begin{aligned}
\int_{5 Q}\left|f_{j, k}(x)\right|^{2} d x & \leq C t^{2(m-j+k)+n} \int_{\mathbb{R}^{n}} \int_{\mathbb{R}^{n}}\left|\nabla_{x}^{k} W_{t^{2 m}}(x, y)\right|^{2} e^{\left(\frac{|x-y|}{c t}\right)^{\nu}} \times \\
& \leq C t^{2(m-j)} \int_{\mathbb{R}^{n}} e^{-\left(\frac{\mu|x(Q)-y|}{c(Q)}\right)^{\nu}} d y \\
& \leq C t^{2(m-j)} \ell(Q)^{n} .
\end{aligned}
$$

We now return to proving the needed estimates. Recall that $f=f_{Q, w}^{\varepsilon}$ in (2.6) and $g$ is defined by (2.9). We have, therefore, $f=e^{-t^{2 m} L} P$ and $g=P-e^{-t^{2 m} L} P$ where $P(x)=\Phi_{Q}(x) \cdot w$ and $t=\varepsilon \ell(Q)$. Also observe that
$p_{\alpha}=w_{\alpha}$ and since $w$ is a unit vector, $\sup \left|p_{\alpha}\right| \leq 1$.
Proof of (2.10) and (2.11). Straightforward from (3.5) with $j=m$ and $j=m-1$.

Proof of (2.18). Write $\nabla^{m}(\mathcal{X} f)=\nabla^{m}\left(\mathcal{X}\left(e^{-t^{2 m} L} P-P\right)\right)+\nabla^{m}(\mathcal{X} P)$. Recall that the components of $\Phi_{Q}(x)$ are $\frac{(x-x(Q))^{\beta}}{\beta!}$ so that it is clear from the choice of $\mathcal{X}$ that $\nabla^{m}(\mathcal{X} P)$ is bounded in $\mathbb{R}^{n}$ with a bound independent of $Q$ and $w$. As for $\nabla^{m}\left(\mathcal{X}\left(e^{-t^{2 m} L} P-P\right)\right)$, we use (3.5) with the Leibniz rule to obtain

$$
\int_{\mathbb{R}^{n}}\left|\nabla^{m}\left(\mathcal{X}\left(e^{-t^{2 m} L} P-P\right)\right)\right|^{2} \leq C\left(1+\varepsilon^{2}+\ldots+\varepsilon^{2 m}\right)|Q|
$$

with $C$ independent of $Q, w$ and $t=\varepsilon \ell(Q)$.
Proof of (2.19). Using the notation above, it amounts to showing the uniform bound

$$
\left|L e^{-t^{2 m} L} P(x)\right| \leq C t^{-m}
$$

for any cube $Q, 0<t \leq \ell(Q)$ where $C$ is independent of $P, Q$ and $t$. Using the same ideas as before, we have

$$
\left(L e^{-t^{2 m} L} P\right)(x)=\left(L e^{-t^{2 m} L}\left(P-P_{z}\right)\right)(x)
$$

where $P_{z}$ is the Taylor polynomial of $P$ at $z$ with degree $m-1$ since $L P_{z}=0$. Letting $z=x$, we obtain

$$
\left(L e^{-t^{2 m} L} P\right)(x)=t^{-m} \sum_{|\alpha|=m} \frac{p_{\alpha}}{\alpha!} \int_{\mathbb{R}^{n}} \widetilde{W}_{t^{2 m}}(x, y)\left(\frac{y-x}{t}\right)^{\alpha} d y
$$

where $\widetilde{W}_{t}(x, y)$ is the kernel of $t L e^{-t L}=-\frac{d e^{-t L}}{d t}$. As this kernel satisfies a similar bound to $W_{t}(x, y)$, the latter integrals are bounded uniformly in $x$ and $t$ and the desired bound follows at once.

Proof of (2.20). Using (2.1) and the definition of $f=f_{Q, w}^{\varepsilon}$, we have

$$
|f(y)| \leq C(\varepsilon \ell(Q))^{-n} \int_{\mathbb{R}^{n}} e^{-\left(\frac{|y-z|}{c \varepsilon \ell(Q)}\right)^{\nu}}|z-x(Q)|^{m} d z
$$

with $C$ independent of $Q, w$ and $\varepsilon$. Using the convolution inequality

$$
\int_{\mathbb{R}^{n}} e^{-\left(\frac{|x(Q)-y|}{b \ell(Q)}\right)^{\nu}} e^{-\left(\frac{\mid y-z z}{c \ell(Q)}\right)^{\nu}} d y \leq C(\ell(Q))^{n} e^{-\left(\frac{|z-x(Q)|}{c^{\ell} \ell(Q)}\right)^{\nu}}
$$

with $C, c^{\prime}$ depending on $n, m, N, b$ and $c$, we obtain

$$
\begin{aligned}
\int_{\mathbb{R}^{n}} e^{-\left(\frac{|x(Q)-y|}{b \ell(Q)}\right)^{\nu}}|f(y)| d y & \leq C \int_{\mathbb{R}^{n}} e^{-\left(\frac{|z-x(Q)| \mid}{c^{\ell} \ell(Q)}\right)^{\nu}}|z-x(Q)|^{m} d z \\
& \leq C(\ell(Q))^{n+m} .
\end{aligned}
$$

Remark. In (2.20), the weight could be weakened to $\left(1+\frac{|x(Q)-y|}{\ell(Q)}\right)^{-n-m-\delta}$ for any $\delta>0$.

## 4 Inhomogeneous operators

Proof of Theorem 1.11. Take $L$ with a representation (1.8) and ellipticity constants $\tilde{\lambda}$ and $\widetilde{\Lambda}$.

First we raise the order of $L$ if necessary. Let

$$
\begin{equation*}
S g=\left(\sum_{|\alpha| \leq m}(-1)^{|\alpha|} \partial^{2 \alpha}\right)^{1 / 2} g, \quad g \in H^{m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right) \tag{4.1}
\end{equation*}
$$

so that $S^{2}$ is an elliptic inhomogeneous operator of order $2 m$ and $L=S B S$ with $B$ bounded, invertible and $\omega$-accretive on $L^{2}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$. Then for $k \in 2 \mathbb{N}$ so that $2 m(1+k) \geq n$, set $L_{k}=S^{k} L S^{k}$. Then $L_{k}$ is an inhomogeneous operator of order $2 m(1+k)$ and it has a representation (1.8) with ellipticity constants $\tilde{\lambda}_{k}$ and $\widetilde{\Lambda}_{k}$ in the inequalities corresponding to (1.9) and (1.10) for $L_{k}$ depending only on $n, m, k, \tilde{\lambda}$ and $\widetilde{\Lambda}$. The square root problems for $L$ and $L_{k}$ are equivalent.

Again, if $2 m(1+k) \geq n, L_{k}$ satisfies the local order $2 m(1+k)$ pointwise upper bound (See [6]) with constants depending on $n, m, N, \tilde{\lambda}$ and $\widetilde{\Lambda}$. The local order $2 m$ pointwise upper bound for an operator $L$ means that there exists $c>0$ such that for all $T>0$ there exists a constant $C$ for which

$$
\begin{equation*}
\left|W_{t^{2 m}}(x, y)\right| \leq C t^{-n} e^{-\left(\frac{|x-y|}{c t}\right)^{\nu}} \tag{4.2}
\end{equation*}
$$

for almost every $(x, y) \in \mathbb{R}^{2 n}$ and all $0<t<T$ where $W_{t}(x, y)$ is again the $\mathcal{M}_{N}(\mathbb{C})$-valued Schwartz kernel of $e^{-t L}$. By the semigroup property, it
suffices to obtain this for one $T>0$ and $C$ usually blows up like $e^{\omega T}$. For simplicity, we choose $T=1$. Hence, we may assume from now on that $L$ has order larger than dimension.

Secondly, let $L_{0}$ be the principal part of $L$, that is the homogeneous part of degree $2 m$ in (1.8). Using interpolation inequalities, it is easy to show that the weak Gårding inequality (1.13) holds for $L_{0}$ with $\lambda=\tilde{\lambda} / 2$ and some $\kappa>0$ depending only on $n, m, N, \tilde{\lambda}$ and $\widetilde{\Lambda}$. We begin with studying the square root of $L_{0}+2 \kappa$, and then return to $L$.

The operator $L_{0}+2 \kappa$ satisfies (1.10) with constant $\lambda^{\sharp}=\inf (\lambda, \kappa)$ and (1.9) with some $\Lambda^{\sharp}$ depending on $n, m, N, \kappa$ and $\widetilde{\Lambda}$. Again, we are interested in establishing only

$$
\left\|\sqrt{L_{0}+2 \kappa} f\right\|_{2} \leq C\|f\|_{H^{m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)}, \quad f \in \mathcal{D}\left(L_{0}\right)
$$

as the inequality for the adjoint is similarly obtained. By the $\mathrm{M}^{\mathrm{c}}$ Intosh-Yagi theorem [16]

$$
\left\|\sqrt{L_{0}+2 \kappa} f\right\|_{2}^{2} \sim \int_{0}^{\infty}\left\|e^{-t^{2 m}\left(L_{0}+2 \kappa\right)} t^{m}\left(L_{0}+2 \kappa\right) f\right\|_{2}^{2} \frac{d t}{t}, \quad f \in \mathcal{D}\left(L_{0}\right)
$$

Since $e^{-t^{2 m}\left(L_{0}+2 \kappa\right)} t^{2 m}\left(L_{0}+2 \kappa\right)$ is uniformly bounded on $L^{2}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$ (or rather, has a bounded extension with uniform bound), we have

$$
\int_{1}^{\infty}\left\|e^{-t^{2 m}\left(L_{0}+2 \kappa\right)} t^{m}\left(L_{0}+2 \kappa\right) f\right\|_{2}^{2} \frac{d t}{t} \leq C\|f\|_{2}^{2}
$$

Since $e^{-t^{2 m}\left(L_{0}+2 \kappa\right)}$ is a contraction on $L^{2}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$ we have

$$
\int_{0}^{1}\left\|e^{-t^{2 m}\left(L_{0}+2 \kappa\right)} t^{m}(2 \kappa) f\right\|_{2}^{2} \frac{d t}{t} \leq \int_{0}^{1} t^{2 m}(2 \kappa)^{2} \frac{d t}{t}\|f\|_{2}^{2} \leq C\|f\|_{2}^{2}
$$

Finally using $e^{-t^{2 m}\left(L_{0}+2 \kappa\right)}=e^{-t^{2 m} 2 \kappa} e^{-t^{2 m} L_{0}}$ and $e^{-u} \leq 1$ for $u \geq 0$, it remains to checking the inequality

$$
\int_{0}^{1}\left\|e^{-t^{2 m} L_{0}} t^{m} L_{0} f\right\|_{2}^{2} \frac{d t}{t} \leq C\left\|\nabla^{m} f\right\|_{2}^{2}
$$

At this point, we are back to the case of homogeneous operators and the same algorithm as in Section 2.1 applies with the restriction that Carleson norms are taken on cubes of sidelength less than 1 since $L_{0}$ satisfies the local upper
bound (because $L_{0}+2 \kappa$ does). All the technical estimates apply provided $t \leq 1$ and $\ell(Q) \leq 1$. We have established the desired inequality for $L_{0}+2 \kappa$.

We return to the operator $L$ and wish to prove at last

$$
\|\sqrt{L} f\|_{2} \leq C\|f\|_{H^{m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)}, \quad f \in \mathcal{D}(L) .
$$

Recall that we just obtained the same inequality with $L_{0}+2 \kappa$. The point is to write $L$ as a perturbation of $L_{0}+2 \kappa$. Set $\Sigma=S^{1 / m}$, with $S$ defined by (4.1) and write

$$
L f=\sum_{k, l \leq m} \Sigma^{k} B_{k, l} \Sigma^{l} f
$$

with

$$
B_{k, l}=\Sigma^{-k}\left((-1)^{k} \sum_{|\alpha|=k,|\beta|=l} \partial^{\alpha}\left(a_{\alpha \beta} \partial^{\beta}\right)\right) \Sigma^{-l}
$$

for all non negative integers $k, l \leq m$ except when $k=l=0$ for which

$$
B_{0,0}=a_{00}-2 \kappa
$$

and $k=l=m$ for which

$$
B_{m, m}=\Sigma^{-m}\left((-1)^{m} \sum_{|\alpha|=m,|\beta|=m} \partial^{\alpha}\left(a_{\alpha \beta} \partial^{\beta}\right)+2 \kappa\right) \Sigma^{-m}=\Sigma^{-m}\left(L_{0}+2 \kappa\right) \Sigma^{-m}
$$

The operators $B_{k, l}$ are bounded operators on $L^{2}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$ and the operator $B_{m, m}$ is nothing but $B$ defined above. The above representation of $L f$ works for $f$ in $H^{m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$ as $L$ is bounded from $H^{m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$ into $H^{-m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$.

It remains to use [6], Chapter 0, Proposition 11, to conclude the proof of Theorem 1.11.

Proof of Proposition 1.14. Under our assumption, the theory of regularly accretive forms of Kato asserts that for all $\kappa^{\prime} \geq \kappa, L+\kappa^{\prime}$ have square roots with same domains. By the preceding result $\mathcal{D}(\sqrt{L+\kappa})=\mathcal{D}(\sqrt{L+2 \kappa})=$ $H^{m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$. Introduce $C(n, m, N, \lambda, \Lambda, \kappa)$ as the best constant $C$ such that

$$
\|\sqrt{L+\kappa} f\|_{2} \leq C\left(\left\|\nabla^{m} f\right\|_{2}^{2}+\kappa\|f\|_{2}^{2}\right)^{1 / 2}
$$

holds for all operators $L$ given by a representation (1.2) with constants $\lambda, \Lambda$ and $\kappa$ in (1.3) and (1.13). A dilation argument with dilation factor $s>0$ in $\mathbb{R}^{n}$ changes $L$ to another homogeneous operator with constants $\lambda, \Lambda, \kappa s^{2 m}$. At the same time, $\left(\left\|\nabla^{m} f\right\|_{2}^{2}+\kappa\|f\|_{2}^{2}\right)^{1 / 2}$ changes to $\left(\left\|\nabla^{m} f\right\|_{2}^{2}+\kappa s^{2 m}\|f\|_{2}^{2}\right)^{1 / 2}$. This implies that $C\left(n, m, N, \lambda, \Lambda, \kappa s^{2 m}\right) \leq C(n, m, N, \lambda, \Lambda, \kappa)$ for all $s, \kappa>0$ so that $C(n, m, N, \lambda, \Lambda, \kappa)$ is independent of $\kappa$. This proves Proposition 1.14.

## 5 Proof of Proposition 1.18

We give a sketch of the proof as it follows closely the ideas in [1]. Let $L$ denote now the operator associated to the $b_{\alpha \beta}$. As this operator is self-adjoint, the weak Gårding inequality is equivalent to the invertibility of

$$
L+\kappa: H^{m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right) \rightarrow H^{-m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)
$$

for some $\kappa>0$. This is done by constructing a parametrix using wavelet expansions.

Let $\Psi_{j, k, \ell}$ be a Schwartz class wavelet basis of $L^{2}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$. Here $j \in \mathbb{Z}$, $k \in \mathbb{Z}^{n}$ and $\ell$ belongs to some finite set with cardinality $N\left(2^{n}-1\right)$. We have $\Psi_{j, k, \ell}(x)=2^{n j / 2} \Psi_{\ell}\left(2^{j} x-k\right)$ where $\Psi_{\ell}$ is valued in $\mathbb{C}^{N}$ and has a Fourier localisation in an annulus $a \leq|\xi| \leq b$. Also one may think that $\Psi_{j, k, \ell}$ has essential support in the dyadic cube $Q_{j, k}$ defined by $2^{j} x-k \in\left[0,1{ }^{n}\right.$ (In fact it can be chosen with rapid decay away from this cube). We index the wavelets by the dyadic cubes and drop the index $\ell$ as it plays no role, so that $\Psi_{j, k, \ell}$ becomes $\Psi_{Q}$ and $j=j(Q)$ is the scale index of $Q$; the larger $j(Q)$, the smaller $Q$. Such a wavelet basis characterizes $B M O$ and the weight function $w$ takes the following form: define

$$
N_{j}(b)=\sup _{R ; j(R) \geq j}|R|^{-1} \sum_{Q \subset R}\left|<b, \Psi_{Q}>\right|^{2} .
$$

Then $b \in B M O$ is characterized by $\sup _{j} N_{j}(b)<+\infty$ and $b \in v m o$ by $b \in B M O$ and $\lim _{j \rightarrow \infty} N_{j}(b)=0$. The condition (1.19) can be rephrased as $\sup _{\alpha, \beta} \lim _{j \rightarrow \infty} N_{j}\left(b_{\alpha \beta}\right) \leq \varepsilon$. See [17].

Define a collection of functions $\theta_{Q}$ by

$$
\theta_{Q}=\left(L_{Q}+\kappa\right)^{-1}\left(\Psi_{Q}\right)
$$

where $L_{Q}$ is the operator with coefficients $b_{\alpha \beta}$ replaced by their averages over $Q$. Hence, $L_{Q}$ have constant coefficients and satisfy ellipticity uniformly. By straightforward Fourier estimates, for $\sigma_{j}=\frac{4^{m j}}{4^{m j}+\kappa}$ and $\kappa \geq 1$, the functions $\left(\sigma_{j(Q)}\right)^{-1}\left(1+2^{m j(Q)}\right) \partial^{\beta} \theta_{Q}$ form a family of "vaguelettes" (which is in the wavelet langage the same as an almost orthogonal family) uniformly with respect to $\kappa$. Hence, $\Psi_{Q} \mapsto\left(\sigma_{j(Q)}\right)^{-1}\left(1+2^{m j(Q)}\right) \partial^{\beta} \theta_{Q}$ is a Calderón-Zygmund operator and

$$
\begin{equation*}
\left\|\sum_{Q} c_{Q}\left(\sigma_{j(Q)}\right)^{-1}\left(1+2^{m j(Q)}\right) \partial^{\beta} \theta_{Q}\right\|_{2}^{2} \leq C \sum_{Q}\left|c_{Q}\right|^{2} \tag{5.1}
\end{equation*}
$$

where $C$ depends only on $n, m, N$. See [17].
Since the $\Psi_{Q}$ 's form an unconditional basis of any $H^{s}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right), s \in \mathbb{R}$, one can define operators by linearity (and continuity once a priori bounds are obtained). Define $P_{\kappa}$ by

$$
P_{\kappa}\left(\Psi_{Q}\right)=\theta_{Q} .
$$

We have

$$
(L+\kappa) P_{\kappa}\left(\Psi_{Q}\right)=\Psi_{Q}-R_{\kappa}\left(\Psi_{Q}\right)
$$

where

$$
R_{\kappa}\left(\Psi_{Q}\right)=(-1)^{\alpha} \partial^{\alpha}\left[\left(b_{\alpha \beta}-m_{Q} b_{\alpha \beta}\right) \partial^{\beta} \theta_{Q}\right] .
$$

The desired invertibility for $L+\kappa$ (or the parametrix property of $P_{\kappa}$ ) follows from

$$
\left\|R_{\kappa}\right\|<1
$$

where the norm is the operator norm on $H^{-m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$. Introducing

$$
T_{\alpha}\left(\Psi_{Q}\right)=\left(b_{\alpha \beta}-m_{Q} b_{\alpha \beta}\right)\left(1+2^{m j(Q)}\right) \partial^{\beta} \theta_{Q}
$$

and $S\left(\Psi_{Q}\right)=2^{m j(Q)} \Psi_{Q}$, we have

$$
R_{\kappa}=(-1)^{\alpha} \partial^{\alpha} T_{\alpha}(1+S)^{-1} .
$$

Since $1+S$ is an isomorphism from $L^{2}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$ onto $H^{-m}\left(\mathbb{R}^{n}, \mathbb{C}^{N}\right)$, it is enough to show that the $L^{2}$ operator norm of $T_{\alpha}$ is small when $\kappa$ is large.

There are two ways of estimating the operator norm of such an operator. To take advantage of this, introduce a threshold $j$ to be chosen and split $T_{\alpha}=T_{\alpha} \pi_{j}+T_{\alpha} \pi_{j}^{\perp}$ where $\pi_{j}$ is the orthogonal projection from $L^{2}$ onto the closed space generated by the $\Psi_{Q}$ for $j(Q)<j$ and $\pi_{j}^{\perp}=I d-\pi_{j}$.

For $T_{\alpha} \pi_{j}$, use (5.1) and the boundedness of the $b_{\alpha \beta}$ through the ellipticity constant $\Lambda$ to obtain that

$$
\left\|T_{\alpha} \pi_{j}\right\| \leq C \Lambda \sup _{j(Q)<j} \sigma_{j(Q)} \leq C \Lambda \sigma_{j} .
$$

For the other part of $T_{\alpha}$, remark that, ignoring $\sigma_{j(Q)}$ as it is close to 1 , the operator $\Psi_{Q} \mapsto\left(1+2^{m j(Q)}\right) \partial^{\beta} \theta_{Q}$ for $j(Q) \geq j$ and $\Psi_{Q} \mapsto 0$ for $j(Q)<j$ is a Calderón-Zygmund operator (with uniform estimates with respect to $j$ and $\kappa$ ). Then, the operator norm of $T_{\alpha} \pi_{j}^{\perp}$ can be estimated using the commutator result of Coifman-Rochberg-Weiss between $B M O$ function and

Calderón-Zygmund operators expressed in the wavelet langage and this gives us

$$
\left\|T_{\alpha} \pi_{j}^{\perp}\right\| \leq C \sup _{\beta} N_{j}\left(b_{\alpha \beta}\right) .
$$

Further details can be found in [1].
Hence,

$$
\left\|T_{\alpha}\right\| \leq C\left(\frac{4^{m j}}{4^{m j}+\kappa}+\sup _{\beta} N_{j}\left(b_{\alpha \beta}\right)\right) .
$$

It remains to choose $j$ so that $N_{j}\left(b_{\alpha \beta}\right)$ is small enough, then choose $\kappa$ large enough.

## References

[1] J.M. Angeletti, S. Mazet, and P. Tchamitchian. Analysis of second order elliptic operators without boundary conditions with VMO or Hölderian coefficients. In Dahmen, Kurdila, Oswald, editor, Multiscale wavelets methods for partial differential equations, volume 6 of Wavelets and its Applications. Acad. Press, 1997.
[2] P. Auscher, S. Hofmann, M. Lacey, A. M ${ }^{c}$ Intosh and Ph. Tchamitchian. The solution of the Kato square root problem for second order elliptic operators on $\mathbb{R}^{n}$, preprint, 2001.
[3] P. Auscher, S. Hofmann, J. Lewis and Ph. Tchamitchian. Extrapolation of Carleson measures and the analyticity of Kato's square root operator. Acta Math., to appear.
[4] P. Auscher, A. Mc Intosh, and A. Nahmod. The square root problem of Kato in one dimension, and first order systems. Indiana Univ. Math. J., 46:659-695, 1997.
[5] P. Auscher and M. Qafsaoui. Equivalence between regularity theorems and heat kernel estimates for higher order operators and systems under divergence form, J. Funct. Anal. 177 (2000), 310-364.
[6] P. Auscher and Ph. Tchamitchian. Square root problem for divergence operators and related topics, volume 249 of Astérisque. Soc. Math. France, 1998.
[7] M. Christ and J.-L. Journé. Polynomial growth estimates for multilinear singular integral operators. Acta Math., 159:51-80, 1987.
[8] R. Coifman, A. M ${ }^{c}$ Intosh, and Y. Meyer. L'intégrale de Cauchy définit un opérateur borné sur $L^{2}(\mathbb{R})$ pour les courbes lipschitziennes. Ann. Math., 116:361-387, 1982.
[9] E.B. Davies. Uniformly elliptic operators with measurable coefficients. J. Funct. Anal., 132:141-169, 1995.
[10] A. Friedman. Partial differential equations., R.E.Krieger publishing company, Florida, 1983.
[11] M. Giaquinta and J. Souček Caccioppoli's inequality and the Legendre-Hadamard condition Math. Ann. 270 (1985) 105-107.
[12] S. Hofmann, M. Lacey and A. Mc Intosh. The solution of the Kato problem for divergence form elliptic operators with Gaussian heat kernel bounds. Preprint, 2000.
[13] T. Kato. Fractional powers of dissipative operators, J. Math. Soc. Japan 13 (1961), p. 246-274.
[14] A. M'Intosh, On the Comparability of $A^{1 / 2}$ and $A^{* 1 / 2}$, Proc. Amer. Math. Soc. 32 (1972), 430-434.
[15] A. M ${ }^{c}$ Intosh and Y. Meyer. Algèbres d'opérateurs définis par des intégrales singulières. C. R. Acad. Sci. Paris, 301, Série 1:395-397, 1985.
[16] A. $\mathrm{M}^{c}$ Intosh and A. Yagi. Operators of type $\omega$ without a bounded $H^{\infty}$ functional calculus. In Miniconference on Operators in Analysis, volume 24 of Center for Math. and Appl., pages 159-174, Canberra, 1989. Australian National Univ.
[17] Y. Meyer. Ondelettes et opérateurs, I $\mathcal{F}$ II, Hermann, 1990.
[18] K. Zhang. A counterexample in the theory of coerciveness for elliptic systems, J. Partial Differential Equations 2 (1989), 79-82.
[19] K. Zhang. On the coercivity of elliptic systems in two dimensional spaces, Bull. Austral. Math. Soc. 54 (1996), 423-430.


[^0]:    *supported by NSF

