Rev. Mat. Iberoam. 40 (2024), no. 4, 1249–1292
DOI 10.4171/RMI/1482

Sums of squares III: Hypoellipticity in the infinitely degenerate regime

Lyudmila Korobenko and Eric Sawyer

Abstract. This is the third paper in a series of three dealing with sums of squares and hypoellipticity in the infinitely degenerate regime. We establish a $C^{2,\delta}$ generalization of M. Christ's smooth sum of squares theorem, and then use a bootstrap argument with the sum of squares decomposition for matrix functions, obtained in our second paper of this series, to prove a hypoellipticity theorem that generalizes some cases of the results of Christ, Hoshiro, Koike, Kusuoka and Stroock and Morimoto for sums of squares, and of Fedĭi and Kohn for degeneracies not necessarily a sum of squares.

1. Introduction

The regularity theory of second order *subelliptic* linear equations with smooth coefficients is well established, see, e.g., [10, 13]. In [13], Hörmander obtained hypoellipticity of sums of squares of smooth vector fields plus a lower order term, whose Lie algebra spans at every point. In [10], Fefferman and Phong considered general nonnegative semidefinite smooth self-adjoint linear operators, and characterized subellipticity in terms of a containment condition involving Euclidean balls and "subunit" balls related to the geometry of the nonnegative semidefinite form associated to the operator. Of course subelliptic operators L with smooth coefficients are hypoelliptic, namely, every distribution solution u of $Lu = \phi$ is smooth when ϕ is smooth. In the converse direction, Hörmander also showed in [13] that a sum of squares of smooth vector fields in $\mathbb{R}^n$, with constant rank Lie algebras, is hypoelliptic if and only if the rank is n. See Trèves [29] for a treatment of further results on characterizing hypoellipticity in certain special cases.

However, the question of hypoellipticity in general remains largely a mystery. A possible form for a characterization involving the effective symbol $\widetilde{\sigma}(x,\xi)$ (when it exists) is given by Christ in [6], motivated by his main hypoellipticity theorem for sums of squares in the infinitely degenerate regime, see Main Theorem 2.3 in [7]. We will generalize this latter theorem of Christ to hold for $C^{2,\delta}$ symbols, which will play a major role in Theorems 2.2 and 2.5 below on hypoellipticity in the infinitely degenerate regime. The difference between these two theorems is that in Theorem 2.5, we assume a partial sum of squares decomposition for the operator L, while in Theorem 2.2, we assume

Mathematics Subject Classification 2020: 35B65 (primary); 35J70, 35H10 (secondary).
Keywords: infinite degeneracy, elliptic operators, hypoellipticity, sums of squares, vector fields.

differential inequalities on the coefficients of L that force this partial sum of squares decomposition. These theorems will be compared to each other and to existing results in the literature below.

Therefore, a basic obstacle to understanding hypoellipticity in general arises when ellipticity degenerates to infinite order in some directions, and we briefly review what is known in this infinite regime here. The theory has only had its surface scratched so far, as evidenced by the results of Fedii [8], Kusuoka and Strook [21], Kohn [15], Koike [16], Korobenko and Rios [17], Korobenko, Rios, Sawyer and Shen [18], Rios, Sawyer and Wheeden [24], Morimoto [22], Akhunov, Korobenko and Rios [1], and the aforementioned paper of Christ [7], to name just a few. In the *rough* infinitely differentiable regime, Rios, Sawyer and Wheeden [24] had earlier obtained in addition results analogous to those above, but where L is "rough" hypoelliptic if every *continuous weak* solution u of $Lu = \phi$ is continuous when ϕ is bounded – continuity was removed in some cases in [18].

In [8], Fedii proved that the two-dimensional operator $\partial/\partial x^2 + f(x)^2\,\partial/\partial y^2$ is hypoelliptic merely under the assumption that f is smooth and positive away from $x = 0$. In [21], Kusuoka and Strook showed using probabilistic methods that under the same conditions on $f(x)$, the three-dimensional analogue $\partial^2/\partial x^2 + \partial^2/\partial y^2 + f(x)^2\,\partial^2/\partial z^2$ of Fedii's operator is hypoelliptic *if and only if*

$$\lim_{x\to 0} x \ln f(x) = 0.$$

Morimoto [22] and Koike [16] introduced the use of nonprobabilistic methods, and further refinements of this approach were obtained in Christ [7], using a general theorem on hypoellipticity of sums of squares of smooth vector fields in the infinite regime, i.e., where the Lie algebra does *not* span at all points. In particular, for the operator $L_3 = \partial^2/\partial x^2 + a^2(x)\,\partial^2/\partial y^2 + b^2(x)\,\partial^2/\partial z^2$ in $\mathbb{R}^3$, Christ proved that if $a, b \in C^\infty$ are even, elliptic, nondecreasing on $[0, \infty)$, and $a(x) \geq b(x)$ for all x, and if in addition $\limsup_{x\to 0}|x \ln a(x)| \neq 0$, and the coefficient b satisfies

$$\lim_{x\to 0} b(x) x\, |\ln a(x)| = 0,$$

then L_3 is hypoelliptic. Moreover, he showed that if some partial derivative of b is nonzero at $x = 0$, then L_3 is hypoelliptic *if and only if* the above condition holds.

On the other hand, the novelty in Kohn [15], which was generalized in [17], and pursued as well in [1], was the absence of any assumption regarding sums of squares of vector fields. This is relevant since it is an open problem whether or not there are smooth nonnegative functions λ on the real line vanishing only at the origin, and to infinite order there, such that they *cannot* be written as a finite sum $\lambda = \sum_{n=1}^{N} f_n^2$ of squares of smooth functions f_n. The existence of such examples are attributed to Paul Cohen in both [5] and [2], but apparently no example has ever appeared in the literature, and the existence of such an example is an open problem, see[1] Remark 5.1 in [23]. This extends moreover to matrices, since if a matrix is a sum of squares (equivalently a sum of positive rank one matrices), then each of its diagonal elements is as well. On the other hand, Kohn makes the additional assumption that $\lambda(x)$ vanishes only at the origin in $\mathbb{R}^m$, something not necessarily

[1] See also https://mathoverflow.net/a/106072, visited on June 12, 2024.

assumed in the other aforementioned works. More importantly, Kohn's theorem applies only to operators of Grushin type $L(x, D) + \lambda(x)L(y, D)$, where the degeneracy $\lambda(x)$ factors out of the operator $\lambda(x)L(y, D)$, a restriction that this paper will in part remove.

Missing then is a treatment of more general smooth operators $L = \nabla A(x)\nabla +$ lower order terms, whose matrix $A(x)$ is *comparable* to an operator in diagonal form of the types considered above – see Definition 1.1 below. Our purpose in this paper is to address this more general case in the following setting of real-valued differential operators. Suppose $1 \le m < p \le n$. Let $L = \nabla A(x)\nabla$, where $A(x) \sim D_{\boldsymbol{\lambda}}(\tilde{x})$, with $\tilde{x} = (x_1, \dots, x_m)$, $x = (x_1, \dots, x_n)$, and where $D_{\boldsymbol{\lambda}}(\tilde{x})$ has C^2 nonnegative diagonal entries $\lambda_1(\tilde{x}), \dots, \lambda_n(\tilde{x})$ depending only on $\tilde{x}$ and positive away from the origin in $\mathbb{R}^m$:

$$A(x) \sim D_{\boldsymbol{\lambda}}(\tilde{x}) = \begin{bmatrix} \mathbb{I}_m & \mathbf{0}_{m\times(p-m-1)} & \mathbf{0}_{m\times(n-p+1)} \\ \mathbf{0}_{(p-m-1)\times m} & D_{\{\lambda_{m+1}(\tilde{x}),\dots,\lambda_{p-1}(\tilde{x})\}} & \mathbf{0}_{(p-m-1)\times(n-p+1)} \\ \mathbf{0}_{(n-p+1)\times m} & \mathbf{0}_{(n-p+1)\times(p-m-1)} & \lambda_p(\tilde{x})\mathbb{I}_{n-p+1} \end{bmatrix}.$$

We will refer to a diagonal matrix having this form for any $m < p \le n$ as a *Grushin matrix function of type* m. Note that the comparability $A(x) \sim D_{\boldsymbol{\lambda}}(\tilde{x})$ implies that $a_{k,k}(x) \approx \lambda_k(\tilde{x})$ for all the diagonal entries, so that $\lambda_k(\tilde{x}) \approx a_{k,k}(\tilde{x}, 0)$ may be assumed smooth without loss of generality. Moreover, $A(x) \sim A_{\text{diag}}(\tilde{x}, 0)$ (see [20], after Definition 10).

All of our theorems will apply to operators L comparable to a Grushin matrix function $A(x)$ of type m as above, that is also positive definite for $\tilde{x} \neq 0$. Moreover, we will require in addition that the intermediate diagonal entries $\{a_{k,k}(\tilde{x})\}_{k=m+1}^{p-1}$ (there will not be any such entries in the case $p = m + 1$) are smooth and *strongly* $C^{4,2\delta}$ (see [19]) for some $\delta > 0$ (we show in [19] that such functions can be written as a sum of squares of $C^{2,\delta}$ functions and, moreover, give a sharp ω-monotonicity criterion for strongly $C^{4,2\delta}$), and that the off diagonal entries of $A(x)$ satisfy certain strongly subordinate inequalities (which are shown to have a weak sharpness property in a certain case, see Theorem 43 in [20]). We emphasize that no additional assumptions are made on the last $n - p + 1$ entries of $D(\tilde{x})$, which are all equal to $\lambda_p(\tilde{x})$.

Our approach is broadly divided into four separate steps, the first and second of which are the subject of the first two papers in this series.

(1) First, a proof that a $C^{3,1}$ function can be written as a finite sum of squares of $C^{1,1}$ functions first appeared in Guan [12], who attributed the result to Fefferman. In [19], we adapted treatments of this result from Tataru [27] and Bony [3] to establish conditions under which a $C^{4,2\delta}$ nonnegative function can be written as a finite sum of squares of C^{2,δ_*} functions for some $\delta, \delta_* > 0$. The methods of Tataru and Bony were in turn modelled on a localized splitting of a nonnegative symbol a, due to Fefferman and Phong [9], who used it to establish a strong form of Gårding's inequality, and is the main idea behind the result of Fefferman appearing in [12]. That splitting used the implicit function theorem to write a nonnegative symbol a as a sum of squares plus a symbol depending on fewer variables, so that induction could be applied. This same scheme was used in [19] to write certain $C^{4,2\delta}$ functions as a sum of squares of C^{2,δ_*} functions, but taking care to arrange assumptions so that the implicit function theorem applied.

(2) Second, in [20], we showed that under analogous conditions on the diagonal entries of a matrix-valued function M, and strong subordinate-type inequalities on the off diagonal entries, M can then be written as a finite sum of squares of $C^{2,\delta}$ vector fields for some $\delta > 0$.

(3) Third, we here extend a theorem of Christ on hypoellipticity of sums of smooth squares of vector fields to the setting of $C^{2,\delta}$ vector fields, with the appropriate notion of gain in a range of Sobolev spaces.

(4) Fourth, we here adapt arguments of Christ together with the above steps to obtain hypoellipticity of linear operators L of the form

$$L = \nabla^{\mathrm{tr}} A(x) \nabla + D(x), \tag{1.1}$$

where the matrix A and the scalar D are smooth functions of $x \in \mathbb{R}^n$, and with $\tilde{x} = (x_1, \ldots, x_m)$, we have

$$A(x) \sim \begin{bmatrix} \mathbb{I}_m & 0 \\ 0 & D_{\boldsymbol{\lambda}}(\tilde{x}) \end{bmatrix},$$

where $\mathbb{I}_m$ is the $m \times m$ identity matrix, and $D_{\_}(\tilde{x})$ is the $(n-m) \times (n-m)$ diagonal matrix with the components of $\boldsymbol{\lambda}(\tilde{x}) = (\lambda_{m+1}(\tilde{x}), \ldots, \lambda_n(\tilde{x}))$ along the diagonal. The component functions $\lambda_\ell(\tilde{x})$ satisfy certain natural conditions described explicitly below.

In the next section, we state our main results on hypoellipticity. Then, in the following section, we use a result on calculus of rough symbols from the 1980's, see [25], to derive a rough version of Christ's hypoellipticity theorem for sums of smooth vector fields in the infinitely degenerate regime, where symbol splitting is inadequate. Finally, in the last sections, we use a bootstrap argument that exploits the $C^{2,\delta}$ regularity of the vector fields, to bring all of these results to bear on proving hypoellipticity for linear partial differential operators L of the form (1.1).

But first we recall the main results from the second paper in this series [20] on sums of squares of matrix functions that we will use here.

Definition 1.1. Let A and B be real symmetric positive semidefinite $n \times n$ matrices. We define $A \preccurlyeq B$ if $B - A$ is positive semidefinite. Let $\beta < \alpha$ be positive constants. A real symmetric positive semidefinite $n \times n$ matrix A is said to be (β, α)-*comparable* to a symmetric $n \times n$ matrix B, written $A \sim_{\beta,\alpha} B$, if $\beta B \preccurlyeq A \preccurlyeq \alpha\, B$, i.e.,

$$\beta\, \xi^{\mathrm{tr}} B \xi \leq \xi^{\mathrm{tr}} A \xi \leq \alpha\, \xi^{\mathrm{tr}} B \xi \quad \text{for all } \xi \in \mathbb{R}^n.$$

We say A is *comparable* to B, written $A \sim B$, if $A \sim_{\beta,\alpha} B$ for some $0 < \beta < \alpha < \infty$.

Note that if A is *comparable* to B, then both A and B are positive semidefinite. Indeed, both $0 \leq (\alpha - \beta) \xi^{\mathrm{tr}} B \xi$ and $0 \leq (1/\beta - 1/\alpha) \xi^{\mathrm{tr}} A \xi$ hold for all $\xi \in \mathbb{R}^n$.

Definition 1.2. A matrix function $\mathbf{A}(x)$ is *subordinate* if $|\frac{\partial \mathbf{A}}{\partial x_k}(x) \cdot \xi|^2 \leq C \xi^{\mathrm{tr}} A(x) \xi$ for all $\xi \in \mathbb{R}^n$, equivalently, $\frac{\partial \mathbf{A}}{\partial x_k}(x)^{\mathrm{tr}} \frac{\partial \mathbf{A}}{\partial x_k}(x) \preccurlyeq C \mathbf{A}(x)$.

Finally, recall the following seminorm from [3]:

$$[h]_{\alpha,\delta}(x) \equiv \limsup_{y,z \to x} \frac{|D^\alpha h(y) - D^\alpha h(z)|}{|y - z|^\delta}.$$

Here the sum of squares decomposition has a quasiconformal block of order $(n-p+1) \times (n-p+1)$, where $1 < p \leq n$. We say that a symmetric matrix function $\mathbf{Q}_p(x)$ is quasiconformal if the eigenvalues $\lambda_i(x)$ of $\mathbf{Q}_p(x)$ are nonnegative and comparable.

Theorem 1.3. *Let* $1 < p \leq n$, $1/4 \leq \varepsilon < 1$, $0 < \delta \leq \delta' < 1/2$, $M \geq 1$. *Define* δ_{n-1} *recursively by* $\delta_0 = \delta$ *and*

$$\frac{\delta_{k+1}}{2+\delta_{k+1}} = \eta \frac{\delta_k}{1+\delta_k}, \quad 0 \leq k \leq n-2,$$

where $\eta = \frac{\delta}{2+\delta}$, *and finally let* δ'' *satisfy*

$$\max\left\{\delta, \frac{\delta_{n-1}}{1-\delta_{n-1}}\right\} \leq \delta'' < \frac{1}{2}.$$

Suppose that $\mathbf{A}(x)$ *is a* $C^{4,2\delta}$ *symmetric* $n \times n$ *matrix function of a variable* $x \in \mathbb{R}^M$, *which is comparable to a diagonal matrix function* $\mathbf{D}(x)$, *hence comparable to its associated diagonal matrix function* $\mathbf{A}_{\mathrm{diag}}(x)$.

Moreover, assume $a_{p,p}(x) \approx a_{p+1,p+1}(x) \approx \cdots \approx a_{n,n}(x)$ *and that the diagonal entries* $a_{1,1}(x), \ldots, a_{p-1,p-1}(x)$ *satisfy the following differential estimates up to fourth order*:

$$(1.2) \qquad \begin{cases} |D^\mu a_{k,k}(x)| \lesssim a_{k,k}(x)^{[1-|\mu|\varepsilon]_+ + \delta'}, & 1 \leq |\mu| \leq 4 \text{ and } 1 \leq k \leq p-1, \\ [a_{k,k}]_{\mu,2\delta}(x) \lesssim 1, & |\mu| = 4 \text{ and } 1 \leq k \leq p-1. \end{cases}$$

Furthermore, assume the off diagonal entries $a_{k,j}(x)$ *satisfy the following differential estimates up to fourth order*:

$$(1.3) \qquad \begin{cases} |D^\mu a_{k,j}| \lesssim \big(\min_{1\leq s\leq j} a_{s,s}\big)^{[1/2+(2-|\mu|)\varepsilon]_+ + \delta''}, & 0 \leq |\mu| \leq 4,\ 1 \leq k < j \leq p-1, \\ [a_{k,j}]_{\mu,2\delta} \lesssim 1, & |\mu| = 4,\ 1 \leq k < j \leq p-1, \\ |D^\mu a_{k,j}| \lesssim \big(\min_{1\leq s\leq k} a_{s,s}\big)^{[1/2+(2-|\mu|)\varepsilon]_+ + \delta''}, & 0 \leq |\mu| \leq 4,\ 1 \leq k \leq p-1 < j \leq n, \\ [a_{k,j}]_{\mu,2\delta} \lesssim 1, & |\mu| = 4,\ 1 \leq k \leq p-1 < j \leq n. \end{cases}$$

Then there is a positive integer $I \in \mathbb{N}$ *such that the matrix function* $\mathbf{A}$ *can be written as a finite sum of squares of* $C^{2,\delta_{n-1}}$ *vectors* $X_{k,j}$, *plus a matrix function* $\mathbf{A}_p$,

$$\mathbf{A}(x) = \sum_{k=1}^{p-1}\sum_{i=1}^{I} X_{k,i}(x) X_{k,i}(x)^{\mathrm{tr}} + \mathbf{A}_p(x), \quad x \in \mathbb{R}^M,$$

where the vectors $X_{k,i}(x)$, $1 \leq k \leq p-1$, $1 \leq i \leq I$, *are* $C^{2,\delta_{n-1}}(\mathbb{R}^M)$,

$$\mathbf{A}_p(x) = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_p(x) \end{bmatrix},$$

and $\mathbf{Q}_p(x) \in C^{4,2\delta}(\mathbb{R}^M)$ *is* quasiconformal. *Moreover, for* $1 \leq k \leq p-1$, *and for* $Z_k \equiv \sum_{i=1}^{I} X_{k,i} X_{k,i}^{\mathrm{tr}} \in C^{4,2\delta}(\mathbb{R}^M)$,

$$(1.4) \qquad \begin{cases} c a_{k,k}\mathbf{e}_k \otimes \mathbf{e}_k \prec Z_k Z_k^{\mathrm{tr}} + \displaystyle\sum_{m=k+1}^{n} a_{m,m}\,\mathbf{e}_m \otimes \mathbf{e}_m \prec C \sum_{m=k}^{n} a_{m,m}\,\mathbf{e}_m \otimes \mathbf{e}_m, \\ \mathbf{Q}_p(x) \sim a_{p,p}(x)\mathbb{I}_{n-p+1}. \end{cases}$$

Finally, if in addition $A(x)$ *is subordinate, then* $\mathbf{Q}_p(x)$ *is also subordinate.*

Remark 1.4. If $a_{k,k}(x) \approx 1$ for $1 \leq k \leq m < p$ in Theorem 1.3, then conditions (1.2) and (1.3) are vacuous for $1 \leq k \leq m$, and moreover the proof shows that the vectors $X_{k,i}$ are actually in $C^{4,2\delta}(\mathbb{R}^M)$ for $1 \leq k \leq m$, $1 \leq i \leq I$.

These remarks yield the following corollary, in which conditions (1.2) and (1.3) play no role.

Corollary 1.5. *Suppose* $\mathbf{A}(x)$ *is a* $C^{4,\delta}(\mathbb{R}^M)$ *symmetric* $n \times n$ *matrix function that is comparable to a diagonal matrix function. In addition, suppose that* $a_{k,k}(x) \approx 1$ *for* $1 \leq k \leq p-1$ *and* $a_{k,k}(x) \approx a_{p,p}(x)$ *for* $p \leq k \leq n$. *Then*

$$\mathbf{A}(x) = \sum_{k=1}^{p-1} X_k(x) X_k(x)^{\mathrm{tr}} + \mathbf{Q}_p(x), \quad x \in \mathbb{R}^M,$$

where $X_k, \mathbf{Q}_p \in C^{4,\delta}(\mathbb{R}^M)$ *and* (1.4) *holds for* $1 \leq k \leq p-1$.

Remark 1.6. If the diagonal entry $a_{k,k}(x)$ is smooth and ω_s-monotone on $\mathbb{R}^n$ for some $s > 1-\varepsilon$, then the diagonal differential estimates (1.2) above hold for $a_{k,k}(x)$ since $|D^\mu a_{k,k}(x)| \leq C_{s,s'} a_{k,k}(x)^{s'}$ for any $s' < s$ (see Theorem 18 in [20]).

Remark 1.7. If in Theorem 1.3, we drop the hypothesis (1.2) that the diagonal entries satisfy the differential estimates, and even slightly weaken the off diagonal hypotheses (1.3), then using the Fefferman–Phong theorem for sums of squares of scalar functions, the proof of Theorem 1.3 shows that the operator $L = \nabla^{\mathrm{tr}} \mathbf{A} \nabla$ can be written as $L = \sum_{j=1}^N X_j^{\mathrm{tr}} X_j$, where the vector fields X_j are $C^{1,1}$ for $j = 1, 2, \ldots, N$. However, unlike the situation for scalar functions, the example in Theorem 38 of [20] shows that we *cannot* dispense entirely with the off diagonal hypotheses (1.3). Moreover, the space $C^{1,1}$ seems not to be sufficient for gaining a positive degree δ of smoothness for solutions to a second order operator, and so this result will neither be used nor proved here.

In this paper, we will apply the sums of squares representations for matrix functions obtained in [20] to a rough generalization of a theorem of Christ, that then leads to our main hypoellipticity theorem via a bootstrap argument.

2. Statement of main hypoellipticity theorems

We begin with the following general hypoellipticity theorem in the infinitely degenerate regime as in Step (4) of the introduction. We emphasize that we make no assumptions regarding the order of vanishing of the matrix function $A(x)$ at the origin. Since we only consider degeneracies at the origin, it is useful to make the following definition.

Definition 2.1. We say that a $q \times q$ matrix function $f : \mathbb{R}^n \to \mathbb{R}^{q^2}$ on $\mathbb{R}^n$ is *elliptical* if $f(x)$ is positive definite for $x \neq 0$. A scalar function f corresponds to the case $q = 1$.

At the end of this section, we will discuss the relationships between the following theorem and earlier work on hypoellipticity. We emphasize again that our operators L are not assumed to have diagonal or even block diagonal matrix $A(x)$.

Theorem 2.2. *Suppose $1 \le m < p \le n$. Let L be a second order real self-adjoint divergence form partial differential operator in $\mathbb{R}^n$ given by*

$$L = \nabla^{\mathrm{tr}} A(x) \nabla + E(x), \tag{2.1}$$

where the matrix A and the scalar E are smooth real functions of $x \in \mathbb{R}^n$, and $A(x)$ is subordinate near the origin, i.e., $|\frac{\partial A}{\partial x_k}(x)\mathbf{u}|^2 \le C\mathbf{u}^{\mathrm{tr}} A(x)\mathbf{u}$ for $1 \le k \le n$, all x in some neighbourhood of the origin, and all unit vectors $\mathbf{u} \in \mathbb{R}^n$.

Suppose further that with $\tilde{x} = (x_1, \dots, x_m)$ we have the following Grushin *assumption*:

$$A(x) \sim \begin{bmatrix} \mathbb{I}_m & 0 \\ 0 & D_{\boldsymbol{\lambda}}(\tilde{x}) \end{bmatrix}, \tag{2.2}$$

where $\mathbb{I}_m$ is the $m \times m$ identity matrix, and $D_{\boldsymbol{\lambda}}(\tilde{x})$ is the $(n-m) \times (n-m)$ diagonal matrix with the components of $\boldsymbol{\lambda}(\tilde{x}) = (\lambda_{m+1}(\tilde{x}), \dots, \lambda_n(\tilde{x}))$ along the diagonal, i.e.,

$$D_\lambda(\xi) = \begin{bmatrix} \lambda_{m+1}(\tilde{x}) & 0 & \cdots & 0 \\ 0 & \lambda_{m+2}(\tilde{x}) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n(\tilde{x}) \end{bmatrix}.$$

(a) *Moreover, we suppose that the component functions λ_ℓ are elliptical in $\mathbb{R}^m$, and that $\lambda_p(\tilde{x}) \approx \lambda_{p+1}(\tilde{x}) \approx \cdots \approx \lambda_n(\tilde{x})$.*

(b) *We also suppose that there are positive numbers $0 < \delta, \delta', \delta'', \delta_{n-1} < 1/2$ and $1/4 \le \varepsilon < 1$ satisfying the conditions in Theorem 1.3, and such that for $k < j \le n$ and $1 \le k \le p-1$, the entries $a_{k,j}(x)$ of $A(x)$ satisfy the differential size inequalities*[2] *in (1.2) and (1.3) for all $x \in \mathbb{R}^n$.*

Then L is hypoelliptic if

$$\lim_{\tilde{x} \to 0} \mu\Big(|\tilde{x}|, \sqrt{\max\{\lambda_{m+1}, \dots, \lambda_p\}(\tilde{x})}\Big) \ln \min\{\lambda_{m+1}, \dots, \lambda_p\}(\tilde{x}) = 0, \tag{2.3}$$

where

$$\mu(t, g) \equiv \max\{g(z)(t - |z|) : 0 \le |z| \le t\}.$$

Moreover, condition (2.3) is necessary for hypoellipticity if in addition $A(x)$ is a diagonal matrix with monotone entries.

Remark 2.3. Note that when $m = 1$, it suffices to assume only *smoothness* of the diagonal entries $\lambda_\ell(\tilde{x})$ in place of (1.2), in view of Bony's sum of squares theorem, [3], Théorème 1.

Remark 2.4. The assumption that $A(x)$ is subordinate is redundant since this is already implied by the Grushin assumption (2.2) together with the off diagonal strong subordinaticity assumptions (1.3)). Indeed, it suffices to show that $|\nabla \mathbf{A}(x) \cdot \mathbf{e}_j|^2 \le C\mathbf{e}_j^{\mathrm{tr}} A(x)\mathbf{e}_j$ for all $1 \le j \le n$. However, from (2.2), we see that the diagonal entries a_{jj} are nonnegative and smooth, and so $|\nabla a_{jj}|^2 \lesssim a_{jj}$ by the inequality of Malgrange, while from (1.3) and the symmetry of $A(x)$ we see that $|\nabla a_{kj}|^2 \lesssim a_{jj}$ for all $1 \le k \ne j \le n$. Finally, $a_{jj} \lesssim \mathbf{e}_j^{\mathrm{tr}} A(x)\mathbf{e}_j$ and so $|\nabla \mathbf{A}(x) \cdot \mathbf{e}_j|^2 \lesssim \mathbf{e}_j^{\mathrm{tr}} A(x)\mathbf{e}_j$.

[2]The diagonal inequalities become more demanding the smaller ε is, while the off diagonal inequalities become less demanding.

The previous remark shows that Theorem 2.2 does not yield any hypoelliptic infinitely degenerate operators $L = \nabla^{\mathrm{tr}} A(x) \nabla$ in which the matrix $A(x)$ is *not* subordinate. Here is a variation, without any special hypotheses on the diagonal entries, that *does* yield hypoellipticity without subordinaticity, and that will be used to prove Theorem 2.2 in conjunction with the sum of squares decomposition in Theorem 1.3. However, the proof of this next result will require a generalization of Christ's sum of squares theorem to include $C^{2,\delta}$ vector fields.

Theorem 2.5. *Let L be a real second order divergence form partial differential operator in $\mathbb{R}^n$ satisfying (2.1). Let $1 \le m < p \le n+1$, and write*

$$x = (x_1, \dots, x_m, x_{m+1}, \dots, x_{p-1}, x_p, \dots, x_n) = (\tilde{x}, \check{x}, \hat{x}) \in \mathbb{R}^m \times \mathbb{R}^{p-m-1} \times \mathbb{R}^{n-p+1},$$

where the middle factor $\mathbb{R}^{p-m-1}$ vanishes if $p = m+1$, and the final factor vanishes if $p = n+1$.

Suppose that there exist $C^{2,\delta}$ vector fields $X_j(x) \in \mathrm{Op}(\mathcal{C}^{2,\delta} S^1_{1,0})$ for $1 \le j \le N$, and an $(n-p+1) \times (n-p+1)$ matrix function $\mathbf{Q}_p(x) \in C^{4,2\delta}$ that is elliptical, quasiconformal and subordinate, such that

$$L = \left(\sum_{j=1}^N X_j^{\mathrm{tr}} X_j + \hat{\nabla}^{\mathrm{tr}} \mathbf{Q}_p(x) \hat{\nabla} \right) + \sum_{j=1}^N A_j X_j + \sum_{j=1}^N X_j^{\mathrm{tr}} \tilde{A}_j + A_0,$$

where $\hat{\nabla} = (\partial_{x_p}, \dots, \partial_{x_n})$ and $A_j, \tilde{A}_j \in \mathrm{Op}(\mathcal{C}^{1,\delta} S^0_{1,0})$, $A_0 \in \mathcal{O}^{-\delta/2+\varepsilon}_{(-\delta/2,\delta/2)}$ for all $\varepsilon > 0$.

Suppose also that there are elliptical scalar functions $\lambda_{m+1}(\tilde{x}), \dots, \lambda_p(\tilde{x}) \in C^2(\mathbb{R}^n)$, with $0 \le \lambda_j \le 1$ for all j, such that $\mathbf{Q}_p(x) \sim \lambda_p(\tilde{x}) \mathbb{I}_{n-p+1}$ and such that the following inequalities hold for all Lipschitz functions v:

$$\text{(2.4)} \quad \begin{cases} \sum_{k=1}^m |\partial_{x_k} v|^2 + \sum_{k=m+1}^{p-1} \lambda_k(\tilde{x}) |\partial_{x_k} v|^2 \lesssim \sum_{j=1}^N |X_j v|^2 + \lambda_p(\tilde{x}) \sum_{k=p}^n |\partial_{x_k} v|^2, \\ \sum_{j=1}^N |X_j v|^2 \lesssim \sum_{k=1}^m |\partial_{x_k} v|^2 + \sum_{k=m+1}^{p-1} \lambda_k(\tilde{x}) |\partial_{x_k} v|^2 + \lambda_p(\tilde{x}) \sum_{k=p}^n |\partial_{x_k} v|^2. \end{cases}$$

Finally, set

$$\Lambda_{\mathrm{sum}}(\tilde{x}) \equiv \sum_{k=m+1}^p \lambda_k(\tilde{x}) \quad \text{and} \quad \Lambda_{\mathrm{product}}(\tilde{x}) \equiv \prod_{k=m+1}^p \lambda_k(\tilde{x}),$$

and define the Koike functional $\mu(t, g)$ for any function $g(\tilde{x})$ by

$$\mu(t, g) \equiv \max\{ g(\tilde{x})(t - |\tilde{x}|) : 0 \le |\tilde{x}| \le t \}.$$

Then the operator L is hypoelliptic if

$$\text{(2.5)} \quad \lim_{x \to 0} \mu\left(|\tilde{x}|, \sqrt{\Lambda_{\mathrm{sum}}}\right) \ln \Lambda_{\mathrm{product}}(\tilde{x}) = 0.$$

This is sharp in the sense that (2.5) holds if L is both hypoelliptic and diagonal with monotone entries.

Here is our rough version, in the setting of sums of squares of real vector fields, of Christ's hypoellipticity theorem as needed in Step (3) of the introduction. Note, in particular, that the vector fields X_j appearing below are only assumed to be $C^{2,\delta}$, while the sum of their squares $\sum_j X_j^{\mathrm{tr}} X_j$ is assumed to be smooth.

Theorem 2.6. *Suppose $1 \le p \le n$ and $N \ge 1$. Let $R \subset T^*V$, the cotangent bundle of an open set $V \subset \mathbb{R}^n$, be any ray, and assume that the operator L has the form*

$$L = \sum_{j=1}^{N} X_j^{\mathrm{tr}} X_j + \sum_{j=1}^{N} A_j X_j + \sum_{j=1}^{N} X_j^{\mathrm{tr}} \tilde{A}_j + R_1 + A_0 + \widehat{\nabla}^{\mathrm{tr}} \cdot \mathbf{Q}_p(x) \widehat{\nabla}, \tag{2.6}$$

where the vector fields X_j, $j = 1, 2, \ldots, N$, are $C^{2,\delta}(\mathbb{R}^n)$ differential operators, and $\mathbf{Q}_p(x)$ is a $C^{4,2\delta}(\mathbb{R}^m)$ $(n-p+1) \times (n-p+1)$ matrix that is subordinate and quasiconformal, and $\widehat{\nabla} = (\partial_{x_p}, \ldots, \partial_{x_n})$.

Assume further that $\mathbf{Q}_p = \mathbf{Q}_p(x) \approx a(x)\mathbb{I}_{n-p+1}$, with $a \in C^{4,2\delta}(\mathbb{R}^n)$ elliptical, $L \in \mathrm{Op}(S^2_{1,0})$, $X_j \in \mathrm{Op}(\mathscr{C}^{2,\delta} S^1_{1,0})$ and $A_j, \tilde{A}_j \in \mathrm{Op}(\mathscr{C}^{1,\delta} S^0_{1,0})$, $A_0 \in \mathcal{O}^{-\delta/2+\varepsilon}_{(-\delta/2,\delta/2)}$ for all $\varepsilon > 0$, in some conic neighbourhood V of R.

(a) *In addition, assume $R_1 = \sum_{k=1}^{n} S_k \Theta_k \circ \widehat{\nabla}$, where each $S_k \in C^{1,\delta}(\mathbb{R}^{m\times m})$ is subunit with respect to $\mathbf{Q}_p$, and $\Theta_k = (\Theta_{kp}, \ldots, \Theta_{kn})$ is a multiplier of order zero.*

(b) *Suppose there exists $w \in C^\infty$, satisfying $w(\xi) \to \infty$ as $|\xi| \to \infty$, such that*

$$\int_{\mathbb{R}^n} w^2(\xi)\, |\hat{u}(\xi)|^2 d\xi \le C \sum_j \|X_j u\|^2 + C \|\sqrt{a}\, \widehat{\nabla} u\|^2 + C\|u\|^2 \tag{2.7}$$

for all $u \in C^1_0(V)$.

(c) *Finally, suppose that for each small conic neighbourhood Γ of R there exist scalar valued symbols ψ, $p \in S^0_{1,0}$ such that ψ is everywhere nonnegative, ψ does not depend on ξ in Γ, $\psi \equiv 0$ in some smaller conic neighbourhood of R, $\psi \ge 1$ on $(T^*V) \setminus \Gamma$, $p \equiv 0$ in a conic neighbourhood of the closure of Γ, and such that for each $\delta > 0$, there exists $C_\delta < \infty$ such that for any relatively compact open subset $U \Subset V$ and for all $u \in C^2_0(U)$ and each index i,*

$$\begin{cases} \|\mathrm{Op}[\log\langle\xi\rangle\{\psi, \sigma(X_i)\}]u\|^2 \\ \qquad \le \delta \sum_j \|X_j u\|^2 + \delta \|\sqrt{a}\, \widehat{\nabla} u\|^2 + C_\delta \|u\|^2 + C_\delta \|\mathrm{Op}(p)u\|^2_{H^1}, \\ \|\sqrt{\mathbf{Q}_p}\, \mathrm{Op}[\log\langle\xi\rangle\{\psi, \widehat{\xi}\}]u\|^2 \\ \qquad \le \delta \sum_j \|X_j u\|^2 + \delta \|\sqrt{a}\, \widehat{\nabla} u\|^2 + C_\delta \|u\|^2 + C_\delta \|\mathrm{Op}(p)u\|^2_{H^1}, \end{cases} \tag{2.8}$$

where $\widehat{\xi} = (\xi_p, \ldots, \xi_n)$.

Then there exists $\gamma > 0$ such that for any $u \in L^2_{\mathrm{loc}}$, we have

$$Lu \in H^\gamma(R) \implies u \in H^\gamma(R).$$

Remark 2.7. Note that the term R_1 arises from the conjugation of $\widehat{\nabla}\cdot\mathbf{Q}_p(x)\widehat{\nabla}$ by $\Lambda_s=(1+|\xi|^2)^{s/2}$, needed in the bootstrap procedure. Indeed, denoting $q_{ij}=(\mathbf{Q}_p)_{ij}$, we have

$$\Lambda_s\widehat{\nabla}\cdot\mathbf{Q}_p(x)\widehat{\nabla}\Lambda_{-s}-\widehat{\nabla}\cdot\mathbf{Q}_p(x)\widehat{\nabla}=\sum_{i,j=p}^{n}[\Lambda_s,q_{ij}]\Lambda_{-s}\partial_{x_i}\partial_{x_j}.$$

Using rough pseudodifferential calculus, we have

$$\sigma([\Lambda_s,q_{ij}]\Lambda_{-s}\partial_{x_i})=-i\sum_{|\alpha|=1}D^\alpha q_{ij}\frac{\xi^\alpha\xi_i}{\langle\xi\rangle^2}=-i\sum_{k=1}^{n}\partial_{x_k}q_{ij}\frac{\xi_k\xi_i}{\langle\xi\rangle^2}\mod \mathcal{O}^{-\varepsilon}_{(-\delta,\delta)}.$$

Denoting

$$S_k=\partial_{x_k}\mathbf{Q}_p\quad\text{and}\quad(\theta_k(\xi))_i=-i\frac{\xi_k\xi_i}{\langle\xi\rangle^2},$$

we have that $R_1=\sum_{k=1}^{n}S_k\Theta_k\circ\widehat{\nabla}$ has the desired properties since $\mathbf{Q}_p$ is subordinate, and

$$\Lambda_s\widehat{\nabla}\cdot\mathbf{Q}_p(x)\widehat{\nabla}\Lambda_{-s}=\widehat{\nabla}\cdot\mathbf{Q}_p(x)\widehat{\nabla}+R_1\mod \mathcal{O}^{-\varepsilon}_{(-\delta,\delta)}.$$

Next, we outline the four steps taken in order to get to the point where we can apply Theorems 1.3, 2.5 and 2.6 to obtain our hypoellipticity Theorem 2.2.

2.1. Summary of the steps

Consider the operator $L=\nabla A(\hat{x})\nabla+E(x)$ with smooth coefficients.

(1) We first apply Theorem 1.3 to write $\nabla A(\hat{x})\nabla=\mathbf{X}^{\mathrm{tr}}\mathbf{X}$ plus a quasiconformal subordinate term $\widehat{\nabla}\cdot\mathbf{Q}_p(x)\widehat{\nabla}$, where the vector fields $\mathbf{X}$ belong to $\mathcal{C}^{2,\delta}S^1_{1,0}$ for some $\delta>0$, and $\mathbf{Q}_p\in C^{4,2\delta}$.

(2) We then use the smooth pseudodifferential calculus to write

$$\Lambda_sL\Lambda_{-s}=L+\widehat{\nabla}\cdot\mathbf{Q}_p(x)\widehat{\nabla}+V\mathbf{X}+\mathbf{X}^{\mathrm{tr}}U+A_0(x,\xi)+R_1,$$

where the pseudodifferential operators $V\mathbf{X},U\mathbf{X}\in\mathrm{Op}\mathcal{C}^{1,\delta}S^1_{1,0}$, and $R_1\in\mathrm{Op}\mathcal{C}^{1,\delta}S^1_{1,0}$ is subunit with respect to the quasiconformal term, and where $A_0\in\mathcal{C}^{0,\delta}S^0_{1,0}$.

(3) We next show that the operator $L=\nabla A(\hat{x})\nabla+E(x)$ is hypoelliptic if and only if for every $s\in\mathbb{R}$, there is $\gamma=\gamma([s])>0$, depending only on the integer part $[s]$ of s, such that

$$u\in H^0\text{ and }\Lambda_sL\Lambda_{-s}u\in H^\gamma\implies u\in H^\gamma.$$

(4) Finally, we apply Theorem 2.6 and Theorem 2.5 to obtain hypoellipticity of L.

Remark 2.8. Note that if we apply *symbol splitting* as in [28] to the vector fields $\mathbf{X}$ to obtain $\mathbf{X}=\mathbf{X}^\natural+\mathbf{X}^\flat$, where $\mathbf{X}^\natural\in\mathrm{Op}S^1_{1,\eta}$ and $\mathbf{X}^\flat\in\mathrm{Op}\mathcal{C}^{2,\delta}S^{1-\eta(2+\delta)}_{1,\eta}$, then the subunit property of the vector field $\mathbf{X}$ is *not* inherited by the smooth vector field $\mathbf{X}^\natural$. Indeed, the definition of $\mathbf{X}^\natural$ shows that it is obtained by applying a mollification of size $2^{-j\eta}$ to a Littlewood–Paley projection onto frequencies of size 2^j, and such mollifications are not comparable when applied to infinitely degenerate fields, even suitably away from the degeneracies.

We end this section on statements of the main hypoellipticity theorems by comparing Theorems 2.2 and 2.5 with previous work in [7, 15, 17, 18, 24]. Comparisons with other work, e.g., [1, 16, 21, 22], are similar and left to the reader.

2.2. Relationship with non-SOS methods

All of the smooth hypoellipticity results in Theorems 2.2 and 2.5 above, in Theorems 1 and 86 of [18], in Theorem 2.18 of [24], and in Theorem 1.3 of [17] assert smoothness of certain rough solutions to an equation of the form

$$Lu(x) = \nabla^{\mathrm{tr}} A(x) \nabla u(x) + E(x), \tag{2.9}$$

where

- the matrix $A(x)$ and the scalar $E(x)$ are smooth real functions of $x \in \mathbb{R}^n$;
- the matrix $A(x)$ is subordinate near the origin, i.e., $|\frac{\partial A}{\partial x_k}(x)\mathbf{u}|^2 \leq C\mathbf{u}^{\mathrm{tr}} A(x)\mathbf{u}$ for x near the origin and all unit vectors $\mathbf{u} \in \mathbb{R}^n$;
- and $A(x)$ is comparable in the sense of quadratic forms to the diagonal matrix

$$D_{\boldsymbol{\lambda}}(x) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_{m+1}(x) & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & \lambda_{m+2}(x) & \ddots & \vdots \\ 0 & 0 & 0 & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & \lambda_n(x) \end{bmatrix},$$

  where $\boldsymbol{\lambda} = (\lambda_{m+1}, \ldots, \lambda_n)$ is a vector of nonnegative smooth functions, and where the matrix $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cdots & 0 \\ 0 & 0 & 1 \end{pmatrix}$ in the top left corner is the $m \times m$ identity matrix $\mathbb{I}_m$ with $1 \leq m < n$. The theorems in [18], [24] and [17] also include first order derivatives, and apply to certain quasilinear equations of this form that are "close" to being linear, and to certain systems as well, but we will restrict our comparisons to the linear case as in (2.9).

One of the main differences in the type of smooth hypoellipticity obtained in these papers, lies in the notion of rough solution that each paper assumes:

(1) the rough solutions in Theorem 2.18 of [24] are assumed to be continuous weak solutions,
(2) the rough solutions in Theorems 1 and 85 of [18] are assumed to be just weak solutions,
(3) the rough solutions in Theorems 2.2 and 2.5 above, and in Theorem 1.3 of [17], are assumed to be merely distribution solutions.

Another difference lies in the fact that Theorems 2.2 and 2.5 above give a "near" characterization in terms of Koike's condition, while the other papers are restricted to broader sufficient conditions.

But perhaps the largest difference of all lies in the geometric assumptions these papers make on the matrices $A(x)$ and $D_{\boldsymbol{\lambda}}(x)$. In Theorem 2.18 of [24], $1 \leq m < n$, and the coefficients λ_k of the diagonal matrix $D_{\boldsymbol{\lambda}}(x)$ are assumed to satisfy:

(2.10) $\lambda_k(x)$ can vanish only on the kth coordinate axis, for $m+1 \leq k \leq n$,

while in Theorem 86 of [18], with $\tilde{x} = (x_1, x_2, \ldots, x_{n-1})$,

(2.11) $m = n-1$ and $\lambda_n(x) = \lambda_n(\tilde{x})$ can vanish only when $\tilde{x} = 0$.

Theorem 1.3 in [17] applies to block diagonal matrices

$$A(x) = \begin{bmatrix} \mathbb{I}_m & 0 & 0 & 0 \\ 0 & B_1(x^1)\lambda_1(\widehat{x}^1) & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & B_N(x^N)\lambda_N(\widehat{x}^N) \end{bmatrix}, \tag{2.12}$$

in which the infinite degeneracy $\lambda_k(\widehat{x}^k)$ is a function of variables not associated with the kth block, and has been factored out of the corresponding subelliptic block $B_k(x^k)$, which is a function of variables associated with the kth block. Finally, in Theorem 2.2 above, with $\tilde{x} = (x_1, x_2, \ldots, x_m)$,

(2.13) $1 \leq m < p \leq n$, and $\lambda_k(x) = \lambda_k(\tilde{x})$ can vanish only when

$$\tilde{x} = 0 \quad \text{for } m+1 \leq k \leq n \quad \text{and} \quad \lambda_p(\tilde{x}) = \lambda_{p+1}(\tilde{x}) = \cdots = \lambda_n(\tilde{x}),$$

and with additional assumptions on the coefficients $a_{jk}(x)$ of the matrix $A(x)$ as described in part (b) of Theorem 2.2, which imply that $A(x)$ can be written as a finite sum of squares of $C^{2,\delta}$ vector fields.

Now we note that if both (2.10) and (2.13) hold, then either $m = n-1$ or L is elliptic. Indeed, if $m \leq n-2$, then at the point $x = t\mathbf{e}_{n-1}$, we have $\lambda_n(t\mathbf{e}_{n-1}) > 0$ by (2.10), and $\lambda_n(t\mathbf{e}_{n-1}) = \lambda_n(0)$ by (2.13), and hence $\lambda_n(0) > 0$. Similarly, if $m+1 \leq k < n$, then at the point $x = t\mathbf{e}_{n-1}$, we have $\lambda_k(t\mathbf{e}_{k+1}) > 0$ by (2.10), and $\lambda_k(t\mathbf{e}_{k+1}) = \lambda_k(0)$ by (2.13), and hence $\lambda_k(0) > 0$.

It is also easy to check that if both (2.12) and (2.13) hold, then $A(x)$ is either elliptic or essentially[3] has the form

$$A(x) = \begin{bmatrix} \mathbb{I}_m & 0 \\ 0 & Q_p(\tilde{x}) \end{bmatrix}, \quad \text{where } Q_p(\tilde{x}) \text{ is conformal,} \tag{2.14}$$

by which we mean that there is $C > 0$ such that for each $\tilde{x}$, the eigenvalues $\lambda_i(\tilde{x})$ of $Q_p(\tilde{x})$ satisfy

$$\frac{1}{C} \leq \frac{\lambda_i(\tilde{x})}{\lambda_j(\tilde{x})} \leq C.$$

In order to compare (2.13) with the other forms, and ignoring those cases in which L is elliptic, we see that

- (2.11) holds when both (2.10) and (2.13) hold, and that
- (2.14) holds when both (2.12) and (2.13) hold.

[3] We are assuming for simplicity here that the final block in (2.12) is elliptic rather than subelliptic.

In these cases, all the papers imply that a rough solution u to the appropriate case of (2.9) is actually smooth, but where in Theorem 2.18 of [24], it is required that u be a continuous weak solution, in Theorem 85 of [18], it is required that u be a weak solution, and in Theorem 1.3 in [17] and Theorem 2.2 above, it is required only that u be a distribution solution. However, in [17], the infinite part of the degeneracy is assumed to be factored out, while this is not assumed in Theorem 2.2 above.

Here are some examples to illustrate these assertions and the scope of Theorems 2.2 and 2.5 in simple situations.

Example 2.9. Theorem 2.18 of [24] and Theorem 1.3 of [17] apply to (certain nonelliptic) diagonal operators of the form

$$L = \frac{\partial^2}{\partial x_1^2} + \lambda_2(x_1, x_3) \frac{\partial^2}{\partial x_2^2} + \lambda_3(x_1, x_2) \frac{\partial^2}{\partial x_3^2},$$

whereas Theorem 2.2 above does not. And Theorem 2.2 above applies to (certain nonelliptic) diagonal operators of the form

$$L = \frac{\partial^2}{\partial x_1^2} + \lambda_2(x_1) \frac{\partial^2}{\partial x_2^2} + \lambda_3(x_1) \frac{\partial^2}{\partial x_3^2},$$

whereas Theorem 2.18 of [24] and Theorem 1.3 of [17] do not.

Next we note that in [20], the notions of hypoellipticity and sums of squares are shown to be incomparable. For example, recall the following special case:

$$A(x) \equiv \begin{bmatrix} 1 & x/2 \\ x/2 & x^2 \end{bmatrix} = \begin{pmatrix} 1 \\ x/2 \end{pmatrix} \begin{bmatrix} 1 & x/2 \end{bmatrix} + \begin{bmatrix} 0 \\ \sqrt{3}x/2 \end{bmatrix} \begin{bmatrix} 0 & \sqrt{3}x/2 \end{bmatrix}$$

This is a finite sum of squares of smooth vector fields which is not subordinate, since $|A'(x)\mathbf{e}_2|^2 = 1/4 + 4x^2$ and $\mathbf{e}_2^{\mathrm{tr}} A(x)\mathbf{e}_2 = x^2$.

On the other hand, $A(x) \equiv \left[\begin{smallmatrix} 1 & 0 \\ 0 & f(x) \end{smallmatrix}\right]$ is subordinate and not a finite sum of squares of smooth vector fields if $f(x)$ is a nonnegative smooth function in $\mathbb{R}^n$ that is not a finite sum of squares of smooth functions. A quadratic polynomial example is given by the 4×4 matrix

$$A_\lambda(x) \equiv \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & x^2+\lambda y^2+2z^2 & -xy & -xz \\ 0 & -xy & y^2+\lambda z^2+2x^2 & -yz \\ 0 & -xz & -yz & z^2+\lambda x^2+2y^2 \end{bmatrix} \equiv \begin{bmatrix} 1 & \mathbf{0}^{\mathrm{tr}} \\ \mathbf{0} & \mathbf{L}_\lambda(x,y,z) \end{bmatrix}$$

that is subordinate when $\lambda > 0$, since $|A_\lambda'(x)\mathbf{u}|^2 \le C|x|^2$ and $\mathbf{u}A_\lambda(x)\mathbf{u} \ge c\lambda|x|^2$ for any unit vector $\mathbf{u}$ in $\mathbb{R}^4$, and by the generalization of a theorem of Cho in Theorem 38 of [20], neither A_λ nor $\mathbf{L}_\lambda$ is a finite sum of squares of even $C^{1,\delta}$ vector fields if $0 < \lambda < 2/81$.

Example 2.10. For any smooth elliptical function $f(x)$ on the real line, the matrix function

$$A(x) \equiv \begin{bmatrix} 1 & \gamma f(x) \\ \gamma f(x) & f(x)^2 \end{bmatrix}$$

is easily seen to be a sum of squares of smooth vector fields that is not subordinate, see, e.g., [20] for a simple proof. On the other hand, Theorem 2.5 shows that $L = \nabla^{\mathrm{tr}} A(x) \nabla$ is hypoelliptic since (2.4) follows immediately from

$$\sum_{k=1}^{m} |\partial_{x_k} v|^2 = \left|\frac{\partial}{\partial x} v\right|^2 + f(x)^2 \left|\frac{\partial}{\partial y} v\right|^2 = \sum_{j=1}^{N} |X_j v|^2,$$

while the Koike condition (2.5) is a consequence of $\Lambda_{\mathrm{sum}}(x) = f(x)^2 = \Lambda_{\mathrm{product}}(x)$. None of the aforementioned results in [17, 18, 24], apply to this simple operator L if f vanishes to infinite order at the origin. Theorem 2.2 only applies if, in addition, the diagonal and off-diagonal estimates (1.2) and (1.3) hold, which in this case means

$$|f'(x)| \lesssim f(x)^{3/2+\delta'''}, \qquad |f''(x)| \lesssim f(x)^{1+\delta'''},$$
$$|f^{(3)}(x)| \lesssim f(x)^{1/2+\delta'''}, \quad |f^{(4)}(x)| \lesssim f(x)^{\delta'''},$$

for some $\delta''' > 0$, since in this case $a_{12} = \gamma\sqrt{a_{22}}$, and the diagonal estimates follow from the off-diagonal ones, which are weakest when $\varepsilon = 1/4$.

Finally, we recall from Section 4.3.2 in [20] a somewhat more complicated example that demonstrates we can obtain hypoellipticity from Theorem 2.5 when L has a simple block form, but where the infinite degeneracy *cannot* be factored out as in [15, 17]. Moreover, neither Theorem 2.2, nor any previous results, such as those in [1, 7, 15, 17, 18, 24] apply. But to see this, we first need some preliminaries. As in [20], we let $\varphi, \psi \colon (0,1) \to (0,1)$ be strictly increasing elliptical flat smooth nearly monotone[4] functions on $(0,1)$, and define the matrix function

$$\mathbf{F}_{\varphi,\psi,h_\rho}(W,t) \equiv \varphi(t)\,\mathbf{L}(W) + (\psi(t) + \eta(t,r))\,\mathbb{I}_3,$$

for $(W,t) \in \Omega \equiv B_{\mathbb{R}^3}(0,1) \times (-1,1)$, where $\mathbf{L}(W) = \mathbf{L}_\lambda(W)$ for some $0 < \lambda < 2/81$, $\mathbb{I}_3$ is the 3×3 identity matrix, $r = |W| = \sqrt{x^2 + y^2 + z^2}$, and $\eta(t,r) = \varphi(r) h_\rho(|t|/r)$, where $h_\rho \colon [0,\infty) \to [0,1]$ is smooth and

$$h_\rho(u) = \begin{cases} 1 & \text{if } 0 \le u \le \rho, \\ 0 & \text{if } 1 \le u < \infty. \end{cases}$$

Then $\mathbf{F}_{\varphi,\psi,h_\rho}$ is a diagonally elliptical flat smooth 3×3 matrix function on $B_{\mathbb{R}^3}(0,1) \times (-1,1)$. By Lemma 40 in [20], $\mathbf{F}_{\varphi,\psi,h_\rho}(W,t)$ cannot be written as a finite sum of squares of $C^{1,\beta}$ vector fields if there is $0 < \beta < 1$ such that

$$\psi(t) = o(\varphi(t) t^4) \quad \text{and} \quad \lim_{t \to 0} \frac{\psi(t)}{\varphi(t)^{2/\beta} t^{4/\beta}} = 0. \tag{2.15}$$

Example 2.11. Now we can give the example of an operator $L = \nabla^{\mathrm{tr}} A \nabla$ in $\mathbb{R}^7$, with a smooth diagonally elliptic subordinate matrix A, that is hypoelliptic by Theorem 2.5, yet

[4] f nearly monotone means that f is ω_s-monotone for all $0 < s < 1$.

it is not a finite sum of squares of $C^{1,1}$ vector fields, and moreover cannot be written in the forms required in either [15] or [17]. Indeed, define the 7×7 matrix function A in block form by

$$A(x,y,z,t,u,v,w)\equiv\begin{bmatrix}\mathbb{I}_4 & \mathbf{0}_{4\times 3}\\ \mathbf{0}_{3\times 4} & \mathbf{F}_{\varphi,\psi}(x,y,z,t)\end{bmatrix},$$

where $\mathbb{I}_4$ is the 4×4 identity matrix, $\mathbf{0}_{m\times n}$ is the $m\times n$ zero matrix, and (2.15) holds. As was shown in [20], $\mathbf{F}_{\varphi,\psi,h_\rho}(x,y,z,t)$ is a smooth subordinate quasiconformal matrix function, and hence Theorem 2.5 applies to show that L is hypoelliptic (recall that Theorem 2.5 does *not* require any off diagonal estimates), yet A is not a finite sum of squares of $C^{1,1}$ vector fields.

We now claim that there is $\mathbf{F}_{\varphi,\psi,h_\rho}$ such that we *cannot* factor $\mathbf{F}_{\varphi,\psi,h_\rho}$ as $\lambda\mathbf{B}$, where λ is smooth (even $C^{4,2\delta}$) and $\mathbf{B}$ is smooth (not even $C^{3,1}$) and elliptic, thus showing that neither [15] nor [17] apply to the operator L.

To see this, let $W=(x,y,z)$, and then using $\mathbf{L}(W)\approx |W|^2\mathbb{I}_3$ as in Theorem 37 of [20], we conclude that

$$\begin{aligned}\operatorname{trace}\mathbf{F}_{\varphi,\psi,h_\rho}(W,t)&=\varphi(t)\operatorname{trace}\mathbf{L}(W)+3(\psi(t)+\eta(t,r))\\ &\approx |W|^2\varphi(t)+\psi(t)+\eta(t,r)\equiv f_{\varphi,\psi,h_\rho}(W,t).\end{aligned}$$

It is not hard to see that we can choose nearly monotone functions φ and ψ satisfying (2.15) such that the quantity $S^{\omega_s}_{\varphi,\psi}(\gamma_\rho+\delta)\equiv\sup_{0<t<1}\frac{\varphi(\gamma t)t^4}{\omega_s(\psi(t))}$ in Remark 5.5 (2) of [19] is finite for some $0<s<1$.

Indeed, with $0<\beta<1$ *in* (2.15) *fixed, and any* $0<s<\beta/3$, *we can take*

$$\varphi(t)=e^{-1/|t|},\quad \psi(t)=e^{-\frac{2}{3s}\frac{1}{|t|}},\quad \rho=\frac{3}{4},\quad \gamma_\rho\equiv\frac{1+\sqrt{1+\rho^2}}{2\rho}=\frac{3}{2},$$

and then

$$\begin{aligned}&\psi(t)=e^{-\frac{2}{3s}\frac{1}{|t|}}=o(e^{-1/|t|}t^4)=o(\varphi(t)t^4),\\ &\lim_{t\to 0}\frac{\psi(t)}{\varphi(t)^{2/\beta}t^{4/\beta}}=\lim_{t\to 0}\frac{e^{-\frac{2}{3s}\frac{1}{|t|}}}{e^{-\frac{2}{\beta}\frac{1}{|t|}}t^{4/\beta}}=\lim_{t\to 0}\frac{e^{(\frac{2}{\beta}-\frac{2}{3s})\frac{1}{|t|}}}{t^{4/\beta}}=0,\\ &\sup_{0<t<1}\frac{\varphi(\gamma t)t^4}{\omega_s(\psi(t))}=\sup_{0<t<1}\frac{e^{-\frac{1}{\gamma_\rho|t|}}t^4}{e^{-\frac{2}{3}\frac{1}{|t|}}}=\sup_{0<t<1}t^4=1.\end{aligned}$$

Then Theorem 5.4 (1) of [19] shows that f_{φ,ψ,h_ρ} is ω_s-monotone. If the factorization $\mathbf{F}_{\varphi,\psi}(W,t)=\lambda(W,t)\mathbf{B}(W,t)$ held, then

$$f_{\varphi,\psi,h}(W,t)\approx\operatorname{trace}\mathbf{F}_{\varphi,\psi}(W,t)=\lambda(W,t)\operatorname{trace}\mathbf{B}(W,t)\approx\lambda(W,t),$$

hence $\lambda\approx f_{\varphi,\psi}$, and so $\lambda(W,t)$ is also ω_s-monotone,[5] and hence λ is a finite sum of squares of $C^{2,\delta_{n-1}}$ functions by Theorem 4.8 in [19].

[5] It is an easy exercise to show that ω_s-monotonicity is preserved under comparability of functions when $0<s<1$.

Moreover, the smooth elliptic matrix $\mathbf{B}(x, y, z, t)$ can be written as a sum of squares of smooth vector fields using the 1-square decomposition[6] of B in the beginning of Section 3 in [20] together with Lemma 34 of [20] and induction. Altogether then, $\lambda(W, t)\mathbf{B}(W, t)$ can be written as a finite sum of squares of $C^{2,\delta_{n-1}}$ vector fields, contradicting the fact that $\mathbf{F}_{\varphi,\psi}(W, t)$ cannot be written as a finite sum of squares of even $C^{1,\beta}$ vector fields if (2.15) holds.

3. A rough variant of Christ's theorem

We now prove our extension of Christ's hypoellipticity theorem, namely, Theorem 2.6, to the case of a sum of squares of *rough* vector fields, whose sum of squares is nevertheless smooth. We will assume the rough symbols are in the classes $\mathcal{C}^{2,\delta} S^{\alpha}_{1,0}$, but we could just as well formulate and prove a variant for the symbol classes $\mathcal{C}^{2,\delta} S^{\alpha}_{\rho,\eta}$, which we leave for the interested reader, as we will not use such a variant in our applications. The proof of this rough theorem is accomplished by adapting the sum of squares argument of Christ [7] in the smooth case. For this we begin with some preliminaries.

3.1. Preliminaries

Here we recall definitions and properties of symbols, Gårding's inequality, parametrices, rough symbols, and wave front sets. We include some proofs for convenience of the reader.

3.1.1. Symbols. We begin by recalling in $\mathbb{R}^n$, the definition of symbols $S^m_{\rho,\eta}$ from Stein (see [26], Chapter VI), the definition of symbols $S^{m,k}_{\rho,\eta}$ and $S^{m+}_{\rho,\eta}$ from Christ [7], and then some results on rough versions of the symbol classes $S^m_{\rho,\eta}$ from [25] and [28]. See also Trèves [29] for symbols defined in open sets $\Omega \subset \mathbb{R}^n$.

Definition 3.1. Let $a(x, \xi)$ be a smooth function on $\mathbb{R}^n \times \mathbb{R}^n$, and let $0 \le \eta < \rho \le 1$ and $-\infty < m < \infty$.

(1) Define $a \in S^m_{\rho,\eta}$, referred to as a *symbol* of type (ρ, η) and order m, if

(3.1) $$|\partial^\alpha_x \partial^\beta_\xi a(x, \xi)| \le C_{\alpha,\beta} \langle\xi\rangle^{m-\rho|\beta|+\eta|\alpha|}, \quad x \in \mathbb{R}^n, \xi \in \mathbb{R}^n, (\alpha, \beta) \in \mathbb{Z}^n_+ \times \mathbb{Z}^n_+.$$

(2) Define $a \in S^{m,k}_{\rho,\eta}$ if $|\partial^\alpha_x \partial^\beta_\xi a(x, \xi)| \le C_{\alpha,\beta} \langle\xi\rangle^{m-\rho|\beta|+\eta|\alpha|} (\log\langle\xi\rangle)^{k+|\alpha|+|\beta|}$.

(3) Define
$$S^{m+}_{\rho,\eta} \equiv \bigcap_{\varepsilon>0} S^{m,\varepsilon}_{\rho-\varepsilon,\eta+\varepsilon}, \quad m \in \mathbb{R}.$$

For a symbol $a \in S^m_{\rho,\eta}$, the associated pseudodifferential operator $A: \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$, also denoted by $A = \mathrm{Op}a$, is defined on the space of rapidly decreasing functions $\mathcal{S}(\mathbb{R}^n)$ on $\mathbb{R}^n$ by
$$Au(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{ix\cdot\xi}\, a(x, \xi)\, \hat{u}(\xi)\, d\xi, \quad x \in \mathbb{R}^n.$$

[6] Which only makes use of square roots and reciprocals of functions that are smooth and positive in this case.

It follows with some work (see, e.g., [26]) that Opa: $\mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is continuous, and, moreover, if a_k converges pointwise to a on $\mathbb{R}^n$, and (3.1) holds for $a = a_k$ uniformly in k, then $a \in S^m_{\rho,\eta}$ as well. By duality Opa: $\mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$ is a continuous map from the space of tempered distributions $\mathcal{S}'(\mathbb{R}^n)$ to itself, and the asymptotic formulas for adjoints and compositions holds without restriction, e.g., if $a \in S^{m_1}_{\rho,\eta}$ and $b \in S^{m_2}_{\rho,\eta}$, then we have Op$a \circ$ Opb = Op$(a \circ b)$, where for all $M \in \mathbb{N}$,

$$a \circ b = \sum_{\ell=0}^{M} \frac{1}{i^\ell \ell!} \hat{\nabla}^\ell_\xi a \cdot \nabla^\ell_x b + E_M, \quad \text{with } E_M \in S^{m_1+m_2-M-1}_{\rho,\eta}.$$

It follows immediately from the definitions that the asymptotic formulas for adjoints and compositions extend to the symbol classes $S^{m+}_{\rho,\eta}$. For example, by uniqueness of the expansions, we have

$$E_M \in S^{m_1+\varepsilon+m_2+\varepsilon-M-1}_{\rho-\varepsilon,\eta+\varepsilon} \subset S^{m_1+m_2-M-1,2\varepsilon}_{\rho-2\varepsilon,\eta+2\varepsilon} \quad \text{for each } \varepsilon > 0,$$

and so

$$E_M \in \bigcap_{\varepsilon>0} S^{m_1+m_2-M-1,2\varepsilon}_{\rho-2\varepsilon,\eta+2\varepsilon} = S^{m_1+m_2-M-1+}_{\rho,\eta}.$$

Now $S^{m+}_{\rho,\eta} \subset S^{m,k}_{\rho,\eta}$, and it turns out that for our purposes, we apply the pseudodifferential calculus to the symbol classes $S^{m+}_{\rho,\eta}$, as well as to the classes $S^{m,k}_{\rho,\eta}$ that arise naturally from the hypotheses of the theorems. We will not necessarily make explicit mention of this distinction in the sequel however.

3.1.2. Parametrices. Let $a(x,\xi) \in S^m_{1,\eta}$ be elliptic of order m, i.e., there are strictly positive continuous functions $\rho(x)$ and $c(x)$ in Ω such that the symbol $a(x,\xi)$ satisfies

$$c(x)|\xi|^m \le |a(x,\xi)|, \quad \xi \in \mathbb{R}^n, \text{with } |\xi| \ge \rho(x),\ x \in \Omega.$$

Proposition 3.2. *Let $a(x,\xi) \in S^m_{1,\eta}(\Omega)$. If $a(x,\xi)$ is elliptic of order m, then there is $b(x,\xi) \in S^{-m}_{1,\eta}$ such that $a \circ b = 1$. Conversely, if there is $b(x,\xi) \in S^{-m}_{1,\eta}$ such that $a \circ b = 1$, then $a(x,\xi)$ is elliptic of order m.*

Proof. Determine recursively symbols b_j from the relations

$$\begin{cases} b_0(x,\xi)a(x,\xi) = 1, \\ b_j(x,\xi)a(x,\xi) = -\displaystyle\sum_{1\le|\alpha|\le j} \frac{1}{\alpha!} \partial^\alpha_\xi a(x,\xi) D^\alpha_x b_{j-|\alpha|}(x,\xi), \quad j \ge 1, \end{cases} \tag{3.2}$$

which make sense only for $|\xi| \ge \rho(x)$. The first three such symbols are given by

$$\begin{aligned} b_0(x,\xi) &= \frac{1}{a(x,\xi)}, \\ b_1(x,\xi) &= -b_0(x,\xi) \sum_{i=1}^{n} \frac{\partial}{\partial \xi_i} a(x,\xi) \frac{1}{i} \frac{\partial}{\partial x_i} b_0(x,\xi) \\ &= -\frac{1}{i} b_0(x,\xi) \nabla_\xi a(x,\xi) \cdot \nabla_x b_0(x,\xi), \end{aligned}$$

$$b_2(x,\xi) = -b_0(x,\xi)\sum_{i=1}^{n}\frac{\partial}{\partial\xi_i}a(x,\xi)\,\frac{1}{i}\,\frac{\partial}{\partial x_i}b_1(x,\xi)$$
$$- b_0(x,\xi)\sum_{|\alpha|=2}\frac{1}{\alpha!}\partial_\xi^\alpha a(x,\xi)D_x^\alpha b_0(x,\xi)$$
$$= -\frac{1}{i}\,b_0(x,\xi)\nabla_\xi a(x,\xi)\cdot\nabla_x b_1(x,\xi) - b_0(x,\xi)\,\frac{1}{2!}\nabla_\xi^2 a(x,\xi)\cdot\nabla_x^2 b_0(x,\xi).$$

To deal with the requirement that $|\xi| \geq \rho(x)$, we select a monotone increasing sequence of continuous functions $\rho_{j+1}(x) > \rho_j(x) > \rho(x)$ and a sequence of smooth cutoff functions $\chi_j(x,\xi) \in C^\infty(\Omega \times \mathbb{R}^n)$ satisfying

$$\chi_j(x,\xi) = \begin{cases} 0 & \text{if } |\xi| \leq \rho_j(x), \\ 1 & \text{if } |\xi| \leq 2\rho_j(x). \end{cases}$$

One can easily prove by induction on j that $\chi_j b_j \in S^{-m-j}(\Omega)$ and, moreover, that for carefully chosen such χ_j, the series $\sum_{j=1}^\infty \chi_j b_j$ converges in $S^{-m}(\Omega)$ to a symbol b satisfying $a \circ b = 1$. Indeed, if $\{K_j\}_{j=1}^\infty$ is a standard exhausting sequence of compact sets for Ω, and if the constants $C_{\alpha,\beta}^{(j)}(K_i)$ satisfy

$$|\partial_\xi^\alpha \partial_x^\beta(\chi_j b_j)| \leq C_{\alpha,\beta}^{(j)}(K_i)\,|\xi|^{-m-j-|\alpha|} \quad \text{for } x \in K_i,\ \xi \in \mathbb{R}^n \setminus \{0\},$$

then we need only require in addition that $\rho_j(x) \geq 2\sup_{i\leq j,|\alpha+\beta|\leq j} C_{\alpha,\beta}^{(j)}(K_i)^{1/j}$.

The converse is an easy exercise using only the consequence

$$a(x,\xi)b(x,\xi) - 1 \in S^{-1}(\Omega),$$

which implies that for every compact set $K \subset \Omega$, there is a constant C_K such that

$$|a(x,\xi)b(x,\xi) - 1| \leq C_K\,\frac{1}{1+|\xi|}. \qquad \blacksquare$$

Corollary 3.3. *Let A belong to $S_{1,0}^m(\Omega)$. Then A is elliptic of order m if and only if there is $B \in S_{1,0}^{-m}(\Omega)$ with*

$$AB = BA = I \bmod S^{-\infty}(\Omega),$$

where $S^{-\infty}(\Omega) = \bigcap_{m\in\mathbb{R}} S_{1,0}^{-m}(\Omega)$.

3.1.3. Rough symbols. The following definitions are taken from [28] and [25].

Definition 3.4. A symbol $\sigma\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ belongs to the rough symbol class $\mathcal{C}^M S_{\rho,\delta}^m$ (where $M \in \mathbb{Z}_+$ and $0 \leq \rho,\delta \leq 1$) if for all multiindices α,β with $|\alpha| \leq M$, there are constants $C_{\alpha,\beta}$ such that

$$|D_x^\alpha D_\xi^\beta \sigma(x,\xi)| \leq C_{\alpha,\beta}(1+|\xi|)^{m+\delta|\alpha|-\rho|\beta|}, \quad x \in \mathbb{R}^n, \xi \in \mathbb{R}^n.$$

If $0 < \mu < 1$, then $\sigma \in \mathcal{C}^{M+\mu} S_{\rho,\delta}^m$ if in addition, for all $x \in \mathbb{R}^n, \xi \in \mathbb{R}^n$, we have

$$\left| D_\xi^\beta \sigma(x+h,\xi) - \sum_{\ell=0}^{M}\frac{(h\cdot\nabla_x)^\ell}{\ell!}D_\xi^\beta\sigma(x,\xi)\right| \leq C_{M,\beta}|h|^{M+\mu}(1+|\xi|)^{m+\delta(M+\mu)-\rho|\beta|}.$$

Definition 3.5. A symbol $\sigma: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ belongs to the operator class $\mathcal{O}_I^m$ if its associated operator

$$(\mathrm{Op}\sigma)u(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{ix\cdot\xi} \sigma(x,\xi) \widehat{u}(\xi)\, d\xi, \quad x \in \mathbb{R}^n,$$

admits a bounded extension from $H^{s+m}_{p,\mathrm{comp}}$ to $H^s_{p,\mathrm{loc}}$ (respectively, $\Lambda^{s+m}_{p,\mathrm{comp}}$ to $\Lambda^s_{p,\mathrm{loc}}$) for $s \in I$ (respectively, $s \in I \cap [0,\infty)$) and all $1 < p < \infty$.

The symbol σ belongs to the operator class $\overline{\mathcal{O}}_I^m$ if in addition $\mathrm{Op}\sigma$ is bounded from $\Lambda^{t+m}_{p,\mathrm{comp}}$ to $\Lambda^t_{p,\mathrm{loc}}$, where t is the right endpoint of the interval I. Here the subscript comp means compactly supported distributions in the space, while the subscript loc means distributions locally in the space.

The following result of Bourdaud is well known, see also Theorem 3 in [25].

Theorem 3.6 ([4]). *For all real m, and all $\nu > 0$ and $0 \le \delta < 1$, we have*

$$\mathcal{C}^\nu S^m_{1,\delta} \subset \overline{\mathcal{O}}^m_{(-(1-\delta)\nu,\nu)}.$$

3.1.4. Rough pseudodifferential calculus. While symbol smoothing is a very effective and relatively simple tool for use in elliptic and finite type situations, it fails to sufficiently preserve the subunit property of vector fields in the infinitely degenerate regime. For this reason, we will instead use the pseudodifferential calculus from [25], to which we now turn.

If $\sigma \in \mathcal{C}^\nu S^{m_1}_{1,\delta_1}$ and $\tau \in \mathcal{C}^{M+\mu+\nu} S^{m_2}_{1,\delta_2}$ have compact support in $\mathbb{R}^n \times \mathbb{R}^n$, then the composition $\mathrm{Op}\sigma \circ \mathrm{Op}\tau$ of the operators $\mathrm{Op}\sigma$ and $\mathrm{Op}\tau$ equals the operator $\mathrm{Op}(\sigma \circ \tau)$, where

$$(\sigma \circ \tau)(x,\eta) \equiv \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{i(x-y)\cdot(\xi-\eta)} \sigma(x,\xi)\tau(y,\eta)\, dy\, d\xi,$$

and the double integral on the right-hand side is absolutely convergent under the compact support assumption, thus justifying the claim. Given such symbols without the assumption of compact support, we may then consider instead the symbols σ_ε and τ_ε, where $a_\varepsilon(x,\xi) \equiv \psi(\varepsilon x, \varepsilon\xi) a(x,\xi)$. Provided $\psi \in C_c^\infty(\mathbb{R}^n \times \mathbb{R}^n)$ is 1 on the unit ball, the symbols a_ε are uniformly in the same symbol class as a, and hence the above formula persists in the limit when the operators are restricted to acting on the space $\mathcal{S}$ of rapidly decreasing functions. Of course it may happen that the resulting symbol $\sigma \circ \tau$ fails to belong to any reasonable rough symbol class $\mathcal{C}^{M+\mu} S^m_{\rho,\delta}$, see Section 5.3 in [25]. Nevertheless, we have the following useful symbol expansion of $\sigma \circ \tau$ valid up to an error operator in an appropriate class $\overline{\mathcal{O}}_I^m$.

Theorem 3.7 ([25, Theorem 4]). *Suppose $\sigma \in \mathcal{C}^\nu S^{m_1}_{1,\delta_1}$ and $\tau \in \mathcal{C}^{M+\mu+\nu} S^{m_2}_{1,\delta_2}$, where M is a nonnegative integer, $0 < \mu, \delta_1, \delta_2 < 1$, $\nu > 0$ and $M + \mu \ge m_1 \ge 0$. Let $\delta \equiv \max\{\delta_1, \delta_2\}$. Then*

$$\sigma \circ \tau = \sum_{\ell=0}^{M} \frac{1}{i^\ell \ell!} \nabla_\xi^\ell \sigma \cdot \nabla_x^\ell \tau + E, \quad E \in \mathcal{O}^{m_1+m_2+(M+\mu)(\delta_2-1)+\varepsilon}_{(-(1-\delta)\nu,\nu)} \quad \textit{for every } \varepsilon > 0.$$

There is an analogous expansion for the symbol of the adjoint operator $(\mathrm{Op}\sigma)^{\mathrm{tr}}$.

3.1.5. Smooth distributions and wave front sets. The following definitions are taken from Trèves [29].

Definition 3.8. A distribution u in an open set $\Omega \subset \mathbb{R}^n$ is said to be C^∞ in some neighbourhood of a point $(x_0, \xi^0) \in \Omega \times (\mathbb{R}^n \setminus \{0\})$ if there is a function $g \in C_c^\infty(\mathbb{R}^n)$ equal to 1 in a neighbourhood of x_0, and an open cone $\Gamma^0 \subset \mathbb{R}^n$ containing ξ^0 such that for every $M > 0$, there is a positive constant C_M satisfying

$$|\widehat{gu}(\xi)| \leq C_M(1 + |\xi|)^{-M}, \quad \xi \in \Gamma^0.$$

Definition 3.9. A distribution u in an open set $\Omega \subset \mathbb{R}^n$ is said to be C^∞ in a conic open subset $\Gamma \subset \Omega \times (\mathbb{R}^n \setminus \{0\})$ if it is C^∞ in some neighbourhood of every point of Γ. The wave front set $WF(u)$ of u is the complement in $\Omega \times (\mathbb{R}^n \setminus \{0\})$ of the union of all conic open sets in which u is C^∞:

$$WF(u) \equiv \Omega \times (\mathbb{R}^n \setminus \{0\}) \setminus \bigcup \big\{\Gamma \text{ conic open} \subset \Omega \times (\mathbb{R}^n \setminus \{0\}) : u \text{ is } C^\infty \text{ in } \Gamma\big\}.$$

For $\gamma \in \mathbb{R}$, the H^γ wave front set of u is defined analogously, where H^γ is the Sobolev space of order γ.

3.2. Proof of Theorem 2.6, the limited smoothness variant of Christ's theorem

Now we begin our proof of the limited smoothness variant Theorem 2.6, in the setting of real vector fields, of Christ's theorem. Let $u \in \mathcal{D}'(V)$ and $0 < \gamma < \delta$ be given. Suppose that the H^γ wave front set of Lu is disjoint from some open conic neighbourhood Γ_0 of a point $(x_0, \xi_0) \in T^*V$. Without loss of generality, we may assume that $u \in \mathcal{E}'(V)$. Fix an integer $K \in \mathbb{Z}$ (possibly quite large) such that $u \in H^{-K}$. We will show that $(x_0, \xi_0) \notin WF_{H^\gamma}(u)$ by first constructing a pseudodifferential operator Λ that is elliptic of order γ in a smaller compact conic neighbourhood Γ_1 of (x_0, ξ_0), and then showing that $\Lambda u \in H^0(\mathbb{R}^d)$.

To do this, let ψ be as in part (c) of Theorem 2.6. Recall the definitions of the symbol classes $S^m_{\rho,\eta}$, $S^{m,k}_{\rho,\eta}$ and $S^{m+}_{\rho,\eta}$ from Definition 3.1. Then, following Christ, we define a symbol of nonconstant order, depending on parameters γ and N_0 by

$$\lambda(x, \xi) = \begin{cases} |\xi|^\gamma e^{-N_0(\log|\xi|)\psi(x,\xi)} & \text{if } |\xi| \geq e, \\ C^\infty \text{ and nonvanishing} & \text{if } |\xi| < e. \end{cases} \tag{3.3}$$

The nonnegativity of ψ implies that $\lambda \in S^{\gamma+}_{1,0}$. Moreover, $\lambda \in S^{\gamma,0}_{1,0}$. With γ fixed, there exists $\theta > 0$ such that for each N_0, we have $\lambda \in S^{-\theta N_0+}_{1,0}$ on the closure of the complement of Γ_1. Now choose N_0 so that $-\theta N_0 < -K$. Then, with

$$\Lambda = \mathrm{Op}(\lambda),$$

we have $\Lambda u \in H^{-K+\theta N_0} \subset H^0$ microlocally on the complement of Γ_1.

Define cutoff functions $\eta_1, \eta_2 \in C_c(\mathbb{R}^d)$ such that $\eta_2 \equiv 1$ in a neighbourhood of the support of u, $\eta_1 \equiv 1$ in a neighbourhood of the support of η_2, and $\mathrm{Supp}\, \eta_1 \subset V$.

Recall that if $a \in S^m_{\rho,\eta}$ and $b \in S^n_{\rho,\eta}$, and $\rho > \eta$, then $\mathrm{Op}(a) \circ \mathrm{Op}(b)$ has a symbol $a \odot b$ with an asymptotic expansion

$$a \odot b(x, \xi) \sim \sum_\alpha c_\alpha \, \partial_\xi^\alpha a(x, \xi) \, \partial_x^\alpha b(x, \xi), \quad c_\alpha = \frac{(-i)^\alpha}{\alpha!}. \tag{3.4}$$

The notation "∼" means that for every N, the operator

$$\mathrm{Op}(a) \circ \mathrm{Op}(b) - \mathrm{Op}\Big(\sum_{\alpha<N} c_\alpha\, \partial_\xi^\alpha a(x,\xi)\, \partial_x^\alpha b(x,\xi)\Big)$$

is smoothing of order $m+n-N(\rho-\eta)$ in the scale of Sobolev spaces. The next lemma is taken verbatim from [7], as it involves only symbols of type $(1,0)$.

Lemma 3.10 (Lemma 4.1 in [7]). *There exists an operator $\Lambda^{-1} \in S_{1,0}^{m+}$, for some $m = m(\gamma)$ depending on γ, such that $\Lambda \circ \Lambda^{-1} - \mathbb{I}$ is smoothing of infinite order. Moreover, such an operator may be constructed with a symbol of the form*

$$(1+f)\lambda^{-1}, \quad f \in S_{1,0}^{-1,2}.$$

Proof. Write $f \sim \sum_{k=1}^{\infty} f_k$. Solve the equation

$$\lambda \odot [(1+f)\lambda^{-1}] \sim 1$$

using the asymptotic expansion (3.4) and the usual iterative procedure as given in (3.2). One obtains $f_1 \in S^{-1,2}$, and by induction, each $f_k \in S_{1,0}^{-k+}$. Choose Λ to be an operator whose full symbol has expansion $\sum_{k=1}^{\infty} f_k$, so that the error is smoothing of all orders in the scale of Sobolev spaces. ∎

To prove an analogue of Lemma 4.2 in [7], we will need an auxiliary lemma.

Lemma 3.11. *Let $P \in \mathrm{Op}(C^\nu S_{1,0}^{m,l})$, where $m, l \in \mathbb{N}$, and let Λ be the operator in (3.3), where we recall that ψ is everywhere nonnegative, vanishes identically in a small conic neighbourhood of (x_0,ξ_0), and it is strictly positive on the complement of Γ_1. Then*

$$\Lambda P \Lambda^{-1} = P + R_1 + R_2 + E,$$

where $R_1 \in \mathrm{Op}(C^{\nu-1} S_{1,0}^{m-1,l+1})$, $R_2 \in \mathrm{Op}(C^{\nu-2} S_{1,0}^{m-2,l+2})$, and $E \in \mathcal{O}_{(-\nu,\nu)}^{m-M-\varepsilon}$ for every $m \le M < \nu$ and some $0<\varepsilon<1$. Moreover, the operator R_1 has the form

$$R_1 = \mathrm{Op}(\{\log\lambda, \sigma(P)\}).$$

Proof. Using Theorem 3.7, we see that the symbol of $\Lambda P - P\Lambda$ divided by λ equals

$$\begin{aligned}
&\frac{1}{\lambda}\{\lambda \odot \sigma(P) - \sigma(P) \odot \lambda\} \\
&\quad = \Big\{\sum_{|\alpha|=1} + \sum_{2\le|\alpha|\le M}\Big\} c_\alpha \Big[\frac{\partial_\xi^\alpha \lambda}{\lambda}\, \partial_x^\alpha \sigma(P) - \partial_\xi^\alpha \sigma(P)\, \frac{\partial_x^\alpha \lambda}{\lambda}\Big] + E \\
&\quad = \sum_{|\alpha|=1} c_\alpha \big[\partial_\xi^\alpha \log\lambda\, \partial_x^\alpha \sigma(P) - \partial_\xi^\alpha \sigma(P)\, \partial_x^\alpha \log\lambda\big] + \text{symbol in } C^{\nu-M} S_{1,0}^{m-2,l+2} + E \\
&\quad = \{\log\lambda, \sigma(P)\} + \text{symbol in } C^{\nu-M} S_{1,0}^{m-2,l+2} + E,
\end{aligned}$$

where $E \in \mathcal{O}_{(-\nu,\nu)}^{m-M-\varepsilon}$ for every $M<\nu$ and some $0<\varepsilon<1$; and $\{\log\lambda, \sigma(P)\}$ is the Poisson bracket of $\log\lambda$ and $\sigma(P)$, and is a symbol in $\mathcal{C}^{\nu-1} S_{1,0}^{m-1,l+1}$. ∎

Define

$$L_1 = \sum_j X_j^{\mathrm{tr}} X_j + \sum_j A_j X_j + \sum_j X_j^{\mathrm{tr}} \tilde{A}_j + A_0, \tag{3.5}$$

so that $L = L_1 + R_1 + \widehat{\nabla} \cdot \mathbf{Q}_p \widehat{\nabla}$. This next lemma is our first analogue of Lemma 4.5 in [7].

Lemma 3.12 (Lemma 4.5 in [7]). *Let Λ be the operator with symbol λ in (3.3). Suppose that L_1 takes the form (3.5). Define*

$$b_j \equiv \mathrm{Op}\{\log\lambda, \sigma(X_j^{\mathrm{tr}})\} \quad \text{and} \quad \widetilde{b}_j \equiv \mathrm{Op}\{\log\lambda, \sigma(X_j)\}.$$

Then there exists a pseudodifferential operator G of the form

$$G = \sum_j B_j \circ X_j + \sum_j X_j^{\mathrm{tr}} \circ \widetilde{B}_j + B_0, \tag{3.6}$$

such that

$$(L_1 + G)\eta_1 \Lambda \eta_2 = \eta_1 \Lambda L \eta_2 + R$$

where

$$\begin{cases} B_j = b_j + c_j \text{ and } \widetilde{B}_j = \widetilde{b}_j + \widetilde{c}_j & \text{for every } j \geq 1, \\ B_0 = \sum_j (b_j \circ \widetilde{b}_j + A_j \widetilde{b}_j + \widetilde{A}_j b_j) \mod \mathrm{Op}(C^{0,\delta} S_{1,0}^{-1,2}), \end{cases} \tag{3.7}$$

where $c_j, \widetilde{c}_j \in \mathrm{Op}(C^{0,\delta} S^{-1,1})$, and $A_j, \widetilde{A}_j \in C^{1,\delta} S_{1,0}^0$ are the coefficients of the differential operator L in (3.5), and $R \in \mathcal{O}_{(-\delta,\delta)}^{-\varepsilon}$.

Proof. In constructing the symbol of G we will work formally, ignoring the cutoff functions η_1 and η_2. This is permissible by pseudolocality, since $\eta_1\eta_2 = \eta_2$. The desired equation $(L_1 + G)\Lambda = \Lambda L + R$ is then equivalent to

$$\begin{aligned} G &= \Lambda L_1 \Lambda^{-1} - L_1 + R\Lambda^{-1} \\ &= \sum_j [\Lambda X_j^{\mathrm{tr}} X_j \Lambda^{-1} - X_j^{\mathrm{tr}} X_j] + R\Lambda^{-1} \\ &\quad + \sum_j \Lambda(A_j X_j + X_j^{\mathrm{tr}} \widetilde{A}_j + A_0)\Lambda^{-1} - \sum_j (A_j X_j + X_j^{\mathrm{tr}} \widetilde{A}_j + A_0) \\ &\equiv G_{\mathrm{top}} + \Lambda G_{\mathrm{lower}} \Lambda^{-1} - G_{\mathrm{lower}}, \end{aligned}$$

where

$$\begin{aligned} G_{\mathrm{top}} &= \sum_j [\Lambda X_j^{\mathrm{tr}} X_j \Lambda^{-1} - X_j^{\mathrm{tr}} X_j] + R\Lambda^{-1} \\ &= \sum_j [(\Lambda X_j^{\mathrm{tr}} \Lambda^{-1})(\Lambda X_j \Lambda^{-1}) - X_j^{\mathrm{tr}} X_j] + R\Lambda^{-1} \end{aligned}$$

and

$$G_{\mathrm{lower}} = \sum_j (A_j X_j + X_j^{\mathrm{tr}} \widetilde{A}_j + A_0).$$

We first consider G_{top}. Using Lemma 3.11, with $P = X_j$, $m = 1$, $l = 0$, we have

$$\begin{aligned}\Lambda X_j \Lambda^{-1} &= X_j + \mathrm{Op}(\{\log\lambda, \sigma(X_j)\}) + \text{symbol in } C^{0,\delta} S^{-1,2}_{1,0} \bmod \mathcal{O}^{-1-\varepsilon}_{(-\delta,\delta)} \\ &= X_j + b_j + c_j \bmod \mathcal{O}^{-1-\varepsilon}_{(-\delta,\delta)},\end{aligned}$$

where $b_j = \mathrm{Op}(\{\log\lambda, \sigma(X_j)\}) \in C^{1,\delta} S^{0,1}_{1,0}$ and c_j has a symbol in $C^{0,\delta} S^{-1,2}_{1,0}$. Since both $\{\log\lambda, \sigma(X_j)\}$ and $\{\log\lambda, \sigma(X_j^{\text{tr}})\}$ belong to $C^{1,\delta} S^{0,1}_{1,0}$, inserting these equations into the identity derived for G_{top} in the preceding paragraph shows that

$$G_{\text{top}} = \sum_j B_j \circ X_j + \sum_j X_j^{\text{tr}} \circ \widetilde{B}_j + B_0,$$

where the operators $B_j, \widetilde{B}_j \in \mathrm{Op}(C^{1,\delta} S^{0,1}_{1,0})$ and $B_0 \in \mathrm{Op}(C^{0,\delta} S^{0,2}_{1,0})$ satisfy (3.7).

Now consider G_{lower}. We can write

$$\begin{aligned}\Lambda G_{\text{lower}} \Lambda^{-1} &= \sum_j \Lambda (A_j X_j + X_j^{\text{tr}} \widetilde{A}_j + A_0) \Lambda^{-1} \\ &= \sum_j (\Lambda A_j \Lambda^{-1} \Lambda X_j \Lambda^{-1} + \Lambda X_j^{\text{tr}} \Lambda^{-1} \Lambda \widetilde{A}_j \Lambda^{-1} + \Lambda A_0 \Lambda^{-1}).\end{aligned}$$

Applying Lemma 3.11 to A_j and X_j, we have

$$\begin{aligned}\Lambda A_j \Lambda^{-1} &= A_j + \text{symbol in } \mathrm{Op}(C^{0,\delta} S^{-1,1}_{1,0}) \bmod \mathcal{O}^{-1-\varepsilon}_{(-\delta,\delta)}, \\ \Lambda X_j \Lambda^{-1} &= X_j + \mathrm{Op}(\{\log\lambda, \sigma(X_j)\}) + \text{symbol in } C^{0,\delta} S^{-1,2}_{1,0} \bmod \mathcal{O}^{-1-\varepsilon}_{(-\delta,\delta)}.\end{aligned}$$

Using Theorem 3.7, this gives

$$\Lambda A_j \Lambda^{-1} \Lambda X_j \Lambda^{-1} = A_j X_j + c_j X_j + \text{ symbol in } C^{0,\delta} S^{0,1}_{1,0} \bmod \mathcal{O}^{-\varepsilon}_{(-\delta,\delta)},$$

with $c_j \in C^{0,\delta} S^{-1,1}_{1,0}$, and where the symbol in $C^{0,\delta} S^{0,1}_{1,0}$ has the form $A_j \tilde{b}_j +$ symbol in $C^{0,\delta} S^{0,1}_{1,0}$ with $\tilde{b}_j = \{\log\lambda, \sigma(X_j)\}$. Analyzing the other terms in $\Lambda G_{\text{lower}} \Lambda^{-1}$ in the same way, we obtain

$$\Lambda G_{\text{lower}} \Lambda^{-1} = B_j X_j + X_j^{\text{tr}} \tilde{B}_j + \tilde{B}_0 \bmod \mathcal{O}^{-\varepsilon}_{(-\delta,\delta)},$$

where $B_j, \tilde{B}_j$ as in (3.7), and $\tilde{B}_0 \in \mathrm{Op}(C^{0,\delta} S^{0,1}_{1,0})$ and has the structure as in (3.7). Combining with the estimate for G_{top}, we obtain the result. ■

Lemma 3.13 (Lemma 4.6 in [7]). *Suppose that L, ψ, p satisfy the hypotheses of Theorem 2.6. Then, for any $N \geq 0$, and for any fixed relatively compact subset $U \subset V$, any $\delta > 0$ and any $f \in C^{\gamma+3}$ supported in U, the operator G constructed in Lemma 3.12 satisfies*

$$|\langle Gf, f\rangle| \leq \delta \sum_j \|X_j f\|^2 + \delta \|\sqrt{a}\,\hat{\nabla} f\|^2 + C_\delta \|f\|^2 + C_\delta \|\mathrm{Op}(p) f\|^2_{H^1}. \tag{3.8}$$

Proof. We first note

$$\sigma(b_j) = \{\log\lambda, \sigma(X_j^{\text{tr}})\} = -N_0 \log|\xi| \{\psi, \sigma(X_j^{\text{tr}})\} + \text{symbol in } C^{1,\delta} S^{0}_{1,0},$$

and similarly for $\widetilde{b}_j$. Using this together with (3.7) and hypothesis (2.8) with $\delta = \delta_0$, we therefore obtain

$$\begin{aligned}|\langle B_j \circ X_j f, f\rangle| &= |\langle (b_j + c_j) \circ X_j f, f\rangle| \le \varepsilon \|X_j f\|^2 + C_\varepsilon \|\widetilde{b}_j f\|^2 + C_\varepsilon \|f\|^2 \\ &\le \varepsilon \|X_j f\|^2 + C_\varepsilon \|\log|\xi|\{\psi, \sigma(X_j^{\mathrm{tr}})\} f\|^2 + C_\varepsilon \|f\|^2 \\ &\le \varepsilon \|X_j f\|^2 + C_\varepsilon \|f\|^2 \\ &\quad + C_\varepsilon \Big(\delta_0 \sum_j \|X_j u\|^2 + \delta_0 \|\sqrt{a}\,\widehat{\nabla} f\|^2 + C_{\delta_0} \|f\|^2 + C_{\delta_0} \|\mathrm{Op}(p) f\|^2_{H^1}\Big).\end{aligned}$$

Choosing $\delta_0 = \varepsilon / C_\varepsilon$, this gives

$$|\langle B_j \circ X_j f, f\rangle| \le \varepsilon \sum_j \|X_j u\|^2 + \varepsilon \|\sqrt{a}\,\widehat{\nabla} f\|^2 + C_\varepsilon \|f\|^2 + C_\varepsilon \|\mathrm{Op}(p) f\|^2_{H^1}.$$

The rest of the terms in (3.6) are handled in the same way, giving (3.8). ■

To handle the Grushin type term $\widehat{\nabla} \cdot \mathbf{Q}_p(x) \widehat{\nabla}$ in (2.6), we will need the following two lemmas.

Lemma 3.14. *There holds*

$$(\widehat{\nabla} \cdot \mathbf{Q}_p \widehat{\nabla} \eta_1 + \mathbf{E}) \Lambda \eta_2 = \eta_1 \Lambda \widehat{\nabla} \cdot \mathbf{Q}_p \widehat{\nabla} \eta_2 + \mathbf{R},$$

where $\mathbf{R} \in \mathcal{O}^{-\varepsilon}_{(-\delta,\delta)}$, *and with* $\widehat{\xi} = (\xi_p, \dots, \xi_n)$, *the matrix operator* $\mathbf{E}$ *takes the form*

$$\begin{aligned}\mathbf{E} &= H \circ \mathbf{Q}_p \widehat{\nabla} + \widehat{\nabla}\Big(\sum_{|\alpha|=1} D^\alpha \mathbf{Q}_p\Big) \circ H_0 + H \circ \Big(\sum_{|\alpha|=1} D^\alpha \mathbf{Q}_p\Big) \circ H_0 \\ &\quad + H_3 \circ \mathbf{Q}_p \widehat{\nabla} + H \circ \mathbf{Q}_p H + \tilde{H}_0\, \mathcal{O}^{-\varepsilon}_{(-\delta,\delta)},\end{aligned} \tag{3.9}$$

where $H = \mathrm{Op}\{\log\lambda, \widehat{\xi}\} \in \mathrm{Op}(S^{0,1}_{1,0})$, $H_0 \in \mathrm{Op}(S^0_{1,0})$, $\tilde{H}_0 \in \mathrm{Op}(C^{0,\delta} S^0_{1,0})$, $H_3 \in \mathrm{Op}(S^{-1,1}_{1,0})$.

Proof. In constructing the symbol of $\mathbf{E}$, we will work formally, ignoring the cutoff functions η_1 and η_2. This is permissible by pseudolocality since $\eta_1 \eta_2 = \eta_2$. Let $\mathbf{L}_2 \equiv \widehat{\nabla} \cdot \mathbf{Q}_p \widehat{\nabla}$. The desired equation $(\mathbf{L}_2 + \mathbf{E})\Lambda = \Lambda \mathbf{L}_2 + \mathbf{R}$ is then equivalent to

$$\begin{aligned}\mathbf{E} &= \Lambda \mathbf{L}_2 \Lambda^{-1} - \mathbf{L}_2 + \mathbf{R}\Lambda^{-1} = \Lambda \widehat{\nabla} \cdot \mathbf{Q}_p \widehat{\nabla} \Lambda^{-1} - \widehat{\nabla} \cdot \mathbf{Q}_p \widehat{\nabla} + \mathbf{R}\Lambda^{-1} \\ &= (\Lambda \widehat{\nabla} \Lambda^{-1}) \cdot (\Lambda \mathbf{Q}_p \widehat{\nabla} \Lambda^{-1}) - \widehat{\nabla} \cdot \mathbf{Q}_p \widehat{\nabla} + \mathbf{R}\Lambda^{-1}.\end{aligned}$$

Next, using Lemma 3.11, we have

$$\Lambda \widehat{\nabla} \Lambda^{-1} = \widehat{\nabla} + \{\log\lambda, \widehat{\xi}\} + \text{symbol in } S^{-1,1}_{1,0} \equiv \widehat{\nabla} + H + H_3,$$

where $H = \{\log\lambda, \widehat{\xi}\} \in \mathrm{Op}(S^{0,1}_{1,0})$ and $H_3 \in \mathrm{Op}(S^{-1,1}_{1,0})$. To estimate $\Lambda \mathbf{Q}_p \widehat{\nabla} \Lambda^{-1}$, we will need a refinement of Lemma 3.11, namely, the estimate obtained in the proof

$$\begin{aligned}\frac{1}{\lambda}\{\lambda \odot \sigma(P) - \sigma(P) \odot \lambda\} &= \Big\{\sum_{|\alpha|=1} + \sum_{|\alpha|=2}\Big\} c_\alpha \Big[\frac{\partial_\xi^\alpha \lambda}{\lambda}\, \partial_x^\alpha \sigma(P) - \partial_\xi^\alpha \sigma(P)\, \frac{\partial_x^\alpha \lambda}{\lambda}\Big] + S \\ &= \{\log\lambda, \sigma(P)\} + \sum_{|\alpha|=2} c_\alpha [\partial_\xi^\alpha \log\lambda\, \partial_x^\alpha \sigma(P) - \partial_\xi^\alpha \sigma(P)\, \partial_x^\alpha \log\lambda] + S,\end{aligned}$$

where $S \in \mathcal{O}^{-1-\varepsilon}_{(-\nu,\nu)}$ for some $0 < \varepsilon < 1$ and $0 < \nu < \delta$. Now, $\sigma(P) = \sigma(\mathbf{Q}_p \widehat{\nabla}) = \mathbf{Q}_p \widehat{\xi}$, so

$$\sum_{|\alpha|=2} c_\alpha [\partial_\xi^\alpha \log\lambda \, \partial_x^\alpha \sigma(P) - \partial_\xi^\alpha \sigma(P) \, \partial_x^\alpha \log\lambda] = \sum_{|\alpha|=2} c_\alpha \partial_\xi^\alpha \log\lambda \, \partial_x^\alpha \mathbf{Q}_p \widehat{\xi},$$

which is a symbol in $\mathcal{C}^{0,\delta} S^{-1,0}_{1,0}$ since ψ does not depend on ξ in Γ, and therefore no logarithmic terms arise from differentiation of $\log\lambda$ with respect to ξ. Altogether, we thus have

$$\begin{aligned}
\Lambda \mathbf{Q}_p \widehat{\nabla} \Lambda^{-1} &= \mathbf{Q}_p \widehat{\nabla} + \mathrm{Op}(\{\log\lambda, \mathbf{Q}_p \widehat{\xi}\}) + \text{symbol in } \mathcal{C}^{0,\delta} S^{-1,0}_{1,0} \\
&= \mathbf{Q}_p \widehat{\nabla} + \sum_{|\alpha|=1} (D^\alpha \mathbf{Q}_p) \widehat{\xi} \, D_\xi^\alpha \log\lambda + \mathbf{Q}_p \cdot \{\log\lambda, \widehat{\xi}\} + \text{symbol in } \mathcal{C}^{0,\delta} S^{-1,0}_{1,0} \\
&= \mathbf{Q}_p \widehat{\nabla} + \sum_{|\alpha|=1} (D^\alpha \mathbf{Q}_p) \cdot \text{symbol in } S^0_{1,0} + \mathbf{Q}_p \cdot H + \text{symbol in } \mathcal{C}^{0,\delta} S^{-1,0}_{1,0},
\end{aligned}$$

where all the equalities hold mod $\mathcal{O}^{-1-\varepsilon}_{(-\delta,\delta)}$. We note that $\widehat{\xi} D_\xi^\alpha \log\lambda \in S^0_{1,0}$ for each α with $|\alpha| = 1$ since ψ does not depend on ξ, and therefore no logarithmic terms arise from differentiation of $\log\lambda$ with respect to ξ. This gives

$$\begin{aligned}
&(\Lambda \widehat{\nabla} \Lambda^{-1}) \cdot (\Lambda \mathbf{Q}_p \widehat{\nabla} \Lambda^{-1}) \\
&\quad = \widehat{\nabla} \cdot \mathbf{Q}_p \widehat{\nabla} + H \circ \mathbf{Q}_p \widehat{\nabla} + H_3 \circ \mathbf{Q}_p \widehat{\nabla} + \widehat{\nabla}\Big(\sum_{|\alpha|=1} D^\alpha \mathbf{Q}_p\Big) \circ H_0 \\
&\qquad + H \circ \Big(\sum_{|\alpha|=1} D^\alpha \mathbf{Q}_p\Big) \circ H_0 + H \circ \mathbf{Q}_p H + \tilde{H}_0 \bmod \mathcal{O}^{-\varepsilon}_{(-\delta,\delta)}. \qquad \blacksquare
\end{aligned}$$

Lemma 3.15. *Let* $\mathbf{E}$ *be a pseudodifferential operator of the form* (3.9). *Then, for any fixed relatively compact subset* $U \subset V$, *any* $\delta > 0$ *and any* $f \in C_c^\infty$ *supported in* U, *we have*

$$|\langle \mathbf{E} f, f\rangle| \le \delta \sum_j \|X_j f\|^2 + \delta \|\sqrt{a}\, \widehat{\nabla} f\|^2 + C_\delta \|f\|^2 + C_\delta \|\mathrm{Op}(p) f\|^2_{H^1}. \tag{3.10}$$

Proof. Here is where we will need to use that the matrix $\mathbf{Q}_p$ is subordinate – in the case $p = n$, $\mathbf{Q}_n$ is simply a scalar and the subordinate inequality is that of Malgrange. We will use (3.9) and the notation $\mathbf{Q}_p' = \sum_{|\alpha|=1} D^\alpha \mathbf{Q}_p$. We have

$$\begin{aligned}
\langle \mathbf{E} f, f\rangle = &-\langle \sqrt{\mathbf{Q}_p}\, H^{\mathrm{tr}} f, \sqrt{\mathbf{Q}_p}\, \widehat{\nabla} f\rangle - \langle H_0 f, \mathbf{Q}_p' \widehat{\nabla} f\rangle + \langle H_0 f, \mathbf{Q}_p' H^{\mathrm{tr}} f\rangle \\
&+ \langle H_3^{\mathrm{tr}} f, \mathbf{Q}_p \nabla' f\rangle + \langle \sqrt{\mathbf{Q}_p} H f, \sqrt{\mathbf{Q}_p}\, H^{\mathrm{tr}} f\rangle + \langle H_0 f, f\rangle.
\end{aligned}$$

Now we use the crucial fact that $\mathbf{Q}_p$ is subordinate, i.e., $|\mathbf{Q}_p'|^2 \le C \mathbf{Q}_p$, and together with Cauchy–Schwarz, this gives

$$|\langle \mathbf{E} f, f\rangle| \le \delta \|\sqrt{\mathbf{Q}_p}\, \widehat{\nabla} f\|^2 + C_\delta \|\sqrt{\mathbf{Q}_p}\, H^{\mathrm{tr}} f\|^2 + C_\delta \|\sqrt{\mathbf{Q}_p}\, H f\|^2 + C_\delta \|f\|^2.$$

Finally, using the definition of λ, we obtain

$$\sigma(H) = \{\log\lambda, \widehat{\xi}\} = -N_0 \log|\xi| \{\psi, \widehat{\xi}\},$$

which together with the fact that $\mathbf{Q}_p \approx a\,\mathbb{I}_{n-p+1}$ shows

$$|\langle \mathbf{E}f, f\rangle| \le \delta\|\sqrt{a}\,\widehat{\nabla} f\|^2 + C_\delta\|\sqrt{a}\,\mathrm{Op}(\log\langle\xi\rangle\{\psi,\widehat{\xi}\})f\|^2 + C_\delta\|f\|^2 .$$

Combining with estimate (2.8), as in the proof of Lemma 3.13, we conclude (3.10). ■

Finally, we obtain an estimate on the subunit term R_1.

Lemma 3.16. *Let $R_1 = \sum_{k=1}^n S_k\Theta_k \circ \widehat{\nabla}$, where each $S_k \in C^{1,\delta}(\mathbb{R}^{m\times m})$ is subunit with respect to $\mathbf{Q}_p$, and $\Theta_k = (\Theta_{kp},\dots,\Theta_{kn})$ is a multiplier of order zero. Then*

$$(R_1\eta_1 + J)\Lambda\eta_2 = \eta_1\Lambda R_1\eta_2 + R,$$

where $J \in \mathrm{Op}(\mathscr{C}^{0,\delta}S^{0,1}_{1,0})$, $R \in \mathcal{O}^{-1-\varepsilon}_{(-\delta,\delta)}$, and

$$|\langle Jf, f\rangle| \le \delta\sum_j \|X_j f\|^2 + \delta\|\sqrt{a}\,\widehat{\nabla} f\|^2 + C_\delta\|f\|^2 + C_\delta\|\mathrm{Op}(p)f\|^2_{H^1}, \tag{3.11}$$

for any $\delta > 0$ and any $f \in C_c^\infty$.

Proof. Proceeding as in the proof of Lemma 3.14, we have

$$\begin{aligned}\Lambda S_k\Theta_k \circ \widehat{\nabla}\Lambda^{-1} &= S_k\Theta_k \circ \widehat{\nabla} + \sum_{|\alpha|=1}(D^\alpha S_k)\widehat{\xi}\,\theta_k(\xi)D^\alpha_\xi \log\lambda + S_k\{\log\lambda, \widehat{\xi}\,\theta_k(\xi)\}\\ &\quad + \text{symbol in } C^{0,\delta}S^{-1,0}_{1,0}\\ &= S_k\Theta_k \circ \widehat{\nabla} + \text{symbol in } C^{0,\delta}S^0_{1,0} + S_kH_k + \text{symbol in } C^{0,\delta}S^{-1,0}_{1,0}\\ &\equiv S_k\Theta_k \circ \widehat{\nabla} + J_k,\end{aligned}$$

where all the equalities hold mod $\mathcal{O}^{-1-\varepsilon}_{(-\delta,\delta)}$ and $H_k \in \mathrm{Op}(S^{0,1}_{1,0})$. Defining $J \equiv \sum_{k=1}^n J_k$ and using the fact that S_k is subunit together with $\mathbf{Q}_p \approx a\mathbb{I}_{n-p+1}$ and (2.8), we obtain (3.11), and the proof is complete. ■

We are now ready to prove a generalization of Lemma 4.4 in [7], which is the main estimate we need.

Lemma 3.17 ([7, Lemma 4.4]). *Let L take the form (2.6) and satisfy (2.7) and (2.8). Let $0 < \gamma < \delta$ be fixed. If N_0 is chosen sufficiently large in the definition of Λ, then for any fixed relatively compact $U \Subset V$ and any $u \in C^{2,\delta}(U)$,*

$$\|\eta_1\Lambda u\|_{L^2(\mathbb{R}^n)} \le C\|\eta_1\Lambda Lu\|_{L^2(\mathbb{R}^n)} + C\|u\|_{H^0(\mathbb{R}^n)}. \tag{3.12}$$

Proof. Recall that

$$\begin{aligned}L &= \sum_j X_j^{\mathrm{tr}}X_j + \sum_j A_jX_j + \sum_j X_j^{\mathrm{tr}}\tilde{A}_j + A_0 + R_1 + \widehat{\nabla}\cdot\mathbf{Q}_p\widehat{\nabla}\\ &\equiv L_1 + L_2 + R_1,\end{aligned}$$

where we used the notation $L_2 = \widehat{\nabla}\cdot\mathbf{Q}_p\widehat{\nabla}$. If we set

$$v \equiv \eta_1\Lambda u \in C^2(\mathbb{R}^n),$$

we have

$$
\begin{aligned}
\langle (L_1+G)v,v\rangle &= \langle L_1 v,v\rangle + \langle Gv,v\rangle \\
&= \sum_j \|X_j v\|_{L^2}^2 + \sum_j \langle A_j \circ X_j v, v\rangle + \sum_j \langle X_j^{\mathrm{tr}} \circ \widetilde{A}_j v, v\rangle + \langle A_0 v,v\rangle + \langle Gv,v\rangle \\
&= \sum_j \|X_j v\|_{L^2}^2 + \sum_j \langle X_j v, A_j^{\mathrm{tr}} v\rangle + \sum_j \langle \widetilde{A}_j v, X_j v\rangle + \langle A_0 v,v\rangle + \langle Gv,v\rangle \\
&= \sum_j \|X_j v\|_{L^2}^2 + O\Big(\sqrt{\sum_j \|X_j v\|_{L^2}^2}\,\|v\|_{L^2} + \|v\|_{L^2}^2\Big) + \langle Gv,v\rangle,
\end{aligned}
$$

since the operators A_j and $\widetilde{A}_j$ have order 0. Similarly,

$$
\langle (L_2+E)v,v\rangle = \langle \nabla'\cdot \mathbf{Q}_p \widehat{\nabla} v, v\rangle + \langle Ev,v\rangle = \int |\sqrt{\mathbf{Q}_p}\widehat{\nabla} v|^2 + \langle Ev,v\rangle,
$$

$$
\langle (R_1+J_0)v,v\rangle = \Big\langle \sum_{i=1}^n S_i \Theta_i \widehat{\nabla} v, v\Big\rangle + \langle Jv,v\rangle \le \delta \int a|\widehat{\nabla} v|^2 + C_\delta \|v\|_{L^2}^2 + \langle J_0 v,v\rangle.
$$

We also have, from Lemmas 3.13, 3.14 and 3.16, that

$$
\begin{aligned}
(L_1+G)v &= (L_1+G)\eta_1\Lambda\eta_2 u = \eta_1\Lambda L_1\eta_2 u + Ru = \eta_1\Lambda L_1 u + Ru,\\
(L_2+E)v &= (L_2+E)\eta_1\Lambda\eta_2 u = \eta_1\Lambda L_2\eta_2 u + Ru = \eta_1\Lambda L_2 u + Ru,\\
(R_1+J_0)v &= (R_1+J)\eta_1\Lambda\eta_2 u = \eta_1\Lambda R_1\eta_2 u + Ru = \eta_1\Lambda R_1 u + Ru,
\end{aligned}
$$

since $\eta_2 u = u$, and hence, adding together,

$$
\begin{aligned}
|\langle (L+G+E+J)v,v\rangle| &\le |\langle \eta_1\Lambda L\eta_2 u, v\rangle| + |\langle Ru,v\rangle| \\
&\le \frac{1}{2}\|\eta_1\Lambda Lu\|_{L^2(\mathbb{R}^n)}^2 + \frac{1}{2}\|Ru\|_{L^2(\mathbb{R}^n)}^2 + \|v\|_{L^2(\mathbb{R}^n)}^2.
\end{aligned}
$$

Thus, from (3.8), (3.10), (3.11) and the above, we conclude that

$$
\begin{aligned}
&\sum_j \|X_j v\|_{L^2}^2 + \|\sqrt{\mathbf{Q}_p}\,\widehat{\nabla} v\|^2 \\
&\quad = \langle (L_1+G)v,v\rangle - \langle Gv,v\rangle + C\Big(\sqrt{\sum_j \|X_j v\|_{L^2}^2}\,\|v\|_{L^2} + \|v\|_{L^2}^2\Big) \\
&\qquad + \langle (L_2+E)v,v\rangle - \langle Ev,v\rangle + \langle (R_1+J)v,v\rangle - \langle J_0 v,v\rangle - \Big\langle \sum_{i=1}^n S_i\Theta_i \widehat{\nabla} v, v\Big\rangle \\
&\quad \le \frac{1}{2}\|\eta_1\Lambda Lu\|_{L^2}^2 + \frac{1}{2}\|Ru\|_{L^2}^2 + C_\delta\|v\|_{L^2}^2 + \delta\sum_j \|X_j f\|_{L^2}^2 + 4\delta\|\sqrt{a}\,\widehat{\nabla} v\|_{L^2}^2 \\
&\qquad + C_\delta \|\mathrm{Op}(p)v\|_{H^1}^2 + C\Big(\sqrt{\sum_j \|X_j v\|_{L^2}^2}\,\|v\|_{L^2} + \|v\|_{L^2}^2\Big).
\end{aligned}
$$

Combining this with the inequality

$$\sqrt{\sum_j \|X_j v\|_{L^2}^2}\,\|v\|_{L^2} \leq \delta \sum_j \|X_j v\|_{L^2}^2 + C_\delta \|v\|_{L^2}^2$$

and the condition $\mathbf{Q}_p \approx a\mathbb{I}_{n-p+1}$, we obtain, choosing δ smaller if necessary,

$$\sum_j \|X_j v\|_{L^2}^2 + \|\sqrt{a}\,\widehat{\nabla} v\|^2 \leq \frac{1}{2}\|\eta_1 \Lambda L u\|_{L^2}^2 + \frac{1}{2}\|Ru\|_{L^2}^2 + C_\delta \|v\|_{L^2}^2 + C_\delta \|\mathrm{Op}(p)v\|_{H^1}^2 \\ + \delta \sum_j \|X_j v\|_{L^2}^2 + \delta \|\sqrt{a}\,\widehat{\nabla} v\|_{L^2}^2.$$

Absorbing the terms $\delta \sum_j \|X_j v\|_{L^2}^2$ and $\delta \|\sqrt{a}\,\widehat{\nabla} v\|_{L^2}^2$ into the left-hand side, and then using that the order of the error term R is $-\varepsilon$, we obtain

$$\sum_j \|X_j v\|_{L^2}^2 + \|\sqrt{a}\,\widehat{\nabla} v\|_{L^2}^2 \leq \|\eta_1 \Lambda L u\|_{L^2(\mathbb{R}^n)}^2 + C\|v\|_{L^2}^2 + C\|u\|_{H^{-\varepsilon}}^2, \tag{3.13}$$

where the term involving the H^1 norm of $\mathrm{Op}(p)\Lambda u$ may be absorbed into $\|u\|_{H^{-\varepsilon}}^2$, since Λ may be made to be regularizing of arbitrary high order in a conic neighbourhood of the symbol p, by choosing N_0 to be sufficiently large.

Next we write

$$\begin{aligned} \|v\|_{L^2}^2 &= \int_{\{\xi\in\mathbb{R}^n:|\xi|\leq N\}} |\widehat{v}(\xi)|^2\,d\xi + \int_{\{\xi\in\mathbb{R}^n:|\xi|>N\}} |\widehat{v}(\xi)|^2\,d\xi \\ &\leq N^{2\gamma}\int_{\{\xi\in\mathbb{R}^n:|\xi|\leq N\}} \langle\xi\rangle^{-2\gamma}|\widehat{v}(\xi)|^2\,d\xi + \frac{1}{w^2(N)}\int_{\{\xi\in\mathbb{R}^n:|\xi|>N\}} w^2(\langle\xi\rangle)|\widehat{v}(\xi)|^2\,d\xi \\ &\leq N^{2\gamma}\|u\|_{H^0}^2 + \frac{1}{w^2(N)}\|w(\langle\xi\rangle)\,\widehat{v}(\xi)\|_{L^2}^2 \\ &\leq N^{2\gamma}\|u\|_{H^0}^2 + \frac{C}{w^2(N)}\Big(\sum_j \|X_j v\|_{L^2}^2 + \|\sqrt{a}\,\widehat{\nabla} v\|_{L^2}^2 + \|v\|_{L^2}^2\Big), \end{aligned}$$

where for the last inequality we used (2.7). Let $\delta = C/w^2(N)$ and note that δ can be made arbitrarily small by choosing N sufficiently large. We combine the above equality with (3.13) to obtain

$$\begin{aligned} \|v\|_{L^2}^2 &\leq C_\delta \|u\|_{H^0}^2 + \delta\Big(\sum_j \|X_j v\|_{L^2}^2 + \|\sqrt{a}\,\widehat{\nabla} v\|_{L^2}^2 + \|v\|_{L^2}^2\Big) \\ &\leq C_\delta \|u\|_{H^0}^2 + \delta(\|\eta_1 \Lambda L u\|_{L^2(\mathbb{R}^n)}^2 + C\|v\|_{L^2}^2 + C\|u\|_{H^{-\varepsilon}}^2). \end{aligned}$$

Choosing δ sufficiently small to absorb the norm $\|v\|_{L^2}^2$ to the left-hand side, we conclude

$$\|\eta_1 \Lambda u\|_{L^2(\mathbb{R}^n)}^2 = \|v\|_{L^2(\mathbb{R}^n)}^2 \leq C_\gamma \|\eta_1 \Lambda L u\|_{L^2(\mathbb{R}^n)}^2 + C_\gamma \|u\|_{H^0(\mathbb{R}^n)}^2,$$

for a constant C_γ depending on γ. ∎

3.2.1. Removal of the smoothness assumption. It remains to remove the smoothness assumption $u \in C^{2,\delta}(U)$ in Lemma 3.17, and to convert the above *a priori* estimate (3.12) to the desired conclusion $\Lambda u \in H^0$ of Theorem 2.6. For this we fix a strictly positive smooth function $r \in C^\infty(\mathbb{R}^n)$ such that

$$r(\xi) \equiv \begin{cases} |\xi|^{-1} & \text{for } |\xi| \geq 2, \\ 1 & \text{for } |\xi| \leq 1, \end{cases}$$

and we fix a large exponent q. For $\varepsilon > 0$ small, define a mollified symbol

$$\lambda_\varepsilon(x,\xi) = r_\varepsilon(\xi) \cdot \lambda(x,\xi) = r(\varepsilon\xi)^q \cdot \lambda(x,\xi), \quad \text{where } r_\varepsilon(\xi) \equiv r(\varepsilon\xi)^q,$$

with $\lambda(x,\xi) = |\xi|^\gamma e^{-N_0(\log|\xi|)\phi(x,\xi)}$ for $|\xi| \geq e$ as in (3.3). Let $\Lambda_\varepsilon = \mathrm{Op}\lambda_\varepsilon$. The symbols $r_\varepsilon(\xi)$ satisfy

$$\frac{|\partial_\xi^\alpha r_\varepsilon|}{r_\varepsilon} \leq C_{\alpha,q}|\xi|^{-|\alpha|} \quad \text{uniformly in } \varepsilon > 0 \text{ and } \xi \in \mathbb{R}^n. \tag{3.14}$$

If q is chosen sufficiently large relative to the order of the distribution u, then $\Lambda_\varepsilon u \in C^2$ for all $\varepsilon > 0$, and since Λ_ε is elliptic of order γ in a conic neighbourhood of (x_0,ξ_0), it suffices to show that the L^2 norm of $\eta_1\Lambda_\varepsilon u$ remains uniformly bounded as $\varepsilon \searrow 0$. However, Lemma 3.17 fails to apply since we do not know that the distribution u is a function in $C^{2,\delta}(U)$, and we now work to circumvent this difficulty.

The parameter N_0 in (3.3) can be chosen sufficiently large so that $\eta_1\Lambda L u \in L^2$ because ϕ is strictly positive in a conic neighbourhood of the H^γ wave front set of u, and hence Λ is regularizing there of order at least $\gamma - \sigma N_0$ for some constant $\sigma > 0$. The L^2 norm of $\eta_1\Lambda_\varepsilon L u$ is bounded uniformly in $\varepsilon > 0$ and tends to the L^2 norm of $\eta_1\Lambda L u$.

As in the proof of Lemma 3.17, we have, for each $\varepsilon > 0$, an operator G_ε and the identities

$$\begin{aligned} (L_1 + G_\varepsilon)\eta_1\Lambda_\varepsilon u &= \eta_1\Lambda_\varepsilon L_1 u + R_\varepsilon u, \\ (L_2 + E_\varepsilon)\eta_1\Lambda_\varepsilon u &= \eta_1\Lambda_\varepsilon L_2 u + R_\varepsilon u, \\ (R_1 + J_\varepsilon)\eta_1\Lambda_\varepsilon u &= \eta_1\Lambda_\varepsilon R_1 u + R_\varepsilon u, \end{aligned}$$

with both sides of the equation in C^2 for each $\varepsilon > 0$. Moreover, the differential inequalities (3.14) ensure that the proof of Lemma 3.17 carries through for each $\varepsilon > 0$ with Λ replaced by Λ_ε, so that G_ε takes the form (3.6), i.e.,

$$G_\varepsilon = \sum_j B_{j,\varepsilon} \circ X_j + \sum_j X_j^{\mathrm{tr}} \circ \widetilde{B}_{j,\varepsilon} + B_{0,\varepsilon},$$

$$B_{0,\varepsilon} \in \mathrm{Op}(\mathscr{C}^{0,\delta} S^{0,2}_{1,\eta}) \quad \text{and} \quad B_{j,\varepsilon}, \widetilde{B}_{j,\varepsilon} \in \mathrm{Op}(\mathscr{C}^{1,\delta} S^{0,1}_{1,\eta}),$$

where the pseudodifferential operator coefficients $B_{0,\varepsilon}$, $B_{j,\varepsilon}$ and $\widetilde{B}_{j,\varepsilon}$ lie uniformly in the indicated operator classes. A similar argument holds for E_ε and J_ε. All functions have sufficient differentiability for the proof of Lemma 3.17 to apply, and this proof, together with the above identity, yield

$$\|\eta_1\Lambda_\varepsilon u\|_{L^2(R)} \leq C\|\eta_1\Lambda_\varepsilon L u\|_{L^2(R)} + C\|u\|_{H^0(R)} \quad \text{uniformly in } \varepsilon > 0.$$

We conclude, as desired, that the L^2 norm of $\eta_1\Lambda_\varepsilon u$ remains bounded as $\varepsilon \searrow 0$.

Thus, we have proved that for any distribution $u \in \mathcal{D}'(V)$, and any $0 < \gamma < \delta$, there is a symbol Λ as in (3.3) that is elliptic of order γ on the conical set Γ, and satisfies

$$\|\eta_1 \Lambda u\|_{L^2(R)} \le C\|\eta_1 \Lambda L u\|_{L^2(R)} + C\|u\|_{H^0(R)}.$$

The proof of Theorem 2.6 is now complete.

Combined with the bootstrapping argument above, this shows that $u \in H^s_{\mathrm{loc}}(R)$ for all $s \in \mathbb{R}$. Indeed, $\eta_2 u \in H^{-M}(R)$ for some M sufficiently large, and thus we can begin the bootstrapping argument at $s = -M$.

4. Proof of Theorem 2.5

We now prove Theorem 2.5. The first step is to use a bootstrapping argument to reduce matters to the level of $L^2(\mathbb{R}^n)$. Consider the general second order divergence form operator

$$Lu(x) \equiv \nabla^{\mathrm{tr}} A(x) \nabla u(x) + D(x)u(x),$$

where A and D are real and smooth, and where $A(x)$ satisfies appropriate form comparability conditions. In order to conclude hypoellipticity of L, it is enough to show that there is $\gamma > 0$ such that for every $s \in \mathbb{R}$, we have the bootstrapping argument

$$u \in H^s_{\mathrm{loc}}(\mathbb{R}^n) \text{ and } Lu \in H^{s+\gamma}_{\mathrm{loc}}(\mathbb{R}^n) \implies u \in H^{s+\gamma}_{\mathrm{loc}}(\mathbb{R}^n) \quad \text{for all } s \in \mathbb{R}.$$

Now, with $\widehat{\Lambda}_s(\xi) \equiv (1+|\xi|^2)^{s/2}$, and $\gamma > 0$ fixed, it suffices to show

$$u \in H^0_{\mathrm{loc}}(\mathbb{R}^n) \text{ and } \Lambda_s L \Lambda_{-s} u \in H^{\gamma}_{\mathrm{loc}}(\mathbb{R}^n) \implies u \in H^{\gamma}_{\mathrm{loc}}(\mathbb{R}^n) \quad \text{for all } s \in \mathbb{R}.$$

The second step is to use the sum of squares assumption in the second paragraph of Theorem 2.5 to show that it is sufficient to establish the conditions of Theorem 2.6. So define

$$\widetilde{G} \equiv [\Lambda_s, L]\Lambda_{-s} = \Lambda_s L \Lambda_{-s} - L, \tag{4.1}$$

and suppose for the moment that the operator L has the simple form

$$L = \sum_j X_j^{\mathrm{tr}} X_j, \tag{4.2}$$

where $L \in S^2_{1,0}$ is smooth and $X_j \in C^{2,\delta}$. We first establish the properties of $\tilde{G}$ we need using the rough version of asymptotic expansion from [25] given in Theorem 3.7 above, which we repeat here for the reader's convenience.

Suppose $\sigma \in \mathcal{C}^{\nu} S^{m_1}_{1,\delta_1}$ and $\tau \in \mathcal{C}^{M+\mu+\nu} S^{m_2}_{1,\delta_2}$, where M is a nonnegative integer, and $0 < \mu, \delta_1, \delta_2 < 1$, $\nu > 0$ and $M + \mu \ge m_1 \ge 0$. Let $\delta \equiv \max\{\delta_1, \delta_2\}$. Then

$$\sigma \circ \tau = \sum_{\ell=0}^{M} \frac{1}{i^{\ell}\ell!} \nabla_\xi^{\ell}\sigma \cdot \nabla_x^{\ell}\tau + E, \quad E \in \mathcal{O}^{m_1+m_2+(M+\mu)(\delta_2-1)+\varepsilon}_{(-(1-\delta)\nu,\nu)} \text{ for every } \varepsilon > 0.$$

Lemma 4.1. *Let L and $\widetilde{G}$ be as in (4.2) and (4.1). Then*

$$\widetilde{G} = \sum_j B_j \circ X_j + \sum_j X_j^{\mathrm{tr}} \circ \widetilde{B}_j + B_0,$$

where $B_0 \in \mathcal{O}^{-\delta/2+\varepsilon}_{(-\delta/2,\delta/2)}$ for every $\varepsilon > 0$ and $B_j, \widetilde{B}_j \in \mathrm{Op}(C^{1,\delta} S^0_{1,0})$.

Proof. First, we note that

$$[\Lambda_s, L] = \sum_j [\Lambda_s, X_j^{\mathrm{tr}}] X_j + X_j^{\mathrm{tr}} [\Lambda_s, X_j],$$

and so we investigate operators $[\Lambda_s, X_j^{\mathrm{tr}}]$ and $[\Lambda_s, X_j]$. The analysis is similar, so we only give details for $[\Lambda_s, X_j]$. Using Theorem (3.7) with $m_1 = s, m_2 = 1, M = 1, \mu = 1 + \delta/2$, $\nu = \delta/2$ and $\delta_1 = \delta_2 = 0$, we have

$$\sigma([\Lambda_s, X_j]) = C \nabla_\xi (1 + |\xi|^2)^{s/2} \cdot \nabla_x \sigma(X_j) + E, \quad \text{where } E \in \mathcal{O}^{1+s-(2+\delta/2)+\varepsilon}_{(-\delta/2,\delta/2)}.$$

Composing with Λ_{-s} and using $\mathrm{Op}(\nabla_\xi (1 + |\xi|^2)^{s/2}) = R^{-1} \circ \Lambda_s$, where $R^{-1} \in S^{-1}_{1,0}$, we obtain

$$X_j^{\mathrm{tr}} [\Lambda_s, X_j] \Lambda_{-s} = X_j^{\mathrm{tr}} \circ \widetilde{B}_j + R,$$

with $\widetilde{B}_j \in C^{1,\delta} S^0_{1,0}$ and $R \in \mathcal{O}^{-\delta/2+\varepsilon}_{(-\delta/2,\delta/2)}$. ∎

Now we start with an operator $L \in S^2_{1,0}$ of the more general form

$$L = \sum_j X_j^{\mathrm{tr}} X_j + A_0 + \widehat{\nabla}^{\mathrm{tr}} \cdot \mathbf{Q}_p(x) \widehat{\nabla}, \tag{4.3}$$

where $X_j \in C^{2,\delta}$ and $A_0 \in S^1_{1,0}$. Using Lemma 4.1 for any operator L in the form (4.3) and Remark 2.7, we can show that the operator $\Lambda_s L \Lambda_{-s}$ has the form

$$\Lambda_s L \Lambda_{-s} = \sum_j X_j^{\mathrm{tr}} X_j + \sum_j B_j X_j + \sum_j X_j^{\mathrm{tr}} \tilde{B}_j + B_0 + R_1 + \widehat{\nabla}^{\mathrm{tr}} \mathbf{Q}_p(x) \widehat{\nabla},$$

where $X_j, B_j, \tilde{B}_j$, and B_0 are as in Lemma 4.1 and R_1 is as in Theorem 2.6. Thus, to show hypoellipticity of the operator (4.3), it is sufficient to show that it satisfies the hypotheses of Theorem 2.6, which completes the second step of the proof.

We prepare for the final step of the proof with an auxiliary lemma (see Lemma 5.1 in [7]), and its corollary to be used later for showing condition (2.7).

Lemma 4.2. *Let $\varphi \in C_0^2(\mathbb{R}^n)$, $f \in C^\infty(\mathbb{R}^n)$ simply positive, and $s > 0$. Then for any $l \in \{1, \ldots, n\}$, there exists a constant C_l independent of s such that*

$$\|\varphi\|^2 \le C_l \Big(\frac{1}{\tau^2 [\min_{|x| \ge s} f(x)]^2} + s^2 \Big) \Big(\|\partial_{x_l} \varphi\|^2 + \int \tau^2 f(x)^2 \varphi(x)^2 \, dx \Big), \tag{4.4}$$

where the minimum is taken over all $x \in \operatorname{supp} \varphi$ such that $|x| \ge s$.

Proof. Fix $s > 0$. For any $x \in \mathbb{R}^n$, we have

$$\varphi(x) = \varphi\Big(x + s\frac{x_l}{|x_l|}\Big) - \int_1^{1+s/|x_l|} \frac{\partial \varphi}{\partial t}(x_1, \dots, x_{l-1}, tx_l, x_{l+1}, \dots, x_n)\, dt,$$

$$\varphi^2(x) \lesssim \varphi^2\Big(x + s\frac{x_l}{|x_l|}\Big) + \Big(\int_1^{1+s/|x_l|} \nabla\varphi(x_1, \dots, x_{l-1}, tx_l, x_{l+1}, \dots, x_n)\cdot(0, \dots, x_l, 0, \dots, 0)\, dt\Big)^2,$$

and thus

$$\begin{aligned}\int_{|x_l|\le s} \varphi^2(x)\, dx &\lesssim \int_{|x_l|\le s} \varphi^2\Big(x + s\frac{x_l}{|x_l|}\Big)\, dx \\ &+ \int_{|x_l|\le s}\Big(\int_1^{1+s/|x_l|} |\partial_l\varphi(x_1, \dots, x_{l-1}, tx_l, x_{l+1}, \dots, x_n) t^{3/4} x_l|^2\, dt \int_1^{1+s/|x_l|} t^{-3/2}\, dt\Big)\, dx \\ &\lesssim \int_{s\le|x_l|\le 2s} \varphi^2(x)\, dx \\ &+ \int_{|x_l|\le s}\int_1^{1+s/|x_l|} |\partial_l\varphi(x_1, \dots, x_{l-1}, tx_l, x_{l+1}, \dots, x_n) t^{3/4} x_l|^2\, dt\, dx.\end{aligned}$$

Switching the order of integration in the last term on the right-hand side and making a change of variables $y = (x_1, \dots, x_{l-1}, tx_l, x_{l+1}, \dots, x_n)$, we obtain

$$\begin{aligned}&\int_{|x_l|\le s}\int_1^{1+s/|x_l|} |\partial_l\varphi(x_1, \dots, x_{l-1}, tx_l, x_{l+1}, \dots, x_n) t^{3/4} x_l|^2\, dt\, dx \\ &\quad\le \int_1^\infty \int_{|y_l|\le 2s} |\partial_l\varphi(y) y_l|^2 t^{-1/2}\frac{dy}{t}\, dt \lesssim s^2\int |\partial_l\varphi(y)|^2\, dy,\end{aligned}$$

which combining with the above gives

$$\int_{|x_l|\le s} \varphi^2(x)\, dx \lesssim \int_{s\le|x|\le 2s} \varphi^2(x)\, dx + s^2\int |\partial_l\varphi(x)|^2\, dx.$$

Finally,

$$\int_{|x_l|\ge s} \tau^2 f(x)^2\varphi(x)^2\, dx \ge \tau^2\big[\min_{|x_l|\ge s} f(x)\big]^2 \int_{|x_l|\ge s} \varphi^2(x)\, dx,$$

and thus altogether

$$\int \varphi^2(x)\, dx \lesssim \frac{1}{\tau^2[\min_{|x_l|\ge s} f(x)]^2}\int_{|x|\ge s} \tau^2 f(x)^2\varphi(x)^2\, dx + s^2\int |\partial_l\varphi(x)|^2\, dx,$$

which implies (4.4). ∎

Lemma 4.3. *Let φ and f as in Lemma 4.2. There exists a strictly positive continuous function w, satisfying $w(\tau) \to \infty$ as $\tau \to \infty$, such that for every $l \in \{1, \dots, n\}$ and some constant $C_l > 0$,*

$$\int w(\tau)^2\varphi(x)^2\, dx \le C_l \int \big(|\partial_l\varphi(x)|^2 + \tau^2 f(x)^2\varphi(x)^2\big)\, dx.$$

Proof. For all $s \geq 0$, define

$$f_0(s) \equiv \min_{x \in \operatorname{supp}\varphi : |x| \geq s} f(x),$$

and note that $f_0(0) = 0$, $f_0(s) > 0$ for $s \neq 0$, and f_0 is nondecreasing on $[0, \infty)$. Let $r = r(\tau) > 0$ be the unique point satisfying

$$\frac{1}{r} = \tau f_0(r). \tag{4.5}$$

Define the function w by

$$w(\tau) = \inf_{0<s<\infty} \Big(\frac{1}{s} + \tau f_0(s) \Big),$$

since $1/s$ is nonincreasing and $f_0(s)$ is nondecreasing in s, we have $w(\tau) \approx 1/r$, where r is given by (4.5). Therefore, $w(\tau) \to \infty$ as $\tau \to \infty$ and using (4.4) with $s = r$, we obtain

$$\begin{aligned} \int w(\tau)^2 \varphi(x)^2 \, dx &\leq C_l \frac{1}{r^2} \Big(\frac{1}{\tau^2 f_0(r)^2} + r^2 \Big) \int \big(|\partial_l \varphi(x)|^2 + \tau^2 f(x)^2 \varphi(x)^2 \big) \, dx \\ &\leq C_l \int (|\partial_l \varphi(x)|^2 + \tau^2 f(x)^2 \varphi(x)^2) \, dx. \end{aligned}$$ ∎

4.1. Sufficiency

We can now proceed to complete the sufficiency part of Theorem 2.5. We note that, without loss of generality, we may assume that the diagonal entries $\lambda_k(\tilde{x})$ are smooth. Indeed, from $A(x) \sim D_{\boldsymbol{\lambda}}(x)$, we obtain $A(x) \sim A_{\mathrm{diag}}(x)$ and hence

$$\lambda_k(\tilde{x}) \approx a_{k,k}(x) \approx a_{k,k}(\tilde{x}, 0, 0), \tag{4.6}$$

where the functions $a_{k,k}(\tilde{x}, 0, 0)$ are smooth for $1 \leq k \leq n$ by assumption.

Proof of sufficiency in Theorem 2.5. Let $(\xi_1, \dots, \xi_m, \eta_{m+1}, \dots, \eta_n)$ denote the dual variables, and denote $\xi = (\xi_1, \dots, \xi_m)$, $\eta = (\eta_{m+1}, \dots, \eta_n)$, $\tilde{x} = (x_1, \dots, x_m)$. Define

$$R = \{(x, \xi, \eta) : x = 0, \xi = 0, \ \eta_{m+1}, \dots, \eta_n > 0\}.$$

The principal symbol of L vanishes on the manifold $\tilde{x} = \xi = 0$, so it suffices to prove that $Lu \in H^s(\mathfrak{N}(R)) \Rightarrow u \in H^s(\mathfrak{N}(R))$ for some conical neighbourhood $\mathfrak{N}(R)$ of the ray R. We start with verifying condition (2.7). Let $\mathcal{F}(u)(\tilde{x}, \eta)$ be the partial Fourier transform of u in $n-m$ variables η. Then, from Lemma 4.3 with $x = \tilde{x}$ and $\varphi(\tilde{x}) = \mathcal{F}(u)(\tilde{x}, \eta)$, we have, for $k = m+1, \dots, p-1$,

$$\int w(\eta_k)^2 \mathcal{F}(u)(\tilde{x}, \eta)^2 \, d\tilde{x} \leq C \int (|\nabla_{\tilde{x}} \mathcal{F}(u)(\tilde{x}, \eta)|^2 + \eta_k^2 \lambda_k(\tilde{x}) \mathcal{F}(u)(\tilde{x}, \eta)^2) \, d\tilde{x},$$

and, for $k = p, \dots, n$,

$$\int w(\eta_k)^2 \mathcal{F}(u)(\tilde{x}, \eta)^2 \, d\tilde{x} \leq C \int (|\nabla_{\tilde{x}} \mathcal{F}(u)(\tilde{x}, \eta)|^2 + \eta_k^2 \lambda_p(\tilde{x}) \mathcal{F}(u)(\tilde{x}, \eta)^2) \, d\tilde{x},$$

where $w(s) \to \infty$ as $s \to \infty$. Adding the inequalities together gives

$$\int w(|\eta|)^2 \mathcal{F}(u)(\tilde{x},\eta)^2 \, d\tilde{x} \\ \leq C \int \Big(|\nabla_{\tilde{x}} \mathcal{F}(u)(\tilde{x},\eta)|^2 + \Big[\sum_{k=m+1}^{p-1} \eta_k^2 \lambda_k(\tilde{x}) + \sum_{k=p}^{n} \eta_k^2 \lambda_p(\tilde{x}) \Big] \mathcal{F}(u)(\tilde{x},\eta)^2 \Big) \, d\tilde{x},$$

where $w(|\eta|) \to \infty$ as $|\eta| \to \infty$. Combining with the first line in (2.4), we obtain

$$\int_{\mathbb{R}^n} w(|(\eta)|)^2 \mathcal{F}(u)(\tilde{x},\eta)^2 \, d\tilde{x} \, d\eta \leq C \sum_j \|X_j u\|^2 + C \|\sqrt{\lambda_p}\, \widehat{\nabla} u\|^2,$$

which gives, upon using the first condition in (2.4) again,

$$\int \min\{|(\xi,\eta)|, w(|(\xi,\eta)|)\}^2 \hat{u}(\xi,\eta)^2 \, d\xi \, d\eta \\ \lesssim \int_{|\xi| \leq |\eta|} w(|\eta|)^2 \, |\hat{u}(\xi,\eta)|^2 \, d\xi \, d\eta + \int_{|\xi| \geq |\eta|} |\xi|^2 \, |\hat{u}(\xi,\eta)|^2 \, d\xi \, d\eta \\ \leq C \sum_j \|X_j u\|^2 + C \|\sqrt{\lambda_p}\, \widehat{\nabla} u\|^2.$$

We proceed to verify (2.8) with $p \equiv 0$ and ψ constructed below. Since the principal symbol of the operator vanishes on $\mathbb{R}^{n-m} \times \mathbb{R}^{n-m}$, namely, when $\tilde{x} = \xi = 0$, we need to localize matters with a cutoff function ψ that enjoys favorable commutation relations with the symbol $\sigma(X_j)$ of the vector field X_j. So, let $p \equiv 0$ and let $\rho > 0$. Let $\psi \in C^\infty(T^*V)$ be homogeneous of degree 0 with respect to (ξ,η) and satisfy

$$\begin{cases} \psi = 1 & \text{if } |(x, \xi/|\eta|)| \geq 3\rho, \\ \psi = 0 & \text{if } |(x, \xi/|\eta|)| \leq \rho, \\ \psi = \psi(x_{m+1},\ldots,x_n) & \text{if } |(\tilde{x}, \xi/|\eta|)| \leq 2\rho. \end{cases}$$

For example, the reader can easily check that for ρ sufficiently small, we can take

$$\psi(x,\xi,\eta) = \widetilde{\psi}(x,\zeta), \quad \text{where } \widetilde{\psi}(x,\zeta) \equiv \begin{cases} \varphi_\varepsilon(x,\zeta) & \text{if } |(x,\zeta)| < 4\rho, \\ 1 & \text{if } |(x,\zeta)| \geq 4\rho, \end{cases}$$

for any $0 < \varepsilon < \rho/2$, where

$$\varphi_\varepsilon(x,\zeta) = \phi_\varepsilon * \mathbf{1}_{\mathbb{R}^{2n} \setminus K_\varepsilon}(x,\zeta), \quad (x,\zeta) \in \mathbb{R}^n \times \mathbb{R}^m,$$

and

$$\phi_\varepsilon(x,\zeta) = \phi\Big(\frac{|(x,\zeta)|}{\varepsilon}\Big), \quad \text{with } \phi(t) = \begin{cases} 1 & \text{if } 0 \leq t < 1/2, \\ 0 & \text{if } 1 \leq t < \infty, \end{cases}$$

and where

$$K_\varepsilon \equiv \{(x,\zeta) : |(\tilde{x},\zeta)| < \rho + \varepsilon \text{ and } |(x,\zeta)| < 2\rho + \varepsilon\}$$

is the intersection of a cylinder of radius $\rho + \varepsilon$ and a ball of radius $2\rho + \varepsilon$.

Thus, $\widetilde{\psi}$ is 1 outside a ball of radius 3ρ, vanishes inside a ball of radius ρ, and within the cylinder, it makes the transition from 0 to 1 while depending *only* on the variables $\tilde{x}$, which will give rise to favourable estimates on commutators below. On the other hand, outside both the region where $|(\xi,\eta)|\le 1$ (the operator L is infinitely smoothing when $|(\xi,\eta)|\le 1$) and the cylinder, the symbol of L is bounded away from 0, hence elliptic.[7]

The main step of Christ's application of his theorem occurs now. We begin by letting $X_j=\sum_{\ell=1}^{n} a_\ell^j(\tilde{x})\,\partial_{x_l}$ for each $j=1,\dots,N$. We now restrict attention to the cylinder $|(\tilde{x},\xi/|\eta|)|\le 2\rho$. Then, since ψ is independent of $\tilde{x}$ in this cylinder, we have

$$\{\psi,\sigma(X_j)\}=i\sum_{\ell=m+1}^{n} a_\ell^j(\tilde{x})\,\partial_{x_l}\psi,$$

with

$$\begin{aligned}|a_\ell^j(\tilde{x})|&\lesssim\sqrt{\lambda_\ell(\tilde{x})},\quad \ell=m+1,\dots,p-1,\\ |a_\ell^j(\tilde{x})|&\lesssim\sqrt{\lambda_p(\tilde{x})},\quad \ell=p,\dots,n,\end{aligned}$$

using conditions (2.4), and

$$\{\psi,\eta\}=i\hat{\nabla}\psi.$$

Using that $|\xi|\le 2\rho|\eta|$ in the cylinder, we have, for each $j=1,\dots,N$,

$$\begin{aligned}&\|\mathrm{Op}[\log\langle(\xi,\eta)\rangle\{\psi,\sigma(X_j)\}]u\|^2\\ &\quad\lesssim\sum_{\ell=m+1}^{p-1}\|\mathrm{Op}[\sqrt{\lambda_\ell(\tilde{x})}\log\langle\eta\rangle]u\|^2+\|\mathrm{Op}[\sqrt{\lambda_p(\tilde{x})}\log\langle\eta\rangle]u\|^2\\ &\quad=\int\Lambda_{\mathrm{sum}}(\tilde{x})\log\langle\eta\rangle^2\mathcal{F}(u)(\tilde{x},\eta)^2\,d\tilde{x}\,d\eta\end{aligned}$$

and

$$\begin{aligned}\|\sqrt{\mathbf{Q}_p}\,\mathrm{Op}[\log\langle(\xi,\eta)\rangle\{\psi,\eta\}]u\|^2&\lesssim\|\sqrt{\lambda_p}\,\mathrm{Op}[\log\langle\eta\rangle]u\|^2\\ &\lesssim\int\Lambda_{\mathrm{sum}}(\tilde{x})\log\langle\eta\rangle^2\mathcal{F}(u)(\tilde{x},\eta)^2\,d\tilde{x}\,d\eta,\end{aligned}$$

upon using the definition of $\Lambda_{\mathrm{sum}}(\tilde{x})$. To show (2.8) it is therefore sufficient to establish the first inequality in the following display (since the second follows directly from (2.4)):

$$\begin{aligned}(4.7)\quad&\int\log\langle\eta\rangle^2\Lambda_{\mathrm{sum}}(\tilde{x})\mathcal{F}(u)(\tilde{x},\eta)^2\,d\tilde{x}\,d\eta\\ &\lesssim\delta\int|\nabla_{\tilde{x}}\mathcal{F}(u)(\tilde{x},\eta,\tau)|^2\,d\tilde{x}\,d\eta\,d\tau\\ &\quad+\delta\int\Big[\sum_{k=m+1}^{p-1}\eta_k^2\lambda_k(\tilde{x})+\sum_{k=p}^{n}\eta_k^2\lambda_p(\tilde{x})\Big]\mathcal{F}(u)(\tilde{x},\eta)^2\,d\tilde{x}\,d\eta+C_\delta\|u\|^2\\ &\lesssim\delta\sum_{j=1}^{N}\|X_ju\|^2+\delta\|\sqrt{\lambda_p}\,\hat{\nabla}u\|^2+C_\delta\|u\|^2.\end{aligned}$$

[7] We identify regions in (x,ζ) with the corresponding regions in (x,ξ,η) under the map $\zeta=(\xi,\eta)$.

Using the definitions of $\Lambda_{\text{sum}}(\tilde{x})$ and $\Lambda_{\text{product}}(\tilde{x})$, we conclude that it is sufficient to show

$$(\log\tau)^2\|\sqrt{\Lambda_{\text{sum}}}\,\varphi\|^2 \le \delta(\tau)\|\nabla_{\tilde{x}}\varphi\|^2 + \delta(\tau)\tau^2\|\sqrt{\Lambda_{\text{product}}}\,\varphi\|^2 \tag{4.8}$$

for all $\varphi \in C_0^1(\mathbb{R}^m)$, where $\delta(\tau) \to 0$ as $\tau \to \infty$. Indeed, (4.8) together with the bound $0 \le \lambda_j \le 1$ implies

$$\begin{aligned}\int \log\langle\eta\rangle^2\Lambda_{\text{sum}}\varphi(\tilde{x})^2\,d\tilde{x} &\le \delta(\langle\eta\rangle)\|\nabla_{\tilde{x}}\varphi\|^2 + \delta(\langle\eta\rangle)\langle\eta\rangle^2\|\sqrt{\Lambda_{\text{product}}}\,\varphi\|^2 \\ &\le \delta(\langle\eta\rangle)\|\nabla_{\tilde{x}}\varphi\|^2 + \delta(\langle\eta\rangle)\Big[\sum_{k=m+1}^{p-1}|\eta_k|^2\|\sqrt{\lambda_k}\,\varphi\|^2 + \sum_{k=p}^{n}|\eta_k|^2\|\sqrt{\lambda_p}\,\varphi\|^2\Big] + C_\delta\|\varphi\|^2.\end{aligned}$$

This implies (4.7) by splitting the region of integration into $|\eta|$ sufficiently large so that $\delta(\langle\eta\rangle) \le \delta$, and the region where $|\eta|$ is bounded, and thus the left-hand side of (4.7) is bounded by $C\|u\|^2$.

To establish (4.8), we first recall for convenience the Koike condition:

$$\lim_{\tilde{x}\to 0}\mu(|\tilde{x}|, \sqrt{\Lambda_{\text{sum}}})\ln\Lambda_{\text{product}}(\tilde{x}) = 0. \tag{4.9}$$

Now let $\phi \in C_0^1(B(0,r))$. Then, with $\phi_{\tilde{y}}(\rho) \equiv \phi(\rho\tilde{y})$, we have

$$\begin{aligned}\int_{|\tilde{x}|\le r}\Lambda_{\text{sum}}(\tilde{x})\phi(\tilde{x})^2\,d\tilde{x} &= \int_{|\tilde{x}|\le r}\Lambda_{\text{sum}}(\tilde{x})(r-|\tilde{x}|)^2\frac{\phi(\tilde{x})^2}{(r-|\tilde{x}|)^2}\,d\tilde{x} \\ &\le \mu(r,\sqrt{\Lambda_{\text{sum}}})^2\int_{|\tilde{x}|\le r}\frac{\phi(\tilde{x})^2}{(r-|\tilde{x}|)^2}\,d\tilde{x} \\ &= \mu(r,\sqrt{\Lambda_{\text{sum}}})^2\int_{\mathbb{S}^{m-1}}\left\{\int_0^r\left(\frac{1}{r-\rho}\int_\rho^r\phi_{\tilde{y}}'(\rho)\right)^2\rho^{m-1}\,d\rho\right\}d\tilde{y} \\ &\le \mu(r,\sqrt{\Lambda_{\text{sum}}})^2\int_{\mathbb{S}^{m-1}}\left\{4\int_0^r\phi_{\tilde{y}}'(\rho)^2\rho^{m-1}\,d\rho\right\}d\tilde{y} \\ &\le 4\mu(r,\sqrt{\Lambda_{\text{sum}}})^2\int|\nabla_{\tilde{x}}\phi(\tilde{x})|^2, \end{aligned} \tag{4.10}$$

where in the last line we have applied Hardy's inequality.

Fix $\varphi \in C_0^1(\mathbb{R}^m)$ as in (4.8). Let $\chi \in C_0^1(\mathbb{R}^1)$ satisfy $\chi(t) = 1$ for $|t| \le 1$ and $\chi(t) = 0$ for $|t| \ge 2$, and define the function

$$\nu(\tilde{x}) \equiv \chi(\tau\Lambda_{\text{product}}(\tilde{x})), \tag{4.11}$$

and the set

$$I(\tau) \equiv \{\tilde{x} \in \operatorname{supp}\varphi : \tau\Lambda_{\text{product}}(\tilde{x}) > 1\}.$$

We can write

$$\begin{aligned}\int\Lambda_{\text{sum}}(\tilde{x})\varphi(\tilde{x})^2\,d\tilde{x} \le{}& 2\int\Lambda_{\text{sum}}(\tilde{x})\nu(\tilde{x})^2\varphi(\tilde{x})^2\,d\tilde{x} \\ &+ 2\int\Lambda_{\text{sum}}(\tilde{x})(1-\nu(\tilde{x}))^2\varphi(\tilde{x})^2\,d\tilde{x}.\end{aligned} \tag{4.12}$$

To estimate the second integral, we notice that it vanishes outside the set $I(\tau)$ and thus

$$(4.13)\quad (\log\tau)^2 \int \Lambda_{\text{sum}}(\tilde{x})(1-\nu(\tilde{x}))^2\,\varphi(\tilde{x})^2\,d\tilde{x} \le (\log\tau)^2 \int_{I(\tau)} \Lambda_{\text{product}}(\tilde{x})\,\varphi(\tilde{x})^2\,d\tilde{x} = \delta(\tau)\tau^2 \int \Lambda_{\text{product}}(\tilde{x})\,\varphi(\tilde{x})^2\,d\tilde{x},$$

where $\delta(\tau) = (\log\tau)^2\tau^{-1} \to 0$ as $\tau \to \infty$.

To estimate the first integral on the right-hand side of (4.12), we define

$$r(\tau) \equiv \sup\{|\tilde{y}| : \tilde{y} \in \operatorname{supp}\varphi : \tau\Lambda_{\text{product}}(\tilde{y}) \le 2\}.$$

Since $\operatorname{supp}\varphi$ is compact, the supremum above is attained at some point $\tilde{z} \in \operatorname{supp}\varphi$, and moreover we have both

$$|\tilde{z}| = r \quad\text{and}\quad \tau = \frac{2}{\Lambda_{\text{product}}(\tilde{z})}.$$

Thus, $\ln\tau \approx \ln 1/\Lambda_{\text{product}}(\tilde{z})$ and so

$$\mu(r(\tau), \sqrt{\Lambda_{\text{sum}}})\ln r(\tau) \approx \mu(|\tilde{z}|, \sqrt{\Lambda_{\text{sum}}})\ln\frac{1}{\Lambda_{\text{product}}(\tilde{z})}.$$

The Koike condition (4.9) now implies

$$(4.14)\qquad \lim_{\tau\to\infty}\mu(r(\tau), \sqrt{\Lambda_{\text{sum}}})\ln r(\tau) = \lim_{\tilde{x}\to 0}\mu(|\tilde{x}|, \sqrt{\Lambda_{\text{sum}}})\ln\frac{1}{\Lambda_{\text{product}}(\tilde{x})} = 0,$$

since $r(\tau) \to 0$ as $\tau \to \infty$. We now need to combine this result with (4.10) to obtain the desired estimate. Let $\phi(\tilde{x}) = \nu(\tilde{x})\varphi(\tilde{x})$. Then, using the definition of $\nu(\tilde{x})$ in (4.11), we obtain

$$\begin{aligned}\int |\nabla_{\tilde{x}}\phi(\tilde{x})|^2\,d\tilde{x} &\le C\int |\nabla_{\tilde{x}}\nu(\tilde{x})|^2\,\varphi(\tilde{x})^2\,d\tilde{x} + C\int \nu(\tilde{x})^2\,|\nabla_{\tilde{x}}\varphi(\tilde{x})|^2\,d\tilde{x}\\ &\le C\tau^2\int_{I(\tau)} |\nabla_{\tilde{x}}\Lambda_{\text{product}}(\tilde{x})|^2\,\varphi(\tilde{x})^2\,d\tilde{x} + C\int |\nabla_{\tilde{x}}\varphi(\tilde{x})|^2\,d\tilde{x}\\ &\le C\tau^2\int_{I(\tau)} \Lambda_{\text{product}}(\tilde{x})\,\varphi(\tilde{x})^2\,d\tilde{x} + C\int |\nabla_{\tilde{x}}\varphi(\tilde{x})|^2\,d\tilde{x},\end{aligned}$$

where in the last inequality we used the Malgrange inequality, see, e.g., Lemme I in [11], applied to $\Lambda_{\text{product}}(\tilde{x}) = \prod_{k=m+1}^{p}\lambda_k(\tilde{x})$, where the functions λ_k are smooth by (4.6). Finally, from the definition of r and (4.11), it follows that

$$\operatorname{supp}\phi \subset \operatorname{supp}\nu \subset \left\{\tilde{y} : \tau < \frac{2}{\Lambda_{\text{product}}(\tilde{y})}\right\} \subset B(0, r(\tau)),$$

since if $|\tilde{x}| > r(\tau)$, then $\tau\Lambda_{\text{product}}(\tilde{y}) > 2$ by the definition of $r(\tau)$.

Combining the above estimate with (4.14) and (4.10), we conclude that

$$(\log\tau)^2\int\Lambda_{\mathrm{sum}}(\tilde x)\nu(\tilde x)^2\,\varphi(\tilde x)^2\,d\tilde x=(\log\tau)^2\int\Lambda_{\mathrm{sum}}(\tilde x)\,\phi(\tilde x)^2 d\tilde x$$
$$\leq\delta(\tau)\Big(\tau^2\int\Lambda_{\mathrm{product}}(\tilde x)\,\varphi(\tilde x)^2\,d\tilde x+\int|\nabla_{\tilde x}\,\varphi(\tilde x)|^2 d\tilde x\Big),$$

with

$$\delta(\tau)=C\mu(r,\sqrt{\Lambda_{\mathrm{sum}}})^2\,(\log\tau)^2\to 0\quad\text{as }\tau\to\infty.$$

Together with (4.13), this gives (4.8). ∎

4.2. Sharpness

We now turn to the sharpness portion of Theorem 2.5. If the Koike condition (4.9) fails, then

$$\begin{aligned}0&<\limsup_{\tilde x\to 0}\mu(|\tilde x|,\sqrt{\Lambda_{\mathrm{sum}}})\ln\frac{1}{\Lambda_{\mathrm{product}}(\tilde x)}\\&=\limsup_{\tilde x\to 0}\mu\Bigg(|\tilde x|,\sqrt{\sum_{k=m+1}^{p}\lambda_k(\tilde x)}\Bigg)\ln\prod_{k=m+1}^{p}\frac{1}{\lambda_k(\tilde x)}\\&\leq\limsup_{\tilde x\to 0}\mu\Bigg(|\tilde x|,\sum_{k=m+1}^{p}\sqrt{\lambda_k(\tilde x)}\Bigg)\sum_{j=m+1}^{p}\ln\frac{1}{\lambda_j(\tilde x)}\\&\leq\sum_{k,j=m+1}^{p}\limsup_{\tilde x\to 0}\mu(|\tilde x|,\sqrt{\lambda_k(\tilde x)})\ln\frac{1}{\lambda_j(\tilde x)}\end{aligned}$$

shows that $p>m+1$ (since $\limsup_{\tilde x\to 0}\mu(|\tilde x|,\sqrt{\lambda_p(\tilde x)})\ln 1/\lambda_p(\tilde x)=0$) and that there is a pair of distinct indices $k,j\in\{m+1,\dots,p\}$ such that

$$\limsup_{\tilde x\to 0}\mu(|\tilde x|,\sqrt{\lambda_k(\tilde x)})\ln\frac{1}{\lambda_j(\tilde x)}>0.$$

Our sharpness assertion in Theorem 2.5 now follows immediately from Proposition 4.5 and Theorem 4.6 below.

To prove these results, we will need the following lemma (see Lemma 2.7 in [14]), whose short proof we include here for the reader's convenience.

Lemma 4.4 ([14]). *Let L be a hypoelliptic operator on $\mathbb{R}^n$. For any multiindex β and any subsets Ω and Ω' of $\mathbb{R}^n$ such that $\Omega'\Subset\Omega$, there exist $N\in\mathbb{N}$ and $C>0$ such that*

$$\|D^\beta u\|^2_{L^2(\Omega')}\leq C\Big(\sum_{|\alpha|\leq N}\|D^\alpha Lu\|^2_{L^2(\Omega)}+\|u\|^2_{L^2(\Omega)}\Big)\quad\textit{for all } u\in C^\infty(\overline{\Omega}).\tag{4.15}$$

Proof. Fix $\Omega'\Subset\Omega$ and consider the set

$$S\equiv\{u\in L^2(\Omega'):D^\alpha Lu\in L^2(\Omega')\text{ for all multiindices }\alpha\}.$$

The family of seminorms $\|u\|_{L^2(\Omega')}$, $\|D^\alpha Lu\|_{L^2(\Omega')}$, $|\alpha| \in \mathbb{N}$, makes it a Fréchet space. Since L is hypoelliptic, we have $S \subset C^\infty(\Omega')$, and in particular $S \subset C^M(\Omega')$ for any $M > 0$. Now consider the inclusion map

$$T : S \to C^M(\Omega');$$

we claim T is closed. Indeed, suppose $\{u_n\} \subset S$ satisfies $u_n \to u$ in S and $u_n \to v$ in $C^M(\Omega')$, in particular, $u_n \to u$ in $L^2(\Omega')$ and $u_n \to v$ in $L^\infty(\Omega')$. Then, for any $n \in \mathbb{N}$,

$$\begin{aligned}\|u - v\|_{L^2(\Omega')} &\le \|u - u_n\|_{L^2(\Omega')} + \|u_n - v\|_{L^2(\Omega')} \\ &\le \|u - u_n\|_{L^2(\Omega')} + \|u_n - v\|_{L^\infty(\Omega')} |\Omega'|^{1/2},\end{aligned}$$

and the right-hand side converges to 0 as $n \to \infty$. This implies $u = v$, i.e., T is closed. By the closed graph theorem, T is continuous, and therefore there exists $N \in \mathbb{N}$ and $C > 0$ such that

$$\|u\|_{C^M(\Omega')} \le C\Big(\sum_{|\alpha| \le N} \|D^\alpha Lu\|^2_{L^2(\Omega')} + \|u\|^2_{L^2(\Omega')}\Big).$$

Since the choice of M was arbitrary, this implies (4.15). ∎

Proposition 4.5. *Fix distinct indices $k, j \in \{m+1, \ldots, p\}$, where $p > m+1$. Define*

$$L_1 \equiv \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_m^2} + \lambda_k(x_1, \ldots, x_m) \frac{\partial^2}{\partial x_k^2} + \lambda_j(x_1, \ldots, x_m) \frac{\partial^2}{\partial x_j^2},$$

$$L_2 \equiv \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_m^2} + \sum_{i=m+1}^{p} \lambda_i(x_1, \ldots, x_m) \frac{\partial^2}{\partial x_i^2} + \sum_{i=p+1}^{n} \lambda_p(x_1, \ldots, x_m) \frac{\partial^2}{\partial x_i^2}.$$

If L_1 is not hypoelliptic in $\mathbb{R}^{m+2}$, then L_2 is not hypoelliptic in $\mathbb{R}^n$.

Proof. Suppose L_1 is not hypoelliptic in $\mathbb{R}^{m+2}$, i.e., there exists a non-smooth function $u = u(x_1, \ldots, x_m, x_k, x_j)$ such that $L_1 u \in C^\infty(\mathbb{R}^{m+2})$. If we define the function v by

$$v(x_1, \ldots, x_n) = u(x_1, \ldots, x_m, x_k, x_j),$$

then v is not smooth since u is not smooth. However,

$$L_2 v(x_1, \ldots, x_n) = L_1 u(x_1, \ldots, x_m, x_k, x_j),$$

and therefore smooth in $\mathbb{R}^n$. ∎

Theorem 4.6. *Suppose that $h, f \in C^\infty(\mathbb{R}^m)$ are* strongly monotone*, i.e.,*

$$f(z) \le f(x) \quad \textit{and} \quad h(z) \le h(x) \quad \textit{for all } z \in B(0, |x|),$$

and satisfy $h(x), f(x) \ge 0$ and $h(0) = f(0) = 0$ for all $x \in \mathbb{R}^m$. Define

$$\mu(t, h) \equiv \max\{h(z)(t - |z|) : 0 \le |z| \le t\}.$$

and suppose in addition that

$$\liminf_{x\to 0} \mu(|x|,h)\ln f(x) \neq 0. \tag{4.16}$$

Then the operator

$$\mathcal{L} \equiv \Delta_x + f^2(x)\frac{\partial^2}{\partial y^2} + h^2(x)\frac{\partial^2}{\partial t^2}$$

fails *to be C^∞-hypoelliptic in $\mathbb{R}^{m+2}$.*

Proof. For $a,\eta > 0$ consider the second order operator $L_\eta \equiv -\Delta_x + f^2(x)\eta^2$ and the eigenvalue problem

$$L_\eta v(x,\eta) = \lambda h^2(x)v(x,\eta), \quad x \in B(0,a), \qquad v(x) = 0, \quad x \in \partial B(0,a).$$

The least eigenvalue is given by the Rayleigh quotient formula:

$$\begin{aligned} \lambda_0(a,\eta) &= \inf_{\varphi(\neq 0)\in C_0^\infty(B)} \frac{\langle L_\eta\varphi,\varphi\rangle_{L^2}}{\langle h^2\varphi,\varphi\rangle_{L^2}} \\ &= \inf_{\varphi(\neq 0)\in C_0^\infty(B)} \frac{\int_B |\nabla\varphi|(x)^2\,dx + \int_B f^2(x)\eta^2\varphi(x)^2\,dx}{\int_B h(x)^2\varphi(x)^2\,dx}. \end{aligned} \tag{4.17}$$

Next, from (4.16), it follows that there exists $\varepsilon > 0$ and sequences $\{a_n\},\{b_n\} \subset \mathbb{R}^m$ such that $|a_n| < |b_n| \le 1$, $b_n \to 0$, and

$$h(a_n)(|b_n| - |a_n|)|\ln f(b_n)| \ge \varepsilon \quad \text{for all } n \in \mathbb{N}. \tag{4.18}$$

Let

$$\eta_n = \frac{1}{f(b_n)} \to \infty \quad \text{as } n \to \infty,$$

By the strong monotonicity of f and h, we have

$$\eta_n f(x) \le 1, \quad h(x) \ge h(a_n) \quad \text{for all } x \in R_n \equiv \{x \in \mathbb{R}^m : |a_n| \le |x| \le |b_n|\}.$$

This implies, using (4.17),

$$\begin{aligned} \lambda_0(|b_n|,\eta_n) &\le \inf_{\varphi(\neq 0)\in C_0^\infty(R_n)} \frac{\langle L_{\eta_n}\varphi,\varphi\rangle_{L^2}}{\langle h^2\varphi,\varphi\rangle_{L^2}} \\ &\le h(a_n)^{-2} \inf_{\varphi(\neq 0)\in C_0^\infty(R_n)} \{(\|\nabla\varphi\|^2 + \|\varphi\|^2)/\|\varphi\|^2\} \\ &\le h(a_n)^{-2}(C(|b_n| - |a_n|)^{-2} + 1) \le C|\ln f(b_n)|^2 = C(\ln\eta_n)^2, \end{aligned}$$

where we used (4.18) and the definition of η_n for the last two inequalities. It also follows, from (4.17) and the fact that $|b_n| \le 1$, that

$$\lambda_0(1,\eta_n) \le \lambda_0(|b_n|,\eta_n) \le C_1(\ln\eta_n)^2. \tag{4.19}$$

Now let $v_0(x,\eta_n)$ be an eigenfunction on the ball $B = B(0,1)$ associated with $\lambda_0(1,\eta_n)$, i.e.,

$$-\Delta v_0(x,\eta_n) = [\lambda_0(1,\eta_n)h^2(x) - f^2(x)\eta_n^2]\,v_0(x,n),$$

and normalized so that

$$\|v_0(\cdot,\eta_n)\|_{L^2(B)} = 1.$$

We first claim that

$$\|v_0(\cdot,\eta_n)\|_{L^2((1/2)B)} \to 1 \quad \text{as } n \to \infty. \tag{4.20}$$

Indeed, we have

$$\begin{aligned}\inf_{1/2<|x|<1} f^2(x)\,\eta_n^2 \int_{1/2<|x|<1} |v_0(x,\eta_n)|^2\,dx &\le \int_B f^2(x)\,\eta_n^2\,|v_0(x,\eta_n)|^2\,dx \\ &\le \int_B |\nabla v_0(x,\eta_n)|^2\,dx + \int_B f^2(x)\,\eta_n^2\,|v_0(x,\eta_n)|^2\,dx \\ &= \lambda_0(1,\eta_n) \int_B h^2(x)\,|v_0(x,\eta_n)|^2\,dx \le C\lambda_0(1,\eta_n).\end{aligned}$$

Dividing both sides by $\inf_{1/2<|x|<1} f^2(x)\,\eta_n^2$ and using (4.19), we obtain that

$$\int_{1/2<|x|<1} |v_0(x,\eta_n)|^2\,dx \to 0 \quad \text{as } n \to \infty,$$

which implies (4.20). Define a sequence of functions

$$u_n(x,y,t) = e^{iy\eta_n + \sqrt{\lambda_0(1,\eta_n)}\,t}\, v_0(x,\eta_n).$$

Then

$$\mathcal{L}u_n = \big(\Delta v_0(x,\eta_n) - \eta_n^2 f^2(x)\,v_0(x,\eta_n) + \lambda_0(1,\eta_n)\,v_0(x,\eta_n)\big)\,e^{iy\eta_n+\sqrt{\lambda_0(1,\eta_n)}\,t} = 0.$$

Now, let $V = B(0,1)\times[-\pi,\pi]\times[-\delta,\delta]$ and $V' = B(0,1/2)\times[-\pi/2,\pi/2]\times[-\delta/2,\delta/2]$ for some $\delta > 0$. We have, using (4.20),

$$\|\partial_y^k u_n\|_{L^2(V')}^2 = \eta_n^{2k}\,\|u_n\|_{L^2(V')}^2 \ge \pi\,\eta_n^{2k} \int_{1/2B}\int_0^{\delta/2} e^{2\sqrt{\lambda_0(1,\eta_n)}t}\,|v_0(x,\eta_n)|^2\,dt\,dx \ge C\,\eta_n^{2k},$$

where the constant C is independent of k and n. On the other hand, using (4.19),

$$\|u_n\|_{L^2(V)}^2 \le Ce^{2\sqrt{\lambda_0(1,\eta_n)}\,\delta} \le C\,\eta_n^{2\sqrt{C_1}\delta}.$$

Since $\eta_n \to \infty$ as $n \to \infty$, these two inequalities contradict (4.15) for $k > \sqrt{C_1}\delta$, and thus, by Lemma 4.4, the operator $\mathcal{L}$ is not hypoelliptic. ∎

5. Proof of Theorem 2.2

Finally, we prove Theorem 2.2 by showing that the requirements of Theorem 2.5 are satisfied. Let L be as in (2.1). We apply Theorem 1.3 to obtain $\mathbf{A} = \sum_{j=1}^N Y_j Y_j^{\mathrm{tr}} + A_p$, and write the second order term in L as

$$\nabla^{\mathrm{tr}}\mathbf{A}\nabla = \sum_{j=1}^N \nabla^{\mathrm{tr}} Y_j Y_j^{\mathrm{tr}}\nabla = \sum_{j=1}^N X_j^{\mathrm{tr}} X_j + \widehat{\nabla}^{\mathrm{tr}}\mathbf{Q}_p\,\widehat{\nabla}, \quad \text{where } X_j = Y_j^{\mathrm{tr}}\nabla,$$

and then note that condition (2.5) is satisfied by the assumption (2.3) of Theorem 2.2. Moreover, condition (2.4) follows from (1.4).

6. Open problems

6.1. First problem

In Theorem 2.2, we have shown that the Koike condition is sufficient for the hypoellipticity of an operator L with $n \times n$ matrix $A(x)$ satisfying certain conditions on both its diagonal and nondiagonal entries. However, in the converse direction we only showed that failure of the Koike condition implies failure of hypoellipticity if in addition L is diagonal with strongly monotone entries. In fact, the proof shows that we need only to assume in addition that $A(x)$ has the block form

$$A(x) = \begin{bmatrix} \begin{bmatrix} a_{1,1}(x) & \cdots & a_{n,1}(x) \\ \vdots & \ddots & \vdots \\ a_{1,n}(x) & \cdots & a_{m,m}(x) \end{bmatrix} & \mathbf{0}_{m\times 1} & \mathbf{0}_{m\times 1} & \cdots & \mathbf{0}_{m\times 1} \\ \mathbf{0}_{1\times m} & a_{m+1,m+1}(x) & 0 & \cdots & 0 \\ \mathbf{0}_{1\times m} & 0 & a_{m+2,m+2}(x) & & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{1\times m} & 0 & 0 & \cdots & a_{n,n}(x) \end{bmatrix},$$

where just $a_{m+1,m+1}(x)$ and $a_{n,n}(x)$ are assumed to be strongly monotone and satisfy (4.16).

Problem 6.1. Is the Koike condition actually necessary and sufficient for hypoellipticity under the assumptions of Theorem 2.2, *without* assuming the above block form for $A(x)$?

6.2. Second problem

Recall that the main theorem in [17] extends Kohn's theorem in [15] to apply with finitely many blocks instead of the two blocks used in [15]. These operators are restricted by being of a certain block form, but they are more general in that the elliptic blocks are multiplied by smooth functions that are positive outside the origin, and have more variables than in our theorems, and furthermore that need not be finite sums of squares of regular functions.

Problem 6.2. Can Theorem 2.5 be extended to more general operators that include the operators appearing in [17]?

6.3. Third problem

What sort of smooth lower order terms of the form $B(x)\nabla$ and $\nabla^{\mathrm{tr}} C(x)$ can we add to the operator L in the main Theorem 2.2? The natural hypothesis to make on the vector fields $B(x)\nabla$ and $C(x)\nabla$ is that they are subunit with respect to $\nabla^{\mathrm{tr}} A(x)\nabla$. However, if we use Theorem 2.5 in the proof, we require more, namely, that $B(x)\nabla$ and $C(x)\nabla$ are linear combinations, with $C^{2,\delta}$ coefficients, of the $C^{2,\delta}$ vector fields $X_j(x)$ arising in the sum of squares Theorem 1.3, something which seems difficult to arrange more generally.

Problem 6.3. Can Theorem 2.2 be extended to operators that include first order terms that are subunit with respect to $\nabla^{\mathrm{tr}} A(x)\nabla$?

Funding. The work of the second author was partially supported by an NSERC grant, number 12409, and the McKay Research Chair grant at McMaster University

References

[1] Akhunov, T., Korobenko, L. and Rios, C.: Hypoellipticity of Fediĭ's type operators under Morimoto's logarithmic condition. *J. Pseudo-Differ. Oper. Appl.* **10** (2019), no. 3, 649–688. Zbl 1423.35084 MR 3990263

[2] Bochnak, J., Coste, M. and Roy, M.-F.: *Géométrie algébrique réelle*. Ergeb. Math. Grenzgeb. (3) 12, Springer, Berlin, 1987. Zbl 0633.14016 MR 949442

[3] Bony, J.-M.: Sommes de carrés de fonctions dérivables. *Bull. Soc. Math. France* **133** (2005), no. 4, 619–639. Zbl 1107.26008 MR 2233698

[4] Bourdaud, G.: L^p estimates for certain nonregular pseudodifferential operators. *Comm. Partial Differential Equations* **7** (1982), no. 9, 1023–1033. Zbl 0499.35097 MR 673825

[5] Brumfiel, G. W.: *Partially ordered rings and semi-algebraic geometry*. London Math. Soc. Lecture Note Ser. 37, Cambridge University Press, Cambridge-New York, 1979. Zbl 0415.13015 MR 553280

[6] Christ, M.: Hypoellipticity: geometrization and speculation. In *Complex analysis and geometry* (*Paris, 1997*), pp. 91–109. Progr. Math. 188, Birkhäuser, Basel, 2000. Zbl 0965.47033 MR 1782661

[7] Christ, M.: Hypoellipticity in the infinitely degenerate regime. In *Complex analysis and geometry* (*Columbus, OH, 1999*), pp. 59–84. Ohio State Univ. Math. Res. Inst. Publ. 9, de Gruyter, Berlin, 2001. Zbl 1015.32032 MR 1912731

[8] Fediĭ, V. S.: On a criterion for hypoellipticity. *Math. USSR-Sb.* **14** (1971), no. 1, 15–45. Zbl 0247.35023 MR 287160

[9] Fefferman, C. and Phong, D. H.: On positivity of pseudo-differential operators. *Proc. Nat. Acad. Sci. U.S.A.* **75** (1978), no. 10, 4673–4674. Zbl 0391.35062 MR 507931

[10] Fefferman, C. and Phong, D. H.: Subelliptic eigenvalue problems. In *Conference on harmonic analysis in honor of Antoni Zygmund, Vol. I, II* (*Chicago, Ill., 1981*), pp. 590–606. Wadsworth Math. Ser., Wadsworth, Belmont, CA, 1983. Zbl 0503.35071 MR 730094

[11] Glaeser, G.: Racine carrée d'une fonction différentiable. *Ann. Inst. Fourier (Grenoble)* **13** (1963), no. 2, 203–210. Zbl 0128.27903 MR 163995

[12] Guan, P.: C^2 a priori estimates for degenerate Monge–Ampère equations. *Duke Math. J.* **86** (1997), no. 2, 323–346. Zbl 0879.35059 MR 1430436

[13] Hörmander, L.: Hypoelliptic second order differential equations. *Acta Math.* **119** (1967), 147–171. Zbl 0156.10701 MR 222474

[14] Hoshiro, T.: Hypoellipticity for infinitely degenerate elliptic and parabolic operators of second order. *J. Math. Kyoto Univ.* **28** (1988), no. 4, 615–632. Zbl 0692.35034 MR 981097

[15] Kohn, J. J.: Hypoellipticity of some degenerate subelliptic operators. *J. Funct. Anal.* **159** (1998), no. 1, 203–216. Zbl 0937.35024 MR 1654190

[16] Koike, M.: A note on hypoellipticity of degenerate elliptic operators. *Publ. Res. Inst. Math. Sci.* **27** (1991), no. 6, 995–1000. Zbl 0804.35025 MR 1145673

[17] Korobenko, L. and Rios, C.: Hypoellipticity of certain infinitely degenerate second order operators. *J. Math. Anal. Appl.* **409** (2014), no. 1, 41–55. Zbl 1388.35013 MR 3095016

[18] Korobenko, L., Rios, C., Sawyer, E. and Shen, R.: Local boundedness, maximum principles, and continuity of solutions to infinitely degenerate elliptic equations with rough coefficients. *Mem. Amer. Math. Soc.* **269** (2021), no. 1311, 130 pp. Zbl 1494.35001 MR 4224718

[19] Korobenko, L. and Sawyer, E.: Sums of squares I: Scalar functions. *J. Funct. Anal.* **284** (2023), no. 6, article no. 109827, 66 p. Zbl 1511.35181 MR 4530899

[20] Korobenko, L. and Sawyer, E.: Sums of squares II: Matrix functions. *Linear Algebra Appl.* **664** (2023), 264–313. Zbl 1526.26002 MR 4546703

[21] Kusuoka, S. and Stroock, D.: Applications of the Malliavin calculus. II. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **32** (1985), no. 1, 1–76. Zbl 0568.60059 MR 783181

[22] Morimoto, Y.: Non-hypoellipticity for degenerate elliptic operators *Publ. Res. Inst. Math. Sci.* **22** (1986), no. 1, 25–30 Zbl 0616.35018 MR 0834345; and Erratum to: "Non-hypoellipticity for degenerate elliptic operators" *Publ. Res. Inst. Math. Sci.* **30** (1994), no. 4, 533–534. Zbl 0813.35012 MR 1308956

[23] Pieroni, F.: On the real algebra of Denjoy–Carleman classes. *Selecta Math. (N.S.)* **13** (2007), no. 2, 321–351. Zbl 1148.26033 MR 2361097

[24] Rios, C., Sawyer, E. T. and Wheeden, R. L.: Hypoellipticity for infinitely degenerate quasilinear equations and the Dirichlet problem. *J. Anal. Math.* **119** (2013), 1–62. Zbl 1310.35081 MR 3043146

[25] Sawyer, E. T.: A symbolic calculus for rough pseudodifferential operators. *Indiana Univ. Math. J.* **45** (1996), no. 2, 289–332. Zbl 0869.35118 MR 1414332

[26] Stein, E. M.: *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*. Princeton Math. Ser. 43, Princeton University Press, Princeton, NJ, 1993. Zbl 0821.42001 MR 1232192

[27] Tataru, D.: On the Fefferman–Phong inequality and related problems. *Comm. Partial Differential Equations* **27** (2002), no. 11-12, 2101–2138. Zbl 1045.35115 MR 1944027

[28] Taylor, M. E.: *Partial differential equations. I, II and III*. Appl. Math. Sci. 115, 116 and 117, Springer, New York, 1996. Zbl 0869.35002 MR 1395148

[29] Trèves, F.: *Introduction to pseudodifferential and Fourier integral operators. Vols. 1 and 2*. University Series in Mathematics, Plenum Press, New York-London, 1980. Zbl 0453.47027 MR 597144

Received March 15, 2023; revised March 26, 2024.

Lyudmila Korobenko
Mathematics and Statistics Department, Reed College
SE Woodstock Blvd, 3203, Portland, OR 97202, USA;
korobenko@reed.edu

Eric Sawyer
Department of Mathematics and Statistics, McMaster University
Main St W, 1280, Hamilton, ON L8S 4L8, Canada;
sawyer@mcmaster.ca