



**HAL**  
open science

# Tunneling estimates and approximate controllability for hypoelliptic equations

Camille Laurent, Matthieu Léautaud

► **To cite this version:**

Camille Laurent, Matthieu Léautaud. Tunneling estimates and approximate controllability for hypoelliptic equations. *Memoirs of the American Mathematical Society*, 2022, 276 (1357), 10.1090/memo/1357. hal-01717587

**HAL Id: hal-01717587**

**<https://hal.science/hal-01717587>**

Submitted on 26 Feb 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Tunneling estimates and approximate controllability for hypoelliptic equations

Camille Laurent\* and Matthieu Léautaud†

October 22, 2017

## Abstract

This article is concerned with quantitative unique continuation estimates for equations involving a “sum of squares” operator  $\mathcal{L}$  on a compact manifold  $\mathcal{M}$  assuming: (i) the Chow-Rashevski-Hörmander condition ensuring the hypoellipticity of  $\mathcal{L}$ , and (ii) the analyticity of  $\mathcal{M}$  and the coefficients of  $\mathcal{L}$ .

The first result is the tunneling estimate  $\|\varphi\|_{L^2(\omega)} \geq Ce^{-\lambda^{\frac{k}{2}}}$  for normalized eigenfunctions  $\varphi$  of  $\mathcal{L}$  from a nonempty open set  $\omega \subset \mathcal{M}$ , where  $k$  is the hypoellipticity index of  $\mathcal{L}$  and  $\lambda$  the eigenvalue.

The main result is a stability estimate for solutions to the hypoelliptic wave equation  $(\partial_t^2 + \mathcal{L})u = 0$ : for  $T > 2 \sup_{x \in \mathcal{M}}(\text{dist}(x, \omega))$  (here,  $\text{dist}$  is the sub-Riemannian distance), the observation of the solution on  $(0, T) \times \omega$  determines the data. The constant involved in the estimate is  $Ce^{c\Lambda^k}$  where  $\Lambda$  is the typical frequency of the data.

We then prove the approximate controllability of the hypoelliptic heat equation  $(\partial_t + \mathcal{L})v = \mathbb{1}_\omega f$  in any time, with appropriate (exponential) cost, depending on  $k$ . In case  $k = 2$  (Grushin, Heisenberg...), we further show approximate controllability to trajectories with polynomial cost in large time.

We also explain how the analyticity assumption can be relaxed, and a boundary  $\partial\mathcal{M}$  can be added in some situations.

Most results turn out to be optimal on a family of Grushin-type operators.

The main proof relies on the general strategy developed by the authors in [LL15].

## Keywords

Stability estimates, Hypoelliptic operators, Eigenfunctions, Wave equation, Heat equation, Control theory.

**2010 Mathematics Subject Classification:** 35B60, 35H10, 35P20, 35L05, 35K05, 93B05, 93B07.

---

\*CNRS UMR 7598 and Sorbonne Universités UPMC Univ Paris 06, Laboratoire Jacques-Louis Lions, F-75005, Paris, France, email: laurent@ann.jussieu.fr

†École Polytechnique, Centre de Mathématiques Laurent Schwartz UMR7640, 91128 Palaiseau cedex France. Most of this research was done when the second author was in CRM, CNRS UMI 3457, Université de Montréal, Case Postale 6128, Succursale Centre-Ville, Montréal (QC) Canada H3C 3J7 and Université Paris Diderot, IMJ-PRG, UMR 7586, Bâtiment Sophie Germain, 75205 Paris Cedex 13 France, email: matthieu.leautaud@polytechnique.edu.

# Contents

<b>1</b>	<b>Introduction and main results</b>	<b>3</b>
1.1	Introduction . . . . .	3
1.2	Main results . . . . .	7
1.3	Comparison to other works . . . . .	16
1.4	Sketch of the proofs and plan of the paper . . . . .	19
1.5	Some remarks and further comments . . . . .	22
<b>2</b>	<b>The quantitative Holmgren-John theorem of [LL15]</b>	<b>26</b>
2.1	A typical quantitative unique continuation result of [LL15] . . . . .	26
2.2	Definitions and tools for propagating the information . . . . .	28
2.3	Semiglobal estimates along foliation by hypersurfaces . . . . .	30
<b>3</b>	<b>The hypoelliptic wave equation, proof of Theorem 1.13</b>	<b>31</b>
3.1	Step 1: Geometric setting and non-characteristic hypersurfaces . . . . .	34
3.2	Step 2: Propagation of smallness . . . . .	36
3.3	Step 3: Energy estimates . . . . .	41
<b>4</b>	<b>The hypoelliptic heat equation</b>	<b>54</b>
4.1	Approximate controllability with polynomial cost in large time: Proof of Theorem 1.20 . . . . .	57
4.2	Approximate controllability in Gevrey-type spaces: Proof of Theorem 1.18 . . . . .	58
4.3	Approximate controllability in natural spaces with exponential cost: Proof of Theorem 1.16 . . . . .	59
4.4	Technical lemmata used for the heat equation . . . . .	61
<b>5</b>	<b>A partially analytic example: Grushin type operators</b>	<b>64</b>
5.1	The geometric context . . . . .	65
5.2	A proof of Estimate (1.24) . . . . .	67
5.3	An observation term in $L^2$ in quantitative unique continuation estimates . . . . .	68
<b>A</b>	<b>On the optimality: Proof of Proposition 1.12</b>	<b>70</b>
<b>B</b>	<b>Subelliptic estimates</b>	<b>71</b>
B.1	$H^s$ subelliptic estimates on compact manifolds . . . . .	71
B.2	Subelliptic estimates for manifolds with boundaries . . . . .	73
<b>C</b>	<b>Sub-Riemannian norm of normal vectors</b>	<b>78</b>

# 1 Introduction and main results

## 1.1 Introduction

Let  $\mathcal{M}$  be a smooth compact connected  $d$ -dimensional manifold without boundary. We denote  $\mathcal{X}^\infty$  the space of smooth vector fields on  $\mathcal{M}$  (with real coefficients), which we identify to derivations on  $\mathcal{M}$ . We assume  $\mathcal{M}$  is endowed with a smooth positive density measure  $ds$ , so that we may integrate functions on  $\mathcal{M}$ <sup>1</sup>. We may then define the space  $L^2(\mathcal{M}) = L^2(\mathcal{M}, ds)$  of square integrable functions with respect to this measure. For  $X \in \mathcal{X}^\infty$ , we define by  $X^*$  its formal dual operator for the duality of  $L^2(\mathcal{M})$ , that is<sup>2</sup>,

$$\int_{\mathcal{M}} X^*(u)(x)v(x)ds(x) = \int_{\mathcal{M}} u(x)X(v)(x)ds(x), \quad \text{for any } u, v \in C^\infty(\mathcal{M}).$$

Given  $m \in \mathbb{N}$  and  $m$  vector fields<sup>3</sup>  $X_1, \dots, X_m \in \mathcal{X}^\infty$ , we are interested in properties of the following (non-positive) second order operator, associated to the  $X_i$ 's (namely the so-called type I Hörmander operator)<sup>4</sup>

$$\mathcal{L} = \sum_{i=1}^m X_i^* X_i. \tag{1.1}$$

Note that this operator is formally symmetric nonnegative, when defined on functions in  $C^\infty(\mathcal{M})$ , since we have

$$(\mathcal{L}u, u)_{L^2(\mathcal{M})} = \sum_{i=1}^m \|X_i u\|_{L^2(\mathcal{M})}^2. \tag{1.2}$$

Both from the geometric control and the operator theoretic points of view, it is in this context natural to consider iterated Lie brackets of the vector fields  $X_i$ . We refer for instance to the following classical article [Bel96] and textbooks [Mon02, Jea14, Rif14, ABB16b].

**Definition 1.1.** For any family  $\mathcal{F}$  of smooth vector fields on  $\mathcal{M}$  and  $\ell \in \mathbb{N}^*$ , we define the subspaces  $\text{Lie}^\ell(\mathcal{F})$  of  $\mathcal{X}^\infty$  by iteration as follows:

- $\text{Lie}^1(\mathcal{F})$  is the space spanned by  $\mathcal{F}$  in  $\mathcal{X}^\infty$ ,
- $\text{Lie}^{\ell+1}(\mathcal{F}) = \text{span} \left( \text{Lie}^\ell(\mathcal{F}) \cup \left\{ [X, Y]; X \in \mathcal{F}, Y \in \text{Lie}^\ell(\mathcal{F}) \right\} \right)$ .

For any point  $x \in \mathcal{M}$ ,  $\ell \in \mathbb{N}^*$ , we denote  $\text{Lie}^\ell(\mathcal{F})(x)$  the set of all tangent vectors  $X(x)$  with  $X \in \text{Lie}^\ell(\mathcal{F})$ .

We shall always assume that the family  $(X_i)$  satisfies the Chow-Rashevski-Hörmander condition (or is “bracket generating”).

<sup>1</sup>See e.g. [Lee13, Chapter 16 p427]: given a local chart  $(U_\kappa, \kappa)$  of  $\mathcal{M}$ , we have  $\int_{U_\kappa} u ds = \int_{\kappa(U_\kappa)} u \circ \kappa^{-1}(y) \varphi^\kappa(y) dy$  for an appropriate smooth positive function  $\varphi^\kappa$ , and for any  $u \in C_c^0(U_\kappa)$ .

<sup>2</sup>Note that in the local chart  $(U_\kappa, \kappa)$  we have  $X^\kappa = \sum_j a_j^\kappa(x) \partial_j$ , and thus  $(X^\kappa)^* = \sum_j -a_j^\kappa(x) \partial_j - \partial_j a_j^\kappa - \frac{\partial_j \varphi^\kappa}{\varphi^\kappa} a_j^\kappa$ , which is a vector field (namely  $-X$ ) plus a multiplication operator (namely  $-\text{div}_{ds}(X)$ , see Remark 1.28 below).

<sup>3</sup>The assumption  $1 \leq m \leq d$  is sometimes made in the references we use, but can always be removed.

<sup>4</sup>See Remark 1.28 below for a discussion on general sub-Riemannian Laplacians.

**Assumption 1.1.** There exists  $\ell \geq 1$  so that for any  $x \in \mathcal{M}$ ,  $\text{Lie}^\ell(X_1, \dots, X_m)(x) = T_x \mathcal{M}$ <sup>5</sup>. Denote then by  $k \in \mathbb{N}^*$  the minimal  $\ell$  for which this holds.

The integer  $k$  will sometimes be referred to as the *hypoellipticity index* of  $\mathcal{L}$ . Assumption 1.1 is central in control theory and operator theory, for it characterizes both the controllability of the controlled ODE driven by the vector fields  $(X_i)$  and the Hypoellipticity of the operator  $\mathcal{L}$ . Let us now recall these two seminal results, namely the Chow-Rashevski theorem and the Hörmander theorem, which we both use in the sequel.

**Theorem 1.2** (Chow [Cho39], Rashevski [Ras38]). *Under Assumption 1.1, the following statement holds: for any  $x_0, x_1 \in \mathcal{M}$ , any  $T > 0$ , there exist  $u_i \in L^1(0, T)$  for  $i \in \{1, \dots, m\}$  such that the unique solution of*

$$\dot{\gamma}(t) = \sum_{i=1}^m u_i(t) X_i(\gamma(t)), \quad \gamma(0) = x_0 \quad (1.3)$$

satisfies  $\gamma(T) = x_1$ .

We refer e.g. to [Jea14, Chapter 1.4], [Rif14, Chapter 1.4] or [Mon02, Chapter 2] for statements and proofs of the Chow-Rashevski theorem, and in particular for the definition of the solution of (1.3). See also [Cor07, Chapter 3] for examples and applications in control theory. This theorem motivates the following definition.

**Definition 1.3** (Horizontal path). We say that an absolutely continuous function  $\gamma : [0, T] \rightarrow \mathcal{M}$  is a horizontal path if there exist  $u_i \in L^1(0, T; \mathbb{R})$  for  $i = 1, \dots, m$  such that for almost every  $t \in [0, T]$ , we have  $\dot{\gamma}(t) = \sum_{i=1}^m u_i(t) X_i(\gamma(t))$ .

Such a trajectory is in particular absolutely continuous and almost everywhere tangent to the so-called horizontal distribution  $\text{span}(X_1, \dots, X_m)$ . The second key role played by Assumption 1.1 in analysis is summarized in the following result.

**Theorem 1.4** (Hörmander [Hör67], Rothschild-Stein [RS76]). *Under Assumption 1.1, the operator  $\mathcal{L}$  in (1.1) is hypoelliptic, that is, for all  $u \in \mathcal{D}'(\mathcal{M})$  and  $x_0 \in \mathcal{M}$ , if  $\mathcal{L}u \in C^\infty$  near  $x_0$  then  $u \in C^\infty$  near  $x_0$ .*

Moreover, it is subelliptic of order  $\frac{1}{k}$ , that is, the following estimates hold: there is  $C > 0$  such that for any  $u \in C^\infty(\mathcal{M})$ , we have

$$\|u\|_{H^{\frac{1}{k}}(\mathcal{M})}^2 \leq C \sum_{i=1}^m \|X_i u\|_{L^2(\mathcal{M})}^2 + C \|u\|_{L^2(\mathcal{M})}^2, \quad (1.4)$$

$$\|u\|_{H^{\frac{1}{k}}(\mathcal{M})}^2 \leq C \|\mathcal{L}u\|_{L^2(\mathcal{M})} + C \|u\|_{L^2(\mathcal{M})}^2, \quad (1.5)$$

$$\|u\|_{H^{\frac{2}{k}}(\mathcal{M})}^2 \leq C \|\mathcal{L}u\|_{L^2(\mathcal{M})}^2 + C \|u\|_{L^2(\mathcal{M})}^2. \quad (1.6)$$

The hypoellipticity was shown by Hörmander [Hör67], who also provides with a subelliptic estimate with loss (see also [Koh78, Koh05] or [HN05, Chapter 2] for a simpler proof). The optimal subelliptic

<sup>5</sup>Note that it is sufficient to assume that for all  $x \in \mathcal{M}$ , there is  $\ell = \ell(x) \in \mathbb{N}$  such that this holds. The upper semi-continuity of  $x \mapsto \ell(x)$  and the compactness of  $\mathcal{M}$  then imply the stronger form of Assumption 1.1 as stated.

estimate (1.4) with gain of  $1/k$  derivatives is proved by [RS76] (see also [FP81, p288] for a different proof, or [BCN82] for a simpler one). More precisely, (even slightly hidden) it is written in [RS76] Theorem 17 and estimate (17.20) p311 in a local form. It is then easy to globalise on the compact manifold  $\mathcal{M}$  to obtain (1.4) (since commutators of  $X_i$  with a smooth cutoff function is a multiplication operator).

Both estimates (1.5) and (1.6) may then be deduced from (1.4). This is clear for (1.5) when recalling (1.2). The proof of (1.6) requires a commutator argument (detailed e.g. in [FP83]) and is proved in Appendix B.1, as well as  $H^s$  variants of (1.5) and (1.6). Note that these subelliptic estimates are also obtained in Fefferman-Phong [FP83] for some wider class of symmetric operators, not necessarily sums of squares, and with a shorter proof.

Since the operator  $\mathcal{L}$  is symmetric non-negative, the hypoellipticity of  $\mathcal{L} + 1$  and the compactness of  $\mathcal{M}$  directly imply that  $\mathcal{L}$  is essentially selfadjoint (see e.g. Reed-Simon [RS80, Theorem X.26]). Hence, it extends uniquely as a selfadjoint operator (its Friedrich extension)

$$\mathcal{L} : D(\mathcal{L}) \subset L^2(\mathcal{M}) \rightarrow L^2(\mathcal{M}),$$

with, according to (1.6),  $H^2(\mathcal{M}) \subset D(\mathcal{L}) \subset H^{\frac{2}{k}}(\mathcal{M})$  (still under Assumption 1.1). The operator  $\mathcal{L}$  is hence selfadjoint on  $L^2(\mathcal{M})$ , with compact resolvent: it admits a Hilbert basis of eigenfunctions  $(\varphi_j)_{j \in \mathbb{N}}$ , associated with the real eigenvalues  $(\lambda_j)_{j \in \mathbb{N}}$ , sorted increasingly, that is

$$\mathcal{L}\varphi_i = \lambda_i\varphi_i, \quad (\varphi_i, \varphi_j)_{L^2(\mathcal{M})} = \delta_{ij}, \quad 0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_j \rightarrow +\infty. \quad (1.7)$$

Note that a bootstrap argument in (1.6) shows that  $\varphi_j \in C^\infty(\mathcal{M})$ . In particular, the spectral decomposition allows to define solutions of the hypoelliptic wave and heat equations (respectively  $(\partial_t^2 + \mathcal{L})v = 0$  and  $(\partial_t + \mathcal{L})u = 0$ ), which we shall consider in this paper.

In addition to Assumption 1.1, we will also assume in the main part of the article that everything is real-analytic. This assumption is not made in Section 5 though, where we give some results in the non-analytic context.

**Assumption 1.2.** The manifold  $\mathcal{M}$ , the density  $ds$ , and the vector fields  $X_i$  are **real-analytic**.

In particular, it implies that the operator  $\mathcal{L}$  has analytic coefficients in any analytic coordinate set compatible with the manifold  $\mathcal{M}$ . Note that under this assumption, the converse of Theorems 1.2 and 1.4 also hold, namely:

- Attainability, in the sense of Theorems 1.2, implies Assumption 1.1, see [Her63, Nag66] (see also [Lob70, Sus73] for generalizations);
- The hypoellipticity of  $\mathcal{L}$ , in the sense of Theorem 1.4, implies Assumption 1.1, see [Der71, Theorem 2.2] if there is no point  $x_0$  where all  $X_i$  cancel.

The analyticity assumption is further discussed in Sections 1.2.4 and 1.3 below.

Before stating our main results, let us provide with classical examples of operators that are considered in the present paper.

**Example 1.5** (Elliptic operators,  $k = 1$ ). In the case  $k = 1$ , then,  $\text{span}(X_1, \dots, X_m)(x) = T_x\mathcal{M}$  for all  $x \in \mathcal{M}$ , and the operator  $\mathcal{L}$  is elliptic. Most of the results stated in this paper (or stronger versions of them) are already known in this situation (and in greater generality), see [LR95, Leb92, LL15]. That all Laplace-Beltrami operators can be written under the form (1.1) is a consequence of Remark 1.1 below.

**Example 1.6** (The Grushin operator,  $k = 2$ ). Consider the torus  $\mathcal{M} = (\mathbb{R}/2\mathbb{Z}) \times (\mathbb{R}/\mathbb{Z})$  (which we identify with  $[-1, 1[ \times [0, 1[$  with periodicity conditions), endowed with the Lebesgue measure  $ds = dx_1 dx_2$  and

$$\mathcal{L} = -(\partial_{x_1}^2 + x_1^2 \partial_{x_2}^2) = X_1^* X_1 + X_2^* X_2, \quad \text{with} \quad X_1 = \frac{\partial}{\partial x_1}, \quad X_2 = x_1 \frac{\partial}{\partial x_2}.$$

We have  $\text{span}(X_1, X_2) = \mathbb{R}^2$  if  $x_1 \neq 0$ , but on the singular set  $x_1 = 0$ , we have  $\text{span}(X_1, X_2) = \mathbb{R}X_1$ . However, we have  $[X_1, X_2] = \frac{\partial}{\partial x_2}$ , so that  $\text{span}(X_1, X_2, [X_1, X_2]) = \mathbb{R}^2$  on the whole  $\mathcal{M}$ , and Assumption 1.1 is satisfied for  $k = 2$ . Remark that  $x_1^2$  is not analytic (not even  $C^1$ ) on the torus  $\mathcal{M}$ ; here it can be replaced e.g. by  $\sin(\pi x_1/2)^2$ , being analytic and satisfying the same Hypoelliptic property. The original Grushin example will also be discussed later with Dirichlet boundary conditions, in which case it is smooth on  $[-1, 1] \times (\mathbb{R}/\mathbb{Z})$  or  $[-1, 1] \times [0, 1]$ .

**Example 1.7** (Higher order Grushin operators,  $k \in \mathbb{N}$ ). Consider again  $\mathcal{M} = (\mathbb{R}/2\mathbb{Z}) \times (\mathbb{R}/\mathbb{Z})$ ,  $ds = dx_1 dx_2$  and, for  $\gamma \in \mathbb{N}$ , set

$$\mathcal{L}_\gamma = -(\partial_{x_1}^2 + x_1^{2\gamma} \partial_{x_2}^2) = X_1^* X_1 + X_2^* X_2, \quad \text{with} \quad X_1 = \frac{\partial}{\partial x_1}, \quad X_2 = x_1^\gamma \frac{\partial}{\partial x_2}. \quad (1.8)$$

Again,  $x_1^{2\gamma}$  may be replaced by  $\sin(\pi x_1/2)^{2\gamma}$  so that  $\mathcal{L}_\gamma$  has analytic coefficients. We have  $\text{span}(X_1, X_2) = \mathbb{R}^2$  if  $x_1 \neq 0$ , but on the singular set  $x_1 = 0$ , we have to use iterated Lie brackets: Since  $\left[ \frac{\partial}{\partial x_1}, x_1^\beta \frac{\partial}{\partial x_2} \right] = \beta x_1^{\beta-1} \frac{\partial}{\partial x_2}$  for all  $\beta \geq 1$ , we have, with  $\mathcal{F} = \{X_1, X_2\}$ , that

- $\text{Lie}^1(\mathcal{F})$  is the space spanned by  $\mathcal{F}$  in  $\mathcal{X}^\infty$ ;
- $\text{Lie}^\ell(\mathcal{F}) = \left\{ f = a \frac{\partial}{\partial x_1} + \sum_{i=\gamma-\ell-1}^\gamma b_i x_1^i \frac{\partial}{\partial x_2} \mid a, b_i \in \mathbb{R} \right\}$  for  $1 \leq \ell \leq \gamma + 1$ ;
- $\text{Lie}^\ell(\mathcal{F}) = \text{Lie}^{\gamma+1}(\mathcal{F}) = \left\{ f = a \frac{\partial}{\partial x_1} + B(x_1) \frac{\partial}{\partial x_2} \mid a \in \mathbb{R}, B \in \mathbb{R}^\gamma[X] \right\}$  if  $\ell \geq \gamma + 1$ .

Hence, for  $x = (0, x_2)$ , we have  $\text{Lie}^\ell(\mathcal{F})(x) = \mathbb{R} \frac{\partial}{\partial x_1}$  if  $\ell < \gamma + 1$  and  $\text{Lie}^{\gamma+1}(\mathcal{F})(x) = \mathbb{R}^2$ . In particular, Assumption 1.1 is fulfilled with  $k = \gamma + 1$ . Note that we recover Example 1.5 in case  $\gamma = 0$  and Example 1.6 in case  $\gamma = 1$ .

**Example 1.8** (The Heisenberg operator on the Heisenberg group). On  $\mathbb{R}^3$  with current point  $w = (x, y, s)$ , the following two vector fields  $X_1 = \partial_x + 2y\partial_s$  and  $X_2 = \partial_y - 2x\partial_s$  constitute the model case for contact geometry. Indeed, we have, with  $\mathcal{F} = \{X_1, X_2\}$ , that

- $\text{Lie}^1(\mathcal{F}) = \text{span}(\mathcal{F})$  is of dimension 2 at any point in  $\mathbb{R}^3$ ;
- $\text{Lie}^2(\mathcal{F}) = \mathbb{R}^3$  at any point in  $\mathbb{R}^3$ , since  $[X_1, X_2] = -4\partial_s$ .

Let us now define a compact context in which these are two analytic vector fields.

First equip  $\mathbb{R}^3$  with the (non-commutative) group law

$$w \bullet w' = (x, y, s) \bullet (x', y', s') = (x + x', y + y', s + s' - 2xy' + 2yx').$$

With this law,  $(\mathbb{R}^3, \bullet)$  (with  $\mathbb{R}^3$  endowed with its canonical differential structure) is a Lie group which we denote by  $G$ . Given  $L > 0$ , the set  $\Gamma = LZ \times LZ \times L^2Z$  is a subgroup of  $G$ , and both vector

fields  $X_1$  and  $X_2$  are left invariant vector fields on  $G$ , i.e. setting  $m_g : G \rightarrow G, w \mapsto g \bullet w$ , we have  $dm_g(X_j(w)) = X_j(m_g(w)) = X_j(g \bullet w)$  for  $j = 1, 2$ . The subgroup  $\Gamma$  being co-compact, the left quotient  $\mathcal{M} := \Gamma \backslash G$  is a compact three dimensional analytic manifold. Moreover, the vector fields  $X_1, X_2$  go to the quotient as analytic vector fields on  $\mathcal{M}$ . From the computation on  $\mathbb{R}^3$ , we obtain  $\dim \text{Lie}^1(\mathcal{F})(w) = 2$  and  $\text{Lie}^2(\mathcal{F})(w) = T_w \mathcal{M}$  for some/any point  $w \in \mathcal{M}$ . The Haar measure turns out to be the Euclidian measure in the coordinates  $(x, y, s)$ . We consider the operator  $\mathcal{L} = X_1^* X_1 + X_2^* X_2 = -X_1^2 - X_2^2 = -\Delta_{\mathbb{H}}$ , where  $\Delta_{\mathbb{H}}$  is the Kohn Laplacian, for which  $k = 2$ . We refer for instance to [BFKG12, Section 1.2] for more on this example.

This last example belongs to the following general class of constant rank sub-Riemannian structures.

**Example 1.9** (Lie Groups). Assume that  $(\mathcal{M}, \bullet)$  is a compact  $d$ -dimensional Lie group. Let  $\mathbb{1}$  be the identity of  $(\mathcal{M}, \bullet)$ , and write  $L := T_{\mathbb{1}} \mathcal{M}$  its Lie algebra. Recall (see e.g. [God82, Tome II, p627]) that there is a unique real-analytic differentiable structure on  $\mathcal{M}$  compatible with the action of  $\bullet$ , with which we endow  $\mathcal{M}$ . We write as in the above example  $m_g : \mathcal{M} \rightarrow \mathcal{M}, x \mapsto g \bullet x$  for the left multiplication. Given  $m < d$  and  $m$  vectors  $(e_1, \dots, e_m) \in L^m$ , we denote by  $(X_1, \dots, X_m)$  the associated  $m$  left-invariant vector fields defined, for  $x \in \mathcal{M}$ , by  $X_j(x) := dm_x(X_j(\mathbb{1})) = dm_x(e_j)$ .

Now, we assume that the vectors  $(e_1, \dots, e_m)$  generate the whole Lie algebra, namely  $\text{Lie}(\mathcal{M}) = L$ , which implies that the vector fields  $(X_1, \dots, X_m)$  satisfy Assumption 1.1, for some  $k$ .

Finally, we remark that, by construction, both the vector fields  $X_j$  and the Haar measure  $ds$  of  $\mathcal{M}$  are real-analytic and left invariant. All our results shall hence apply to the associated operator  $\mathcal{L}$ .

Finally, let us mention that hypoelliptic operators appear naturally in several physical and mathematical contexts such as stochastic processes and the theory of functions of several complex variables. We refer to [Bra14, Chapter 2] for a presentation of some of these applications.

## 1.2 Main results

Our main results under Assumptions 1.1 and 1.2 are of three different types:

1. Tunneling estimates for eigenfunctions  $\varphi_j$  of  $\mathcal{L}$  (Section 1.2.1);
2. Quantitative approximate observability (and associated controllability) of the hypoelliptic wave equation  $(\partial_t^2 + \mathcal{L})v = 0$  from a subset  $\omega \subset \mathcal{M}$  (Section 1.2.2);
3. Quantitative approximate observability (and associated controllability) of the hypoelliptic heat equation  $(\partial_t + \mathcal{L})u = 0$  from  $\omega$  (Section 1.2.3);.

Also, we provide with a class of examples (which are generalizations of those considered in Example (1.7)) where all these results hold as well without the analyticity Assumption 1.2 (Section 1.2.4).

All results obtained depend explicitly on the hypoellipticity index  $k$  of the operator considered, i.e. the minimal number of iterated brackets necessary to span the whole tangent space, given by Assumption 1.1. We finally prove with an example that the results are optimal in general.

### 1.2.1 Eigenfunction tunneling

Our first result is the following.



**Theorem 1.10.** *Let  $\omega$  be a nonempty open subset of  $\mathcal{M}$ . Then, there is  $C, c > 0$  such that every eigenfunction  $\varphi_i$  of  $\mathcal{L}$  associated to the eigenvalue  $\lambda_i$  satisfies*

$$\|\varphi_j\|_{L^2(\mathcal{M})} \leq C e^{c\lambda_j^{k/2}} \|\varphi_j\|_{L^2(\omega)}. \quad (1.9)$$

This estimate may be read as  $\|\varphi_j\|_{L^2(\omega)} \geq \frac{1}{C} e^{-c\lambda_j^{k/2}}$  for all normalized eigenfunctions, and hence quantizes the possible vanishing rate of eigenfunctions on any subdomain  $\omega$ .

In the case  $k = 1$ , i.e. when  $\mathcal{L}$  is an elliptic operator, the analyticity assumption 1.2 is not needed and the result follows from the Donnelly-Fefferman paper [DF88]. In this situation, it also holds on a manifold with boundary for Dirichlet eigenfunctions [LR95] (see also [LR97] for other boundary conditions).

We shall also deduce from estimates of [BCG14, Section 2.3] that the tunneling estimate (1.9) is optimal in the following particular setting (close to Example (1.7)).

**Example 1.11** (Higher order Grushin operators on the rectangle). Consider the manifold *with boundary*  $\mathcal{M} = [-1, 1] \times [0, 1]$  or  $\mathcal{M} = [-1, 1] \times (\mathbb{R}/\mathbb{Z})$ , endowed with the Lebesgue measure  $dx$ , and for  $\gamma > 0$ , define the operator  $\mathcal{L}_\gamma = -(\partial_{x_1}^2 + x_1^{2\gamma} \partial_{x_2}^2)$  as in (1.8) with Dirichlet conditions on  $\partial\mathcal{M}$ . If  $\gamma \in \mathbb{N}$ , then the operator  $\mathcal{L}_\gamma$  is hypoelliptic of order  $k = \gamma + 1$  (i.e. Assumption 1.1 is fulfilled with  $k = \gamma + 1$ ).

**Proposition 1.12.** *Consider, for  $\gamma > 0$  the situation of Example 1.11. Assume that  $\bar{\omega} \cap \{x_1 = 0\} = \emptyset$ . Then there exists  $C, c_0 > 0$  and a sequence  $(\lambda_j, \varphi_j)$  of eigenvalues and associated eigenfunctions of  $\mathcal{L}_\gamma$  with  $\lambda_j \rightarrow +\infty$  such that*

$$\|\varphi_j\|_{L^2(\omega)} \leq C e^{-c_0 \lambda_j^{\frac{\gamma+1}{2}}} \|\varphi_j\|_{L^2(\mathcal{M})}.$$

We recall that if  $\gamma \in \mathbb{N}^*$ , then  $\mathcal{L}_\gamma$  is hypoelliptic of order  $k = \gamma + 1$ , so that Proposition 1.12 shows that, in general, one cannot expect a better estimate than that of Theorem 1.10. We shall also prove (see Section 1.2.4) that Estimate (1.9) holds as well in a setting containing those of Example 1.7 and Example 1.11, thus providing a genuine converse of Proposition 1.12 (for  $\gamma \in \mathbb{N}^*$ ).

Note that in the analytic context, the *qualitative* uniqueness:

$$\left( \mathcal{L}\varphi = \lambda\varphi \text{ on } \mathcal{M}, \quad \varphi = 0 \text{ on } \omega \right) \implies \varphi \equiv 0 \text{ on } \mathcal{M},$$

was proved by Bony [Bon69], as a consequence of the Holmgren-John theorem. Removing the analyticity assumption, even for such a qualitative unique continuation property, remains a very subtle issue, as discussed in Section 1.3.1 below.

## 1.2.2 Quantitative approximate observability of the hypoelliptic wave equation

To state our main result here, we need to introduce the appropriate notions of Sobolev spaces and sub-Riemannian distance, which are adapted to the analysis of the operator  $\mathcal{L}$ .

All along the paper, we shall use the functional calculus given, for appropriate functions  $f$  and  $u$ , by

$$f(\mathcal{L})u = \sum_{j \in \mathbb{N}} f(\lambda_j)(u, \varphi_j)_{L^2(\mathcal{M})} \varphi_j. \quad (1.10)$$

This allows for instance to define the operators  $(1 + \mathcal{L})^{\frac{s}{2}} : C^\infty(\mathcal{M}) \rightarrow C^\infty(\mathcal{M})$ , which, by duality, may be extended as operators  $(1 + \mathcal{L})^{\frac{s}{2}} : \mathcal{D}'(\mathcal{M}) \rightarrow \mathcal{D}'(\mathcal{M})$ . We next define the Sobolev spaces

$$\mathcal{H}_{\mathcal{L}}^s = \{u \in \mathcal{D}'(\mathcal{M}), (1 + \mathcal{L})^{\frac{s}{2}} u \in L^2(\mathcal{M})\}, \quad s \in \mathbb{R},$$

and associated norms

$$\|u\|_{\mathcal{H}_{\mathcal{L}}^s} = \|(1 + \mathcal{L})^{\frac{s}{2}} u\|_{L^2(\mathcal{M})}, \quad s \in \mathbb{R}.$$

Let us now also introduce basic notions of sub-Riemannian geometry needed to formulate our main result. We refer to [Bel96, Mon02, Jea14, Rif14, ABB16a, ABB16b] for an introductions to sub-Riemannian geometry, as well as for further developments. The so-called sub-Riemannian metric associated to the vector fields  $(X_1, \dots, X_m)$  is defined, for  $x \in \mathcal{M}$  and  $v \in T_x \mathcal{M}$ , by

$$g(x, v) := \begin{cases} \inf \left\{ \sum_{i=1}^m u_i^2 \mid (u_1, \dots, u_m) \in \mathbb{R}^m, \sum_{i=1}^m u_i X_i(x) = v \right\} & \text{if } v \in \text{span}(X_i(x), i \in \{1, \dots, m\}), \\ +\infty & \text{if not.} \end{cases} \quad (1.11)$$

This defines for any  $x \in \mathcal{M}$  a positive definite quadratic form  $g(x, \cdot)$  on the the horizontal space

$$\text{span}(X_1(x), \dots, X_m(x)).$$

Remark that, if finite, the infimum is in fact a minimum, and is realized by a unique vector  $(u_1, \dots, u_m) \in \mathbb{R}^m$ . Given  $\gamma : [0, 1] \rightarrow \mathcal{M}$  an absolutely continuous path, we define its length accordingly by

$$\text{length}(\gamma) := \int_0^1 \sqrt{g(\gamma(t), \dot{\gamma}(t))} dt.$$

The fact that this quantity is finite implies that  $\dot{\gamma}(t) \in \text{span}(X_1(\gamma(t)), \dots, X_m(\gamma(t)))$  for almost all  $t \in [0, 1]$ . Also, it is always finite if  $\gamma$  is a horizontal path (in the sense of Definition 1.3).

Then, this allows to define a sub-Riemannian (also called Carnot-Carathéodory) distance on  $\mathcal{M}$  by

$$d_{\mathcal{L}}(x_0, x_1) = \inf \{ \text{length}(\gamma) \mid \gamma \text{ horizontal path, } \gamma(0) = x_0, \gamma(1) = x_1 \}.$$

The Chow-Rashevski Theorem 1.2 implies that, under Assumption 1.1, the distance defined by  $d_{\mathcal{L}}$  is always finite.

With these definitions in hand, we may now state our main result, which concerns the quantitative unique continuation (or quantitative approximate observability) for the Hypoelliptic wave equation

$$\begin{cases} \partial_t^2 u + \mathcal{L}u &= 0 \\ (u, \partial_t u)|_{t=0} &= (u_0, u_1). \end{cases} \quad (1.12)$$

**Theorem 1.13.** *Let  $\mathcal{L}$  as above satisfying Assumptions 1.1 and 1.2. Assume that  $\omega$  is a non empty open set of  $\mathcal{M}$  and let  $T > \sup_{x \in \mathcal{M}} d_{\mathcal{L}}(x, \omega)$ . Then, there exist  $\kappa, C, \mu_0 > 0$  such that we have*

$$\|(u_0, u_1)\|_{L^2 \times \mathcal{H}_{\mathcal{L}}^{-1}} \leq C e^{\kappa \mu^k} \|u\|_{L^2([ -T, T ] \times \omega)} + \frac{1}{\mu} \|(u_0, u_1)\|_{\mathcal{H}_{\mathcal{L}}^1 \times L^2} \quad (1.13)$$

for all  $\mu \geq \mu_0$  and for any  $(u_0, u_1) \in \mathcal{H}_{\mathcal{L}}^1 \times L^2$ , and associated solution  $u$  solution of (1.12) on  $] -T, T[$ .

Note first that this estimate could be stated equivalently for all  $\mu > 0$  (see e.g. [LL15, Lemma A.3]). We chose to keep the above formulation to underline the interesting case (being  $\mu$  large). Note also that this theorem can be equivalently rewritten under one of the following two formulations (see e.g. [LL15, Lemma A.3])

$$\|(u_0, u_1)\|_{\mathcal{H}_{\mathcal{L}}^1 \times L^2} \leq C e^{c\Lambda^k} \|u\|_{L^2((0, T) \times \omega)}, \quad \text{with } \Lambda = \frac{\|(u_0, u_1)\|_{\mathcal{H}_{\mathcal{L}}^1 \times L^2}}{\|(u_0, u_1)\|_{L^2 \times \mathcal{H}_{\mathcal{L}}^{-1}}}, \quad (1.14)$$

or

$$\|(u_0, u_1)\|_{L^2 \times \mathcal{H}_{\mathcal{L}}^{-1}} \leq C \frac{\|(u_0, u_1)\|_{\mathcal{H}_{\mathcal{L}}^1 \times L^2}}{\log \left( \frac{\|(u_0, u_1)\|_{\mathcal{H}_{\mathcal{L}}^1 \times L^2}}{\|u\|_{L^2((0, T) \times \omega)}} + 1 \right)^{\frac{1}{k}}}, \quad (1.15)$$

where, in the last expression, the function  $x \mapsto (\log(1 + \frac{1}{x}))^{-\frac{1}{k}}$  has to be extended by zero at  $x = 0^+$ .

Again, in the particular situation of Example 1.11, i.e. for the operators (1.8), the sequence of eigenfunctions of Proposition 1.12 shows that the exponent  $e^{\kappa\mu^k}$  in (1.13) (resp.  $e^{c\Lambda^k}$  in (1.14) and  $\log^{-\frac{1}{k}}$  in (1.15)) cannot be improved in general.

**Remark 1.14.** That  $d_{\mathcal{L}}$  is the relevant distance function in view of the study of the Hypocoelliptic partial differential operator  $\mathcal{L}$  comes from the fact that the sub-Riemannian metric  $g(x, v)$  and the principal symbol  $\ell(x, \xi)$  of the operator  $\mathcal{L}$  are linked through the Legendre transform  $\frac{1}{2}g(x, v) = \max_{\xi \in T_x^* \mathcal{M}} (\langle \xi, v \rangle - \frac{1}{2}\ell(x, \xi))$  (see [Bel96, Section 1.2] or Appendix C):  $d_{\mathcal{L}}$  is thus the appropriate distance when analyzing properties of  $\mathcal{L}$ .

Moreover, the assumption on the time  $T > \sup_{x \in \mathcal{M}} d_{\mathcal{L}}(x, \omega)$  is optimal because of the finite speed of propagation satisfied by equation (1.12). Indeed, Hypocoelliptic wave equations also satisfy a finite speed of propagation similar to the classical wave equation but with the Riemannian distance replaced by the sub-Riemannian distance  $d_{\mathcal{L}}$ . This (not obvious) fact was proved by Melrose [Mel86] (see also the remarks in [JSC87, Section 4] for the link between the distance defined in [Mel86] and  $d_{\mathcal{L}}$ ).

As a corollary of this result (see [Rob95] or [LL17, Appendix]), we obtain the approximate controllability of the Hypocoelliptic wave equation, as well as an estimate of the cost of approximate controls. Here, we only state approximate controllability to zero, which is equivalent to approximate controllability to the whole state space  $\mathcal{H}_{\mathcal{L}}^1 \times L^2$  on account to the reversibility of the equation.

**Corollary 1.15** (Cost of approximate control). *For any  $T > 2 \sup_{x \in \mathcal{M}} d_{\mathcal{L}}(x, \omega)$ , there exist  $C, c > 0$  such that for any  $\varepsilon > 0$  and any  $(u_0, u_1) \in \mathcal{H}_{\mathcal{L}}^1 \times L^2$ , there exists  $g \in L^2((0, T) \times \omega)$  with*

$$\|g\|_{L^2((0, T) \times \omega)} \leq C e^{\frac{c}{\varepsilon^k}} \|(u_0, u_1)\|_{\mathcal{H}_{\mathcal{L}}^1 \times L^2},$$

such that the solution of

$$\begin{cases} (\partial_t^2 + \mathcal{L})u = \mathbb{1}_{\omega} g & \text{in } (0, T) \times \mathcal{M}, \\ (u, \partial_t u)|_{t=0} = (u_0, u_1), & \text{in } \mathcal{M}, \end{cases}$$

satisfies  $\|(u, \partial_t u)|_{t=T}\|_{L^2 \times \mathcal{H}_{\mathcal{L}}^{-1}} \leq \varepsilon \|(u_0, u_1)\|_{\mathcal{H}_{\mathcal{L}}^1 \times L^2}$ .

To the authors' knowledge, these results are the first ones concerning the approximate observability/controllability of hypoelliptic waves. They furnish not only the approximate observability/controllability but also an (optimal in general) estimate of the cost.

In the elliptic case  $k = 1$ , these can be obtained by the theory developed by Lebeau in [Leb92] (even on a manifold with boundary). However, in this (elliptic) case, the analyticity assumption can be removed, as proved by the authors in [LL15]. This followed a long series of papers concerning the *qualitative* unique continuation [RT73, Ler88, Rob91, Hör92, Tat95] (see also [RZ98, Hör97, Tat99b] for more general operators), i.e. the property:

$$\left( (\partial_t^2 + \mathcal{L})u = 0 \text{ on } (-T, T) \times \mathcal{M}, \quad u = 0 \text{ on } (-T, T) \times \omega \right) \implies u \equiv 0,$$

and another series of papers [Rob95, Phu10, Tat99a] concerning variants of Estimate (1.13) (still in the elliptic case  $k = 1$ ) which are not optimal with respect to the minimal time and the exponent of  $\mu$ . We refer to the introductions of [LL15, LL16] for a more detailed discussion on this issue. Here, in the analytic context, we directly prove the quantitative result but, to our knowledge, even the qualitative result was not known.

We shall see that we prove actually a more general statement in which the term  $\|(u_0, u_1)\|_{\mathcal{H}_x^1 \times L^2}$  in the right-handside of Estimate (1.13) can be changed into  $\|(u_0, u_1)\|_{\mathcal{H}_x^s \times \mathcal{H}_x^{s-1}}$  for any  $s > 0$ , if changing the power of  $\mu$  accordingly, see Theorem 1.31 below.

### 1.2.3 Quantitative approximate observability of the hypoelliptic heat equation

We now turn to the study of observability properties for solutions of the hypoelliptic heat equation

$$\begin{cases} \partial_t y + \mathcal{L}y = 0, & \text{in } (0, T) \times \mathcal{M}, \\ y(0) = y_0 & \text{in } \mathcal{M}, \end{cases} \quad (1.16)$$

from a subdomain  $\omega \subset \mathcal{M}$ . By duality, we are equivalently concerned here with different controllability properties of the following system

$$\begin{cases} (\partial_t + \mathcal{L})u = \mathbb{1}_\omega g, & \text{in } (0, T) \times \mathcal{M}, \\ u(0) = u_0, & \text{in } \mathcal{M}. \end{cases} \quad (1.17)$$

We provide with three main results, still under Assumptions 1.1 and 1.2:

1. For any  $k \in \mathbb{N}^*$ , we prove an approximate observability result in any time  $T > 0$  with a frequency-dependent constant of order  $Ce^{\Lambda^k}$ , where  $\Lambda = \frac{\|y_0\|_{\mathcal{H}_x^1}}{\|y_0\|_{L^2}}$ , or, equivalently, approximate controllability with cost  $e^{\frac{C}{\varepsilon^k}}$ . These are the analogues of Theorem 1.13 and Corollary 1.15 for parabolic equations.
2. If we moreover assume the data to be sufficiently smooth (in some Gevrey-type norm with respect to the spectral decomposition of  $\mathcal{L}$ ), then the cost of approximate observability can be improved to a polynomial one, i.e. of the form  $\frac{C}{\varepsilon^\beta}$  for some  $\beta > 0$ ; this yields approximate controllability in a much weaker topology, but with a much lower cost.
3. Finally, in the very particular case  $k = 2$  (including Grushin and Heisenberg operators), we prove an approximate observability/controllability property *to trajectories* in large time with a polynomial cost. This may be interpreted as a counterpart of the exact controllability to trajectories for the heat equation [LR95, F196] (case  $k = 1$ ). There is no similar result if  $k > 2$ .

The first result we obtain provides the cost of approximate observability of the whole state space  $L^2(\mathcal{M})$ . There is no restriction for the hypoellipticity index  $k$ , but the (exponential) cost depends on this parameter.

**Theorem 1.16.** *For all  $T > 0$ , there exist  $C, c > 0$  such that for any  $y_0 \in \mathcal{H}_{\mathcal{L}}^1$  and associated solution  $y$  of (1.16), we have*

$$\|y_0\|_{L^2}^2 \leq C e^{c\Lambda^k} \int_0^T \int_{\omega} |y(t, x)|^2 dx dt, \quad \Lambda = \frac{\|y_0\|_{\mathcal{H}_{\mathcal{L}}^1}}{\|y_0\|_{L^2}}, \quad (1.18)$$

and, for any  $\mu > 0$ ,

$$\|y_0\|_{L^2}^2 \leq C e^{c\mu^k} \int_0^T \int_{\omega} |y(t, x)|^2 dx dt + \frac{1}{\mu^2} \|y_0\|_{\mathcal{H}_{\mathcal{L}}^1}^2. \quad (1.19)$$

That (1.18) and (1.19) are equivalent comes for instance from [LL15, Lemma A.3]. Again, in the particular situation of Example 1.11, i.e. for the operators (1.8), the sequence of eigenfunctions of Proposition 1.12 shows that the exponent  $e^{c\mu^k}$  in (1.19) (resp.  $e^{c\Lambda^k}$  in (1.18)) cannot be improved in general.

This theorem generalizes the results of Fernandez-Cara-Zuazua and Phung [FCZ00, Phu04] in the elliptic case  $k = 1$ . Yet, in this framework, the analyticity was not necessary (as in all above stated results in the case  $k = 1$ ) and the setting can be relaxed (uniform dependence of the constants with respect to lower order terms and to the time  $T$ , boundary value problems...).

As a corollary (see [LL17, Appendix]), we obtain, given an initial state and a target state both belonging to the space  $L^2(\mathcal{M})$ , and given a precision  $\varepsilon$ , the existence of a control function bringing the initial state in an  $\varepsilon$ -neighborhood of the target (in appropriate topology). We obtain as well an estimate of the cost of the control.

**Corollary 1.17** (Cost of approximate control to the state space). *For any  $T > 0$ , there exist  $C, c > 0$  such that for any  $\varepsilon > 0$  and any  $u_0 \in L^2(\mathcal{M}), u_1 \in L^2(\mathcal{M})$ , there exists  $g \in L^2((0, T) \times \omega)$  with*

$$\|g\|_{L^2((0, T) \times \omega)} \leq C e^{\frac{c}{\varepsilon^k}} \|e^{-T\mathcal{L}}u_0 - u_1\|_{L^2(\mathcal{M})},$$

such that the solution of (1.17) issued from  $u_0$  satisfies

$$\|u(T) - u_1\|_{\mathcal{H}_{\mathcal{L}}^{-1}} \leq \varepsilon \|e^{-T\mathcal{L}}u_0 - u_1\|_{L^2(\mathcal{M})}.$$

To state our second result concerning the hypoelliptic heat equation, we need to introduce the following spectral Gevrey-type norms for functions defined on  $\mathcal{M}$ : For  $\alpha > 0, \theta \in \mathbb{R}$ , we set

$$\|u\|_{\alpha, \theta}^2 = \sum_{j \in \mathbb{N}} e^{2\theta\lambda_j^\alpha} |u_j|^2, \quad \text{with } u = \sum_{j \in \mathbb{N}} u_j \varphi_j. \quad (1.20)$$

For  $\theta \geq 0$ , we define  $H^{\alpha, \theta}$  to be the subspace of  $L^2(\mathcal{M})$  consisting in functions  $u$  such that  $\|u\|_{\alpha, \theta} < \infty$ . For  $\theta < 0$ , we let  $H^{\alpha, \theta}$  be the completed of linear combinations of eigenfunctions for this norm. Remark that taking as usual  $L^2(\mathcal{M})$  as a pivot space, the space  $H^{\alpha, -\theta}$  is identified to  $(H^{\alpha, \theta})'$  for  $\theta \geq 0$ . Also, according to the hypoellipticity Assumption 1.1 (see Corollary B.2) we have  $H^{\alpha, \theta} \subset C^\infty(\mathcal{M})$  for  $\theta > 0$ , so that  $H^{\alpha, -\theta}$  is larger than spaces of distributions on  $\mathcal{M}$  (and its topology weaker).

We obtain the following result, which assumes the data to be extremely regular and then yields approximate observability with a polynomial cost only. The regularity needed is linked to the hypoellipticity index  $k$ .

**Theorem 1.18.** Fix any  $k \in \mathbb{N}^*$ . There exists  $\theta_0 > 0$  such that for any  $T > 0$  and any  $\theta > \theta_0$ , there exist  $C > 0$  so that for  $\varepsilon > 0$ , we have for any solution  $y$  to (1.16),

$$\|y_0\|_{L^2}^2 \leq \frac{C}{\varepsilon^{\frac{\theta_0}{\theta-\theta_0}}} \int_{T/2}^T \int_{\omega} |y(t, x)|^2 dt dx + \varepsilon \|y(0)\|_{k/2, \theta}^2. \quad (1.21)$$

Again Proposition 1.12 shows that a polynomial cost is optimal for data in  $H^{\theta, \frac{k}{2}}$  in the situation of Example 1.11. Note that we provide with an explicit dependence of the polynomial cost (i.e. the power  $\frac{\theta_0}{\theta-\theta_0}$ ) with respect to the regularity of the data; in particular, we see how it improves when  $\theta$  becomes larger. As a Corollary of Theorem 1.18, we obtain an approximate controllability result with polynomial cost, but the target is well approximated in a very weak topology.

**Corollary 1.19** (Cost of approximate control to the state space in very weak topology). *With  $\theta_0 > 0$  given as in Theorem 1.18 (depending only on  $\mathcal{M}, \omega, \mathcal{L}$ ), for any  $T > 0$  and  $\theta > \theta_0$ , there exist  $C > 0$  such that for any  $\varepsilon > 0$  and any  $u_0 \in L^2(\mathcal{M}), u_1 \in L^2(\mathcal{M})$ , there exists  $g \in L^2((0, T) \times \omega)$  with*

$$\|g\|_{L^2((0, T) \times \omega)} \leq \frac{C}{\varepsilon^{\frac{\theta_0}{\theta-\theta_0}}} \|e^{-T\mathcal{L}}u_0 - u_1\|_{L^2(\mathcal{M})},$$

such that the solution of (1.17) issued from  $u_0$  satisfies

$$\|u(T) - u_1\|_{k/2, -\theta} \leq \varepsilon \|e^{-T\mathcal{L}}u_0 - u_1\|_{L^2(\mathcal{M})}.$$

We are not aware of any such results, even for the usual heat equation (i.e. with  $k = 1$ ). In this case, our proof also works in the  $C^\infty$  context, and in the presence of boundaries, starting from the estimates obtained in [LL15] or the spectral estimates of [LR95].

Our last main result concerning the hypoelliptic heat equation is, as opposed to the first two ones, concerned with final state approximate observability (or equivalently an approximate controllability to trajectories) with a *polynomial* cost, and is restricted to the case  $k = 2$ .

**Theorem 1.20.** Assume that  $k = 2$ . There exist  $T_0, C > 0$  such that for all  $\eta > 0$ , all  $T > T_0 + \eta$  and all  $\varepsilon > 0$ , we have for any solution  $y$  to (1.16),

$$D \|y(T)\|_{L^2}^2 \leq \frac{1}{\varepsilon^\beta} \int_{T-\eta}^T \int_{\omega} |y(t, x)|^2 dt dx + \varepsilon \|y(0)\|_{L^2}^2, \quad (1.22)$$

with  $D = \min\{\frac{e^{-C/\eta}}{C}, 1\}$  and  $\beta = \frac{T_0}{T-(T_0+\eta)}$ .

In particular, we obtain an explicit estimate on how the cost improves as  $T$  increases. This result gives directly the following corollary concerning approximate controllability to trajectories (or, equivalently, to zero) at a polynomial cost (see again [LL17, Appendix]).

**Corollary 1.21** (Cost of approximate control to trajectories if  $k = 2$ ). *Assume that  $k = 2$ , and let  $T_0 > 0$  as in Theorem 1.20. For all  $\eta > 0$ , all  $T > T_0 + \eta$  and all  $\varepsilon > 0$ , we have the following statement: for any  $u_0, \tilde{u}_0 \in L^2$ , there exists  $g \in L^2((0, T) \times \omega)$  with*

$$\|g\|_{L^2((0, T) \times \omega)} \leq \frac{\tilde{C}}{\varepsilon^\beta} \|u_0 - \tilde{u}_0\|_{L^2},$$

such that the associated solution  $u$  of (1.17) satisfies

$$\|u(T) - e^{-T\mathcal{L}}\tilde{u}_0\|_{L^2(\mathcal{M})} \leq \varepsilon \|u_0 - \tilde{u}_0\|_{L^2},$$

where  $\beta = \frac{T_0}{T - (T_0 + \eta)}$  and  $\tilde{C} = \tilde{C}(\eta, T_0, T)$ .

Remark that these two results only hold in the case  $k = 2$ . Once again, in the particular situation of Example 1.11, i.e. for the operators (1.8), the sequence of eigenfunctions of Proposition 1.12 shows that this result cannot hold if  $k \geq 3$  (or even, if  $\gamma > 1$ ). Let us here be more precise: assume in the context of Example 1.11 that an estimate of the form (1.22) is satisfied for a cost function  $\Phi(\varepsilon)$ , that is

$$\|y(T)\|_{L^2}^2 \leq \Phi(\varepsilon) \int_0^T \int_{\omega} |y(t, x)|^2 dt dx + \varepsilon \|y(0)\|_{L^2}^2, \quad \text{for all } \varepsilon > 0,$$

and test it with  $y(t) = e^{-\lambda_j t} \varphi_j$  (solution of (1.16)), where  $\varphi_j$  is given by Proposition 1.12. Then we have for all  $\varepsilon > 0$ ,

$$e^{-2\lambda_j T} \leq \Phi(\varepsilon) \frac{e^{-2c_0 \lambda_j^{k/2}}}{\lambda_j} + \varepsilon, \quad \text{for all } \varepsilon > 0.$$

Fixing then  $\varepsilon = \varepsilon_j := \frac{e^{-2\lambda_j T}}{2} \rightarrow 0^+$  and taking logarithm yields  $2(c_0 \lambda_j^{k/2} - \lambda_j T) \leq \log \Phi(\varepsilon_j)$ . That  $\Phi(\varepsilon) \leq C/\varepsilon^\beta$  (i.e. having a polynomial cost) implies  $k \leq 2$ . In the case  $k = 2$ , this implies  $\Phi(\varepsilon_j) \geq \frac{1}{(2\varepsilon_j)^{\frac{c_0}{T}-1}}$ , so that for  $T < c_0$ , the polynomial cost cannot be improved. Unfortunately, the constant  $T_0$  in the above result is much larger than  $c_0$ , so that this discussion does not imply neither that the polynomial cost is the optimal one, nor that a minimal time is necessary. However, as we shall see in Section 1.3.2 below, it may happen that exact controllability holds for no time  $T > 0$ , which may indicate that a polynomial cost is not far from being sharp.

#### 1.2.4 Relaxing the analyticity assumption

In this section, we provide with a simple family of examples for which the analyticity Assumption 1.2 can be partially removed. Still this family contains those of Examples 1.7 and 1.11. In this context, most above results hold as well. The motivation is both to relax the analyticity assumption (and replace it with analyticity with respect to one set of variables), and to include the setting of the article [BCG14].

**Example 1.22** (Partially analytic Grushin-type operators). Consider the manifold with boundary  $\mathcal{M} = [-1, 1] \times (\mathbb{R}/\mathbb{Z})$ , endowed with the Lebesgue measure  $dx$ , and define, for  $f \in C^\infty([-1, 1] \times (\mathbb{R}/\mathbb{Z}))$ , the Grushin type operator

$$\mathcal{L} = X_1^* X_1 + X_2^* X_2, \quad X_1 = \partial_{x_1}, \quad X_2 = f(x_1, x_2) \partial_{x_2}, \quad (1.23)$$

that is

$$\mathcal{L} = -\partial_{x_1}^2 - f^2 \partial_{x_2}^2 - (2f \partial_{x_2} f) \partial_{x_1},$$

with Dirichlet conditions on  $\partial\mathcal{M}$ . We further assume that

- $f(x_1, x_2)$  is *analytic* in the variable  $x_2$  (that is, for any point  $x = (x_1, x_2) \in ]-1, 1[ \times (\mathbb{R}/\mathbb{Z})$ ,  $f$  is equal to its partial Taylor expansion at  $x_2$  with respect to the variable  $x_2$  uniformly in a neighborhood of  $x$  in  $] -1, 1[ \times (\mathbb{R}/\mathbb{Z})$ );

- there exists  $\varepsilon > 0$  such that  $f(x_1, x_2) = f(x_1)$  does not depend on  $x_2$ , and  $f(x_1) > 0$  for all  $x_1 \in [-1, -1 + \varepsilon] \cup [1 - \varepsilon, 1]$ ;
- $X_1$  and  $X_2$  satisfy the Chow-Rashevski-Hörmander Assumption 1.1;
- $f(x_1, x_2)$  does not depend on  $x_2$  in a neighborhood of  $\omega$ .

Note that under these assumptions, the operator  $\mathcal{L}$  is elliptic near the boundary  $\partial\mathcal{M}$ .

For instance if  $f(x_1, x_2) = f(x_1) \in C^\infty([-1, 1])$  does not depend on the variable  $x_2$ , and we have :

$$f(x_1) \neq 0, \text{ for } x_1 \neq 0 \quad f^{(\alpha)}(0) = 0, \text{ for all } \alpha \leq k - 2, \quad \text{and } f^{(k-1)}(0) \neq 0,$$

then, the operator  $\mathcal{L}$  defined by (1.23), namely  $\mathcal{L} = -\partial_{x_1}^2 - f(x_1)^2 \partial_{x_2}^2$ , satisfies all assumptions of Example 1.22 (in particular, it is hypoelliptic of order  $k$ ). This contains the situation of Examples 1.7 and 1.11 (for  $\gamma \in \mathbb{N}$  of course).

We prove the following result.

**Theorem 1.23.** *In the context of Example 1.22, all results of Theorems 1.10, 1.18 and 1.20 still hold, as well as their corollaries.*

*Theorem 1.13 is true with the following estimate instead*

$$\|(u_0, u_1)\|_{L^2 \times \mathcal{H}_{\mathcal{L}}^{-1}} \leq C e^{\kappa\mu} \|u\|_{L^2(\cdot, -T, T[\times\omega])} + \frac{1}{\mu} \|(u_0, u_1)\|_{\mathcal{H}_{\mathcal{L}}^k \times \mathcal{H}_{\mathcal{L}}^{k-1}}. \quad (1.24)$$

*Theorem 1.16 is still true but with the estimates*

$$\|y_0\|_{L^2}^2 \leq C e^{c\Lambda_k} \int_0^T \int_{\omega} |y(t, x)|^2 dx dt, \quad \Lambda_k = \frac{\|y_0\|_{\mathcal{H}_{\mathcal{L}}^k}}{\|y_0\|_{L^2}}, \quad (1.25)$$

$$\|y_0\|_{L^2}^2 \leq C e^{c\mu} \int_0^T \int_{\omega} |y(t, x)|^2 dx dt + \frac{1}{\mu^2} \|y_0\|_{\mathcal{H}_{\mathcal{L}}^k}^2. \quad (1.26)$$

Note that since  $f$  does not vanish near  $\partial\mathcal{M}$ , the metric  $g$  defined as in (1.11) is Riemannian near  $\partial\mathcal{M}$  and the notions of length and distance defined above can be extended up to the boundary.

We explain in Section 5 how the proofs in the completely analytic case need to be modified. The version that are true in the partially analytic are actually some particular cases of general estimates with all Sobolev scales for measuring the typical frequency. For instance, (1.24) is a particular case of Theorem 1.31 with  $s = k$ . We refer to the discussion in Section 1.5 below.

**Remark 1.24.** Under appropriate assumption on  $f$ , it is classical to extend Theorem 1.23 to the same situation as in Example 1.22, but on the domain  $\mathcal{M} = [-1, 1]_{x_1} \times [-1, 1]_{x_2}$  with Dirichlet boundary conditions by using symmetry arguments. Also, the case of the domain  $\mathcal{M} = (\mathbb{R}/\mathbb{Z})^2$  is simpler.

**Remark 1.25.** All observability results of Theorem 1.23 also hold if the internal observation term  $\|u\|_{L^2(\cdot, -T, T[\times\omega])}$  is replaced by a boundary observation  $\|\partial_n u\|_{L^2(\cdot, -T, T[\times\Gamma])}$ , where  $\Gamma$  is a nonempty open subset of  $\partial\mathcal{M}$  and  $\partial_n$  denotes the normal derivative to  $\partial\mathcal{M}$ . See [LL15, Section 5]. In turn, they imply their boundary controllability counterparts. We did not state these results for the sake of brevity.



### 1.3 Comparison to other works

#### 1.3.1 Previous works on unique continuation for sum-of-squares operators

Observability inequalities (as those provided by Theorems 1.10, 1.13, 1.16, 1.18 and 1.20 above, or as (1.27) below) are quantitative estimates of the unique continuation property for the operator involved (namely  $\mathcal{L} - \lambda$ ,  $\partial_t^2 + \mathcal{L}$ , and  $\partial_t + \mathcal{L}$  respectively). Hence, when studying such inequalities, it is natural to compare our results with the known unique continuation properties for such operators. When the ellipticity condition is dropped, i.e., when  $k > 1$ , this property seems to be a very intricate problem, even under the simplest form

$$(\mathcal{L}u = 0 \text{ on } \mathcal{M}, \quad u = 0 \text{ on } \omega) \implies u = 0 \text{ in a neighborhood of } \omega \text{ in } \mathcal{M}.$$

To our knowledge, the most general such result was proved by Bony [Bon69], and holds under both the Chow-Rashevski-Hörmander condition and the assumption that the coefficients of the operator are *analytic*. Therefore, our assumptions 1.1 and 1.2 (except in the partially analytic case of Theorem 1.23) are essentially the same as in this paper. In particular, Theorem 1.10 could be read as quantification of Bony’s result. Also, the proof of the result of Bony mainly relies on the Holmgren-John theorem and is quite indirect. Here, we need to make a new proof of his result, that we are also able to quantify.

Some attempts have been done to relax this analyticity assumption. Watanabe [Wat82] proved the unique continuation property for  $C^\infty$  coefficients in dimension  $d = 2$ . Yet, later on, Bahouri [Bah86] proved a surprising general non-uniqueness result: for a large class of sum-of-squares operators  $\mathcal{L}$  with  $C^\infty$  coefficients, and satisfying Assumption 1.1, there is  $C^\infty$  potentials  $V$  such that  $\mathcal{L} + V$  does not satisfy the local unique continuation property. These counterexamples to unique continuation contain for instance in dimension  $d = 3$  and  $d = 4$  the case where the dimension spanned by the vector fields is of dimension  $d - 1$  (Heisenberg-like situations). Moreover, this result suggests that a classical Carleman estimate approach cannot work for all hypoelliptic operators. Also, it strongly suggest that the (complete or partial) analyticity assumptions that we make are not completely artificial.

This analyticity assumptions might be completely removed in some specific situations where the operator is elliptic outside of a submanifold, see the comments of Bahouri [Bah86, p140]. This was proved in the paper [Gar93] by Garofalo for specific examples. Colombini-Del Santo-Zuily [CDSZ93] also treated some related classes of degenerate elliptic operators having a specific form with respect to a hypersurface. Nevertheless, even in these situations, the quantitative estimates that we obtain are optimal as stated in Proposition 1.12.

All these results are concerned with the unique continuation property for operators like  $\mathcal{L}$  (“degenerate elliptic operators”). We are not aware of works studying the unique continuation property for operators like  $\partial_t^2 + \mathcal{L}$  or  $\partial_t + \mathcal{L}$  (“hyperbolic, resp. parabolic operators with a degenerate elliptic part”), except in the context of control theory, that we review in the next section.

#### 1.3.2 Previous works on the controllability of the hypoelliptic heat equation

The investigation of the controllability of hypoelliptic operators is for the moment quite at an early stage and has been mainly restricted to some specific operators or classes of operators. A striking result concerning the parabolic observation problem (1.16), where  $\mathcal{L} = \mathcal{L}_\gamma$  is given by Example 1.11 (i.e. higher order Grushin operators on the rectangle, with Dirichlet boundary conditions), was proved by Beauchard, Cannarsa and Guglielmi [BCG14]. The authors are interested in the following observability inequality,

equivalent to the controllability to zero (and hence to trajectories)

$$\|y(T)\|_{L^2(\mathcal{M})}^2 \leq C \int_0^T \int_{\omega} |y(t, x)|^2 dt dx, \quad \text{for all } y_0 \in L^2(\mathcal{M}) \text{ and } y \text{ solution of (1.16)}. \quad (1.27)$$

**Theorem 1.26** (Beauchard, Cannarsa and Guglielmi [BCG14]). *Assume  $\mathcal{L} = \mathcal{L}_\gamma$  is given by Example 1.11.*

1. *If  $\gamma \in [0, 1[$ , then the observability inequality (1.27) holds true for any nonempty open set  $\omega \subset \mathcal{M}$  in any time  $T > 0$ .*
2. *If  $\gamma = 1$  and if  $\omega = ]a, b[ \times ]0, 1[$  where  $0 < a < b < 1$ , then there exists  $T^* \geq a^2/2$  such that*
  - *for every  $T > T^*$  the observability inequality (1.27) holds true,*
  - *for every  $T < T^*$  the observability inequality (1.27) is false.*
3. *If  $\gamma > 1$  and  $\omega \subset (0, 1) \times (0, 1)$ , then the observability inequality (1.27) never holds true, in any time  $T > 0$ .*

In the case  $\gamma = 1$  (i.e.  $k = 2$ ) and with a symmetric observation region  $\omega = (] - b, -a[ \cup ]a, b[) \times ]0, 1[$  with  $0 < a < b \leq 1$ , it has been recently proved by Beauchard, Miller and Morancey [BMM15] that  $T^* = a^2/2$  is actually the critical time. This result is quite surprising since parabolic type equations often display an infinite speed of propagation. The controllability of parabolic evolutions thus usually holds in an arbitrary small time independent on the geometry; appearance of a minimum controllability time is hence unusual. Yet, the proof uses a lot the specific geometry of  $\omega$  as a vertical strip. Indeed, another very striking result was recently proved by Koenig in the case  $\gamma = 1$  (i.e. for the Grushin operator): if  $\omega$  is disjoint from an horizontal strip, null-controllability never holds (in any time).

**Theorem 1.27** (Koenig [Koe17]). *Let  $\mathcal{L} = \mathcal{L}_\gamma$  be given by Example 1.11 with  $\gamma = 1$ . Assume that there is  $0 < c < d < 1$  such that  $\omega \cap (] - 1, 1[ \times ]c, d]) = \emptyset$ . Then, for any  $T > 0$  the observability inequality (1.27) is false.*

A remarkable consequence of this result, when compared with the two abovementioned ones is that a *geometric condition* on the set  $\omega$  is needed for the observability estimate 1.27 to hold.

Hence, the best result one can then expect in a general situation is a final state approximate observability result with a cost function  $\Phi(\varepsilon) \xrightarrow{\varepsilon \rightarrow 0^+} +\infty$  (or equivalently an approximate controllability to trajectories with cost  $\Phi(\varepsilon)$ ), which is precisely our Theorem 1.20 and Corollary 1.21 with  $\Phi(\varepsilon) = \varepsilon^{-\beta}$ .

Finally, let us also underline that all these result hold in the context of Example 1.11, that is for the operator  $-(\partial_{x_1}^2 + x_1^{2\gamma} \partial_{x_2}^2)$ , which coefficients are analytic (indeed constant) with respect to the variable  $x_2$ . As such, they fit into the framework of Theorem 1.23 as long as  $\gamma \in \mathbb{N}$ .

More recently, there has been some study of the control of the Heat equation on the Heisenberg group by Beauchard and Cannarsa [BC17]. It still corresponds to the case  $k = 2$  as described in Example 1.8. Some phenomenon similar to the Grushin case seem to occur with the existence of a minimal time if the observability is made on a cylinder. This strenghtens the fact suggested by our result that the important parameter is the hypoelliptic index  $k$ .

Yet, in both cases of Grushin and Heisenberg (or more generally when  $k = 2$ ), it remains to understand what is the geometric property on the observation set  $\omega$  that makes the difference between the polynomial cost that is provided by our result without any assumption on  $\omega$  (which is likely to be optimal in general as suggested by Theorem 1.27) and the exact controllability that requires some geometrical assumptions on  $\omega$ .

### 1.3.3 Controllability of other equations driven by degenerate elliptic operators

To conclude this section, let us mention different works related to the controllability of parabolic equations driven by hypoelliptic or degenerate elliptic operators, that do not fit in the framework of the present article.

First, the paper [Mor15] by Morancey treats the approximate controllability (or the unique continuation property) for the heat equation associated with the Laplace Beltrami operator of the Grushin sub-Riemannian metric defined in [BL13]. This operator is equal to the Grushin operator discussed in Example 1.6 plus a singular potential on the singular set  $x_1 = 0$ . Hence, the analysis of the cost associated to approximate controls is much beyond the scope of the present paper.

Second, we only considered here type I (selfadjoint) Hörmander operators, that is  $\mathcal{L} = \sum_{i=1}^m X_i^* X_i$ . Another classical class of hypoelliptic operators consists in type II Hörmander operators, namely  $\mathcal{L} = \sum_{i=1}^m X_i^* X_i + X_0$ , where the vector field (the drift)  $X_0$  is necessary to span the full tangent space with iterated Lie brackets. These are no longer selfadjoint operators. The simplest example is the so-called Kolmogorov (or Fokker-Planck) operator  $\mathcal{L} = -\partial_v^2 + v\partial_x$ . Our results do not apply in this setting, especially because our main theorems only see the principal symbol of the operator. Yet, recent progress has been made to analyse the observability/controllability of parabolic equations driven by such operators (mainly for some variants of the Kolmogorov operator, though). We quote for instance the papers [BZ09, Bea14, BHR15]. The observability/controllability problem has also been considered on the whole space  $\mathbb{R}^d$ . This led to other geometrical problems about how the domain is "spread out" at infinity, see for instance Le Rousseau-Moyano [LRM16] for the Kolmogorov equation and Beauchard-Pravda-Starov [BPS16a, BPS16b] for some more general class of quadratic operators. In these last papers, the idea of the proof is to combine some observability of low frequency (where the frequency are the usual Fourier ones or related to the harmonic oscillator) with some decay and regularizing properties of the semigroup.

It would be very interesting to understand the common features and differences of our results and methods with these ones. At first sight, it seems that in both cases, one important idea (which goes back to Lebeau-Robbiano [LR95]) is to compare the decay rate of the heat equation with the cost of observability (or control) of low frequency. But, in our paper, we define "frequency" with respect to the hypoelliptic operator  $\mathcal{L}$  and then, the decay is natural for high frequency. The hard part is then to understand the cost of observability of low frequencies. In these papers, it seems that they have chosen to define the frequency as the usual Euclidian one (or with respect to a fixed well known operator as the harmonic oscillator in [BPS16b]). In this situation, the observation at low (Euclidian) frequency does not really see the hypoelliptic operator and follows from more usual Carleman estimates. Yet, the decay rate of high frequency (see for instance Proposition 2.2 of [BPS16a]) and the understanding of the commutation with Fourier cutoff turns out to be much more complicated and to reflect deeply the hypoelliptic properties of the operator.

Note also that there has been several studies of the unique continuation property for type II Hörmander operators (see e.g. [LZ82] for related operators).

Finally, other types of degeneracies have also been studied, as for instance elliptic operators with coefficients vanishing near the boundary of a domain  $\mathcal{M} \subset \mathbb{R}^d$ . In this case, adaptations of the usual Carleman estimates (combined with appropriate Hardy inequalities) are sometimes tractable. The literature is vast, and we simply mention the recent memoir [CMV16] and refer the reader to the references therein.

## 1.4 Sketch of the proofs and plan of the paper

Even though this is not explicit in the discussion above, the cornerstone result of this paper is Theorem 1.13, concerning the hypoelliptic wave equation. All results concerning eigenfunctions (Theorem 1.10) or the hypoelliptic heat equation (Theorems 1.16, 1.18 and 1.20) are then deduced from Theorem 1.13. The proof in the partially analytic case (Theorem 1.23) shall be discussed afterwards. Let us hence first comment the proof of Theorem 1.13.

The proof of Theorem 1.13 is based on the general strategy developed by the authors in [LL15] for quantifying and propagating unique continuation properties. From [LL15], we only use here (except for the partially analytic situation of Theorem 1.23) the “Holmgren-John” case, i.e. when the operator has analytic coefficients. It states basically

- that an appropriate quantitative (low frequency) estimate holds across any non-characteristic hypersurface;
- that such local estimates can be propagated, leading towards global ones.

In Section 2, we review results and tools developed in [LL15]; for sake of readability, we specify the latter to the very particular case of *second order* operators that are elliptic when restricted to  $\zeta_a = 0$  (the cotangent variable to the analytic variable, called  $\xi_a$  in [LL15]), which includes all operators studied in the present article.

Here, when compared to the case of the classical wave equation, two more difficulties arise: one of geometric nature, and one related to the compatibility of energy space associated to  $\mathcal{L}$  and those dealt with in [LL15].

Let us first describe the geometric difficulty. The proof is inspired by the case of the classical wave equation given in [LL15, Section 6.1]: the idea is, given a point  $x_0 \in \mathcal{M}$ , to take any path  $\gamma : [0, 1] \rightarrow \mathcal{M}$  with  $\gamma(0) = x_0$  and  $\gamma(1) \in \omega$  (observation set), of length sufficiently small, and then to construct a family of appropriate noncharacteristic hypersurfaces in these coordinates near  $[-T, T] \times \gamma$ . There, we apply the general theorem of [LL15], which allows to bound the solution  $u$  to  $(\partial_t^2 - \Delta)u = 0$  in a neighborhood of  $(t, x) = (0, x_0)$  by  $u$  in  $[-T, T] \times \omega$ .

Here, due to the non definiteness/ellipticity of the operator  $\mathcal{L}$ , we are not able to construct *global* coordinates near *any* path  $\gamma$  together with appropriate noncharacteristic hypersurfaces, in which to apply the results of [LL15]. To overcome this difficulty, we do not consider *any* path between  $x_0$  and  $\omega$ , but rather only so called *normal geodesics*, that is, projections on  $\mathcal{M}$  of hamiltonian curves of the principal symbol of the operator  $\mathcal{L}$ . The existence of such paths  $\gamma$  (minimizing the sub-Riemannian distance) from any point  $x_0$  to  $\omega$  is a well-known result in sub-Riemannian geometry, proved by Rifford and Trélat [RT05]. Then, locally near a point of  $\gamma$ , the introduction of normal geodesic coordinates allows us to define local coordinates in which to apply a *local* version of our results in [LL15].

A new difficulty, linked to the methods used in [Tat95, Tat99b, RZ98, Hör97, LL15], then arises: the whole setting of these papers relies on a splitting of space into analytic and non-analytic coordinates. Hence, most “patchable estimates” (linked to a relation  $\triangleleft$ , see Section 2.2) produced in [LL15] require the analytic variable to be global and straight, which is obviously not the case here. To solve this problem we do not rely on the main (neither global, nor local) result of [LL15], but rather on the specific result of [LL15, Theorem 4.11], which takes into account the possible changes of variables. Having these results in hand allows to prove an estimate of the form

$$\|u\|_{L^2(\mathbb{J}_{-\varepsilon, \varepsilon} \times \mathcal{M})} \leq C e^{\kappa\mu} \|u\|_{L^2(\mathbb{J}_{-T, T} \times \omega)} + \frac{C}{\mu} \|u\|_{H^1(\mathbb{J}_{-T, T} \times \mathcal{M})}, \quad (1.28)$$

for  $\mu$  large and  $u$  solution to  $(\partial_t^2 + \mathcal{L})u = 0$ . This estimate is the same as that obtained in [LL15] for the wave equation.

This leads us to the second main difficulty we have to face in the proof of Theorem 1.13. Whereas the left hand-side of (1.28) is bounded from below by the natural  $L^2 \times \mathcal{H}_{\mathcal{L}}^{-1}$  norm of the data, the right hand-side is not directly linked to their  $\mathcal{H}_{\mathcal{L}}^1 \times L^2$  norm. More precisely, the hypoelliptic estimates Rothschild and Stein [RS76] (see Theorem 1.4 above and Appendix B.1) imply that  $\|u\|_{H^1(\cdot)_{-T,T}(\times \mathcal{M})} \leq C \|(u_0, u_1)\|_{\mathcal{H}_{\mathcal{L}}^k \times \mathcal{H}_{\mathcal{L}}^{k-1}}$ . This provides a weaker version of Theorem 1.13 which has exactly the same form as in the case of the wave equation (cost  $e^{\kappa\mu}$ ), but with the norm  $\|(u_0, u_1)\|_{\mathcal{H}_{\mathcal{L}}^k \times \mathcal{H}_{\mathcal{L}}^{k-1}}$  in the right hand-side. This weaker version is however interesting for itself since the proof is much less involved, and we prove it in Section 3.3.1.

To obtain the estimate of Theorem 1.13 (and in fact, a family of such estimates with any  $\mathcal{H}_{\mathcal{L}}^s \times \mathcal{H}_{\mathcal{L}}^{s-1}$ ,  $s > 0$ , in the right hand-side, see Theorem 1.31 below), we thus need to work with a version of (1.28) still containing frequency cutoff localization and an  $e^{-c\mu}$  small remainder (instead of the  $1/\mu$  one). These low-frequency-with-exponentially-small-remainder estimates are then combined with the spectral representation of solutions to  $(\partial_t^2 + \mathcal{L})u = 0$  in order to gain back derivatives in the remainder term. Such estimates are close to those we prove in [LL17] for the classical wave equation. These final energy estimates are performed in Section 3.3.3, and conclude the proof of Theorem 1.13.

Starting from Theorem 1.13, let us now explain how to deduce the other results of the paper, namely Theorems 1.10, 1.16, 1.18 and 1.20. First of all, Theorem 1.10 is simply deduced from Theorem 1.13 by using a particular solution to the wave equation (1.12), namely  $u(t, x) = \cos(\sqrt{\lambda_j}t)\varphi_j(x)$ . See Section 3.3.2.

Section 4 is devoted to the proofs of Theorems 1.16, 1.18 and 1.20, which follow the general idea that the controllability/observability properties for hyperbolic equations implies controllability/observability properties for their parabolic counterpart, see [Rus73, Mil06a, EZ11a, EZ11b] (see also [LR95]). This has been named as “transmutation methods” by Luc Miller [Mil06a]. Here, we use the method developed in [EZ11a]. In that paper, Ervedoza and Zuazua deduced the (exact final time) observability of the heat equation (known from [LR95, FI96]) from the approximate observability estimate for waves (namely the analogue of Theorem 1.13) as proved in [Phu10] (with loss) or [LL15] (without loss). Their proof consists in constructing an appropriate kernel  $k_T(t, \cdot)$  such that if  $y(t)$  is a solution to the usual heat equation,  $u(t) = \int_0^T k_T(t, s)y(s)ds$  is a solution to the usual wave equation, to which we can apply the analogue of Theorem 1.13. Because of the exponential cost in term of the frequency ( $e^{c\Lambda}$ ), the resulting estimates are only useful at low frequency: for data having (spectral) frequencies  $\sqrt{\lambda_j} \leq \sqrt{\lambda}$ , one then obtain observability (or controllability if we think about the dual problem) at cost  $e^{c\sqrt{\lambda}}$  as in [LR95]. The proof of final state observability then follows from comparing this cost with the heat dissipation for frequencies  $\sqrt{\lambda_j} \geq \sqrt{\lambda}$ , namely  $e^{-t\lambda}$  as in the original proof [LR95] (see also [LRL12] or the simplified argument of [Mil10]).

Here, we follow the approach of [EZ11a] (in particular, we use the same kernel  $k$  and its properties) in the proofs of Theorems 1.16, 1.18 and 1.20, with the following modifications.

The proof of Theorems 1.18 and 1.20 are very close to that of [EZ11a]. However, application of the method of [EZ11a] yields that the observability of low frequencies  $\sqrt{\lambda_j} \leq \sqrt{\lambda}$  costs  $e^{c\lambda^{k/2}}$ , see Lemma 4.2 (remark that Proposition 1.12 implies that this is optimal in general). This cost has to be compared to the dissipation for high frequencies  $\sqrt{\lambda_j} \geq \sqrt{\lambda}$ , namely  $e^{-t\lambda}$ . Hence, we see that the case  $k = 1$  (classical heat equation, already discussed),  $k = 2$ , and  $k > 2$  display very different features:

1. In case  $k = 2$ , the cost of observation of low frequencies  $e^{c\lambda}$  and the parabolic dissipation for high frequencies  $e^{-t\lambda}$  have the same strength: in this case, we need to wait a time long enough so that the dissipation “beats” the cost of the observability (essentially  $t > c$ ). Moreover, the iterative procedure devised in [LR95] in order to control/observe *all* frequencies in finite time cannot converge here: each step would need a time  $t > c$ . Therefore, we only obtain the approximate controllability result of Theorem 1.20, with a cost improving as time increases. See Section 4.1 for the proof of Theorem 1.20.
2. In case  $k > 2$ , the dissipation for high frequencies  $e^{-t\lambda}$  has no chance to compete with the cost of observation of low frequencies  $e^{c\lambda^{k/2}}$ . Assuming that the initial data are in the Gevrey-type space  $H^{\theta, \alpha}$  with  $\alpha = k/2$  allows to compensate for the cost of low frequencies  $e^{c\lambda^{k/2}}$  (the  $\theta$  having to be compared to  $c$ ), leading to Theorem 1.18. Note that parabolic dissipation at high frequencies does not play any role here: low frequencies are observed thanks to transmutation and high-frequencies are absorbed by the Gevrey norm. The cases  $k = 1, 2$  (in which the Gevrey norm is relatively weaker) in Theorem 1.18 are a little different and require elements similar to those used in the proof of Theorem 1.16. Note that this type of result seems to be new for the classical heat equation as well (in which case our proof also holds in a much more general setting). Proof of Theorem 1.18 is performed in Section 4.2.

Finally, the proof of Theorem 1.16 in Section 4.3 relies on the same transmutation technique. However, we do not split the solution into low and high-frequencies, but rather apply the transmutation kernel  $k_T(t, s)$  to the full solution  $y$  to the heat equation:  $u(t) = \int_0^T k_T(t, s)y(s)ds$  is a solution to the wave equation. We then prove a fine asymptotics analysis of  $\int_0^T k_T(0, s)e^{-\lambda s}ds$  for high frequencies together with convexity estimates to bound the frequency function of  $u(0)$  by the frequency function of  $y(0)$ , namely  $\Lambda = \frac{\|y(0)\|_{\mathcal{H}_{\mathcal{L}}^1}}{\|y(0)\|_{L^2}}$ . The proof of this result via a direct transmutation method seems to be new, even for the classical heat equation. The usual proofs [FCZ00, Phu04] rather rely on the exact final time observability estimate, which does not hold here in general. However, as opposed to [FCZ00, Phu04], we do not recover uniform estimates in terms of the control time  $T$  as  $T \rightarrow 0^+$ .

Finally, in Section 5, we prove the partially analytic result of Theorem 1.23. Only the analogue of Theorem 1.13 at regularity  $\mathcal{H}_{\mathcal{L}}^k$  needs to be proved (namely estimate (1.24)), since, as discussed above, all results of Theorems 1.10, 1.16, 1.18 and 1.20 (under the appropriate form) are corollaries of that of Theorem 1.13. The situations is almost the same as that of Theorem 1.13 except for four main differences. First, the presence of the boundary makes it complicated to apply globally the geometric result of Rifford and Trélat [RT05], and we only rely on a local version of it. Second, the partial analyticity assumption does not allow to make changes of variables. This difficulty is overcome by the very simple geometry of  $[-1, 1]_{x_1} \times \mathbb{T}_{x_2}$ , in which we barely do not perform any change of variable. Third, the application of the results in [LL15] yields an observation term in a mixed  $L^2 - H^1$  norm; we have to refine this estimate to recover the  $L^2$  observation term. Finally, the available hypoelliptic estimates do not apply directly in the presence of boundary and we have to patch hypoelliptic estimates in the interior with elliptic estimates at the boundary.

The paper ends with three appendices, the first of which, Appendix A is devoted to the proof of the optimality result of Proposition 1.12 using some estimates of [BCG14, Section 2.3]. The second part, Appendix B provides the proof of several subelliptic estimates that are used throughout the paper. They are consequences of Theorem 1.4. Finally, Appendix C contains a technical computation.

## 1.5 Some remarks and further comments

This section contains several remarks concerning the setting of the present paper and the results we obtain.

**Remark 1.28** (Sub-Riemannian Laplacians). Here, we explain why the assumption that  $\mathcal{L}$  writes as a sum of  $X_j^* X_j$ , although seemingly restrictive, contains in fact a general family of intrinsically defined sub-Riemannian Laplacians.

We first define here the sub-Riemannian Laplacian  $\Delta_{(U,f),ds}$  associated to a sub-Riemannian structure  $(U, f)$  on  $\mathcal{M}$  and a smooth density  $ds$ . We then explain why it can be rewritten under the form (1.1) for some (sufficiently many) vector fields  $X_1, \dots, X_m$ .

First, we assume that  $\mathcal{M}$  is equipped with a general sub-Riemannian structure  $(U, f)$ , see [Bel96, Definition 1.3] or [ABB16a, Definition 3.2] with  $U$  a Euclidean bundle with base  $\mathcal{M}$  and  $f : U \rightarrow T\mathcal{M}$  a smooth map being linear on fibers. This allows to define first a sub-Riemannian metric, that is, a metric on the horizontal distribution  $\mathcal{D}$  with  $\mathcal{D}_x = f(U_x) \subset T_x\mathcal{M}$  by  $g(x, v) = \inf\{|u|, u \in U_x, v = f(x, u)\}$  (where  $|\cdot|$  denotes the Euclidean norm in  $U$ ). Second, this provides a sub-Riemannian gradient  $\nabla_{(U,f)}$  on  $\mathcal{M}$ : namely, for  $u \in C^\infty(\mathcal{M})$ ,  $\nabla_{(U,f)}u(x)$  is the unique vector in  $\mathcal{D}_x$  such that for all  $v \in \mathcal{D}_x$ , we have  $d_x u(v) = \tilde{g}(x, \nabla_{(U,f)}u(x), v)$  (where  $\tilde{g}$  is the bilinear form associated to  $g$ ).

Next, the smooth density  $ds$  allows to define the divergence  $\operatorname{div}_{ds}$  of a vector field  $X \in \mathcal{X}^\infty$  by

$$\left. \frac{d}{dt} (e^{tX})^*(ds) \right|_{t=0} = \operatorname{div}_{ds}(X)ds,$$

where  $e^{tX}$  denotes the flow of  $X$  (or, equivalently, by the formula  $X^* = -X - \operatorname{div}_{ds}(X)$ ). Hence, a natural definition of the sub-Riemannian Laplacian  $\Delta_{(U,f),ds}$  is

$$\Delta_{(U,f),ds}u = \operatorname{div}_{ds}(\nabla_{(U,f)}u), \quad u \in C^\infty(\mathcal{M}).$$

Now, according to [ABB16a, Corollary 3.26], the sub-Riemannian structure  $(U, f)$  is equivalent to a free one, that is, there exist  $m \in \mathbb{N}$  and  $m$  vector fields  $X_1, \dots, X_m$  on  $\mathcal{M}$  such that the horizontal distribution at  $x \in \mathcal{M}$  is given by  $\mathcal{D}_x = \operatorname{span}(X_1(x), \dots, X_m(x))$ , and the metric on this distribution is defined by (1.11). A computation similar to that in Appendix C shows that the sub-Riemannian gradient  $\nabla_{(U,f)}$  of a function  $u$  is then given by

$$\nabla_{(U,f)}u = \sum_{i=1}^m (X_i u) X_i.$$

Hence, the formula  $\operatorname{div}_{ds}(uX) = u \operatorname{div}_{ds}(X) + Xu$  for  $u \in C^\infty(\mathcal{M})$  and  $X \in \mathcal{X}^\infty$  yields

$$\Delta_{(U,f),ds}u = \operatorname{div}_{ds} \left( \sum_{i=1}^m (X_i u) X_i \right) = \sum_{i=1}^m \operatorname{div}_{ds}(X_i) X_i u + X_i^2 u = - \sum_{i=1}^m X_i^* X_i u.$$

As a consequence, all results presented in this article remain valid for general, intrinsically defined sub-Riemannian Laplacians  $\Delta_{(U,f),ds}$ .

In the above discussion, we assume the density  $ds$  to be given: the sub-Riemannian Laplacian  $\Delta_{(U,f),ds}$  then depends both on the sub-Riemannian structure  $(U, f)$  and the density. One may also wonder whether, given the sub-Riemannian structure  $(U, f)$  only, there is an associated intrinsic choice of density  $ds$ , as in the Riemannian case. This question is an object of current research. While it seems that in the equiregular case, i.e., when the growth vector does not depends on the point, there is a natural intrinsic measure (namely

the Popp measure, see e.g. [Mon02]), there is no consensus for what should be the natural one in the general case. For instance, in the Grushin case of Example 1.6 the metric  $g$  defined in (1.11) is Riemannian outside of  $\{x_1 = 0\}$ . Hence, a natural choice would be to take the Riemannian density outside of  $\{x_1 = 0\}$ . The associated Laplacian is equal to that of Example 1.6 plus a potential which is singular on  $\{x_1 = 0\}$ . The analysis in [BL13] shows that the zone  $\{x_1 = 0\}$  creates a barrier which the information cannot cross. This is in strong contrast with results of the present paper.

To conclude this remark, let us also notice that the whole class of operators studied by Fefferman and Phong in [FP83] is not contained in the class of sub-Riemannian Laplacians  $\Delta_{(U,f),d_s}$  defined above. It would be interesting to investigate the questions of the present paper for such hypoelliptic operators.

**Remark 1.29** (Norm of the observation term). Note that in the right hand-side of (1.13), the observation term only comes with a  $L^2$  norm (which is not the case in most results in [LL15]). This is due to the fact that we are in the context of operators with analytic coefficients with respect to all variables. In the case of partially analytic operators as described in Subsection 1.2.4, we are able to get observability in  $L^2$  using a refined argument (see Section 5.3). Similar arguments are also applied to the classical wave equation in a forthcoming companion paper [LL17].

**Remark 1.30** (Other levels of  $\mathcal{H}_{\mathcal{L}}^s$  regularity). As already mentionned, Theorem 1.13 is a particular case of general estimates where all Sobolev scales are possible for measuring the typical frequency of the initial datum. Indeed, we prove the following more general result.

**Theorem 1.31.** *Let  $\mathcal{L}$  as above satisfying Assumptions 1.1 and 1.2. Assume that  $\omega$  is a non empty open set of  $\mathcal{M}$  and  $T > \sup_{x \in \mathcal{M}} d_{\mathcal{L}}(x, \omega)$ . Then, for any  $s > 0$ , there exist  $\kappa, C, \mu_0$  so that for all  $\mu \geq \mu_0$ , and all  $u$  solution of (1.12), we have*

$$\|(u_0, u_1)\|_{L^2 \times \mathcal{H}_{\mathcal{L}}^{-1}} \leq C e^{\kappa \mu^k} \|u\|_{L^2([-T, T] \times \omega)} + \frac{1}{\mu^s} \|(u_0, u_1)\|_{\mathcal{H}_{\mathcal{L}}^s \times \mathcal{H}_{\mathcal{L}}^{s-1}} \quad (1.29)$$

Theorem 1.13 is the case  $s = 1$  of Theorem 1.31 and we believed that the frequency functions  $\Lambda$  used in Theorem 1.13 is the more natural presentation. Yet, it turns out that Theorem 1.31 is actually easier to prove for  $s = k$ . Below, we first prove this simpler case  $s = k$  (Section 3.3.1); then we need to prove refined estimates for the general case (Section 3.3.3). The second part of the proof seems to require additional arguments in the partially analytic case described in Section 1.2.4. That is the reason why the estimate (1.24) of Theorem 1.23 is restricted to the case  $s = k$ . Nevertheless, as already explained, most of the results about eigenfunction and the hypoelliptic heat equation only use the easier case  $s = k$ .

Remark also that, in inequalities such as (1.29), deducing the  $\mathcal{H}_{\mathcal{L}}^{s'}$  case from the  $\mathcal{H}_{\mathcal{L}}^s$  case follows from an interpolation argument if  $s' > s$ . Indeed, for  $0 \leq s < s'$ , denoting  $\tilde{\mathcal{H}}^s = \mathcal{H}_{\mathcal{L}}^s \times \mathcal{H}_{\mathcal{L}}^{s-1}$  we have

$$\|U\|_{\tilde{\mathcal{H}}^s} \leq \|U\|_{\tilde{\mathcal{H}}^0}^{1-\frac{s}{s'}} \|U\|_{\tilde{\mathcal{H}}^{s'}}^{\frac{s}{s'}} \leq (1 - \frac{s}{s'}) A \|U\|_{\tilde{\mathcal{H}}^0} + \frac{s}{s'} A^{1-\frac{s}{s'}} \|U\|_{\tilde{\mathcal{H}}^{s'}}, \quad \text{for all } A > 0.$$

Taking then  $A = \frac{1}{2} \mu^s$  and putting this into (1.29) yields the same estimate with  $s$  replaced by  $s'$ . Hence, the difficulty in Theorem 1.31 when compared to the case  $s = k$  (which proof is simpler) is only for small  $s > 0$ .

Finally, let us mention that the result of Theorem 1.16 remains valid as well with the  $\mathcal{H}_{\mathcal{L}}^1$ -norm replaced by any  $\mathcal{H}_{\mathcal{L}}^s$ -norm,  $s > 0$ , when modifying the powers accordingly.

**Remark 1.32** (Constants). Note that there are mostly two relevant constants in Theorem 1.13, namely the minimal time  $2 \sup_{x \in \mathcal{M}} d_{\mathcal{L}}(x, \omega)$ , and the constant  $\kappa$  in the exponent of 1.13. All constants appearing in the



results of Section 1.2.3 can be explicitly formulated in terms of these two. For instance, in Theorem 1.20, the constant  $T_0$  can be taken as  $T_0 > \kappa$ , where  $\kappa$  is the exponent in (1.13) for some  $S > \sup_{x \in \mathcal{M}} d_{\mathcal{L}}(x, \omega)$ . We refer to Remark 4.3 below for more on this subject.

**Remark 1.33** (Interpolation spaces, see [Leb92, LL17]). Notice that Estimate (1.22) may be reformulated (after an optimization in  $\varepsilon$ ) as the following interpolation inequality, for  $y$  solution to (1.16)

$$\|y(T)\|_{L^2} \leq C \|y\|_{L^2((0,T) \times \omega)}^{\frac{T-(T_0+\eta)}{T-\eta}} \|y(0)\|_{L^2}^{\frac{T_0}{T-\eta}},$$

while (1.21) can be written

$$\|y(0)\|_{L^2}^2 \leq C \|y\|_{L^2((0,T) \times \omega)}^{\frac{\theta-\theta_0}{\theta}} \|y(0)\|_{k/2, \theta}^{\frac{\theta_0}{\theta}}.$$

More generally, in both cases, there exists  $\alpha \in ]0, 1[$  so that, we have the estimates

$$\|y_0\|_{F_1} \leq C \|y_0\|_{F_{obs}}^\alpha \|y_0\|_{F_0}^{1-\alpha}, \quad (1.30)$$

where  $F_0$ ,  $F_T$  and  $F_{obs}$  which are defined as the spaces of data obtained as the completion of linear combinations of eigenfunctions of  $\mathcal{L}$  for the norms

$$\begin{aligned} \|y_0\|_{F_1} &= \|e^{-T\mathcal{L}}y_0\|_{L^2}, & \text{resp. } \|y_0\|_{F_1} &= \|y_0\|_{L^2}, \\ \|y_0\|_{F_0} &= \|y_0\|_{L^2}, & \text{resp. } \|y_0\|_{F_0} &= \|y_0\|_{k/2, \theta}, \\ \|y_0\|_{F_{obs}} &= \|y\|_{L^2((0,T) \times \omega)} = \|e^{-t\mathcal{L}}y_0\|_{L^2((0,T) \times \omega)} & \text{in both cases.} \end{aligned} \quad (1.31)$$

The latter are proper norms as a consequence of uniqueness, backward uniqueness (consequence e.g. of Lemma 4.6 below) and unique continuation property for the hypoelliptic heat equation (1.16). Note that we have  $F_0 \subset F_{obs}, F_1$  in the first case, and  $F_0 \subset F_1 \subset F_{obs}$  in the second.

As a consequence of (1.30) (see for instance [Leb92, Appendix, Lemma 1]), there exists  $\delta > 0$  such that

$$[F_0, F_{obs}]_\delta \subset F_1,$$

where  $[F_0, F_{obs}]_\delta$  is the space of interpolation between  $F_0$  and  $F_{obs}$ . As in Lebeau [Leb92, Section 3], this yields

$$F'_1 \subset [F'_0, F'_{obs}]_{1-\delta}. \quad (1.32)$$

Now, the duality between (1.16) and (1.17) will allow to identify the spaces  $F'_0, F'_1, F'_{obs}$ , to deduce properties of the controllable and the attainable sets.

First, the duality between (1.16) and (1.17) writes

$$\int_0^T (\mathbb{1}_\omega y(T-t), g)_{L^2(\mathcal{M})} dt = (y_0, u(T))_{L^2(\mathcal{M})} - (y(T), u_0)_{L^2(\mathcal{M})}. \quad (1.33)$$

We define

$$\begin{aligned} E_{att} &= \{u_1 \in L^2(\mathcal{M}), \text{ there exists } g \in L^2((0, T) \times \omega), \\ &\text{s.t. the associated solution } u \text{ to (1.17) with } u(0) = 0 \text{ satisfies } u(T) = u_1\} \end{aligned}$$

the space of attainable data from zero with  $L^2$  control, endowed with the norm

$$\|u_1\|_{E_{att}} = \inf \left\{ \|g\|_{L^2((0,T) \times \omega)}, g \in L^2((0,T) \times \omega) \right. \\ \left. \text{s.t. the associated solution } u \text{ to (1.17) with } u(0) = 0 \text{ satisfies } u(T) = u_1 \right\}.$$

From (1.33), we obtain for all  $u_1 \in E_{att} \subset L^2$  and all  $y_0 \in L^2$

$$|(y_0, u_1)_{L^2(\mathcal{M})}| = \left| \int_0^T (\mathbb{1}_\omega y(T-t), g)_{L^2(\mathcal{M})} dt \right| \leq \|y_0\|_{F_{obs}} \|u_1\|_{E_{att}}.$$

Hence, the  $L^2(\mathcal{M})$  scalar product extends uniquely as a duality product  $\langle y_0, u_1 \rangle_{F_{obs}, E_{att}}$ , allowing to identify  $F'_{obs}$  with  $E_{att}$ . With the identification, we have as well  $F_1 = e^{T\mathcal{L}}L^2(\mathcal{M})$  and  $F_0 = L^2(\mathcal{M})$  (resp.  $F'_1 = L^2(\mathcal{M})$  and  $F'_0 = H^{k/2, \theta}$ ) so that  $F_1 \approx e^{-T\mathcal{L}}L^2(\mathcal{M})$  and  $F'_0 \approx L^2(\mathcal{M})$  (resp.  $F'_1 \approx L^2(\mathcal{M})$  and  $F'_0 \approx H^{k/2, -\theta}$ ). With (1.32), this yields

$$e^{-T\mathcal{L}}L^2 \subset [E_{att}, L^2]_{1-\delta}, \\ \text{resp. } L^2 \subset [E_{att}, H^{k/2, -\theta}]_{1-\delta}.$$

We also define  $E_{cont}$  the (abstract) space of data that can be controled towards zero as the completion of the space

$$\{u_0 \in L^2, \text{ there exists } g \in L^2((0,T) \times \omega), \text{ s.t. the associated solution } u \text{ to (1.17) satisfies } u(T) = 0\}$$

for the norm

$$\|u_0\|_{E_{cont}} = \inf \left\{ \|g\|_{L^2((0,T) \times \omega)}, g \in L^2((0,T) \times \omega) \right. \\ \left. \text{s.t. the associated solution } u \text{ to (1.17) satisfies } u(T) = 0 \right\}.$$

From (1.33), we obtain for all  $u_1 \in E_{cont} \subset L^2$  and all  $y_0 \in L^2$

$$|(y(T), u_1)_{L^2(\mathcal{M})}| = \left| \int_0^T (\mathbb{1}_\omega y(T-t), g)_{L^2(\mathcal{M})} dt \right| \leq \|y_0\|_{F_{obs}} \|u_1\|_{E_{cont}}.$$

Similarly, the scalar product  $\langle y_0, u_1 \rangle = (e^{-T\mathcal{L}}y_0, u_1)_{L^2(\mathcal{M})}$  extends uniquely as a duality product  $\langle y_0, u_1 \rangle_{F_{obs}, E_{att}}$ , allowing to identify  $F'_{obs}$  with  $E_{cont}$ . With this same identification, we have as well  $F'_1 \approx L^2(\mathcal{M})$  and  $F'_0 \approx e^{-T\mathcal{L}}L^2(\mathcal{M})$  (resp.  $F'_1 \approx e^{-T\mathcal{L}}L^2(\mathcal{M})$  and  $F'_0 \approx e^{-T\mathcal{L}}H^{k/2, -\theta}$ ). Similarly, (1.32) also yields

$$L^2 \subset [E_{cont}, e^{-T\mathcal{L}}L^2]_{1-\delta} \\ e^{-T\mathcal{L}}L^2 \subset [E_{cont}, e^{-T\mathcal{L}}H^{k/2, -\theta}]_{1-\delta}.$$

Note that when  $k > 2$ ,  $e^{-T\mathcal{L}}H^{k/2, -\theta} \approx H^{k/2, -\theta}$  which is not a distributional set whatever the time  $T$  is. Yet, if  $k = 2$ ,  $e^{-T\mathcal{L}}H^{k/2, -\theta} \approx H^{1, -\theta+T}$ , which is a space of very regular functions if  $T > \theta$ .

The next remark concerns the results of Theorem 1.20 and Corollary 1.21.

**Remark 1.34** (Large time approximate controllability with polynomial cost of “critical anomalous diffusion”). The proof of Theorem 1.20 and Corollary 1.21 also applies to any positive selfadjoint operator satisfying spectral estimates (or similar estimates for the control of the heat equation for spectrally localized initial data) like

$$\|w\|_{L^2(\mathcal{M})} \leq C e^{c\lambda} \|w\|_{L^2(\omega)}, \quad \text{for all } w = \sum_{\lambda_j \leq \lambda} w_j \varphi_j. \quad (1.34)$$

This is in particular the case for the square root of the Laplacian  $\sqrt{-\Delta_g}$  where  $\Delta_g$  is the elliptic Laplace-Beltrami operator on a compact Riemannian manifold (even without the analyticity assumption and with Dirichlet boundary condition), see [LR95]. Therefore, for the associated evolution operator (so called “critical anomalous diffusion”)  $\partial_t + \sqrt{-\Delta_g}$  (studied in [Mil06b]), the same approximate controllability result with a polynomial cost holds. Note also that even for the one dimensional case (namely the operator  $\partial_t + |\partial_x|$  on the circle), it has been proved by Koenig [Koe17] that exact controllability in finite time  $T > 0$  never holds (as long as the control domain is not the whole circle). In particular, it suggests that approximate controllability at polynomial cost might be the best to obtain under general spectral assumptions like (1.34).

*Acknowledgements.* We wish to thank warmly Davide Barilari and Emmanuel Trélat for very helpful discussions on sub-Riemannian geometry. In particular, Emmanuel Trélat brought our attention to the references [RT05] and [Der71], and the properties of Example 1.9, Remark 1.28 and Appendix C were explained to us by Davide Barilari. The first author is partially supported by the Agence Nationale de la Recherche under grant SRGI ANR-15-CE40-0018 and IPROBLEMS ANR-13-JS01-0006. The second author is partially supported by the Agence Nationale de la Recherche under grant GERASIC ANR-13-BS01-0007-01.

## 2 The quantitative Holmgren-John theorem of [LL15]

In this section, we briefly review some results obtained in [LL15], that will be at the core of the proof of the present paper. We shall only consider a very particular class of operators, namely second order operators with real principal symbol. Also, we shall only consider non-characteristic surfaces. This assumption can be also relaxed (see e.g. [LL15] Definition 1.7 and Remark 1.9), even though we are not aware of any application of the refined result.

The interest of taking such operators and surfaces is that, in this context, several assumptions and formulations of the results in [LL15] are simplified. In this section, we state results for an operator  $P$  in  $\mathbb{R}^n$ , where, in the application in Section 3 below, we shall mainly consider  $P$  as a local version of  $\partial_t^2 + \mathcal{L}$  on  $\mathbb{R} \times \mathbb{R}^d$  (recall that  $\dim(\mathcal{M}) = d$ ), that is  $n = d + 1$ . For this reason (and as opposed to the notation of [LL15]), we shall denote by  $z \in \mathbb{R}^n$  the running variable and  $\zeta \in \mathbb{R}^n$  its cotangent variable. In the applications in the next section, we will have  $z = (t, x)$  and  $\zeta = (\tau, \xi)$ .

### 2.1 A typical quantitative unique continuation result of [LL15]

A typical instance (in the situation describe above) of the main result of [LL15] may be stated as follows (see [LL15, Theorem 1.11] together with [LL15, Remark 1.10]).

**Geometric setting:** (see Figure 1) We first fix two splittings of  $\mathbb{R}^n$ :

- $\mathbb{R}^n = \mathbb{R}^{n_a} \times \mathbb{R}^{n_b}$ , where  $n_a + n_b = n$ . We denote  $z = (z_a, z_b)$  the global variable and  $\zeta = (\zeta_a, \zeta_b)$  the associated cotangent variable.
- and  $\mathbb{R}^n = \mathbb{R}_{z'}^{n-1} \times \mathbb{R}_{z_n}$ ,

possibly in two different bases. We let  $D$  be a bounded open subset of  $\mathbb{R}^{n-1}$  with smooth boundary and  $G = G(z', \varepsilon)$  a  $C^2$  function defined in a neighborhood of  $\overline{D} \times [0, 1]$ , such that

1. For all  $\varepsilon \in (0, 1]$ , we have  $\{z' \in \mathbb{R}^{n-1}, G(z', \varepsilon) \geq 0\} = \overline{D}$ ;
2. for all  $z' \in D$ , the function  $\varepsilon \mapsto G(z', \varepsilon)$  is strictly increasing;
3. for all  $\varepsilon \in (0, 1]$ , we have  $\{z' \in \mathbb{R}^{n-1}, G(z', \varepsilon) = 0\} = \partial D$ .

We set  $G(z', 0) = 0$ ,  $S_0 = \overline{D} \times \{0\}$  and, for  $\varepsilon \in (0, 1]$ ,

$$S_\varepsilon = \{(z', z_n) \in \mathbb{R}^n, z_n \geq 0 \text{ and } G(z', \varepsilon) = z_n\} = (\overline{D} \times \mathbb{R}) \cap \{(z', z_n) \in \mathbb{R}^n, G(z', \varepsilon) = z_n\};$$

$$K = \{z \in \mathbb{R}^n, 0 \leq z_n \leq G(z', 1)\}.$$

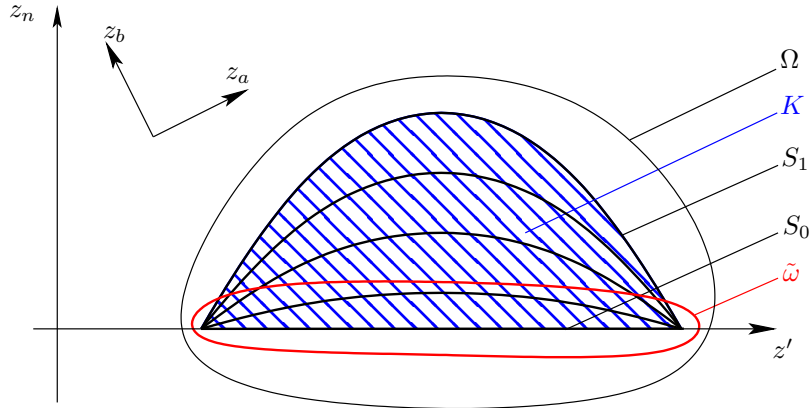


Figure 1: Geometric setting of Theorem 2.1

We recall that the local surface  $S := \{\varphi = 0\} \ni z_0$ ,  $d\varphi(z_0) \neq 0$  is called *non-characteristic* at  $z_0$  for the differential operator  $P$  with principal symbol  $p$  if  $p(z_0, d\varphi(z_0)) \neq 0$ , and that this is a property of the sole surface  $S$  (together with the point  $z_0$  and the principal symbol of the operator  $p$ ) and not its defining function  $\varphi$ .

Note also that in the main part of the paper, the operators are analytic with respect to all variables, In this case, Theorem 2.1 is a quantitative version of the Holmgren-John theorem (for second order operators, see [LL15] in the general case), and may be seen as a generalization of [Leb92], which concern the (analytic) wave operator.

**Theorem 2.1.** *In the above geometric setting, we moreover let  $\Omega \subset \mathbb{R}^{n_a} \times \mathbb{R}^{n_b}$  be a bounded open neighborhood of  $K$ , and  $P$  be a differential operator of order 2 on  $\Omega$  such that*

- all coefficients of  $P$  are smooth and depend analytically on the variable  $z_a$ ,
- the principal symbol of  $P$ , namely  $p(z, \zeta) = Q_z(\zeta)$ , is a  $z$ -family of real quadratic forms such that  $\zeta_b \mapsto Q_z(0, \zeta_b)$  is definite on  $\mathbb{R}^{n_b}$  for any  $z \in \Omega$ .

Assume also that, for any  $\varepsilon \in [0, 1 + \eta)$ ,  $\eta > 0$ , the surface  $S_\varepsilon$  is non-characteristic for  $P$  at each point of  $S_\varepsilon$ .

Then, for any open neighborhood  $\tilde{\omega} \subset \Omega$  of  $S_0$ , there exists a neighborhood  $U$  of  $K$ , and constants  $\kappa, C, \mu_0 > 0$  such that for all  $\mu \geq \mu_0$  and  $u \in C_0^\infty(\mathbb{R}^n)$ , we have

$$\|u\|_{L^2(U)} \leq C e^{\kappa\mu} \left( \|u\|_{H_b^1(\tilde{\omega})} + \|Pu\|_{L^2(\Omega)} \right) + \frac{C}{\mu} \|u\|_{H^1(\Omega)}, \quad (2.1)$$

where we have denoted  $\|u\|_{H_b^1(\tilde{\omega})} = \sum_{|\beta| \leq 1} \|D_b^\beta u\|_{L^2(\tilde{\omega})}$ .

Unfortunately, in the present paper, this global result does not apply under this form. In order to state the refined (and more technical) version, used in the main part of the paper, we shall need some definitions taken from [LL15, Section 2.3].

## 2.2 Definitions and tools for propagating the information

We first define the following regularization process for functions  $f$  defined on  $\mathbb{R}^n$ , by  $f \mapsto f_\lambda$  with

$$f_\lambda := e^{-\frac{|D_a|^2}{\lambda}} f = \mathcal{F}_a^{-1} \left( e^{-\frac{|\zeta_a|^2}{\lambda}} \mathcal{F}_a(f)(\zeta_a, z_b) \right) (z_a),$$

where  $\mathcal{F}_a$  denotes the Fourier transform in the variable  $z_a$  only, or, equivalently

$$f_\lambda(z_a, z_b) = \left( \frac{\lambda}{4\pi} \right)^{\frac{n_a}{2}} \left( e^{-\frac{\lambda}{4}|\cdot|^2} *_{\mathbb{R}^{n_a}} f(\cdot, z_b) \right) (z_a) = \left( \frac{\lambda}{4\pi} \right)^{\frac{n_a}{2}} \int_{\mathbb{R}^{n_a}} f(y_a, z_b) e^{-\frac{\lambda}{4}|z_a - y_a|^2} dy_a.$$

Then, we also need to introduce frequency localization functions, i.e. appropriately smoothed Fourier multipliers. Let  $m(\zeta_a)$  be a smooth radial function (i.e. depending only on  $|\zeta_a|$ ), compactly supported in  $|\zeta_a| < 1$  such that  $m(\zeta_a) = 1$  for  $|\zeta_a| < 3/4$ . We denote by  $M^\mu$  the Fourier multiplier  $M^\mu u = m\left(\frac{D_a}{\mu}\right) u$ , that is

$$(M^\mu u)(z_a, z_b) = \mathcal{F}_a^{-1} \left( m\left(\frac{\zeta_a}{\mu}\right) \mathcal{F}_a(u)(\zeta_a, z_b) \right) (z_a).$$

Given  $\lambda, \mu > 0$ , we shall denote by  $M_\lambda^\mu$  the Fourier multiplier of symbol  $m_\lambda^\mu(\zeta_a) = m_\lambda\left(\frac{\zeta_a}{\mu}\right)$ , i.e.  $M_\lambda^\mu = m_\lambda^\mu(D_a) = m_\lambda\left(\frac{D_a}{\mu}\right)$  or

$$(M_\lambda^\mu u)(z_a, z_b) = \mathcal{F}_a^{-1} \left( m_\lambda\left(\frac{\zeta_a}{\mu}\right) \mathcal{F}_a(u)(\zeta_a, z_b) \right) (z_a),$$

with, according to the above notation for the subscript  $\lambda$ ,

$$m_\lambda(\zeta_a) = \left( \frac{\lambda}{4\pi} \right)^{\frac{n_a}{2}} \int_{\mathbb{R}^{n_a}} m(\eta_a) e^{-\frac{\lambda}{4}|\zeta_a - \eta_a|^2} d\eta_a. \quad (2.2)$$

Note that in this definition, the symbol is first regularized and then dilated. We stress the fact that these Fourier multipliers only act in the variable  $z_a$ .

The typical local estimate of [LL15, Theorem 3.1], which is the building block for semiglobal statements like that of Theorem 2.1, reads as

$$\|M_\mu^{\beta\mu}\sigma_\mu u\|_1 \leq Ce^{\kappa\mu} \left( \|M_\mu^{\alpha\mu}\vartheta_\mu u\|_1 + \|Pu\|_{L^2(B(0,R))} \right) + Ce^{-\kappa'\mu} \|u\|_1,$$

for all  $\mu \geq \mu_0$  and  $u \in C_0^\infty(\mathbb{R}^n)$ , where  $\sigma$  is a cutoff function in a small ball  $B(0,r)$ ,  $r < R$ , whereas  $\vartheta$  is a cutoff in only one side (the one where the information is taken) of the hypersurface passing through zero (and being non-characteristic).

Here, and below, the norm  $\|\cdot\|_1$  is the norm  $\|\cdot\|_{H^1(\mathbb{R}^n)}$ .

Such an estimate only provides information on the low frequency part of the function, through the frequency cutoff  $M_\mu^{\beta\mu}$ , with an exponentially small  $Ce^{-\kappa'\mu}$  remainder (as opposed to the  $1/\mu$  remainder term in (2.1)). Iterating this result allows us to propagate the low frequency information. In this section, we recall some tools, used in [LL15, Section 4], for this iterative procedure. They are aimed at describing how information on the low frequency part of the solution can be deduced from one subregion to another one.

The following is [LL15, Definition 4.4], given here in the context of second order operators.

**Definition 2.2.** Fix  $\Omega$  be an open set of  $\mathbb{R}^n = \mathbb{R}^{n_a} \times \mathbb{R}^{n_b}$  and  $P$  a differential operator of order 2 defined in  $\Omega$ , and  $(V_j)_{j \in J}$  and  $(U_i)_{i \in I}$  two finite collections of bounded open sets of  $\mathbb{R}^n$ . We say that  $(V_j)_{j \in J}$  is **under the strong dependence** of  $(U_i)_{i \in I}$ , denoted by

$$(V_j)_{j \in J} \triangleleft (U_i)_{i \in I},$$

if there exists  $W_i \Subset U_i$  such that for any  $\vartheta_i \in C_0^\infty(\mathbb{R}^n)$  such that  $\vartheta_i(z) = 1$  on a neighborhood of  $\overline{W_i}$ , for any  $\tilde{\vartheta}_j \in C_0^\infty(V_j)$  and for all  $\kappa, \alpha > 0$ , there exist  $C, \kappa', \beta, \mu_0 > 0$  such that for all  $(\mu, u) \in [\mu_0, +\infty) \times C_0^\infty(\mathbb{R}^n)$ , we have

$$\sum_{j \in J} \|M_\mu^{\beta\mu}\tilde{\vartheta}_{j,\mu} u\|_1 \leq Ce^{\kappa\mu} \left( \sum_{i \in I} \|M_\mu^{\alpha\mu}\vartheta_{i,\mu} u\|_1 + \|Pu\|_{L^2(\Omega)} \right) + Ce^{-\kappa'\mu} \|u\|_1.$$

If the cardinal of  $I$  is one, writing  $U$  the single set of the family  $(U_i)_{i \in I}$ , we simply denote  $(V_j)_{j \in J} \triangleleft U$ . We use the same convention for  $V$  in case the cardinal of  $J$  is one. The norm  $\|\cdot\|_1$  is taken in  $\mathbb{R}^n$ .

We summarize the properties of this relation in the following proposition [LL15, Proposition 4.5].

**Proposition 2.3.** *We have the following properties*

1. If  $(V_j)_{j \in J} \triangleleft (U_i)_{i \in I}$  with  $U_i = U$  for all  $i \in I$ , then  $(V_j)_{j \in J} \triangleleft U$ .
2. If  $V_i \Subset U_i$  for any  $i \in I$ , then,  $(V_i)_{i \in I} \triangleleft (U_i)_{i \in I}$ .
3. If  $V_i \Subset U_i$  for any  $i \in I$ , then  $\bigcup_{i \in I} V_i \triangleleft (U_i)_{i \in I}$ .
4. If for any  $i \in I$ ,  $V_i \triangleleft U_i$ , then  $(V_i)_{i \in I} \triangleleft (U_i)_{i \in I}$ . In particular, if for any  $i \in I$ ,  $U_i \triangleleft U$ , then  $(U_i)_{i \in I} \triangleleft U$ .

5. The relation  $\triangleleft$  is transitive, that is

$$[(V_j)_{j \in J} \triangleleft (U_i)_{i \in I} \text{ and } (U_i)_{i \in I} \triangleleft (W_k)_{k \in K}] \implies (V_j)_{j \in J} \triangleleft (W_k)_{k \in K}.$$

Note that we do not always have  $U \triangleleft U$ .

**Remark 2.4.** We stress the fact that the definition of  $\triangleleft$  actually depends on the set  $\Omega$ , the splitting  $\mathbb{R}^n = \mathbb{R}^{n_a} \times \mathbb{R}^{n_b}$  and the operator  $P$ . The dependence of  $\triangleleft$  upon these objects will be mentioned when needed. For the applications, it is important that the function  $u$  is not necessarily supported in  $\Omega$ .

In the following, we will only need to use this relation  $\triangleleft$  in some appropriate coordinate charts. However, it will not be a problem for what we want to prove, even on a compact manifold. Indeed, we will fix some coordinate chart on an open set  $\Omega \subset \mathbb{R}^n$  close to a point or close to a trajectory. Then, we will use the relation  $\triangleleft$  related to  $\Omega$  to finally obtain some estimates which will be invariant by changes of coordinates.

We will also use the following proposition, [LL15, Proposition 4.9], which allows to iterate local propagation results towards global ones.

**Proposition 2.5.** *Assume that there exists some open sets  $U_0, U_{i,j}, \omega_{i,j}, V_{i,j}$ , with  $j \in \llbracket 1, N \rrbracket$  and  $i \in I_j$  ( $I_j$  finite) such that we have*

$$\begin{aligned} U_{i,j} \triangleleft V_{i,j} \quad \text{and} \quad \omega_{i,j} \Subset U_{i,j}, \quad \text{for all } j \in \llbracket 1, N \rrbracket \text{ and } i \in I_j; \\ V_{m,l+1} \Subset \left[ U_0 \cup \bigcup_{j \in \llbracket 1, l \rrbracket} \bigcup_{i \in I_j} \omega_{i,j} \right], \quad \text{for all } l \in \llbracket 0, N-1 \rrbracket, \text{ and } m \in I_{l+1}, \end{aligned}$$

where we consider the union  $\bigcup_{j \in \llbracket 1, l \rrbracket}$  empty if  $l = 0$ . Then, we have  $\left[ U_0 \cup \bigcup_{j \in \llbracket 1, N \rrbracket} \bigcup_{i \in I_j} \omega_{i,j} \right] \triangleleft V_0$  for any open set  $V_0$  such that  $U_0 \Subset V_0$ .

In this proposition, the local propagation results is  $U_{i,j} \triangleleft V_{i,j}$  but the iteration is made by packets. Roughly speaking, we use all sets corresponding to indices  $i \in I_j, j \leq l$  to deduce the information on the sets with indices  $i \in I_{l+1}$ .

### 2.3 Semiglobal estimates along foliation by hypersurfaces

Now, we formulate the results of [LL15] in the form they will be used in the next section, which is different from Theorem 2.1 with two respects:

- First, we keep the formulation with  $\triangleleft$ ; this means that we keep a frequency cutoff in both handsides of the estimate, as well as an exponentially small remainder (it is a low frequency estimate only). This allows to patch estimates together (which is no longer the case when the high frequencies have been taken into account, i.e. when estimates take the form of (2.1)). The high frequencies will then be taken into account to close the estimates with two different methods in Section 3.3.
- Second, we allow the linear change of variables between the two splittings (namely  $(z_a, z_b)$  for the analytic dependence and  $(z', z_n)$  for the geometry) to be replaced by a diffeomorphism, which shall be very useful in the following.

We give a first statement that is a low frequency formulation of Theorem 2.1, using the notation  $\triangleleft$  (see [LL15, Theorem 4.7]).

**Theorem 2.6.** *Under the assumptions of Theorem 2.1, there exists an open neighborhood  $U$  of  $K$  such that*

$$U \triangleleft \tilde{\omega}.$$

This essentially means that Estimate (2.1) may be replaced by the following: for all  $\chi \in C_0^\infty(U)$  and  $\varphi \in C_0^\infty(\Omega)$  such that  $\varphi = 1$  on a neighborhood of  $\tilde{\omega}$ , we have

$$\|M_\mu^{\beta\mu}\chi_\mu u\|_1 \leq C e^{\kappa\mu} \left( \|M_\mu^\mu \varphi_\mu u\|_1 + \|Pu\|_{L^2(\Omega)} \right) + C e^{-\kappa'\mu} \|u\|_1, \quad (2.3)$$

(for any  $\kappa > 0$ , there exist  $C, \beta, \kappa', \mu_0 > 0$  such that for  $\mu \geq \mu_0$ ) i.e. keep the frequency cutoff and the exponentially small remainder.

A remaining drawback of this statement, given by the geometric framework of Theorems 2.1, is that the hypersurfaces are described by graphs in some coordinates (namely  $(z', z_n)$ ). This choice of description is mainly convenient to make the foliation more effective and order the hypersurfaces more easily, but is too rigid for the application in the present paper. Now, we give a slight variant of Theorem 2.6, more adapted to possible changes of variables.

**Theorem 2.7.** *Let  $\Omega \subset \mathbb{R}^n = \mathbb{R}^{n_a} \times \mathbb{R}^{n_b}$  and  $P$  be a differential operator of order 2 on  $\Omega$  such that*

- *all coefficients of  $P$  are smooth and depend analytically on the variable  $z_a$ ,*
- *the principal symbol of  $P$  namely  $p(z, \zeta) = Q_z(\zeta)$  is a  $z$ -family of real quadratic forms, such that  $\zeta_b \mapsto Q_z(0, \zeta_b)$  is definite on  $\mathbb{R}^{n_b}$  for any  $z \in \Omega$ .*

*Let  $\Phi$  be a diffeomorphism of class  $C^2$  from  $\Omega$  to  $\tilde{\Omega} = \Phi(\Omega)$ . Assume that the Geometric Setting of Theorem 2.1 is satisfied for some  $D, G, K, S_\varepsilon$  on  $\tilde{\Omega}$  (and not on  $\Omega$ ). Assume further that for any  $\varepsilon \in [0, 1 + \eta)$ ,  $\eta > 0$ , the surface  $\Phi^{-1}(S_\varepsilon)$  (well defined on  $\Omega$ ) is non-characteristic for  $P$  (at every point of  $\Phi^{-1}(S_\varepsilon)$ ).*

*Then, for all neighborhood  $\omega$  of  $\Phi^{-1}(S_0)$ , there exists an open neighborhood  $U \subset \Omega$  of  $\Phi^{-1}(K)$  such that*

$$U \triangleleft \omega,$$

*where  $\triangleleft = \triangleleft_{\Omega, P}$  is related to the operator  $P$  defined on  $\Omega$  (see Remark 2.4).*

In this result, the operator  $P$  has the appropriate form in  $\Omega \subset \mathbb{R}^{n_a} \times \mathbb{R}^{n_b}$  whereas the geometry of the surfaces is defined in  $\tilde{\Omega}$ , both being linked by a diffeomorphism.

With this theorem in hand, we may now prove the results presented in Section 1.2.

### 3 The hypoelliptic wave equation, proof of Theorem 1.13

The main goal of this section is to prove Theorem 1.13. The proof is inspired by the case of the classical wave equation (see [LL15, Section 6.1]) with mainly two differences:

- We are not able to construct global coordinates near any path  $\gamma$ , as in the case of the wave equation. However, if  $\gamma$  is a normal geodesic (see definition 3.3 below), we are able to do this construction locally. Then, this local result needs to be iterated.



- The  $H^1$  norm is no longer equivalent to the energy norm for the hypoelliptic operator. We thus need to use hypoelliptic estimates instead.

Therefore, the proof is divided in two parts, the first of which concerning the geometric iteration process, and the second the energy estimates.

Let us start by introducing geometrical definitions and facts used all along the proofs. First, denote by  $\ell = \ell(x, \xi) \in C^\infty(T^*\mathcal{M})$  the principal symbol of the operator  $\mathcal{L}$ , that is

$$\ell(x, \xi) = \sum_{i=1}^m \langle \xi, X_i(x) \rangle^2. \quad (3.1)$$

where  $\langle \xi, X(x) \rangle = \langle \xi, X(x) \rangle_{T_x^*\mathcal{M}, T_x\mathcal{M}}$  is the duality bracket.

**Remark 3.1.** In view of unique continuation results, note that a local hypersurface  $\{\varphi = 0\}$  at  $x_0 \in \mathcal{M}$  (where  $\varphi : \mathcal{M} \rightarrow \mathbb{R}$  with  $\varphi(x_0) = 0$  and  $d\varphi(x_0) \neq 0$ ) is characteristic for the operator  $\mathcal{L}$  if  $\ell(x_0, d\varphi(x_0)) = 0$ , that is, according to (3.1), if

$$\langle d\varphi(x_0), X_i(x) \rangle = 0 \quad \text{for all } i \in \{1, \dots, m\}.$$

**Definition 3.2.** The Hamiltonian curve of the symbol  $\ell$  issued from  $\rho_0 \in T^*\mathcal{M}$  is the unique maximal solution  $\rho(s) = (\gamma(s), \xi(s))$  of the ODE

$$\dot{\rho}(s) = H_\ell(\rho(s)), \quad \rho(0) = \rho_0, \quad (3.2)$$

where  $H_\ell$  is the Hamiltonian vector field associated to the Hamiltonian  $\ell$ . In local charts, this is

$$\begin{cases} \dot{\gamma}(s) &= \frac{\partial}{\partial \xi} \ell(\gamma(s), \xi(s)), \\ \dot{\xi}(s) &= -\frac{\partial}{\partial x} \ell(\gamma(s), \xi(s)). \end{cases}$$

Such a curve is smooth (even real analytic since  $\mathcal{M}$  and  $\ell$  are). Moreover, the first equation writes

$$\dot{\gamma}(s) = \sum_{j=1}^m 2 \langle \xi(s), X_j(\gamma(s)) \rangle X_j(\gamma(s)), \quad (3.3)$$

so that the projection on  $\mathcal{M}$  of a Hamiltonian curve is a smooth horizontal curve (see Definition 1.3). Note also that a Hamiltonian curve, locally defined according to the Cauchy-Lipschitz theorem, is actually globally defined (see e.g. [Rif14, Proposition 2.3.2]). Note finally that the Hamiltonian  $\ell$  is preserved along a Hamiltonian curve of  $\ell$ , i.e.  $\ell(\rho(s)) = \ell(\rho_0)$  for every  $\rho_0 \in T^*\mathcal{M}$ ,  $s \in \mathbb{R}$ , where  $\rho$  is the solution to (3.2).

Given a Hamiltonian curve  $\rho = (\gamma, \xi) : [0, S] \rightarrow T^*\mathcal{M}$ , one may compute the speed along the horizontal geodesic  $\gamma : [0, S] \rightarrow \mathcal{M}$ . Namely, we have (see Lemma C.1 for a proof of the first identity)

$$g(\gamma(s), \dot{\gamma}(s)) = \sum_{j=1}^m 4 \langle \xi(s), X_j(\gamma(s)) \rangle^2 = 4\ell(\gamma(s), \xi(s)) = 4\ell(\rho(0)), \quad \text{for all } s \in [0, S].$$

We can hence compute the length of the horizontal path  $\gamma$ , namely,

$$\text{length}(\gamma) = \int_0^S \sqrt{g(\gamma(s), \dot{\gamma}(s))} ds = 2S \sqrt{\ell(\rho(0))}.$$

This motivates the following definition:

**Definition 3.3.** We say that a horizontal curve  $\gamma : [0, L_0] \rightarrow \mathcal{M}$  is a normal geodesic if there exists  $\xi(s) \in T_{\gamma(s)}^* \mathcal{M}$  such that  $s \mapsto (\gamma(s), \xi(s))$  is a Hamiltonian curve of the symbol  $\ell$  with  $\ell(\gamma(s), \xi(s)) = \frac{1}{4}$ .

As a consequence of this definition, such a curve  $\gamma : [0, L_0] \rightarrow \mathcal{M}$  has unit speed:

$$\sqrt{g(\gamma(s), \dot{\gamma}(s))} = 2\sqrt{\ell(\gamma(s), \xi(s))} = 2\sqrt{\ell(\rho(0))} = 1, \quad \text{for all } s \in [0, L_0],$$

(it is hence parametrized by arclength) and length  $L_0$ .

**Definition 3.4.** We say that a curve  $\gamma : [0, L] \rightarrow \mathcal{M}$  is a minimizing geodesic path between  $x$  and  $y$  if  $\gamma(0) = x$ ,  $\gamma(L) = y$ , if  $\gamma$  is a horizontal curve and if we have  $d_{\mathcal{L}}(x, y) = \text{length}(\gamma)$  together with  $g(\gamma(s), \dot{\gamma}(s))$  constant.

See [Rif14, Chapter 2] or [ABB16a, Section 3.3]. Note that all above definitions are intrinsic. The following key result in our proofs is [RT05, Theorem 1.1].

**Theorem 3.5** (Rifford-Trélat [RT05]). *For all  $x_1 \in \mathcal{M}$ , there exists a dense subset  $N_{x_1} \subset \mathcal{M}$  such that for all  $x \in N_{x_1}$ , there is a (unique) minimizing geodesic path between  $x_1$  and  $x$ . Moreover, this path is a normal geodesic path.*

As a direct corollary, we obtain the following result, which is a key step in the proof of Theorem 1.13 (in particular for obtaining the minimal time).

**Corollary 3.6.** *Let  $\omega$  a nonempty open subset of  $\mathcal{M}$  and  $T > \sup_{x \in \mathcal{M}} d_{\mathcal{L}}(x, \omega)$ . Then, for any  $x_1 \in \mathcal{M}$ , there exists  $x_0 \in \omega$  and a normal geodesic path  $\gamma : [0, L] \rightarrow \mathcal{M}$  of length  $L \in (0, T)$ , so that  $\gamma(L) = x_1$ ,  $\gamma(0) = x_0$ , and  $\gamma$  is also a minimizing geodesic path.*

Note that this path being minimizing, it is in particular non self-intersecting.

*Proof.* According to the definition of  $T$ , the set  $\mathcal{O}_{x_1} := \omega \cap \{x \in \mathcal{M}, d_{\mathcal{L}}(x, x_1) < T\}$  is open (the continuity of  $d_{\mathcal{L}}$  is a consequence of the Chow-Rashevski Theorem 1.2) and nonempty. Hence, it intersects the dense set  $N_{x_1}$  given by Theorem 3.5. Taking any  $x_0 \in \mathcal{O}_{x_1} \cap N_{x_1}$ , there is a normal geodesic curve  $\gamma : [0, L] \rightarrow \mathcal{M}$  joining  $x_0$  and  $x_1$  of length  $L \in (0, T)$  which is also a minimizing geodesic.  $\square$

Now, one of the main purposes of the present Section 3 is to give a proof of the following proposition (which essentially amounts to (1.28)). Indeed, Theorem 1.31 in the simple case  $s = k$  will follow. The proof of the full range of  $s$  in Theorem 1.31 will require more work.

**Proposition 3.7.** *Let  $T > 0$ ,  $x_0, x_1 \in \mathcal{M}$ , and assume that there is a normal geodesic path of length  $L \in (0, T)$  between  $x_0$  and  $x_1$ . Then, for any  $\varepsilon > 0$ , there exists  $\tilde{\varepsilon} > 0$ , there is  $C, \kappa, \mu_0 > 0$  such that for all  $u \in H^1(\cdot - T, T[\times \mathcal{M}])$  solution of  $(\partial_t^2 + \mathcal{L})u = 0$  on  $\cdot - T, T[\times \mathcal{M}$ , and for all  $\mu \geq \mu_0$ , we have*

$$\|u\|_{L^2(\cdot - \tilde{\varepsilon}, \tilde{\varepsilon}[\times B(x_1, \tilde{\varepsilon}))} \leq C e^{\kappa \mu} \|u\|_{L^2(\cdot - T, T[\times B(x_0, \varepsilon))} + \frac{C}{\mu} \|u\|_{H^1(\cdot - T, T[\times \mathcal{M})}.$$

Note here that the  $H^1(\cdot - T, T[\times \mathcal{M})$  norm is the usual one issued from the structure of compact manifold. Also, since  $\varepsilon$  and  $\tilde{\varepsilon}$  are arbitrary small, the balls could be defined according to any metric on  $\mathcal{M}$  defining an equivalent topology (balls could equivalently be replaced by neighborhoods). Yet, we will use the distance induced by the sub-Riemannian geometry since it is the important one in other parts of the proof (for defining the distance  $d_{\mathcal{L}}$  for instance) and to avoid any confusion.

Using Proposition 3.7, together with Corollary 3.6 and a compactness argument directly yields the following global estimate, which is the main result of this step.

**Corollary 3.8.** *Let  $\omega$  a nonempty open set  $\mathcal{M}$  and  $T > \sup_{x \in \mathcal{M}} d_{\mathcal{L}}(x, \omega)$ . Then, there exist  $\tilde{\varepsilon}, C, \kappa, \mu_0 > 0$  such that for all  $u \in H^1(\cdot - T, T[\times \mathcal{M})$  solution of  $(\partial_t^2 + \mathcal{L})u = 0$  on  $\cdot - T, T[\times \mathcal{M}$ , and for all  $\mu \geq \mu_0$ , we have,*

$$\|u\|_{L^2(\cdot - \tilde{\varepsilon}, \tilde{\varepsilon}[\times \mathcal{M})} \leq C e^{\kappa \mu} \|u\|_{L^2(\cdot - T, T[\times \omega)} + \frac{C}{\mu} \|u\|_{H^1(\cdot - T, T[\times \mathcal{M})}.$$

The last step towards the proof of Theorem 1.31 (in the case  $s = k$ ; the general case requires a slightly more precise version of this result), performed in Section 3.3, will be to transfer the time-space information carried by this inequality into some Sobolev norm  $\mathcal{H}_{\mathcal{L}}^s$  related to the operator  $\mathcal{L}$ . This will be the object of the next paragraph.

Sections 3.1 and 3.2 are devoted to the proof of Proposition 3.7. For this, the main point is to apply the quantitative unique continuation result adapted to changes of variable, namely Theorem 2.7. The main drawback of this result is that it only works in a subset  $\Omega$  of  $\mathbb{R}^n$ , i.e. it is not invariant by diffeomorphism. More precisely, all estimates obtained from the results of Section 2.3 can only be patched together in straight coordinates.

As a consequence, in the present context, we need to introduce *global* coordinates near the trajectory between the points  $x_0$  and  $x_1$ . Then, locally, we shall define hypersurfaces to match the geometric setting of Section 2.1. This will be done in another set of coordinates (not necessarily analytic), which is allowed by the precise formulation of Theorem 2.7.

Finally, the energy estimates needed to conclude the proof of Theorem 1.31 are performed in Section 3.3.

### 3.1 Step 1: Geometric setting and non-characteristic hypersurfaces

We now consider  $x_1, x_0 \in \mathcal{M}$  (recall that  $\dim(\mathcal{M}) = d$ ), together with a normal geodesic  $\gamma$  as in Corollary 3.6. The curve  $\gamma$  is non self-intersecting, so that there exists an open neighborhood  $\mathcal{N}_\gamma$  of  $\gamma$  in  $\mathcal{M}$ , an open set  $\Omega_\gamma \subset \mathbb{R}^d$  and an *analytic* diffeomorphism

$$\phi_\gamma : \mathcal{N}_\gamma \rightarrow \Omega_\gamma \subset \mathbb{R}^d \tag{3.4}$$

such that  $\phi_\gamma(\gamma([0, L])) \subset \Omega_\gamma$ . From now on, **we shall only work in the chart**  $(\Omega_\gamma, \phi_\gamma)$ . For the sake of readability, we shall keep the same notation for all objects pulled back from  $\mathcal{N}_\gamma \subset \mathcal{M}$  to  $\Omega_\gamma \subset \mathbb{R}^d$ . For instance, we shall still denote  $X_j$  instead of  $(\phi_\gamma^{-1})^* X_j$ ,  $\mathcal{L}$  instead of  $(\phi_\gamma^{-1})^* \mathcal{L} \phi_\gamma^*$ ,  $\gamma \subset \Omega_\gamma$  instead of  $(\phi_\gamma^{-1})^* \gamma$  etc... Recall that all above definitions (in particular Definitions 3.2, 3.3 and 3.4) are intrinsic, so that, in particular, the whole sub-Riemannian structure may be transported to  $\Omega_\gamma$ , and the curve  $\gamma$  is still a normal geodesic in  $\Omega_\gamma$ .

Now, we define other local coordinates in which to construct the (local) noncharacteristic surfaces in order to apply Theorem 2.7. We first need the following lemma.

**Lemma 3.9** (local coordinates). *Given  $\gamma : [0, L] \rightarrow \Omega_\gamma$  a normal geodesic path, for any point  $x_0 = \gamma(s_0)$  on this curve, there exists an open neighborhood  $V_{x_0}$  of  $x_0$ , and appropriate coordinates, denoted  $x = (\tilde{x}, x_d) \in \mathbb{R}^{d-1} \times \mathbb{R}$  (with associated cotangent variables  $\xi = (\tilde{\xi}, \xi_d) \in \mathbb{R}^{d-1} \times \mathbb{R}$ ) in which*

- the symbol  $\ell$  can be written

$$\ell(x, \xi) = \xi_d^2 + r(x, \tilde{\xi}), \tag{3.5}$$

where  $r(x, \tilde{\xi})$  is a non negative quadratic form in  $\tilde{\xi}$ ;

- the point  $x_0$  is sent to  $(0, \dots, 0, s_0)$  and  $\gamma(s) = (0, \dots, 0, s)$  for  $s$  close enough to  $s_0$ .

*Proof.* This is e.g. a consequence of [Hör85, Appendix C.5]. More precisely, we denote by  $s \mapsto (\gamma(s), \xi(s)) \in T^*\Omega \setminus 0$  the Hamiltonian curve associated to the normal geodesic  $\gamma$ . We let  $\varphi$  be a real-valued function defined locally in a neighborhood of  $x_0$ , such that  $\varphi(x_0) = 0$  and  $d\varphi(x_0) = \xi(s_0)$ . Then,  $\varphi$  is a non-characteristic function for  $\ell$  at  $x_0$  since  $\ell(x_0, d\varphi(x_0)) = \ell(\gamma(s_0), \xi(s_0)) = \frac{1}{4} \neq 0$ , according to the definition of a normal geodesic path. According to [Hör85, Corollary C.5.3], there are local coordinates  $(\check{x}, x_d) \in \mathbb{R}^{d-1} \times \mathbb{R}$ , defined in a neighborhood of 0 in which

- $x_0$  is sent to 0,
- the surface  $\{\varphi = 0\}$  is given by  $\{x_d = 0\}$ ,
- the first item of the lemma holds.

We now just have to check that the second item of the lemma holds in these coordinates. First remark that  $d\varphi(x_0) = \xi(s_0)$  is sent to  $(0, \xi_d)$  for some  $\xi_d \in \mathbb{R}^*$ , so that  $\ell(\gamma(s_0), \xi(s_0)) = \frac{1}{4}$  implies, with the form of  $\ell$  in (3.5), that  $\xi_d^2 = \frac{1}{4}$ . Up to changing  $x_d \mapsto -x_d$  (without changing any of the three properties described in the above items), we may further assume that  $\xi_d > 0$ . Hence  $d\varphi(x_0) = \xi(s_0)$  is sent to  $(\check{\xi}, \xi_d) = (0, \frac{1}{2})$ . Second, the form of  $\ell$  in (3.5) yields that the Hamiltonian curves of  $\ell$  satisfy in these coordinates:

$$\dot{\check{x}} = \partial_{\check{\xi}} r(x, \check{\xi}), \quad \dot{x}_d = 2\xi_d, \quad \dot{\check{\xi}} = -\partial_{\check{x}} r(x, \check{\xi}), \quad \dot{\xi}_d = -\partial_{x_d} r(x, \check{\xi}). \quad (3.6)$$

The Hamiltonian curve associated to the normal geodesic  $\gamma$  in these coordinates is the unique curve of (3.6) passing through  $x = 0$  and  $(\check{\xi}, \xi_d) = (0, \frac{1}{2})$ . But the function  $(\check{x}, x_d, \check{\xi}, \xi_d)(s) = (0, s - s_0, 0, \frac{1}{2})$  solves (3.6) since  $\partial_{\check{x}} r, \partial_{x_d} r$  are quadratic in  $\check{\xi}$  and  $\partial_{\check{\xi}} r$  is linear in  $\check{\xi}$  (and, in particular, all vanish at  $\check{\xi} = 0$ ). It also starts at time  $s_0$  at  $(0, 0, 0, \frac{1}{2})$ , so that  $(0, s - s_0, 0, \frac{1}{2})$  is the sought Hamiltonian curve. As a consequence, the normal geodesic  $\gamma$  is given by  $(\check{x}, x_d)(s) = (0, s - s_0)$  in these coordinates. This concludes the proof after the linear change of variable  $(\check{x}, x_d) \mapsto (\check{x}, x_d + s_0)$ .  $\square$

**Lemma 3.10** (Construction of non characteristic hypersurfaces in normal coordinates). *Assume that for some  $r_0, l_0 > 0$ , the symbol  $\ell$  is given by (3.5) in coordinates  $(\check{x}, x_d) \in \tilde{\Omega} := B(0, r_0) \times ]-l_0, 2l_0[ \subset \mathbb{R}_{\check{x}}^{d-1} \times \mathbb{R}_{x_d}$ . Then, for any  $t_0 > l_0$  and  $0 < r_1 < r_0$ , there exists  $D, G, K, S_\varepsilon$  satisfying items 1-2-3 of the Geometric Setting of Section 2.1 in  $\mathbb{R}^n = \mathbb{R}^{d+1}$  in the coordinates*

$$(z', z_n) = (t, \check{x}, x_d), \quad \text{with} \quad z' = (t, \check{x}) \quad \text{and} \quad z_n = x_d, \quad (3.7)$$

together with

4.  $D \subset [-t_0, t_0] \times \overline{B}(0, r_1)$ , that is  $S_0 \subset [-t_0, t_0]_t \times \overline{B}(0, r_1)_{\check{x}} \times \{0\}_{x_d} \subset \mathbb{R}_t \times \subset \mathbb{R}_{\check{x}}^{d-1} \times \{0\}_{x_d}$ ;
5.  $\{0\}_t \times \{0\}_{\check{x}} \times [0, l_0]_{x_d} \subset K$ ;
6. for any  $\varepsilon \in [0, 1 + \eta)$ , the surface  $S_\varepsilon$  is non-characteristic for  $P = \partial_t^2 + \mathcal{L}$  at each point of  $S_\varepsilon$ .

*Proof.* The principal symbol of the operator  $P = \partial_t^2 + \mathcal{L}$  in the coordinates of Lemma 3.9 is given by

$$p(\check{x}, x_d, \tau, \check{\xi}, \xi_d) = -\tau^2 + \ell(\check{x}, x_d, \tau, \check{\xi}, \xi_d) = -\tau^2 + \xi_d^2 + r(x, \check{\xi}), \quad \xi = (\check{\xi}, \xi_d), \quad (3.8)$$

To match the geometric setting of Section 2.1, we define the coordinates  $(z', z_n)$  according to (3.7), as well as

$$D = \left\{ (t, \check{x}) \left| \left( \frac{\check{x}}{r_1} \right)^2 + \left( \frac{t}{t_0} \right)^2 < 1 \right. \right\}, \quad G(t, \check{x}, \varepsilon) = \varepsilon l_0 \psi \left( \sqrt{\left( \frac{\check{x}}{r_1} \right)^2 + \left( \frac{t}{t_0} \right)^2} \right),$$

$$\phi_\varepsilon(t, \check{x}, x_d) := G(t, \check{x}, \varepsilon) - x_d, \quad S_\varepsilon = \{\phi_\varepsilon = 0\}, \quad \varepsilon \in [0, 1 + \eta],$$

with  $r_1, \eta > 0$  small to be fixed, where  $\psi$  is such that, for some  $\eta_0, \eta_1 > 0$ ,

$$\psi : [-1 - \eta_0, 1 + \eta_0] \rightarrow [-\eta_1, 1], \quad \text{smooth and even,} \quad \psi(\pm 1) = 0, \quad \psi(0) = 1,$$

$$\psi(s) \geq 0, \quad \text{if and only if } s \in [-1, 1], \quad \text{and } |\psi'| \leq \alpha \text{ on } [-1 - \eta_0, 1 + \eta_0],$$

with  $1 < \alpha < \frac{t_0}{l_0}$ . This is possible since  $\frac{t_0}{l_0} > 1$ .

Note first that Item 4 is satisfied according to the definition of  $D$ . Note also that the point  $(t, \check{x}, x_d) = (0, 0, l_0)$  belongs to  $S_1 = \{\phi_1 = 0\}$ . Hence, Item 5 is satisfied since  $\{0\}_t \times \{0\}_{\check{x}} \in D$  and  $G(0_{t, \check{x}}, 1) = l_0$ , so that we have  $0 \leq x_d \leq G(0_{t, \check{x}}, 1) = l_0$  if  $x_d \in [0, l_0]$ .

Let us now check Item 6. We have

$$d\phi_\varepsilon(t, \check{x}, x_d) = \varepsilon l_0 \left( \left( \frac{\check{x}}{r_1} \right)^2 + \left( \frac{t}{t_0} \right)^2 \right)^{-1/2} \psi' \left( \sqrt{\left( \frac{\check{x}}{r_1} \right)^2 + \left( \frac{t}{t_0} \right)^2} \right) \left( \frac{tdt}{t_0^2} + \frac{\check{x}d\check{x}}{r_1^2} \right) - dx_d.$$

Given the form of the principal symbol of the operator  $P$  in these coordinates (see (3.8)), we obtain

$$-p(\check{x}, x_d, d\phi_\varepsilon(t, \check{x}, x_d)) = \varepsilon^2 l_0^2 \frac{t^2}{t_0^4} \left( \left( \frac{\check{x}}{r_1} \right)^2 + \left( \frac{t}{t_0} \right)^2 \right)^{-1} |\psi'|^2 - l_0^2 \frac{\varepsilon^2}{r_1^4} r(x, \check{x}) \left( \left( \frac{\check{x}}{r_1} \right)^2 + \left( \frac{t}{t_0} \right)^2 \right)^{-1} |\psi'|^2 - 1$$

where  $|\psi'|^2$  is taken at the point  $\left( \sqrt{\left( \frac{\check{x}}{r_1} \right)^2 + \left( \frac{t}{t_0} \right)^2} \right)$ . Since  $r$  is non negative, we get

$$-p(\check{x}, x_d, d\phi_\varepsilon(t, \check{x}, x_d)) \leq \varepsilon^2 l_0^2 \frac{t^2}{t_0^4} \left( \left( \frac{\check{x}}{r_1} \right)^2 + \left( \frac{t}{t_0} \right)^2 \right)^{-1} |\psi'|^2 - 1$$

Since  $|\psi'| \leq \alpha$  and  $\varepsilon \in [0, 1 + \eta]$ , we obtain for any  $(t, \check{x}, x_d) \in \overline{D} \times [0, l_0]$ ,

$$\begin{aligned} -p(\check{x}, x_d, d\phi_\varepsilon(t, \check{x}, x_d)) &\leq \frac{\varepsilon^2}{t_0^2} l_0^2 \left( \frac{t}{t_0} \right)^2 \left( \frac{t}{t_0} \right)^{-2} \alpha^2 - 1 \\ &\leq (1 + \eta)^2 \frac{l_0^2}{t_0^2} \alpha^2 - 1 < 0, \end{aligned}$$

the last constant being negative for  $\eta$  small enough because  $\alpha < \frac{t_0}{l_0}$ . Therefore, the surface  $S_\varepsilon = \{\phi_\varepsilon = 0\}$  is noncharacteristic for any  $\varepsilon \in [0, 1 + \eta]$ , which concludes the proof of Item 6, and hence of the lemma.  $\square$

### 3.2 Step 2: Propagation of smallness

We recall that  $z = (t, x)$  and  $n = d + 1$ , and introduce the notation

$$\text{Vois}(K, r) := \bigcup_{z \in K} B(z, r) \quad \text{for } K \subset \mathbb{R}^n.$$

**Lemma 3.11** (local version near a piece of a normal geodesic). *Given  $\gamma : [0, L] \rightarrow \Omega_\gamma$  a normal geodesic path and fix  $\tilde{s} \in [0, L]$ . Then, there exists  $\tilde{L}_s, \tilde{r}_s$  small such that for any  $s_0 \in [0, L]$ ,  $L_0 > 0$  so that  $[s_0, s_0 + L_0] \subset ]\tilde{s} - \tilde{L}_s, \tilde{s} + \tilde{L}_s[$ , for all  $T > L_0$  and  $0 < r_1 < \tilde{r}_s$ , there exists  $r_2 > 0$  so that*

$$] - r_2, r_2[ \times \text{Vois}(\gamma([s_0, s_0 + L_0]), r_2) \triangleleft ] - T, T[ \times \text{Vois}(\gamma(s_0), r_1), \quad (3.9)$$

where  $\triangleleft$  is related to the operator  $P = \partial_t^2 + \mathcal{L}$  in the set  $] - T - \varepsilon, T + \varepsilon[ \times \Omega_\gamma$ .

See Figure 2 for a picture of the sets involved in (3.9).

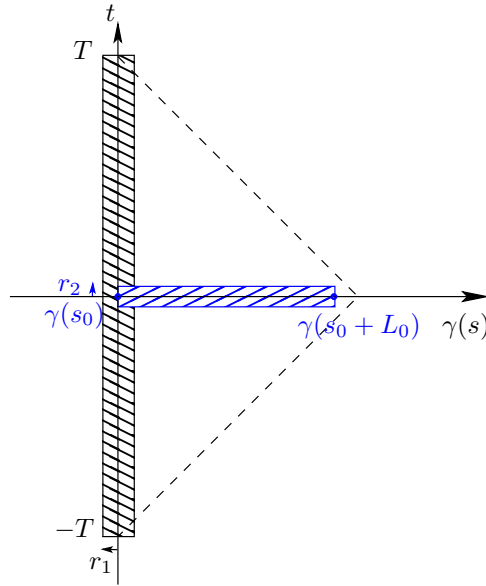


Figure 2: Schematic representation of the sets involved in Lemma 3.11.

Recall that  $\text{length}(\gamma([s_0, s_0 + L_0])) = L_0 < T$ ; the dashed line represents the boundary of the wave cone.

The proof is almost the same as in Theorem 6.3 of [LL15]. The only difference is that the coordinates where we have a nice diagonal form for the operator  $P$  are not global and are not those where we want to apply the local result. Note that this would not have been a problem if we had proved that the relation  $\triangleleft$  is invariant by change of coordinates. Now, we perform the following steps:

- use Lemma 3.9 to obtain nice coordinates in a neighborhood of  $\gamma(s)$ ;
- construct the non characteristic hypersurfaces in these coordinates according to Lemma 3.10;
- apply Theorem 2.7 in these coordinates, w.r.t. those surfaces, keeping in mind that the fact to be non characteristic is invariant by changes of coordinates.

Note also that the presence of  $\tilde{s}$  and  $s_0$  in the statement may seem strange and it would look simpler to consider only  $s_0$  and intervals  $[s_0, s_0 + L_0]$ . Yet, this will be useful later in compactness and covering arguments where we have to subtract and cut some intervals.

*Proof.* By Lemma 3.9, we can find a diffeomorphism from a neighborhood of  $\gamma(\tilde{s})$  in  $\Omega_\gamma$  to  $B_{\mathbb{R}^{d-1}}(0, r_0) \times ]\tilde{s} - 4\tilde{L}_s, \tilde{s} + 4\tilde{L}_s[\ni (\tilde{x}, x_d)$  in which the symbol  $\ell$  is as in (3.5) and  $\gamma(s)$  is given by  $(\tilde{x}, x_d) = (0, \tilde{s})$ . By translation and using  $[s_0, s_0 + L_0] \subset ]\tilde{s} - \tilde{L}_s, \tilde{s} + \tilde{L}_s[$ , we have a diffeomorphism  $\Phi$  from a neighborhood  $V$  of  $\gamma(s_0)$  in  $\Omega_\gamma$  onto  $B_{\mathbb{R}^{d-1}}(0, r_0) \times ]-L_0, 2L_0[\ni (\tilde{x}, x_d)$  where  $\Phi(\gamma(s_0)) = 0$ .

For later purposes, fix  $r_3$  so that

$$\Phi^{-1}(B_{\mathbb{R}^d}(0, r_3)) \Subset B_{\mathbb{R}^d}(\gamma(s_0), r_1). \quad (3.10)$$

We will keep denoting  $\Phi$  the same diffeomorphism acting on  $] -T - \varepsilon, T + \varepsilon[ \times V$  leaving the  $t$  variable unchanged, and set  $\tilde{\Omega} := ] -T - \varepsilon, T + \varepsilon[ \times B_{\mathbb{R}^{d-1}}(0, r_0) \times ] -L_0, 2L_0[$ .

Now, construct  $D, G, K, S_\varepsilon$  according to Lemma 3.10. In particular, all surfaces  $S_\varepsilon$  are non characteristic for  $P$  (or, more precisely, for  $(\Phi^*)^{-1}P\Phi^*$  in  $\tilde{\Omega}$ ). By change of coordinates, the surface  $\Phi^{-1}(S_\varepsilon)$  is non characteristic for  $P$  at each point of  $\Phi^{-1}(S_\varepsilon)$ .

For  $0 < L_0 < t_0 < T$ , take  $\omega = \Phi^{-1}(] -t_0, t_0[ \times B(0, r_3))$  so that Item 4 of Lemma 3.10 implies that  $\omega$  is a neighborhood of  $\Phi^{-1}(S_0)$ .

The assumptions of Theorem 2.7 are fulfilled (with  $n_a = n = d + 1$ , i.e. in the Holmgren-John case), so there exists an open neighborhood  $U \subset \Omega$  of  $\Phi^{-1}(K)$  such that  $U \triangleleft \omega$ . Note here that the strict application of Theorem 2.7 yields this result for the relation  $\triangleleft ] -T - \varepsilon, T + \varepsilon[ \times V, P$ , but the latter then implies the same property for the relation  $\triangleleft ] -T - \varepsilon, T + \varepsilon[ \times \Omega_\gamma, P$ . We will keep the notation  $\triangleleft$ .

Moreover, Item 5 of Lemma 3.10 implies, after having applied  $\Phi^{-1}$ , that

$$\Phi^{-1}(\{0\}_t \times \{0\}_{\tilde{x}} \times [0, l_0]_{x_d}) \subset \Phi^{-1}(K) \subset U.$$

Using the form of  $P$  on  $\tilde{\Omega}$  and that  $\gamma$  is a normal geodesic, we obtain that  $\Phi^{-1}(\{0\}_t \times \{0\}_{\tilde{x}} \times [0, L_0]_{x_d}) = \{0\}_t \times \gamma([s_0, s_0 + L_0])$ . In particular since  $U$  is open and the previous set is compact, we can find  $r_2 > 0$  so that  $] -r_2, r_2[ \times \text{Vois}(\gamma([s_0 - r_2, s_0 + L_0]), r_2) \Subset U$ . Items 2 and 5 of Proposition 2.3 imply

$$] -r_2, r_2[_{\mathbb{R}_t} \times \text{Vois}(\gamma([s_0 - r_2, s_0 + L_0]), r_2) \triangleleft \omega.$$

Finally, the definition of  $\omega$ ,  $T > t_0$  and (3.10) imply  $\omega \Subset ] -T, T[ \times \text{Vois}(\gamma(s_0), r_1)$ . This gives the final result by applying again Item 2 and 5 of Proposition 2.3.  $\square$

We can iterate the previous local result to get a more global one, which will be the main step for Proposition 3.7.

**Proposition 3.12** (global version near a normal geodesic). *Given  $\gamma : [0, L] \rightarrow \Omega_\gamma$  a normal geodesic path, and let  $0 < L < T$ . Then, there exists  $r_0$  small, such that for any  $0 < r_1 < r_0$ , there exists  $r_2 > 0$  such that*

$$] -r_2, r_2[ \times \text{Vois}(\gamma([0, L]), r_2) \triangleleft ] -T, T[ \times \text{Vois}(\gamma(0), r_1). \quad (3.11)$$

*Proof of Proposition 3.12.* We prove the result for another  $\tilde{T} > T$ . It gives the result since it is arbitrary.

For any  $\tilde{s} \in [0, L]$ , Lemma 3.11 provides  $\tilde{L}_s$  and  $\tilde{r}_s$  and an interval  $]\tilde{s} - \tilde{L}_s, \tilde{s} + \tilde{L}_s[$  with the appropriate conclusion. By compactness of  $\gamma([0, L])$ , we can extract a finite covering such that  $\gamma([0, L]) \subset \bigcup_{j \in [1, N]} \gamma(]\tilde{s}_j - \tilde{L}_j, \tilde{s}_j + \tilde{L}_j[)$ . Then, the issue is that  $\sum_{j=1}^N 2\tilde{L}_j$  may be very large with respect to  $2T$ . To overcome this difficulty, starting from this covering, we can always obtain (for this, suppress some intervals and cut them when they overlap too much) a finite number of intervals  $[s_j, s_j + L_j]$  and associate times  $T_j$  that satisfy the following properties:

- $[0, L] \subset \cup_{j=1}^N ]s_j, s_j + L_j[$ ,
- $]s_j, s_j + L_j] \subset ]\tilde{s}_j - \tilde{L}_j, \tilde{s}_j + \tilde{L}_j[$ , for all  $j \in \llbracket 1, N \rrbracket$ ,
- $s_1 = 0$ ,
- $s_{j+1} \in ]s_j, s_j + L_j[$ , for all  $j \in \llbracket 1, N-1 \rrbracket$ ,
- $L < \sum_{j=1}^N L_j < T$ ,
- $L_j < T_j$  and  $\sum_{j=1}^N T_j < T$ .

So, for any  $j \in \llbracket 1, N \rrbracket$ , since  $]s_j, s_j + L_j] \subset ]\tilde{s}_j - \tilde{L}_j, \tilde{s}_j + \tilde{L}_j[$ , Lemma 3.11 can be applied to the path  $\gamma([s_j, s_j + L_j])$  and gives the existence of  $\tilde{r}_s$  associated to  $\tilde{s}_j$ , which we here denote  $r_0^j$ . We also denote by  $r_0$  the minimum of all  $r_0^j$ ,  $j \in \llbracket 1, N \rrbracket$ , so that the conclusion of Lemma 3.11 remains true with any choice of  $r_1^j < r_0$ . We next define  $r_1^j$  and  $r_2^j$  recursively in the following way:

- $r_1^1 = \min(r_0, r_1)/2$  and  $r_2^1$  is given by the Lemma 3.11 for the interval  $[s_1, s_1 + L_1] = [0, L_1]$ .
- We choose  $r_1^{j+1} = \min(r_0, r_2^j)/4$  and  $r_2^{j+1}$  is given by Lemma 3.11 applied to the path  $]s_j, s_j + L_j] \subset ]\tilde{s}_j - \tilde{L}_j, \tilde{s}_j + \tilde{L}_j[$  and the time  $T_j > L_j$ .

The conclusion of Lemma 3.11 is then

$$] - r_2^j, r_2^j[_{\mathbb{R}^t} \times \text{Vois}(\gamma([s_j, s_j + L_j]), r_2^j) \triangleleft ] - T_j, T_j[_{\mathbb{R}^t} \times \text{Vois}(\gamma(s_j), r_1^j). \quad (3.12)$$

Now, for any  $l \in \llbracket 1, N \rrbracket$ , consider a sequence of time  $(t_i^l)_{i \in I_l}$  such that  $(]t_i^l - r_2^l, t_i^l + r_2^l])_{i \in I_l}$  is a finite covering of  $] - T + \sum_{j=1}^l T_j, T - \sum_{j=1}^l T_j[$ . One can also impose  $t_i^l \in ] - T + \sum_{j=1}^l T_j, T - \sum_{j=1}^l T_j[$ .

We want to apply Proposition 2.5 with the following definitions for  $j \in \llbracket 1, N \rrbracket$ ,  $i \in I_j$

- $U_{i,j} = ]t_i^j - r_2^j, t_i^j + r_2^j[_{\times} \text{Vois}(\gamma([s_j, s_j + L_j]), r_2^j)$ ,
- $\omega_{i,j} = ]t_i^j - r_2^j/2, t_i^j + r_2^j/2[_{\times} \text{Vois}(\gamma([s_j, s_j + L_j]), r_2^j/2)$ ,
- $V_{i,j} = ]t_i^j - T_j, t_i^j + T_j[_{\times} \text{Vois}(\gamma(s_j), r_1^j)$ ,
- $U_0 = ] - T, T[_{\times} \text{Vois}(\gamma(0), 2r_1^1)$ .

Since Lemma 3.11 is invariant by translation in time, (3.12) and the choices of  $r_1^j, r_2^j$  give  $U_{i,j} \triangleleft V_{i,j}$ . We also have  $\omega_{i,j} \Subset U_{i,j}$ . So, the main point to check is

$$V_{m,l+1} \Subset \left[ U_0 \cup \bigcup_{j \in \llbracket 1, l \rrbracket} \bigcup_{i \in I_j} \omega_{i,j} \right], \quad \text{for all } m \in I_{l+1}, \quad \text{and } l \in \llbracket 1, N-1 \rrbracket. \quad (3.13)$$

We first check the degenerate case  $l = 0$ , which amounts to proving that  $V_{m,1} \Subset U_0$  for  $m \in I_1$ . Since  $t_m^1 \in ] - T + T_1, T - T_1[$  (by choice), we have  $]t_m^1 - T_1, t_m^1 + T_1[ \Subset ] - T, T[$ . Moreover, since  $s_1 = 0$ , we have  $\text{Vois}(\gamma(s_1), r_1^1) \Subset \text{Vois}(\gamma(0), 2r_1^1)$ , and  $V_{m,1} \Subset U_0$  by definition.



Concerning the case  $l \in \llbracket 1, N-1 \rrbracket$ , we prove the stronger property

$$V_{m,l+1} \Subset \bigcup_{i \in I_l} \omega_{i,l}, \quad \text{for all } m \in I_{l+1}. \quad (3.14)$$

First, we have by definition

$$\bigcup_{i \in I_l} \omega_{i,l} = \left[ \bigcup_{i \in I_l} ]t_i^l - r_2^l/2, t_i^l + r_2^l/2[ \right] \times \text{Vois}(\gamma([s_l, s_l + L_l]), r_2^l/2).$$

Since  $(]t_i^l - r_2^l/2, t_i^l + r_2^l/2[)_{i \in I_l}$  is a finite covering of  $] -T + \sum_{j=1}^l T_j, T - \sum_{j=1}^l T_j[$ , we have

$$] -T + \sum_{j=1}^l T_j, T - \sum_{j=1}^l T_j[ \times \text{Vois}(\gamma([s_l, s_l + L_l]), r_2^l/2) \subset \bigcup_{i \in I_l} \omega_{i,l}. \quad (3.15)$$

We also have  $t_i^{l+1} \in ] -T + \sum_{j=1}^{l+1} T_j, T - \sum_{j=1}^{l+1} T_j[$ , so that

$$]t_i^{l+1} - T_{l+1}, t_i^{l+1} + T_{l+1}[\Subset ] -T - T_{l+1} + \sum_{j=1}^{l+1} T_j, T + T_{l+1} - \sum_{j=1}^{l+1} T_j[$$

that is

$$]t_i^{l+1} - T_{l+1}, t_i^{l+1} + T_{l+1}[\Subset ] -T + \sum_{j=1}^l T_j, T - \sum_{j=1}^l T_j[. \quad (3.16)$$

Moreover,  $\gamma(s_{l+1}) \in \gamma([s_l, s_l + L_l])$  and  $r_1^{l+1} < r_2^l/2$  by construction, so

$$\text{Vois}(\gamma(s_{l+1}), r_1^{l+1}) \Subset \text{Vois}(\gamma([s_l, s_l + L_l]), r_2^l/2). \quad (3.17)$$

Combining the definition of  $V_{i,l+1} = ]t_i^{l+1} - T_{l+1}, t_i^{l+1} + T_{l+1}[ \times \text{Vois}(\gamma(s_{l+1}), r_1^{l+1})$ , (3.16), (3.17) and (3.15), we obtain  $V_{i,l+1} \Subset \bigcup_{i \in I_l} \omega_{i,l}$ . This finishes the proof of (3.14) and therefore (3.13), so that all assumptions of Proposition 2.5 are satisfied.

The conclusion of this proposition can then be written as

$$\left[ U_0 \cup \bigcup_{j \in \llbracket 1, N \rrbracket} \bigcup_{i \in I_j} \omega_{i,j} \right] \triangleleft V_0 \text{ for any } V_0 \text{ such that } U_0 \Subset V_0. \quad (3.18)$$

Now, we pick  $r_2 < \min(T - \sum_{j=1}^N T_j, r_2^l/2)$ . Using (3.15) and then the covering property  $[0, L] \subset \bigcup_{j \in \llbracket 1, N \rrbracket} ]s_j, s_j + L_j[$ , we obtain

$$\begin{aligned} ] -r_2, r_2[ \times \text{Vois}(\gamma([s_j, s_j + L_j]), r_2/2) &\subset \bigcup_{i \in I_j} \omega_{i,j}, \\ ] -r_2, r_2[ \times \text{Vois}(\gamma([0, L]), r_2/2) &\subset \bigcup_{j \in \llbracket 1, N \rrbracket} \bigcup_{i \in I_j} \omega_{i,j}. \end{aligned}$$

In particular, we have

$$] - r_2/4, r_2/4[ \times \text{Vois}(\gamma([0, L]), r_2/4) \Subset \bigcup_{j \in \llbracket 1, N \rrbracket} \bigcup_{i \in I_j} \omega_{i,j}.$$

Now, since  $\tilde{T} > T$  and  $r_1^1 \leq r_1/2$ , we have  $U_0 \Subset ] - \tilde{T}, \tilde{T}[ \times \text{Vois}(\gamma(0), 2r_1) := V_0$ . Combining this together with (3.18) and the use of Proposition 2.3 (several times), we finally obtain

$$] - r_2/4, r_2/4[ \times \text{Vois}(\gamma([0, L]), r_2/4) \Subset ] - \tilde{T}, \tilde{T}[ \times \text{Vois}(\gamma(0), 2r_1).$$

This concludes the proof of Proposition 3.12.  $\square$

From Proposition 3.12, let us now briefly explain the proof of Proposition 3.7. We proceed exactly as in [LL15, Section 4.2], in the proof that Theorem 4.7 implies Theorem 1.11 (which, in Section 2 of the present paper, corresponds to the proof that Theorem 2.6 implies Theorem 2.1). Note that it only consists in getting rid of the frequency cutoffs appearing in the definition of  $\llcorner$ , i.e. considering all frequencies, at the cost of replacing the  $e^{-\kappa\mu}$  exponentially small remainder by a  $\frac{1}{\mu}$ . This concludes the proof of Proposition 3.7.

For an application in the context of Section 5 (partially analytic operators with a boundary), we need to relax the condition that  $\gamma$  is globally a normal geodesic.

**Remark 3.13.** The proof of Step 2 took the following structure.

- Lemma 3.11 proves some relations of dependence (3.9) for some local region around some small part of a normal geodesic.
- Proposition 3.12 uses the fact that the relations of dependence that we obtained in Lemma 3.11 can be iterated to be around a global normal geodesic to get some relation of dependence that has the same form but globally, namely (3.11).

Therefore, with exactly the same iteration process as Proposition 3.12, except that we invoke Proposition 3.12 itself instead of Lemma 3.11, we can obtain that the same result as Proposition 3.12 is true if  $\gamma$  is only geodesic by piece. This is the following Proposition.

**Proposition 3.14** (global version near a piecewise normal geodesic). *The same result as Proposition 3.12 is true if  $\gamma$  is only normal geodesic by piece.*

### 3.3 Step 3: Energy estimates

#### 3.3.1 Simple energy estimates concluding the proof of Theorem 1.31 with $s = k$

As precised earlier in the introduction, Theorem 1.31 is easier to prove in the specific case  $s = k$ . To conclude this proof from Corollary 3.8, it only remains to prove for solutions of (1.12) the two estimates:

$$\|(u_0, u_1)\|_{L^2 \times \mathcal{H}_Z^{-1}} \leq C_{\tilde{\varepsilon}} \|u\|_{L^2(\llcorner_{-\tilde{\varepsilon}, \tilde{\varepsilon}} \times \mathcal{M})}, \quad (3.19)$$

$$\|u\|_{H^1(\llcorner_{-T, T} \times \mathcal{M})} \leq C_T \|(u_0, u_1)\|_{\mathcal{H}_Z^k \times \mathcal{H}_Z^{k-1}}. \quad (3.20)$$

On the one hand, (3.19) is a "straightforward observability estimate", the proof of which is exactly the same as inequality (6.9) in the proof of Theorem 6.1 in [LL15] for the classical wave equation.

On the other hand, the proof of (3.20) relies on the (optimal) subelliptic estimates stated in Theorem 1.4 and Corollary B.1. First define the energies

$$\mathcal{E}_s(u) = \frac{1}{2} \|\partial_t u\|_{\mathcal{H}_L^{s-1}}^2 + \frac{1}{2} \|u\|_{\mathcal{H}_L^s}^2 = \frac{1}{2} \|(u, \partial_t u)\|_{\mathcal{H}_L^s \times \mathcal{H}_L^{s-1}}^2, \quad s \in \mathbb{R}. \quad (3.21)$$

Then, rewriting (1.12) as  $(\partial_t^2 + \mathcal{L} + 1)u = u$  and taking the inner product with  $(\mathcal{L} + 1)^{s-1} \partial_t u$  yields

$$\frac{d}{dt} \mathcal{E}_s(u)(t) = \left( (\mathcal{L} + 1)^{\frac{s-1}{2}} u, (\mathcal{L} + 1)^{\frac{s-1}{2}} \partial_t u \right)_{L^2}, \quad \text{and hence } -\mathcal{E}_s(u)(t) \leq \frac{d}{dt} \mathcal{E}_s(u)(t) \leq \mathcal{E}_s(u)(t),$$

so that we have

$$\mathcal{E}_s(u)(t) \leq C_T \mathcal{E}_s(u)(0) \quad \text{for } t \in [-T, T]. \quad (3.22)$$

Also, according to Corollary B.2, we have  $\|u\|_{H^1(\mathcal{M})} \leq C \|u\|_{\mathcal{H}_L^k}$ . This estimate yields

$$\|u\|_{H^1([-T, T] \times M)}^2 = \int_{-T}^T \left( \|\partial_t u\|_{L^2(\mathcal{M})}^2 + \|u\|_{H^1(\mathcal{M})}^2 \right) dt \leq C \int_{-T}^T \left( \mathcal{E}_1(u) + \|u\|_{\mathcal{H}_L^k}^2 \right) dt.$$

Since  $\mathcal{E}_s(u) \leq \mathcal{E}_\sigma(u)$  for  $s \leq \sigma$ , we obtain

$$\|u\|_{H^1([-T, T] \times M)}^2 \leq C \int_{-T}^T \mathcal{E}_k(u)(t) dt \leq C_T \mathcal{E}_k(u)(0) = 2C_T \|(u_0, u_1)\|_{\mathcal{H}_L^k \times \mathcal{H}_L^{k-1}}^2.$$

This proves (3.20), which, combined with the estimate of Corollary 3.8 and (3.19) implies

$$\|(u_0, u_1)\|_{L^2 \times \mathcal{H}_L^{-1}} \leq C_s e^{\kappa \mu} \|u\|_{L^2([-T, T] \times \omega)} + \frac{C}{\mu} \|(u_0, u_1)\|_{\mathcal{H}_L^k \times \mathcal{H}_L^{k-1}}.$$

This finally proves the estimate (1.29) with  $s = k$ , up to changing  $\mu$  by  $\mu^k/C$ , and  $\mu_0$  and  $\kappa$  accordingly. Estimate (1.14) is then a direct consequence of [LL15, Lemma A.3] and the inequality  $\|u\|_{L^2([-T, T] \times \omega)} \leq C \|(u_0, u_1)\|_{\mathcal{H}_L^k \times \mathcal{H}_L^{k-1}}$ .

**Remark 3.15.** Note that the previous energy estimates do not use any analyticity property of the solution and are equally true in the partially analytic case.

**Remark 3.16.** Until this point, all proofs work as well if  $\mathcal{L}$  is replaced by  $\mathcal{L} + V$  where  $V$  is a time-dependent nonnegative complex-valued analytic potential. Beware that in Section 3.3.3 below, the spectral theory used restricts the discussion to time-independent real-valued analytic potentials.

### 3.3.2 Interlude: eigenfunction tunneling, a proof of Theorem 1.10

Denoting  $v(t, x) = \cos(\sqrt{\lambda_j} t) \varphi_j$ , we remark that  $v$  is solution to

$$\begin{cases} \partial_t^2 v + \mathcal{L}v &= 0 \\ (v, \partial_t v)_{t=0} &= (\varphi_j, 0). \end{cases}$$

Therefore, thanks to Theorem 1.31 with  $s = k$ , we have the estimate

$$\|\varphi_j\|_{\mathcal{H}_L^1} \leq C e^{c\Lambda} \|v\|_{L^2([-T, T] \times \omega)} = C_T e^{c\Lambda} \|\varphi_j\|_{L^2(\omega)},$$

with  $\Lambda = \frac{\|\varphi_j\|_{\mathcal{H}_L^k}}{\|\varphi_j\|_{L^2}} = (\lambda_j + 1)^{k/2}$ . This proves Theorem 1.10 from Theorem 1.31 with  $s = k$  (or any given  $s > 0$ ).

### 3.3.3 End of the proof of Theorem 1.31 in the general case

We now prove the appropriate energy estimates to conclude the proof of Theorem 1.31 for any  $s > 0$ . Recall again that  $z = (t, x)$  and that all variables are analytic, that is  $z = z_a$ . Hence, we have  $\zeta = \zeta_a = (\tau, \xi)$  together with  $D = D_a = D_{t,x}$ .

**Remark 3.17.** The main idea of this section is that since we are only dealing with low frequency estimates, the remainder can be taken in any "arbitrary weak" norm; this gives then a lot of flexibility for the the norms that we can finally take.

Yet, in our estimates, the remainder term is actually in  $H^1$ . It would certainly be possible to obtain a weaker norm in the general Theorem of Quantitative Unique Continuation of [LL15], but it would require to revisit all proofs of [LL15], and even the Carleman estimate itself. Instead, we could try to only apply our general estimate to a function with low frequency, for example  $u = m^\mu(D_{t,x})v$  where  $v$  is solution of  $Pv = 0$ . But in this case, to obtain good estimates of  $Pu$ , we need some nice exponential estimate for the commutator  $[P, m^\mu(D_{t,x})]$ . This would certainly be possible but quite lengthy and would require some analytic regularity properties for  $P$ . Instead, for the specific case of the wave type equations, regularity in time morally implies regularity in space. That is why it is more convenient to consider a regularizer  $m^\mu(D_t)$  which commutes exactly with the wave operator. Remark that the method may extend to other *evolution* equations.

**Remark 3.18.** In this section, we are in the case where all variables are analytic, that is  $z_a = z = (t, x)$ . Therefore, the proof below does not a priori apply to the partially analytic case described in Section 1.2.4. This explains why the statement of Theorem 1.23 is slightly less general. The same results might be true in the partially analytic case but would certainly require more work and additional arguments.

In the relation  $\triangleleft$ , the remainder terms are always measured in the norm  $H^1$ . In this section, we describe how, in the case of solutions of the wave equation, the remainder term can be chosen in any weak Sobolev norm.

The starting point of the proof is Proposition 3.12, which implies the following statement (without the use of the notation  $\triangleleft$ ). Let  $\gamma : [0, L] \rightarrow \Omega_\gamma$  a normal geodesic path such that  $\gamma(0) \in \omega$ , and let  $0 < L < T$ . Then, there exists  $\varepsilon > 0$  such that for any  $\vartheta \in C_0^\infty(\mathbb{R}^{d+1})$  equal to one in a neighborhood of  $] - L, L[ \times \{\gamma(0)\}$  and  $\tilde{\vartheta} \in C_0^\infty(\mathbb{R}^{d+1})$ , supported in  $] - 2\varepsilon, 2\varepsilon[ \times B(\gamma(L), 2\varepsilon)$ , we have: for all  $\kappa, \alpha > 0$ , there exist  $C, \kappa', \beta, \mu_0 > 0$  such that for all  $(\mu, v) \in [\mu_0, +\infty) \times C_0^\infty(\mathbb{R}^{d+1})$ , we have

$$\left\| M_\mu^{\beta\mu} \tilde{\vartheta}_\mu v \right\|_1 \leq C e^{\kappa\mu} \left( \left\| M_\mu^{\alpha\mu} \vartheta_\mu v \right\|_1 + \|Pv\|_{L^2(\tilde{\Omega}_\gamma)} \right) + C e^{-\kappa'\mu} \|v\|_1. \quad (3.23)$$

where  $\tilde{\Omega}_\gamma = ] - T, T[ \times \Omega_\gamma$ . In particular, we may (and we shall) take  $\vartheta$  supported in  $] - T, T[ \times \omega$  and  $\tilde{\vartheta} = 1$  in a neighborhood of  $] - \varepsilon, \varepsilon[ \times B(x_0, \varepsilon)$ . Notice also that we can impose  $\beta < \alpha < 1$ ; indeed, the estimate for a smaller  $\beta$  is actually worse (than that for a larger  $\beta$ ) up to an error term of the form  $C e^{-\alpha\mu} \|v\|_1$  (see e.g. Lemma 3.24, Item 3 below).

Up to choosing the coordinate charts  $\Omega_\gamma$  smaller, we can still select some other open set  $\Omega \subset \mathbb{R}^d$  with  $\Omega_\gamma \Subset \Omega$  so that there is an analytic diffeomorphism from an open neighborhood of  $\gamma$  to  $\Omega$  and such that this diffeomorphism coincides with the  $\phi_\gamma$  defined in (3.4) on  $\Omega_\gamma$ .

From this starting point, the plan of the proof of Theorem 1.31 is to apply (3.23) to the function

$$v = \chi^0(t) \chi_\mu^1\left(\frac{D_t}{\mu}\right) \chi_\mu^2(t) \chi^3(x) u, \quad (3.24)$$

where  $u$  is a solution of

$$\begin{cases} Pu = (\partial_t^2 + \mathcal{L})u &= 0 \\ (u, \partial_t u)_{t=0} = (u_0, u_1) &= U_0, \end{cases} \quad (3.25)$$

pulled back to the local coordinate chart  $\Omega$  (we however keep the notation  $u$  to lighten the notation).

The time, frequency, and spatial cut off functions  $\chi^i$  are chosen as follows:

- $\chi^2 \in C_0^\infty(]-T-\varepsilon, T+\varepsilon[)$  so that  $\chi^2(t) = 1$  on  $]-T, T[$ ; we also write  $\chi_\mu^2 = (\chi^2)_\mu$ ,
- $\chi^0 \in C_0^\infty(]-T-\varepsilon, T+\varepsilon[)$  so that  $\chi^0(t) = 1$  in an neighborhood of  $\text{supp}(\chi^2)$ ,
- $\chi^3 \in C_0^\infty(\Omega)$  so that  $\chi^3(x) = 1$  for  $x \in \Omega_\gamma$ ,
- $\chi^1 \in C_0^\infty(\mathbb{R})$  is supported in  $]-2, 2[$  and such that  $\chi^1(\tau) = 1$  for  $\tau \in ]-1, 1[$ ; we also write  $\chi_\mu^1 = (\chi^1)_\mu$ .

**Remark 3.19.** Note that since  $\chi^3 \in C_0^\infty(\Omega)$  and  $\chi^0 \in C_0^\infty(\mathbb{R})$ ,  $v$  is a well defined function in  $C_0^\infty(\mathbb{R}^{d+1})$  if  $U_0$  and hence  $u$  is smooth. So, (3.23) is applicable. Moreover,  $v$  is not a local function in terms of  $u$  in the time variable; indeed, it depends on all values of  $u(t)$  for  $t \in \mathbb{R}$ . However, since  $u$  is a solution to Equation (3.25),  $u$  may rather be seen as a function of the data  $U_0$  only.

Our task is now to estimate each term in (3.23) in terms of the observation  $\|u\|_{L^2(]-T, T[ \times \omega)}$  and appropriate (weak) norms on  $U_0$  (this is done respectively in Lemmata 3.20, 3.21 and 3.22 below). Finally, to conclude the proof, it will remain to estimate also the high frequency part of the solution (this is done in Lemma 3.23). Several technical lemmata and estimates used in this proof are postponed to the next section for the sake of readability.

All along this section, we shall use the following natural product norms for  $s \in \mathbb{R}$ ,

$$\|U_0\|_{s, \times} = \|U_0\|_{\mathcal{H}_x^s \times \mathcal{H}_x^{s-1}} = \sqrt{2\mathcal{E}_s(U_0)}.$$

**Lemma 3.20.** *For all  $s \geq 0$ , there is  $C, N, c, \mu_0 > 0$  such that for all  $\mu \geq \mu_0$  and for all  $u$  solution to (3.25) and  $v$  defined accordingly in (3.24), we have*

$$\|M_\mu^{\alpha\mu} \vartheta_\mu v\|_1 \leq C\mu^N \|u\|_{L^2(]-T, T[ \times \omega)} + Ce^{-c\mu} \|U_0\|_{-s, \times}.$$

*Proof of Lemma 3.20.* First note that, according to (3.36) below, we have (for some  $N \in \mathbb{N}$ , changing from line to line),

$$m_\mu(\eta) \leq C\mu^N \langle \eta \rangle^N e^{-\frac{\mu}{4} \text{dist}(\eta, \text{supp}(m))^2}.$$

Hence, recalling the definition of  $m$  in Section 2.2, we obtain, for all  $s \geq 0$ , for some  $C, N, c > 0$ ,

$$\begin{aligned} |\eta|^{s+1} m_\mu(\eta) &= |\eta|^{s+1} m_\mu(\eta) \mathbf{1}_{|\eta| \leq 3/2} + |\eta|^{s+1} m_\mu(\eta) \mathbf{1}_{|\eta| \geq 3/2} \\ &\leq C \mathbf{1}_{|\eta| \leq 3/2} + C\mu^N \langle \eta \rangle^N e^{-\frac{\mu}{2} (|\eta|-1)^2} \mathbf{1}_{|\eta| \geq 3/2} \\ &\leq C \mathbf{1}_{|\eta| \leq 3/2} + C\mu^N \langle \eta \rangle^N e^{-\frac{\mu}{8} |\eta|^2} \mathbf{1}_{|\eta| \geq 3/2} \\ &\leq C \mathbf{1}_{|\eta| \leq 3/2} + Ce^{-c\mu}. \end{aligned}$$

As a consequence, since  $m = 1$  on a neighborhood of  $[0, 3/4]$ , we obtain, for all  $s \geq 0$ , the existence of  $C, N, c > 0$  such that for  $\mu \geq \mu_0$ , we have

$$\begin{aligned} \langle \zeta \rangle m_\mu \left( \frac{\zeta}{\alpha \mu} \right) &\leq C \mu^N \langle \zeta \rangle^{-s} \mathbf{1}_{\frac{|\zeta|}{\alpha \mu} \leq 3/2} + C \langle \zeta \rangle^{-s} e^{-c\mu} \\ &\leq C \mu^N \langle \zeta \rangle^{-s} m_\mu \left( \frac{\zeta}{2\alpha \mu} \right) + C \langle \zeta \rangle^{-s} e^{-c\mu}. \end{aligned}$$

This implies

$$\|M_\mu^{\alpha\mu} \vartheta_\mu v\|_1 \leq C \mu^N \|M_\mu^{2\alpha\mu} \vartheta_\mu v\|_{-s} + C e^{-c\mu} \|\vartheta_\mu v\|_{-s}. \quad (3.26)$$

Concerning the last term, we have

$$\|\vartheta_\mu v\|_{-s} \leq C \|v\|_{-s} = C \left\| \chi^0(t) \chi_\mu^1 \left( \frac{D_t}{\mu} \right) \chi_\mu^2(t) \chi^3(x) u \right\|_{-s} \leq C \|\chi_\mu^2(t) \chi^3(x) u\|_{-s} \leq C \|U_0\|_{-s, \times},$$

where, in the last inequality, we used (3.38).

Concerning the first term in the right hand-side of (3.26), we write

$$\begin{aligned} \|M_\mu^{2\alpha\mu} \vartheta_\mu v\|_{-s} &= \left\| M_\mu^{2\alpha\mu} \vartheta_\mu \chi^0(t) \chi_\mu^1 \left( \frac{D_t}{\mu} \right) \chi_\mu^2(t) \chi^3(x) u \right\|_{-s} \\ &\leq \left\| M_\mu^{2\alpha\mu} \vartheta_\mu \chi^0(t) \left( 1 - \chi_\mu^1 \left( \frac{D_t}{\mu} \right) \right) \chi_\mu^2(t) \chi^3(x) u \right\|_{-s} + \|M_\mu^{2\alpha\mu} \vartheta_\mu \chi^0(t) \chi_\mu^2(t) \chi^3(x) u\|_{-s} \end{aligned} \quad (3.27)$$

Concerning the first term in (3.27), we have

$$\begin{aligned} &\left\| M_\mu^{2\alpha\mu} \vartheta_\mu \chi^0(t) \left( 1 - \chi_\mu^1 \left( \frac{D_t}{\mu} \right) \right) \chi_\mu^2(t) \chi^3(x) u \right\|_{-s} \\ &\leq \left\| M_\mu^{2\alpha\mu} \vartheta_\mu \left( 1 - \chi_\mu^1 \left( \frac{D_t}{\mu} \right) \right) \chi_\mu^2(t) \chi^3(x) u \right\|_{-s} + \left\| M_\mu^{2\alpha\mu} \vartheta_\mu \left( 1 - \chi^0(t) \right) \left( 1 - \chi_\mu^1 \left( \frac{D_t}{\mu} \right) \right) \chi_\mu^2(t) \chi^3(x) u \right\|_{-s} \\ &\leq C e^{-c\mu} \|\chi_\mu^2(t) \chi^3(x) u\|_{-s}, \end{aligned}$$

where we have used Lemma 3.29 for the first term and then that  $\chi^0(t) = 1$  on  $\text{supp } \vartheta$  and Lemma 3.26 for the second. That this term is bounded by  $C e^{-c\mu} \|U_0\|_{-s, \times}$  then follows from to (3.38).

The second term in (3.27) is simpler to handle. Consider  $\theta_\omega$  a smooth cutoff function supported in  $] -T, T[ \times \omega$  and equal to one on a neighborhood of  $\text{supp } \vartheta$ . We have

$$\begin{aligned} \|M_\mu^{\alpha\mu} \vartheta_\mu \chi^0(t) \chi_\mu^2(t) \chi^3(x) u\|_{-s} &\leq \|\vartheta_\mu \theta_\omega \chi^0(t) \chi_\mu^2(t) \chi^3(x) u\|_{-s} + \|\vartheta_\mu (1 - \theta_\omega) \chi^0(t) \chi_\mu^2(t) \chi^3(x) u\|_{-s} \\ &\leq \|u\|_{L^2(-T, T] \times \omega)} + C e^{-c\mu} \|\chi_\mu^2(t) \chi^3(x) u\|_{-s} \\ &\leq \|u\|_{L^2(-T, T] \times \omega)} + C e^{-c\mu} \|U_0\|_{-s, \times}, \end{aligned}$$

where we have used Lemma 3.26 and then again Estimate (3.38) in the last step. The last two estimates combined with (3.26) and (3.27) conclude the proof of the lemma.  $\square$

The proofs of the following two lemmata are based on the spectral representation (3.28) of the solution  $u$  as

$$u(t, x) = \sum_{j \in \mathbb{N}} \left( a_j^+ e^{i\sqrt{\lambda_j}t} + a_j^- e^{-i\sqrt{\lambda_j}t} \right) \varphi_j(x), \quad (t, x) \in \mathbb{R} \times \mathcal{M}. \quad (3.28)$$

This explicit expression allows to prove that a time-frequency cutoff  $\chi(\frac{D_t}{\mu})$  truncates also space-frequencies (see in particular the use of Lemma 3.30 below).

**Lemma 3.21.** *For all  $s \geq 0$ , there is  $C, c > 0$  such that for all  $\mu \geq 1$  and for all  $u$  solution to (3.25) and  $v$  defined accordingly in (3.24), we have*

$$\|Pv\|_{L^2(\tilde{\Omega}_\gamma)} \leq C e^{-c\mu} \|U_0\|_{-s, \times}.$$

*Proof of Lemma 3.21.* Since  $Pu = 0$ ,  $\chi^3(x) = 1$  on  $\Omega_\gamma$  and  $\chi^0(t) = 1$  on  $] - T, T[$ , we have on  $\tilde{\Omega}_\gamma = ] - T, T[ \times \Omega_\gamma$

$$Pv = (\partial_t^2 + \mathcal{L})\chi^0(t)\chi_\mu^1\left(\frac{D_t}{\mu}\right)\chi_\mu^2(t)\chi^3(x)u = \chi^0(t)\chi_\mu^1\left(\frac{D_t}{\mu}\right)[\partial_t^2, \chi_\mu^2(t)]\chi^3(x)u, \quad (3.29)$$

with  $[\partial_t^2, \chi_\mu^2(t)] = (\partial_t^2 \chi^2)_\mu(t) + 2(\partial_t \chi^2)_\mu \partial_t$  (the terms  $[\chi^0, \partial_t^2]$  and  $[\chi^3, \mathcal{L}]$  being supported outside of  $\tilde{\Omega}_\gamma$ ). We only treat the second term in this commutator, the first one being simpler to handle.

We split  $u$  given in (3.28) into high and low frequencies as  $u = u_{\leq} + u_{>}$  with

$$u_{\leq} := \mathbf{1}_{\sqrt{\mathcal{L}} \leq 8\mu} u = \sum_{\sqrt{\lambda_j} \leq 8\mu} \dots, \quad u_{>} := \mathbf{1}_{\sqrt{\mathcal{L}} > 8\mu} u = \sum_{\sqrt{\lambda_j} > 8\mu} \dots. \quad (3.30)$$

We also write  $f(t) = \partial_t \chi^2(t)$  and  $f_\mu(t) = (\partial_t \chi^2)_\mu(t)$ . We first estimate the low frequencies as

$$\begin{aligned} \left\| \chi^0(t)\chi_\mu^1\left(\frac{D_t}{\mu}\right)f_\mu\chi^3(x)\partial_t u_{\leq} \right\|_{L^2(\tilde{\Omega}_\gamma)} &\leq C \left\| \mathbf{1}_{]-T, T[} \chi_\mu^1\left(\frac{D_t}{\mu}\right)f_\mu\chi^3(x)\partial_t u_{\leq} \right\|_{L^2(\mathbb{R}_t \times \Omega_\gamma)} \\ &\leq C \left\| \mathbf{1}_{]-T, T[} \chi_\mu^1\left(\frac{D_t}{\mu}\right)f_\mu\chi_\mu^0(t)\chi^3(x)\partial_t u_{\leq} \right\|_0 \\ &\quad + C \left\| \mathbf{1}_{]-T, T[} \chi_\mu^1\left(\frac{D_t}{\mu}\right)f_\mu(1 - \chi_\mu^0(t))\chi^3(x)\partial_t u_{\leq} \right\|_0. \end{aligned}$$

Concerning the first term in the right hand-side, we have

$$\begin{aligned} \left\| \mathbf{1}_{]-T, T[} \chi_\mu^1\left(\frac{D_t}{\mu}\right)f_\mu\chi_\mu^0(t)\chi^3(x)\partial_t u_{\leq} \right\|_{L^2(\mathbb{R}_t \times \Omega_\gamma)} &\leq \left\| \mathbf{1}_{]-T, T[} \chi_\mu^1\left(\frac{D_t}{\mu}\right)f_\mu \right\|_{L^2(\mathbb{R}_t) \rightarrow L^2(\mathbb{R}_t)} \left\| \chi_\mu^0(t)\chi^3(x)\partial_t u_{\leq} \right\|_0 \\ &\leq C e^{-c\mu} \|(u_{\leq}, \partial_t u_{\leq})|_{t=0}\|_{1, \times} \leq C e^{-c\mu} \mu^s \|U_0\|_{-s, \times}, \end{aligned}$$

after having used Lemma 3.27 together with (3.39). Concerning the second term in the right hand-side, we write

$$\begin{aligned} \left\| \mathbf{1}_{]-T, T[} \chi_\mu^1\left(\frac{D_t}{\mu}\right)f_\mu(1 - \chi_\mu^0(t))\chi^3(x)\partial_t u_{\leq} \right\|_{L^2(\mathbb{R}_t \times \Omega_\gamma)} \\ \leq \left\| \mathbf{1}_{]-T, T[} \chi_\mu^1\left(\frac{D_t}{\mu}\right)(1 - \chi_\mu^0(t)) \right\|_{L^2(\mathbb{R}_t) \rightarrow L^2(\mathbb{R}_t)} \left\| f_\mu(t)\chi^3(x)\partial_t u_{\leq} \right\|_0 \end{aligned}$$

and we conclude with the same arguments. Hence, we have the low frequency estimate

$$\left\| \chi^0(t) \chi_\mu^1 \left( \frac{D_t}{\mu} \right) f_\mu \chi^3(x) \partial_t u \leq \right\|_{L^2(\tilde{\Omega}_\gamma)} \leq C e^{-c\mu} \|U_0\|_{-s, \times}. \quad (3.31)$$

Concerning now the high-frequency part, still denoting  $f = \partial_t \chi^2$ , we have

$$\left\| \chi^0(t) \chi_\mu^1 \left( \frac{D_t}{\mu} \right) f_\mu(t) \chi^3(x) \partial_t u > \right\|_{L^2(\tilde{\Omega}_\gamma)} \leq \left\| \chi_\mu^1 \left( \frac{D_t}{\mu} \right) f_\mu(t) \chi^3(x) \partial_t u > \right\|_{L^2(\mathbb{R}_t \times \mathcal{M})},$$

where

$$\chi_\mu^1 \left( \frac{D_t}{\mu} \right) f_\mu(t) \chi^3(x) \partial_t u > = \sum_{\sqrt{\lambda_j} > 8\mu} i\sqrt{\lambda_j} \left[ \chi_\mu^1 \left( \frac{D_t}{\mu} \right) f_\mu(t) \left( a_j^+ e^{i\sqrt{\lambda_j}t} - a_j^- e^{-i\sqrt{\lambda_j}t} \right) \right] [\chi^3 \varphi_j](x).$$

As a consequence of the triangular inequality, the Plancherel theorem and the fact that  $(\varphi_j)$  is an orthonormal family in  $L^2(\mathcal{M})$ , we obtain

$$\begin{aligned} & \left\| \chi^0(t) \chi_\mu^1 \left( \frac{D_t}{\mu} \right) f_\mu(t) \chi^3(x) \partial_t u > \right\|_{L^2(\tilde{\Omega}_\gamma)} \\ & \leq \sum_{\sqrt{\lambda_j} \geq 8\mu} \sqrt{\lambda_j} \left\| \chi_\mu^1 \left( \frac{D_t}{\mu} \right) f_\mu(t) \left( a_j^+ e^{i\sqrt{\lambda_j}t} - a_j^- e^{-i\sqrt{\lambda_j}t} \right) \right\|_{L^2(\mathbb{R}_t)} \|\chi^3(x) \varphi_j\|_{L^2(\mathcal{M})} \\ & \leq \sum_{\sqrt{\lambda_j} \geq 8\mu} \sqrt{\lambda_j} \left\| \chi_\mu^1 \left( \frac{\tau}{\mu} \right) \left( \widehat{f}_\mu(\tau - \sqrt{\lambda_j}) a_j^+ - \widehat{f}_\mu(\tau + \sqrt{\lambda_j}) a_j^- \right) \right\|_{L^2(\mathbb{R}_\tau)} \\ & \leq \sum_{\sqrt{\lambda_j} \geq 8\mu} \sqrt{\lambda_j} |a_j^+| \left\| \chi_\mu^1 \left( \frac{\tau}{\mu} \right) \widehat{f}_\mu(\tau - \sqrt{\lambda_j}) \right\|_{L^2(\mathbb{R}_\tau)} + \sqrt{\lambda_j} |a_j^-| \left\| \chi_\mu^1 \left( \frac{\tau}{\mu} \right) \widehat{f}_\mu(\tau + \sqrt{\lambda_j}) \right\|_{L^2(\mathbb{R}_\tau)}. \end{aligned}$$

Using now Lemma 3.30, this yields

$$\begin{aligned} \left\| \chi^0(t) \chi_\mu^1 \left( \frac{D_t}{\mu} \right) f_\mu(t) \chi^3(x) \partial_t u > \right\|_{L^2(\tilde{\Omega}_\gamma)} & \leq \sum_{\sqrt{\lambda_j} \geq 8\mu} \sqrt{\lambda_j} C e^{-c\sqrt{\lambda_j}} (|a_j^+| + |a_j^-|) \\ & \leq C \left( \sum_{\sqrt{\lambda_j} \geq 8\mu} e^{-c\sqrt{\lambda_j}} (|a_j^+|^2 + |a_j^-|^2) \right)^{1/2} \leq C e^{-c\mu} \|U_0\|_{-s, \times}, \end{aligned}$$

after having used the Cauchy-Schwarz inequality in  $\ell^2(\mathbb{N})$ . This together with (3.31) and (3.29) concludes the proof of the lemma.  $\square$

The following Lemma will be used to estimate the last term in (3.23).

**Lemma 3.22.** *For all  $s \geq 0$ , there is  $C, N > 0$  such that for all  $\mu \geq 1$  and for all  $u$  solution to (3.25) and  $v$  defined accordingly in (3.24), we have*

$$\|v\|_1 \leq C \mu^N \|U_0\|_{-s, \times}.$$



*Proof.* We first write

$$\|v\|_1 \leq \left\| \chi_\mu^1 \left( \frac{D_t}{\mu} \right) \chi_\mu^2(t) \chi^3(x) u \right\|_{H^1(\mathbb{R}_t \times \mathcal{M})},$$

and decompose  $u$  in (3.28) as  $u = u_\leq + u_\gt$  with  $u_\leq = \mathbb{1}_{\sqrt{\mathcal{L}} \leq 8\mu} u$  and  $u_\gt = \mathbb{1}_{\sqrt{\mathcal{L}} > 8\mu} u$  being defined as in (3.30). We have

$$\left\| \chi_\mu^1 \left( \frac{D_t}{\mu} \right) \chi_\mu^2(t) \chi^3(x) u \right\|_{H^1(\mathbb{R}_t \times \mathcal{M})} \leq \left\| \chi_\mu^1 \left( \frac{D_t}{\mu} \right) \chi_\mu^2(t) \chi^3(x) u_\leq \right\|_{H^1(\mathbb{R}_t \times \mathcal{M})} + \left\| \chi_\mu^1 \left( \frac{D_t}{\mu} \right) \chi_\mu^2(t) \chi^3(x) u_\gt \right\|_{H^1(\mathbb{R}_t \times \mathcal{M})}.$$

Concerning the first term (low-frequencies), we simply use (3.40) to write

$$\begin{aligned} \left\| \chi_\mu^1 \left( \frac{D_t}{\mu} \right) \chi_\mu^2(t) \chi^3(x) u_\leq \right\|_{H^1(\mathbb{R}_t \times \mathcal{M})} &\leq C \left\| \chi_\mu^2(t) \chi^3(x) u_\leq \right\|_{H^1(\mathbb{R}_t \times \mathcal{M})} \leq C \|(u_\leq, \partial_t u_\leq)|_{t=0}\|_{k, \times} \\ &\leq C \mu^{k+s} \|U_0\|_{-s, \times}. \end{aligned}$$

Concerning the second term (high-frequencies), we use  $\|f(t)g(x)\|_{H^1(\mathbb{R} \times \mathcal{M})} \leq C \|f\|_{H^1(\mathbb{R})} \|g\|_{H^1(\mathcal{M})}$  and proceed as in the above proof of Lemma 3.21. This yields

$$\begin{aligned} &\left\| \chi_\mu^1 \left( \frac{D_t}{\mu} \right) \chi_\mu^2(t) \chi^3(x) u_\gt \right\|_{H^1(\mathbb{R}_t \times \mathcal{M})} \\ &\leq C \sum_{\sqrt{\lambda_j} \geq 8\mu} \left\| \langle \tau \rangle \chi_\mu^1 \left( \frac{\tau}{\mu} \right) \left( \widehat{\chi}_\mu^2(\tau - \sqrt{\lambda_j}) a_j^+ + \widehat{\chi}_\mu^2(\tau + \sqrt{\lambda_j}) a_j^- \right) \right\|_{L^2(\mathbb{R}_\tau)} \|\chi^3(x) \varphi_j\|_{H^1(\mathcal{M})} \\ &\leq C \sum_{\sqrt{\lambda_j} \geq 8\mu} \lambda_j^{N(k,s)} \left\| \chi_\mu^1 \left( \frac{\tau}{\mu} \right) \left( \widehat{\chi}_\mu^2(\tau - \sqrt{\lambda_j}) a_j^+ + \widehat{\chi}_\mu^2(\tau + \sqrt{\lambda_j}) a_j^- \right) \right\|_{L^2(\mathbb{R}_\tau)} \\ &\leq C \sum_{\sqrt{\lambda_j} \geq 8\mu} C e^{-c\sqrt{\lambda_j}} (|a_j^+| + |a_j^-|) \leq C \|U_0\|_{-s, \times}, \end{aligned}$$

where we have used  $\|\varphi_j\|_{H^s(\mathcal{M})} \leq C(\lambda_j + 1)^{\frac{ks}{2}}$  (which follows directly from Corollary B.2), Lemma 3.30 and the Cauchy-Schwarz inequality in  $\ell^2(\mathbb{N})$ .  $\square$

Before concluding the proof of Theorem 1.31, we need to explain how to estimate the high-frequency part of the solution (recall indeed that our starting point, Estimate (3.23), is a low-frequency estimate only). This is the aim of the following lemma, which proof is close to the proof that Theorem 4.7 implies Theorem 1.11 in [LL15].

**Lemma 3.23.** *For all  $s \in [0, k]$ , there is  $C, \mu_0 > 0$  such that for all  $\mu \geq \mu_0$  and for all  $u$  solution to (3.25), we have*

$$\left\| (1 - M_\mu^{\beta\mu}) \widetilde{\vartheta}_\mu \chi^0(t) \chi_\mu^2(t) \chi^3(x) u \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^d)} \leq \frac{C}{\mu^{s/k}} \|U_0\|_{s, \times}.$$

Note that this estimate is almost

$$\left\| (1 - M_\mu^{\beta\mu}) \widetilde{\vartheta}_\mu v \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^d)} \leq \frac{C}{\mu^{s/k}} \|U_0\|_{s, \times}.$$

(which is also true, but not used here), the difference being that  $v$  contains an additional time-frequency cutoff  $\chi_\mu^1(\frac{D_t}{\mu})$  (which does not play any role in the estimates below).

The proof below only gives the endpoint case  $s = k$ , the intermediate situations being deduced by interpolation. A direct proof of intermediate estimates would follow the same lines, yet being slightly longer.

*Proof.* We first notice that it is enough to treat the case  $s = k$ . Indeed,  $s = 0$  is direct by standard energy estimates for  $v$  (see for instance (3.22)) and uniform bound on  $\chi_\mu^2$  and  $\tilde{\vartheta}_\mu$ . Hence, since all operators involved are linear, the result for  $s \in [0, k]$  follows by interpolation, see for instance [Tar07, Chapter 23].

Concerning the case  $s = k$ , we write  $w = \chi^0(t)\chi_\mu^2(t)\chi^3(x)u$ , together with

$$\left\| (1 - M_\mu^{\beta\mu})\tilde{\vartheta}_\mu w \right\|_0 \leq C \sup_{\eta \in \mathbb{R}^{d+1}} \left| \langle \eta \rangle^{-1} (1 - m_\mu) \left( \frac{\eta}{\beta\mu} \right) \right| \left\| \tilde{\vartheta}_\mu w \right\|_1.$$

In the range  $|\eta| \geq \beta\mu/2$  with  $\mu \geq \mu_0$ , we have the loose estimate  $\left| \langle \eta \rangle^{-1} (1 - m_\mu) \left( \frac{\eta}{\beta\mu} \right) \right| \leq \frac{C}{\mu}$  whereas in the range  $|\eta| \leq \beta\mu/2$ , using  $\text{dist}(\text{supp}(1 - m(\frac{\cdot}{\beta})), \{|\eta| \leq \beta/2\}) > 0$ , we have  $\left| (1 - m_\mu) \left( \frac{\cdot}{\beta\mu} \right) \right| \leq Ce^{-c\mu}$  according to estimate (3.36) below. This implies, for  $\mu \geq \mu_0$  that

$$\left\| (1 - M_\mu^{\beta\mu})\tilde{\vartheta}_\mu w \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^d)} \leq \frac{C}{\mu} \left\| \tilde{\vartheta}_\mu w \right\|_{H^1(\mathbb{R}_t \times \mathbb{R}_x^d)}. \quad (3.32)$$

Next, we have that

$$\left\| \tilde{\vartheta}_\mu w \right\|_{H^1(\mathbb{R}_t \times \mathbb{R}_x^d)} \leq \|w\|_{H^1(\mathbb{R}_t \times \mathbb{R}_x^d)} \leq \left\| \chi_\mu^2(t)\chi^3(x)u \right\|_{H^1(\mathbb{R}_t \times \mathbb{R}_x^d)}. \quad (3.33)$$

uniformly for  $\mu \geq 1$ , since all derivatives of  $\tilde{\vartheta}_\mu$  are uniformly bounded for  $\mu \geq 1$ . We write for  $\tilde{w} = \chi_\mu^2(t)\chi^3(x)u$

$$\|\tilde{w}\|_{H^1(\mathbb{R}_t \times \mathbb{R}_x^d)} \leq \|\tilde{w}\|_{H^1(\mathbb{R}_t; L^2(\mathbb{R}_x^d))} + \|\tilde{w}\|_{L^2(\mathbb{R}_t; H^1(\mathbb{R}_x^d))},$$

and estimate each term separately. Concerning the first term in this estimate, we have

$$\begin{aligned} \|\tilde{w}\|_{H^1(\mathbb{R}_t; L^2(\mathbb{R}_x^d))} &\leq \|\tilde{w}\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^d)} + \|\partial_t \tilde{w}\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^d)} \\ &\leq \left\| \chi_\mu^2(t)\chi^3(x)u \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^d)} + \left\| (\partial_t \chi^2)_\mu(t)\chi^3(x)u \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^d)} + \left\| \chi_\mu^2(t)\chi^3(x)\partial_t u \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^d)} \\ &\leq C \|U_0\|_{k, \times}. \end{aligned}$$

after having used (3.38), (3.39) and  $k \geq 1$ . Similarly, the second term is estimated as

$$\|\tilde{w}\|_{L^2(\mathbb{R}_t; H^1(\mathbb{R}_x^d))} = \left\| \chi_\mu^2(t)\chi^3(x)u \right\|_{L^2(\mathbb{R}_t; H^1(\mathbb{R}_x^d))} \leq C \|U_0\|_{k, \times},$$

as a direct consequence of (3.40). The above three estimates together with (3.32) and (3.33) conclude the proof of the lemma.  $\square$

With the above four lemmata in hand, we can now conclude the proof of Theorem 1.31.

*Proof of Theorem 1.31.* We first prove the result for  $s \in ]0, k]$ , the conclusion for all  $s > 0$  being then a consequence of Remark 1.30.

Starting from Estimate (3.23), combined with Lemmata 3.20, 3.21 and 3.22 to bound the terms in the right hand side, we first obtain the intermediate estimate

$$\left\| M_\mu^{\beta\mu} \tilde{\vartheta}_\mu v \right\|_1 \leq C e^{c\mu} \|u\|_{L^2(]-T, T[\times\omega)} + C e^{-c'\mu} \|U_0\|_{-s, \times}. \quad (3.34)$$

Note that in order to obtain this inequality, we have chosen  $\kappa$  in (3.23) to be small enough compared to the constant  $c$  appearing in Lemmata 3.20 and 3.21. Now, recalling that  $v = \chi^0(t) \chi_\mu^1(\frac{D_t}{\mu}) \chi_\mu^2(t) \chi^3(x) u$ , we decompose

$$\begin{aligned} \tilde{\vartheta}_\mu \chi^0(t) \chi_\mu^2(t) \chi^3(x) u &= M_\mu^{\beta\mu} \tilde{\vartheta}_\mu v + M_\mu^{\beta\mu} \tilde{\vartheta}_\mu \chi^0(t) \left(1 - \chi_\mu^1\left(\frac{D_t}{\mu}\right)\right) \chi_\mu^2(t) \chi^3(x) u \\ &\quad + (1 - M_\mu^{\beta\mu}) \tilde{\vartheta}_\mu \chi^0(t) \chi_\mu^2(t) \chi^3(x) u. \end{aligned} \quad (3.35)$$

The first term is estimated in (3.34), the last one is estimated in Lemma 3.23 so that it only remains to estimate the second one. We have

$$\begin{aligned} M_\mu^{\beta\mu} \tilde{\vartheta}_\mu \chi^0(t) \left(1 - \chi_\mu^1\left(\frac{D_t}{\mu}\right)\right) \chi_\mu^2(t) \chi^3(x) u &= M_\mu^{\beta\mu} \tilde{\vartheta}_\mu \left(1 - \chi_\mu^1\left(\frac{D_t}{\mu}\right)\right) \chi_\mu^2(t) \chi^3(x) u \\ &\quad + M_\mu^{\beta\mu} \tilde{\vartheta}_\mu (1 - \chi^0(t)) \left(1 - \chi_\mu^1\left(\frac{D_t}{\mu}\right)\right) \chi_\mu^2(t) \chi^3(x) u. \end{aligned}$$

According to Lemma 3.29 and (3.38) below, the first of these two terms satisfies

$$\left\| M_\mu^{\beta\mu} \tilde{\vartheta}_\mu \left(1 - \chi_\mu^1\left(\frac{D_t}{\mu}\right)\right) \chi_\mu^2(t) \chi^3(x) u \right\|_0 \leq C e^{-c\mu} \left\| \chi_\mu^2(t) \chi^3(x) u \right\|_0 \leq C e^{-c\mu} \|U_0\|_{0, \times},$$

whereas, according to Item 2 of Lemma 3.24, we have  $\left\| \tilde{\vartheta}_\mu (1 - \chi^0(t)) \right\|_{L^2(\mathbb{R}^n) \rightarrow L^2(\mathbb{R}^n)} \leq C e^{-c\mu}$  so that the second one is bounded as well as

$$\left\| M_\mu^{\beta\mu} \tilde{\vartheta}_\mu (1 - \chi^0(t)) \left(1 - \chi_\mu^1\left(\frac{D_t}{\mu}\right)\right) \chi_\mu^2(t) \chi^3(x) u \right\|_0 \leq C e^{-c\mu} \left\| \chi_\mu^2(t) \chi^3(x) u \right\|_0 \leq C e^{-c\mu} \|U_0\|_{0, \times}.$$

Coming back to the decomposition (3.35), using the above two estimates together with (3.34) and Lemma 3.23, we now have (for  $s \in ]0, k]$ )

$$\left\| \tilde{\vartheta}_\mu \chi^0(t) \chi_\mu^2(t) \chi^3(x) u \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^d)} \leq C e^{\kappa\mu} \|u\|_{L^2(]-T, T[\times\omega)} + \frac{C}{\mu^{s/k}} \|U_0\|_{s, \times}.$$

Now, since  $\tilde{\vartheta} = 1$  in a neighborhood of  $] - \varepsilon, \varepsilon[ \times B(x_0, \varepsilon)$ , for  $\mu$  large enough, we have  $\tilde{\vartheta}_\mu \geq 1/2$  on  $] - \varepsilon, \varepsilon[ \times B(x_0, \varepsilon)$ , and the same also holds for  $\chi_\mu^2(t)$ . Hence, we obtain

$$\|u\|_{L^2(]-\varepsilon, \varepsilon[ \times B(x_0, \varepsilon))} \leq C e^{\kappa\mu} \|u\|_{L^2(]-T, T[\times\omega)} + \frac{C}{\mu^{s/k}} \|U_0\|_{s, \times}.$$

Since  $x_0 \in \mathcal{M}$  is arbitrary, a compactness argument allows to obtain other constants still denoted  $C, \kappa, \mu_0, \varepsilon > 0$  such that for  $\mu \geq \mu_0$ ,

$$\|u\|_{L^2(]-\varepsilon, \varepsilon[ \times \mathcal{M})} \leq C e^{\kappa\mu} \|u\|_{L^2(]-T, T[\times\omega)} + \frac{C}{\mu^{s/k}} \|U_0\|_{s, \times}.$$

We conclude the proof of Theorem 1.31, in the case  $s \in ]0, k]$ , by using Estimate (3.19) and changing  $\mu$  into  $\mu^k$ . In the case  $s > k$ , the proof follows from the estimate with  $s = k$  and an interpolation argument, as explained in Remark 1.30.  $\square$

### 3.3.4 Technical lemmata used in the previous section (only)

In this section, we collect some technical results used in the above Section 3.3.3 (and in that section only). We first state results that are either directly taken from [LL15], or direct consequences of these. Second, we prove three lemmata using these results.

**Estimates taken from [LL15].** The following estimate is [LL15, Equation (2.9)] in a simple situation: given a continuous function  $f$  on  $\mathbb{R}^{n_a}$ , we have, for all  $\zeta \in \mathbb{R}^{n_a}$

$$|f_\lambda(\zeta)| \leq C \langle \lambda \rangle^{\frac{n_a}{2}} \|f\|_{L^\infty} \langle \text{dist}(\zeta, \text{supp}(f)) \rangle^{n_a-1} e^{-\frac{\lambda}{4} \text{dist}(\zeta, \text{supp}(f))^2} \quad (3.36)$$

**Lemma 3.24** (Lemma 2.3 of [LL15]). *The following three statements holds true.*

1. For any  $d > 0$ , there exist  $C, c > 0$  such that for any  $f_1, f_2 \in L^\infty(\mathbb{R}^n)$  such that

$$\text{dist}(\text{supp}(f_1), \text{supp}(f_2)) \geq d,$$

and all  $\lambda \geq 0$ , we have

$$\|f_{1,\lambda} f_2\|_{L^\infty} \leq C e^{-c\lambda} \|f_1\|_{L^\infty} \|f_2\|_{L^\infty}, \quad \|f_{1,\lambda} f_{2,\lambda}\|_{L^\infty} \leq C e^{-c\lambda} \|f_1\|_{L^\infty} \|f_2\|_{L^\infty}.$$

2. If moreover  $f_1, f_2 \in C^\infty(\mathbb{R}^n)$  have bounded derivatives, then for all  $k \in \mathbb{N}$ , there exist  $C, c > 0$  such that for all  $\lambda \geq 1$ , we have

$$\|f_{1,\lambda} f_2\|_{H^k(\mathbb{R}^n) \rightarrow H^k(\mathbb{R}^n)} \leq C e^{-c\lambda}.$$

3. Let  $f_1, f_2 \in L^\infty(\mathbb{R}^{n_a})$  such that  $\text{dist}(\text{supp}(f_1), \text{supp}(f_2)) > 0$ . Then there exist  $C, c > 0$  such that for all  $\lambda \geq 1$ , for all  $k \in \mathbb{N}$ , for all  $\mu \geq 1$ , we have

$$\begin{aligned} \|f_{1,\lambda}(D_a/\mu) f_2(D_a/\mu)\|_{H^k(\mathbb{R}^n) \rightarrow H^k(\mathbb{R}^n)} &\leq C e^{-c\lambda}, \\ \|f_{1,\lambda}(D_a/\mu) f_{2,\lambda}(D_a/\mu)\|_{H^k(\mathbb{R}^n) \rightarrow H^k(\mathbb{R}^n)} &\leq C e^{-c\lambda}. \end{aligned}$$

**Lemma 3.25** (Lemma 2.9 of [LL15]). *Let  $k \in \mathbb{N}$  and  $f \in C_0^\infty(\mathbb{R}^n)$ . Then, there exist  $C, c$  such that, for any  $\lambda, \mu > 0$ , we have*

$$\begin{aligned} \left\| M_\lambda^\mu f_\lambda (1 - M_\lambda^{2\mu}) \right\|_{H^k(\mathbb{R}^n) \rightarrow H^k(\mathbb{R}^n)} &\leq C e^{-c\frac{\mu^2}{\lambda}} + C e^{-c\lambda}, \\ \left\| (1 - M_\lambda^{2\mu}) f_\lambda M_\lambda^\mu \right\|_{H^k(\mathbb{R}^n) \rightarrow H^k(\mathbb{R}^n)} &\leq C e^{-c\frac{\mu^2}{\lambda}} + C e^{-c\lambda}. \end{aligned}$$

**Lemma 3.26.** *Let  $f_1, f_2 \in C^\infty(\mathbb{R})$  with all derivatives bounded and such that  $\text{supp}(f_1) \cap \text{supp}(f_2) = \emptyset$ . Then, for any  $s \in \mathbb{N}$ , there is  $C, c > 0$  such that for all  $w \in H^{-s}(\mathbb{R})$ , we have*

$$\|f_1 f_{2,\mu} w\|_{-s} \leq C e^{-c\mu} \|w\|_{-s}.$$

Lemma 3.26 is obtained by duality from Item 2 of Lemma 3.24.

**Lemma 3.27.** *Let  $f \in C^\infty(\mathbb{R})$  bounded such that  $\text{supp}(f) \cap [-T, T] = \emptyset$ . Then, there is  $C, c > 0$  such that*

$$\|\mathbb{1}_{]-T, T}[\chi_\mu^1(D_t/\mu)f_\mu\|_{L^2(\mathbb{R}_t) \rightarrow L^2(\mathbb{R}_t)} \leq Ce^{-c\mu}.$$

Lemma 3.27 is a particular case of [LL15, Lemma 2.10].

**A few estimates using the above lemmata.** Recalling the definition of  $\mathcal{E}_s$  in (3.21) we now refine the rough Estimate (3.22). Indeed, on account to the spectral theory of  $\mathcal{L}$ , if we denote  $\Pi_0$  the spectral projector on  $\ker(\mathcal{L}) = \text{span}_{L^2}(1)$ , we have, for all  $u \in \mathcal{H}_{\mathcal{L}}^s$ ,

$$\|u\|_{\mathcal{H}_{\mathcal{L}}^s} \simeq \|\mathcal{L}^{\frac{s}{2}}u\|_{L^2(\mathcal{M})} + |\Pi_0 u|.$$

Notice now that the energy  $\|\mathcal{L}^{\frac{s}{2}}u\|_{L^2}^2 + \|\mathcal{L}^{\frac{s-1}{2}}\partial_t u\|_{L^2}^2$  is preserved by the equation (3.25), and that the equation for the zero frequency is  $\partial_t^2 \Pi_0 u = 0$  hence growing at most linearly. As a consequence, we finally obtain that for all  $s \in \mathbb{R}$ , there is  $C > 0$  such that for all solution  $u$  of (3.25), we have

$$\mathcal{E}_s(u)(t)dt \leq C(1 + |t|)\mathcal{E}_s(u)(0), \quad \text{for all } t \in \mathbb{R}. \quad (3.37)$$

This estimate is now used to bound some integrals of  $\mathcal{E}_s(u)$ .

**Lemma 3.28.** *Let  $\eta \in C_0^\infty(\mathbb{R})$ , denote  $\chi^3(x)$  as above and fix  $s \geq 0$ . Then, there is  $C > 0$  such that for all  $\mu \geq 1$  and for all solution  $u$  of (3.25), we have*

$$\|\eta_\mu(t)\chi^3(x)u\|_{-s} \leq C\|U_0\|_{-s, \times}, \quad (3.38)$$

$$\|\eta_\mu(t)\chi^3(x)\partial_t u\|_0 \leq C\|U_0\|_{1, \times}, \quad (3.39)$$

$$\|\eta_\mu(t)\chi^3(x)u\|_1 \leq C\|U_0\|_{k, \times}, \quad (3.40)$$

$$\|\eta_\mu(t)\chi^3(x)u\|_{L^2(\mathbb{R}_t; H^{\frac{k}{2}}(\mathbb{R}_x^d))} \leq C\|U_0\|_{s, \times}. \quad (3.41)$$

*Proof.* We first remark (see e.g. (3.36)) that there is  $C, c > 0$  such that for  $\mu \geq 1$ , we have  $0 \leq \eta_\mu(t) \leq Ce^{-c|t|}$  for all  $t \in \mathbb{R}$ . To prove (3.38), we now simply write

$$\begin{aligned} \|\eta_\mu(t)\chi^3(x)u\|_{-s}^2 &\leq \int_{\mathbb{R}} Ce^{-c|t|} \|\chi^3(x)u\|_{H^{-s}(\Omega_\gamma)}^2 dt = \int_{\mathbb{R}} Ce^{-c|t|} \|\chi^3(x)u\|_{H^{-s}(\mathcal{M})}^2 dt \\ &\leq \int_{\mathbb{R}} Ce^{-c|t|} \|u\|_{H^{-s}(\mathcal{M})}^2 dt \leq \int_{\mathbb{R}} Ce^{-c|t|} \|u\|_{\mathcal{H}_{\mathcal{L}}^{-s}}^2 dt, \end{aligned}$$

where we used (B.5) and  $s \geq 0$  in the last inequality. Recalling the definition of  $\mathcal{E}_s$  in (3.21) together with estimate (3.37), we now have

$$\|\eta_\mu(t)\chi^3(x)u\|_{-s}^2 \leq \int_{\mathbb{R}} Ce^{-c|t|} \mathcal{E}_{-s}(u)(t)dt \leq C \left( \int_{\mathbb{R}} e^{-c|t|}(1 + |t|)dt \right) \mathcal{E}_{-s}(u)(0) = C\|U_0\|_{\mathcal{H}_{\mathcal{L}}^{-s} \times \mathcal{H}_{\mathcal{L}}^{-s-1}}^2,$$

which concludes the proof of (3.38). The proof of (3.39) is the same, except that we use  $\|\partial_t u\|_{L^2(\mathcal{M})}^2 \leq 2\mathcal{E}_1(u)$  instead of (B.5). The proof of (3.40) is similar: after using the chain rule, each term is either of the form  $\eta_\mu(t)\chi^3(x)u$  (treated in (3.38)) or of the form  $\eta_\mu(t)\chi^3(x)\partial_t u$  (treated in (3.39)) or of the form  $\eta_\mu(t)\chi^3(x)\partial_x u$ , for which the proof is the same using  $\|u\|_{H^{\frac{k}{2}}(\mathcal{M})}^2 \leq 2\mathcal{E}_s(u)$ , consequence of Corollary (B.2), instead of (B.5). The proof of (3.41) is the same, still using Corollary (B.2).  $\square$

**Lemma 3.29.** Fix  $\alpha < 1$ . Let  $\chi^1 \in C_0^\infty(\mathbb{R})$  such that  $\chi^1(\tau) = 1$  for  $\tau \in ]-1, 1[$  and  $\vartheta \in C_0^\infty(\mathbb{R}^{1+d})$ . Then, for any  $s \in \mathbb{N}$ , there is  $C, c > 0$  such that for all  $w \in H^{-s}(\mathbb{R}^{1+d})$ , we have

$$\left\| M_\mu^{\alpha\mu} \vartheta_\mu \left(1 - \chi_\mu^1\left(\frac{D_t}{\mu}\right)\right) w \right\|_{-s} \leq C e^{-c\mu} \|w\|_{-s}.$$

Note that in the proofs above, the parameter  $\alpha$  of Lemma 3.29 is both taken to be the parameters  $\alpha$  or  $\beta$  appearing estimate (3.23). This is the reason why we assumed  $\alpha, \beta < 1$  there.

*Proof.* This lemma is very close to (and a consequence of) Lemma 3.25 except that  $\chi_\mu^1(D_t)$  is a Fourier cutoff in  $D_t$  only whereas  $M_\mu^{\alpha\mu}$  are Fourier cutoffs in the whole  $D_{t,x}$  (and that the Sobolev orders are negative). Recall that  $M_\mu^{\alpha\mu} = m_\mu\left(\frac{D}{\alpha\mu}\right)$ ,  $D = D_{t,x}$ , where  $m$  is compactly supported in  $|\xi| < 1$  and  $m(\xi) = 1$  for  $|\xi| < 3/4$  (see the beginning of Section 2.2).

Let  $\tilde{m}$  be a radial smooth function on  $\mathbb{R}^{1+d}$  such that  $\tilde{m}(\xi) = 1$  in a neighborhood of  $|\xi| \leq \alpha$  and  $\tilde{m}(\xi) = 0$  in a neighborhood of  $|\xi| \geq 1$ . Then we have  $1 - \chi^1 = 0$  on the support of  $\tilde{m}$ . For  $s \in \mathbb{N}$ , we write

$$\begin{aligned} \|(1 - \chi_\mu^1(D_t)) \vartheta_\mu M_\mu^{\alpha\mu} w\|_s &\leq \|(1 - \chi_\mu^1(D_t/\mu))(1 - \tilde{m}_\mu(D/\mu)) \vartheta_\mu M_\mu^{\alpha\mu} w\|_s \\ &\quad + \|(1 - \chi_\mu^1(D_t)) \tilde{m}_\mu(D/\mu) \vartheta_\mu M_\mu^{\alpha\mu} w\|_s \\ &\leq C \|(1 - \tilde{m}_\mu(D/\mu)) \vartheta_\mu M_\mu^{\alpha\mu} w\|_s + C \|(1 - \chi_\mu^1(D_t)) \tilde{m}_\mu(D/\mu) \vartheta_\mu M_\mu^{\alpha\mu} w\|_s. \end{aligned}$$

According to Lemma 3.25 and the respective supports of  $\tilde{m}$  and  $m(\frac{\cdot}{\alpha})$ , we have

$$\|(1 - \tilde{m}_\mu(D/\mu)) \vartheta_\mu M_\mu^{\alpha\mu} w\|_s \leq C e^{-c\mu} \|w\|_s.$$

Also, according to Item 3 of 3.24, and the respective supports of  $\tilde{m}$  and  $\chi^1$ , we have

$$\|(1 - \chi_\mu^1(D_t/\mu)) \tilde{m}_\mu(D/\mu)\|_{H^s \rightarrow H^s} \leq C e^{-c\mu},$$

and hence

$$\|(1 - \chi_\mu^1(D_t/\mu)) \tilde{m}_\mu(D/\mu) \vartheta_\mu M_\mu^{\alpha\mu} w\|_s \leq C e^{-c\mu} \|w\|_s$$

This finally yields for  $s \in \mathbb{N}$

$$\|(1 - \chi_\mu^1(D_t/\mu)) \vartheta_\mu M_\mu^{\alpha\mu} w\|_s \leq C e^{-c\mu} \|w\|_s,$$

and the sought estimate by a duality argument.  $\square$

**Lemma 3.30.** Let  $\chi \in C_0^\infty(\mathbb{R})$  and  $m \in C_0^\infty(]-1, 1[)$ , and define

$$f_{\mu,\lambda}(\tau) = m_\mu\left(\frac{\tau}{\mu}\right) \widehat{\chi}_\mu(\tau - \lambda).$$

Then, for all  $\sigma \in \mathbb{R}$ , there is  $C, c > 0$  so that we have

$$\|\langle \tau \rangle^\sigma f_{\mu,\lambda}(\tau)\|_{L^2} \leq C e^{-c|\lambda|}, \quad \text{for all } \lambda \in \mathbb{R}, \mu \geq 0 \text{ such that } |\lambda| \geq 4\mu. \quad (3.42)$$

*Proof.* We decompose  $f_{\mu,\lambda} = f_{\mu,\lambda}^1 + f_{\mu,\lambda}^2$  with  $f_{\mu,\lambda}^1(\tau) = f_{\mu,\lambda}(\tau)\mathbb{1}_{|\tau|\leq 2\mu}$  and  $f_{\mu,\lambda}^2(\tau) = f_{\mu,\lambda}(\tau)\mathbb{1}_{|\tau|>2\mu}$ . Using that  $m_\mu$  is uniformly bounded, we have

$$\begin{aligned} \|\langle \tau \rangle^\sigma f_{\mu,\lambda}^1\|_{L^2}^2 &\leq C\langle 2\mu \rangle^\sigma \int_{-2\mu}^{2\mu} e^{-\frac{|\tau-\lambda|^2}{\mu}} |\widehat{\chi}(\tau-\lambda)|^2 d\tau = C\langle 2\mu \rangle^\sigma \int_{-2\mu-\lambda}^{2\mu-\lambda} e^{-\frac{|\tau|^2}{\mu}} |\widehat{\chi}(\tau)|^2 d\tau \\ &\leq C\langle \lambda \rangle^\sigma \int_{|\tau|\geq |\lambda|/2} e^{-\frac{|\tau|^2}{\mu}} |\widehat{\chi}(\tau)|^2 d\tau \leq C\langle \lambda \rangle^\sigma e^{-\frac{|\lambda|^2}{4\mu}} \int_{|\tau|\geq |\lambda|/2} |\widehat{\chi}(\tau)|^2 d\tau \leq Ce^{-|\lambda|}, \end{aligned}$$

where we have used that  $|\lambda| \geq 4\mu$  implies  $|\lambda|/2 \leq |\lambda| - 2\mu$  (and in particular  $\tau \in [-2\mu - \lambda, 2\mu - \lambda]$  implies  $|\tau| \geq |\lambda|/2$ ) and  $\frac{|\lambda|^2}{4\mu} \geq |\lambda|$ .

Concerning now  $f_{\mu,\lambda}^2$ , remark that  $|s| \geq 2$  implies  $\text{dist}(s, [-1, 1]) \geq |s|/2$ . Hence, using (3.36) together with the support of  $m$ , we have uniformly

$$|m_\mu(s)\mathbb{1}_{|s|\geq 2}| \leq C\langle \mu \rangle^{1/2} e^{-\frac{\mu s^2}{16}} \mathbb{1}_{|s|\geq 2}.$$

Using this with  $s = \tau/\mu$ , we obtain

$$\|\langle \tau \rangle^\sigma f_{\mu,\lambda}^2\|_{L^2}^2 \leq C\langle \mu \rangle^{1/2} \int_{|\tau|\geq 2\mu} \langle \tau \rangle^\sigma e^{-\frac{\tau^2}{16\mu}} e^{-\frac{|\tau-\lambda|^2}{\mu}} |\widehat{\chi}(\tau-\lambda)|^2 d\tau \leq C\langle \mu \rangle^{1/2} e^{-\frac{\lambda^2}{5\mu}} \int_{\mathbb{R}} \langle \tau \rangle^\sigma |\widehat{\chi}(\tau-\lambda)|^2 d\tau,$$

where we have used the estimate  $\frac{\tau^2}{16} + |\tau-\lambda|^2 \geq \lambda^2 \min\{s^2/16 + (s-1)^2 \mid s \in \mathbb{R}\} \geq c\lambda^2$  with  $c > 0$ . Using now that  $|\lambda| \geq 4\mu$ , we have

$$\|\langle \tau \rangle^\sigma f_{\mu,\lambda}^2\|_{L^2}^2 \leq C\langle \mu \rangle^{1/2} e^{-c\frac{\lambda^2}{\mu}} \langle \lambda \rangle^\sigma \int_{\mathbb{R}} \langle \tau \rangle^\sigma |\widehat{\chi}(\tau)|^2 d\tau \leq C\langle \lambda \rangle^{\sigma+1/2} e^{-4c|\lambda|},$$

which concludes the proof of the lemma.  $\square$

## 4 The hypoelliptic heat equation

This section is devoted to the proofs of Theorems 1.16, 1.18 and 1.20, which all rely on the methods of [EZ11a, Propositions 1 and 2] (proved in [EZ11b, Section 3]). We summarize these results in the next proposition for readability.

**Proposition 4.1** ([EZ11a, EZ11b]). *Let  $T, S > 0$  and  $\alpha > 2S^2$ . Then, there exists some kernel function  $k_T(t, s)$  such that*

- if  $y$  is solution of the heat equation (1.16), then  $w(s) = \int_0^T k_T(t, s)y(t)dt$  is solution of

$$\begin{cases} \partial_s^2 w + \mathcal{L}w = 0, & \text{for } s \in ]-S, S[, \\ (w, \partial_s w)|_{s=0} = \left(0, \int_0^T \partial_s k_T(t, 0)y(t)dt\right) = \left(0, \int_0^T e^{-\alpha\left(\frac{1}{t} + \frac{1}{T-t}\right)} y(t)dt\right); \end{cases} \quad (4.1)$$

- for all  $\delta \in ]0, 1[$ , there is  $C > 0$  such that for all  $(t, s) \in ]0, T[ \times ]-S, S[$ ,  $k_T$  satisfies

$$|k_T(t, s)| \leq C|s| \exp\left(\frac{1}{\min\{t, T-t\}} \left(\frac{s^2}{\delta} - \frac{\alpha}{(1+\delta)}\right)\right). \quad (4.2)$$

Note that this last estimate is most useful for  $\delta$  sufficiently close to one so that  $\alpha \geq S^2(1 + \frac{1}{\delta})$ .

The proof of Theorems 1.20 and 1.18 then follows the Lebeau-Robbiano transmutation method, as implemented in [EZ11a], splitting high and low frequencies. The proof of Theorems 1.16 is slightly different and does not rely on this splitting. For the purposes of Theorems 1.20 and 1.18, we define

$$E_\lambda = \text{span} \{ \varphi_j, \lambda_j \leq \lambda \},$$

where  $(\lambda_j, \varphi_j)$  are the spectral elements of  $\mathcal{L}$ , defined in (1.7). The first step of the proofs of Theorems 1.20 and 1.18 is to show, using the above transmutation technique, that we can transfer estimates obtained for solutions of the wave equation to solutions of the heat equation. More precisely, we first prove the following low-frequency observability estimate, with a precise estimation of the observability constant with respect to the cutoff frequency.

**Lemma 4.2.** *There exist  $C, \gamma > 0$  such that for any  $T > 0, \lambda \geq 0$ , for every  $y_0 \in E_\lambda$  and associated solution  $y$  to (1.16), we have*

$$\|y(T)\|_{L^2}^2 \leq \frac{C}{T} e^{(2\gamma\lambda^{k/2} + \frac{C}{T})} \int_0^T \int_\omega |y(t, x)|^2 dt dx. \quad (4.3)$$

Moreover, there exists  $c_0 > 0$  such that for any  $T > 0$  there exists  $C = C_T > 0$  such that for any  $\lambda \geq 0$ , any  $y_0 \in E_\lambda$  and associated solution  $y$  to (1.16), we have

$$\|y_0\|_{L^2}^2 \leq C e^{2c_0\lambda^{k/2}} \int_0^T \int_\omega |y(t, x)|^2 dt dx. \quad (4.4)$$

**Remark 4.3.** • The constant  $\gamma$  appearing in the exponent in (4.3) may exactly be taken as  $\gamma = \kappa + \varepsilon$  for any  $\varepsilon > 0$  where  $\kappa$  appears in the exponent in Estimate (1.13), Theorem 1.13 for some  $S > \sup_{x \in \mathcal{M}} d_{\mathcal{L}}(x, \omega)$ . In this case, the constant  $C > 0$  in front of the exponential also depend on  $\varepsilon$ .

- The constant  $c_0$  appearing in the exponent in (4.4) may also be taken as  $c_0 = \kappa + \varepsilon$  for any  $\varepsilon > 0$  (where  $\kappa$  appears in the exponent in Estimate (1.13), Theorem 1.13 for some  $S > \sup_{x \in \mathcal{M}} d_{\mathcal{L}}(x, \omega)$ ) in the case where  $k \geq 2$ , but only  $c_0 = \kappa + 2\sqrt{\alpha} + \varepsilon$  for any  $\varepsilon > 0$  in the case  $k = 1$ , which is the classical (elliptic) heat equation (where  $\alpha$  is any constant  $> \sqrt{2} \sup_{x \in \mathcal{M}} d_{\mathcal{L}}(x, \omega)$ ).
- This is exactly the cost of controlling low frequencies, following [LR95]. For instance, (4.3) implies that for all  $y_0 \in \Pi_\lambda L^2(\mathcal{M}) = E_\lambda$  ( $\Pi_\lambda$  being the orthogonal projector associated to the spectral space of  $\mathcal{L}$  with eigenvalues lower than  $\lambda$ ), there exists  $f \in L^2((0, T); L^2(\omega))$  with  $\|f\|_{L^2((0, T); L^2(\omega))}^2 \leq \frac{C}{T} e^{(2\gamma\lambda^{k/2} + \frac{C}{T})} \|y_0\|_{L^2}^2$  such that the solution to

$$\begin{cases} \partial_t y + \mathcal{L}y &= \Pi_\lambda \mathbf{1}_\omega f \\ y(0) &= y_0 \end{cases} \quad (4.5)$$

satisfies  $y(T) = 0$ . Note that this finite dimensional observability/controllability property is interesting in itself. For the time being and to the authors' knowledge, it is now understood in few situations, i.e. essentially in case  $\mathcal{L}$  is an elliptic selfadjoint second order operator [LR95], the bi-Laplace operator [LR05], the Stoke operator [CSL16], and in case of some lower order perturbation of such operators [Léa10].

Again, the situation of Example 1.11, the exponent  $\lambda^{k/2}$  is optimal in general, as can be seen when testing on eigenfunctions and using Proposition 1.12.



In the proofs of Estimate (4.4) and Theorem 1.16, we shall moreover need the following definitions:

$$\mathcal{I}(T, \lambda) = \int_0^T e^{-\alpha(\frac{1}{t} + \frac{1}{T-t})} e^{-\lambda t} dt, \quad (4.6)$$

and

$$\mathcal{I}(T, \mathcal{L})u = \sum_{j \in \mathbb{N}} \mathcal{I}(T, \lambda_j) a_j \varphi_j = \sum_{j \in \mathbb{N}} \left( \int_0^T e^{-\alpha(\frac{1}{t} + \frac{1}{T-t})} e^{-\lambda_j t} dt \right) a_j \varphi_j, \quad \text{for } u = \sum_{j \in \mathbb{N}} a_j \varphi_j. \quad (4.7)$$

*Proof of Lemma 4.2.* The proofs of (4.3) and (4.4) are similar. Let us start with that of (4.3). We start by using Theorem 1.31 in the simpler case  $s = k$ . Now, we fix any  $S > \sup_{x \in \mathcal{M}} d_{\mathcal{L}}(x, \omega)$ : Estimate (1.13) yields the existence of  $C, \kappa, \mu_0 > 0$  such that for all  $W_0 = (w, \partial_t w)|_{s=0}$  (note that all constants then depend on these, and hence on the chosen  $S > 0$ ), the associated solution to (4.1) satisfies

$$\|W_0\|_{L^2 \times \mathcal{H}_{\mathcal{L}}^{-1}} \leq C e^{\kappa \mu} \|w\|_{L^2(\cdot-S, S[\times \omega])} + \frac{1}{\mu} \|W_0\|_{\mathcal{H}_{\mathcal{L}}^k \times \mathcal{H}_{\mathcal{L}}^{k-1}}, \quad \mu \geq \mu_0. \quad (4.8)$$

Note that (4.8) implies the same estimate for all  $\mu > 0$ , in which case  $\kappa$  has to be replaced by a bigger constant.

Assume now that  $w(s)$  is associated to  $y$  as  $w(s) = \int_0^T k_T(t, s) y(t) dt$ , where  $y$  is the solution to (1.16) with initial datum  $y_0 \in E_\lambda$ . Then, in (4.1),  $W_0$  is of the particular form  $W_0 = \left(0, \int_0^T e^{-\alpha(\frac{1}{t} + \frac{1}{T-t})} y(t) dt\right)$ , so that a calculation (see [EZ11a, Equation (3.3)]) yields

$$\begin{aligned} \|W_0\|_{L^2 \times \mathcal{H}_{\mathcal{L}}^{-1}}^2 &\geq (1 + \lambda)^{-1} \|W_0\|_{\mathcal{H}_{\mathcal{L}}^1 \times L^2}^2 = (1 + \lambda)^{-1} \left\| \int_0^T e^{-\alpha(\frac{1}{t} + \frac{1}{T-t})} y(t) dt \right\|_{L^2}^2 \\ &\geq \frac{(1 + \lambda)^{-1} T^2}{9} e^{-\frac{9\alpha}{T}} \|y(T)\|_{L^2}^2. \end{aligned} \quad (4.9)$$

Moreover, we have  $W_0 \in E_\lambda \times E_\lambda$  so that

$$\frac{\|W_0\|_{\mathcal{H}_{\mathcal{L}}^k \times \mathcal{H}_{\mathcal{L}}^{k-1}}}{\|W_0\|_{L^2 \times \mathcal{H}_{\mathcal{L}}^{-1}}} \leq (1 + \lambda)^{\frac{k}{2}}. \quad (4.10)$$

As a consequence, (4.8) implies

$$\left(1 - \frac{(1 + \lambda)^{\frac{k}{2}}}{\mu}\right) \|W_0\|_{L^2 \times \mathcal{H}_{\mathcal{L}}^{-1}} \leq C e^{\kappa \mu} \|w\|_{L^2(\cdot-S, S[\times \omega])}, \quad \mu \geq \mu_0,$$

and hence, choosing  $\mu = (1 + \lambda)^{\frac{k}{2}}(1 + \varepsilon)$  for  $\varepsilon \in (0, 1)$ , this is

$$\varepsilon \|W_0\|_{L^2 \times \mathcal{H}_{\mathcal{L}}^{-1}} \leq C(1 + \varepsilon) e^{\kappa(1+\lambda)^{\frac{k}{2}}(1+\varepsilon)} \|w\|_{L^2(\cdot-S, S[\times \omega])}, \quad \lambda \geq \lambda_0 = \mu_0^{\frac{2}{k}},$$

and  $\|W_0\|_{L^2 \times \mathcal{H}_{\mathcal{L}}^{-1}} \leq C_\varepsilon e^{(\kappa+\varepsilon)\lambda^{\frac{k}{2}}} \|w\|_{L^2(\cdot-S, S[\times \omega])}$  for all  $\varepsilon > 0$  (different from that in the previous line). Using Cauchy-Schwarz inequality together with (4.2) (with  $\delta$  sufficiently close to one) we obtain that, for

some  $C > 0$  depending only on  $S, \alpha, \delta$ , but not on  $T$ , we have

$$\begin{aligned} \|w\|_{L^2(\cdot)_{-S, S[\times\omega]}}^2 &\leq \left( \int_{]0, T[\times]-S, S[} k_T(t, s)^2 dt ds \right) \int_0^T \int_{\omega} |y(t, x)|^2 dx dt \\ &\leq CT e^{\frac{C}{T}} \int_0^T \int_{\omega} |y(t, x)|^2 dx dt, \end{aligned} \quad (4.11)$$

which then gives (4.3) for  $\lambda \geq \lambda_0$ . The estimate for  $\lambda \in [0, \lambda_0]$  remains valid up to changing the constant  $C$ .

To prove (4.4), we follow the same lines, except for the lower bound (4.9), which we replace by an estimate of Corollary 4.5 below. Namely, with the notation (4.7), we have  $W_0 = (0, \mathcal{I}(T, \mathcal{L})y_0)$ , so that, according to Corollary 4.5, we have for all  $T > 0$  and  $s \in \mathbb{R}$ , the existence of  $C = C_{\alpha, T, s} > 0$  such that

$$\|W_0\|_{L^2 \times \mathcal{H}_c^{-1}} = \|\mathcal{I}(T, \mathcal{L})y_0\|_{\mathcal{H}_c^{-1}} \geq C \|y_0\|_{\frac{1}{2}, -2\sqrt{\alpha}, -1-\frac{3}{2}} = C \left\| (\mathcal{L} + 1)^{-\frac{5}{4}} e^{-2\sqrt{\alpha}\mathcal{L}} y_0 \right\|_{L^2}. \quad (4.12)$$

Recalling that  $y_0 \in E_\lambda$ , this implies

$$\|W_0\|_{L^2 \times \mathcal{H}_c^{-1}} \geq C \left\| (\mathcal{L} + 1)^{-\frac{5}{4}} e^{-2\sqrt{\alpha}\mathcal{L}} y_0 \right\|_{L^2} \geq C \left\| e^{-(2+\varepsilon)\sqrt{\alpha}\mathcal{L}} y_0 \right\|_{L^2} \geq C e^{-(2+\varepsilon)\sqrt{\alpha}\lambda} \|y_0\|_{L^2}.$$

Applying then (4.8) (for any  $\mu > 0$ ) with this lower bound, together with (4.11) and (4.10) as above and the fact that  $e^{(2+\varepsilon)\sqrt{\alpha}\lambda} \leq C_\varepsilon e^{\varepsilon\lambda^{\frac{k}{2}}}$  for any  $k \geq 2$  (in case  $k = 1$ , the constant  $(2 + \varepsilon)\sqrt{\alpha}$  has to be taken into account), concludes the proof of (4.4), and hence of the lemma.  $\square$

Note that in the proof of (4.4), and in the case  $k > 2$ , we could simply replace (4.12) by the rough estimate

$$\|y_\lambda(0)\|_{L^2}^2 \leq e^{2\lambda T} \|y_\lambda(T)\|_{L^2}^2,$$

which would be enough for the purpose of Estimate (4.4). This is not possible at all in case  $k = 1$ , and in case  $k = 2$ , would require  $c_0$  to depend (linearly) on  $T$ .

#### 4.1 Approximate controllability with polynomial cost in large time: Proof of Theorem 1.20

From the low-frequency Lemma 4.2, Estimate (4.3), the proof of the theorem follows the spirit of [LR95, Mil10, EZ11a] but is simpler. It combines the cost of controllability of low frequencies, of order  $e^{\gamma\lambda^{\frac{k}{2}}} = e^{\gamma\lambda}$  ( $k = 2$  in this part) and the dissipation of the heat at high frequency, of order  $e^{-t\lambda}$ . However, here, we do not perform the usual iterative procedure since it does not seem to improve the estimates.

*Proof of Theorem 1.20.* For  $y \in L^2(\mathcal{M})$ , we decompose  $y = y_\lambda + r_\lambda$  with  $y_\lambda \in E_\lambda$  and  $r_\lambda \in E_\lambda^\perp$ .

On the one hand, using Lemma 4.2 Estimate (4.3) for  $y_\lambda$  on the time interval  $(T - \eta, T)$  (the problem being time invariant), we obtain, uniformly with respect to  $T > 0, \eta \in ]0, T[, \lambda > 0$ ,

$$\|y_\lambda(T)\|_{L^2}^2 \leq C e^{(2\gamma\lambda^{k/2} + \frac{C}{\eta})} \int_{(T-\eta)}^T \int_{\omega} |y_\lambda(t, x)|^2 dt dx. \quad (4.13)$$

On the other hand, we have

$$\begin{aligned} \|r_\lambda(t)\|_{L^2} &\leq e^{-\lambda t} \|r_\lambda(0)\|_{L^2} \leq e^{-\lambda t} \|y(0)\|_{L^2}, \\ \int_{T-\eta}^T \int_\omega |r_\lambda(t, x)|^2 dt dx &\leq \frac{1}{2\lambda} e^{-2\lambda(T-\eta)} \|y(0)\|_{L^2}^2. \end{aligned} \quad (4.14)$$

The last estimate gives

$$\int_{(T-\eta)}^T \int_\omega |y_\lambda(t, x)|^2 dt dx \leq 2 \int_{(T-\eta)}^T \int_\omega |y(t, x)|^2 dt dx + \frac{1}{\lambda} e^{-2\lambda(T-\eta)} \|y(0)\|_{L^2}^2.$$

So, using successively (4.14), (4.13) and the last estimate, we finally obtain for  $T > 0, \eta \in ]0, T[, \lambda > 0$ ,

$$\begin{aligned} \|y(T)\|_{L^2}^2 &= \|y_\lambda(T)\|_{L^2}^2 + \|r_\lambda(T)\|_{L^2}^2 \\ &\leq \|y_\lambda(T)\|_{L^2}^2 + e^{-2\lambda T} \|y(0)\|_{L^2}^2 \\ &\leq C e^{(2\gamma\lambda^{k/2} + \frac{C}{\eta})} \int_{T-\eta}^T \int_\omega |y(t, x)|^2 dt dx + C \left( e^{-2\lambda T} + e^{(2\gamma\lambda^{k/2} + \frac{C}{\eta} - 2\lambda(T-\eta))} \right) \|y(0)\|_{L^2}^2 \\ &\leq C e^{(2\gamma\lambda^{k/2} + \frac{C}{\eta})} \int_{T-\eta}^T \int_\omega |y(t, x)|^2 dt dx + 2C e^{(2\gamma\lambda^{k/2} + \frac{C}{\eta} - 2\lambda(T-\eta))} \|y(0)\|_{L^2}^2. \end{aligned}$$

Now, we recall that we assume  $k = 2$  (for  $k > 2$ , the diffusion cannot compete with the cost of controllability of low frequencies). Consequently, we obtain, for all  $T > 0, \eta \in ]0, T[, \lambda > 0$ ,

$$\|y(T)\|_{L^2}^2 \leq C e^{C/\eta} \left( e^{2\gamma\lambda} \int_{T-\eta}^T \int_\omega |y(t, x)|^2 dt dx + e^{2\lambda(\gamma+\eta-T)} \|y(0)\|_{L^2}^2 \right). \quad (4.15)$$

This now provides information if  $T$  is sufficiently large. Namely, setting  $\varepsilon = e^{-2\lambda}$ , we obtain the existence of  $C > 0$  such that for all  $\eta > 0, T \geq \gamma + \eta$ , all  $\varepsilon \in ]0, 1[$ , we have

$$\|y(T)\|_{L^2}^2 \leq C e^{C/\eta} \left( \frac{1}{\varepsilon^\gamma} \int_{T-\eta}^T \int_\omega |y(t, x)|^2 dt dx + \varepsilon^{T-(\gamma+\eta)} \|y(0)\|_{L^2}^2 \right).$$

Changing  $\varepsilon^{T-(\gamma+\eta)}$  into  $\varepsilon$ , this implies the existence of  $C > 0$  such that for all  $\eta > 0$ , all  $T > \gamma + \eta$  and all  $\varepsilon \in ]0, 1[$ , we have

$$\frac{e^{-C/\eta}}{C} \|y(T)\|_{L^2}^2 \leq \frac{1}{\varepsilon^{\frac{\gamma}{T-(\gamma+\eta)}}} \int_{T-\eta}^T \int_\omega |y(t, x)|^2 dt dx + \varepsilon \|y(0)\|_{L^2}^2.$$

This concludes the proof of Theorem 1.20 with  $T_0 := \gamma$  after having remarked that the parabolic dissipation yields  $\|y(T)\|_{L^2}^2 \leq \|y(0)\|_{L^2}^2$ , and hence the case  $\varepsilon \geq 1$ .  $\square$

## 4.2 Approximate controllability in Gevrey-type spaces: Proof of Theorem 1.18

The proof of Theorem 1.18 follows the same lines as Theorem 1.20, decomposing into low and high frequencies, but uses Estimate (4.4) instead of (4.3).

*Proof of Theorem 1.18.* For  $y \in L^2(\mathcal{M})$  arbitrary, we again write the decomposition  $y = y_\lambda + r_\lambda$  with  $y_\lambda \in E_\lambda$  and  $r_\lambda \in E_\lambda^\perp$ . Note that, using the fact that  $y_\lambda$  is solution of the heat equation in  $E_\lambda$ , we obtain from Lemma 4.2 that

$$\|y_\lambda(0)\|_{L^2}^2 \leq C_T e^{2c_0\lambda^{k/2}} \int_0^T \int_\omega |y_\lambda(t, x)|^2 dt dx.$$

Moreover, we have (recall that the norm  $\|\cdot\|_{\alpha, \theta}$  is defined in (1.20);  $\alpha$  will be eventually taken equal to  $k/2$ )

$$\begin{aligned} \|r_\lambda(t)\|_{L^2} &\leq e^{-\lambda t} \|r_\lambda(0)\|_{L^2} \leq e^{-\lambda t} e^{-\theta\lambda^\alpha} \|r_\lambda(0)\|_{\alpha, \theta} \leq e^{-\lambda t - \theta\lambda^\alpha} \|y(0)\|_{\alpha, \theta} \\ \int_{T-\eta}^T \int_\omega |r_\lambda(t, x)|^2 dt dx &\leq \frac{1}{2\lambda} e^{-2\lambda(T-\eta)} e^{-2\theta\lambda^\alpha} \|y(0)\|_{\alpha, \theta}^2. \end{aligned}$$

From this last estimate, we obtain

$$\begin{aligned} \int_0^T \int_\omega |y_\lambda(t, x)|^2 dt dx &\leq 2 \int_0^T \int_\omega |y(t, x)|^2 dt dx + 2 \int_0^T \int_\omega |r_\lambda(t, x)|^2 dt dx \\ &\leq 2 \int_0^T \int_\omega |y(t, x)|^2 dt dx + \lambda^{-1} e^{-2\theta\lambda^\alpha} \|y(0)\|_{\alpha, \theta}^2. \end{aligned}$$

So, combining all these estimates, we finally obtain for  $\lambda > 0$

$$\begin{aligned} \|y(0)\|_{L^2}^2 &= \|y_\lambda(0)\|_{L^2}^2 + \|r_\lambda(0)\|_{L^2}^2 \\ &\leq \|y_\lambda(0)\|_{L^2}^2 + e^{-2\theta\lambda^\alpha} \|y(0)\|_{\alpha, \theta}^2 \\ &\leq C_T e^{2c_0\lambda^{k/2}} \left( 2 \int_0^T \int_\omega |y(t, x)|^2 dt dx + e^{-2\theta\lambda^\alpha} \|y(0)\|_{\alpha, \theta}^2 \right) + e^{-2\theta\lambda^\alpha} \|y(0)\|_{\alpha, \theta}^2. \end{aligned} \quad (4.16)$$

Now, for  $\alpha = k/2$ , we find for all  $\lambda > 0$  that

$$\|y(0)\|_{L^2}^2 \leq C e^{2c_0\lambda^{k/2}} \int_0^T \int_\omega |y(t, x)|^2 dt dx + C e^{-2(\theta-c_0)\lambda^{k/2}} \|y(0)\|_{k/2, \theta}^2.$$

Assuming  $\theta > c_0$ , and setting  $\varepsilon = C e^{-2(\theta-c_0)\lambda^{k/2}} \in ]0, 1[$ , this is precisely (1.21) with  $\theta_0 = c_0$ . The full range of  $\varepsilon > 0$  follows from the simple estimate  $\|y(0)\|_{L^2}^2 \leq \|y(0)\|_{k/2, \theta}^2$ .  $\square$

### 4.3 Approximate controllability in natural spaces with exponential cost: Proof of Theorem 1.16

Let us now proceed to the proof of Theorem 1.16. It does not rely on frequency cutoff (we do not distinguish between low and high frequencies), and hence on Lemma 4.2. Instead, we directly apply the transmutation result of Proposition 4.1 to the full solution and use precise properties of the operator  $\mathcal{I}(T, \mathcal{L})$  defined in (4.7) (which we already used in the proof of Estimate (4.4)), proved in the next section. Note also that here, as opposed to the above two sections, we need to use the strong version of Theorem 1.13.

*Proof of Theorem 1.16.* We apply directly the transmutation kernel to the solution. Using Theorem 1.13, we obtain

$$\|W_0\|_{L^2 \times \mathcal{H}_L^{-1}} \leq C e^{\kappa \mu^k} \|w\|_{L^2(\cdot)_{-S, S[\cdot \times \omega]}} + \frac{1}{\mu} \|W_0\|_{\mathcal{H}_L^1 \times L^2}, \quad \mu > 0, \quad (4.17)$$

see e.g. [LL15, Lemma A.3] to obtain the range  $\mu \in [0, \mu_0]$  (however deteriorating the constant  $\kappa$ ). Then, we recall that, with the notation (4.7), we have

$$W_0 = (0, \mathcal{I}(T, \mathcal{L})y_0),$$

so that, according to Corollary 4.5, we have for all  $T > 0$  and  $s \in \mathbb{R}$ , the existence of  $C = C_{\alpha, T, s} > 0$  such that

$$C^{-1} \|y_0\|_{\frac{1}{2}, -2\sqrt{\alpha}, s - \frac{3}{2}} \leq \|W_0\|_{\mathcal{H}_L^{s+1} \times \mathcal{H}_L^s} = \|\mathcal{I}(T, \mathcal{L})y_0\|_{\mathcal{H}_L^s} \leq C \|y_0\|_{\frac{1}{2}, -2\sqrt{\alpha}, s - \frac{3}{2}},$$

where  $\|y_0\|_{\frac{1}{2}, -2\sqrt{\alpha}, s} = \left\| (\mathcal{L} + 1)^{\frac{s}{2}} e^{-2\sqrt{\alpha}\mathcal{L}} y_0 \right\|_{L^2}$  (see (4.19) for the definition of the norms). In particular, this implies

$$\frac{\|W_0\|_{\mathcal{H}_L^1 \times L^2}}{\|W_0\|_{L^2 \times \mathcal{H}_L^{-1}}} \leq C \frac{\|y_0\|_{\frac{1}{2}, -2\sqrt{\alpha}, -\frac{3}{2}}}{\|y_0\|_{\frac{1}{2}, -2\sqrt{\alpha}, -1 - \frac{3}{2}}} = C \Lambda_1(H(\mathcal{L})y_0) \leq 2C \Lambda_1(y_0), \quad \Lambda_1(y_0) = \frac{\|y_0\|_{\mathcal{H}_L^1}}{\|y_0\|_{L^2}},$$

where  $H(\lambda) = (\lambda + 1)^{-\frac{5}{4}} e^{-2\sqrt{\alpha}\lambda}$ , which is positive, decreasing to zero, and we have thus used Corollary 4.9 in the last inequality.

When combined with (4.11) (still valid in this context), we now obtain, for all  $\mu > 0$ ,

$$\|W_0\|_{L^2 \times \mathcal{H}_L^{-1}}^2 \leq C e^{2\kappa \mu^k} \int_0^T \int_{\omega} |y(t, x)|^2 dx dt + \frac{C}{\mu^2} \Lambda_1(y_0)^2 \|W_0\|_{L^2 \times \mathcal{H}_L^{-1}}^2.$$

Writing  $\Lambda = \Lambda_1(y_0)$ , taking  $\mu = \sqrt{2C}\Lambda$ , and recalling that  $C^{-1} \|y_0\|_{\frac{1}{2}, -2\sqrt{\alpha}, -1 - \frac{3}{2}} \leq \|W_0\|_{L^2 \times \mathcal{H}_L^{-1}}$ , this gives after absorption

$$\|y_0\|_{\frac{1}{2}, -2\sqrt{\alpha}, -1 - \frac{3}{2}}^2 \leq C e^{c\Lambda^k} \int_0^T \int_{\omega} |y(t, x)|^2 dx dt, \quad \Lambda = \frac{\|y_0\|_{\mathcal{H}_L^1}}{\|y_0\|_{L^2}}. \quad (4.18)$$

To conclude, we recall that  $\|y_0\|_{\frac{1}{2}, -2\sqrt{\alpha}, -1 - \frac{3}{2}} = \left\| (\mathcal{L} + 1)^{-\frac{5}{4}} e^{-2\sqrt{\alpha}\mathcal{L}} y_0 \right\|_{L^2} \geq C \left\| e^{-3\sqrt{\alpha}(\mathcal{L}+1)} y_0 \right\|_{L^2}$  and we use Lemma 4.6 with  $F(s) = s + 1$  and  $G(s) = e^{-3\sqrt{\alpha}s}$  which is convex, to finally obtain

$$\|y_0\|_{L^2} \leq C e^{c\Lambda_1(y_0)} \left\| e^{-3\sqrt{\alpha}(\mathcal{L}+1)} y_0 \right\|_{L^2} \leq C e^{c\Lambda_1(y_0)} \|y_0\|_{\frac{1}{2}, -2\sqrt{\alpha}, -1 - \frac{3}{2}}.$$

Together with (4.18), this concludes the proof of (1.18). Now, to prove (1.19), take any  $\mu > 0$ . Either  $\Lambda_1(y_0) = \frac{\|y_0\|_{\mathcal{H}_L^1}}{\|y_0\|_{L^2}} \geq \mu$ , and (1.19) holds (without the observation term), or else  $\Lambda_1(y_0) \leq \mu$ , and (1.18) yields (1.19) (without additional term on the right hand-side). This concludes the proof of the Theorem.  $\square$

#### 4.4 Technical lemmata used for the heat equation

In this section, we collect three technical lemmata that we used in the proofs of Theorems 1.16, 1.18 and 1.20 above.

First, we need an asymptotic expansion of the integral  $\mathcal{I}(T, \lambda)$  defined in (4.6) as  $\lambda \rightarrow +\infty$ .

**Lemma 4.4.** *For all  $\alpha > 0$  and  $T > 0$ , there exists  $C_T, \lambda_0 > 0$  such that for all  $\lambda \geq \lambda_0$ , there is  $R(T, \lambda) \in \mathbb{R}$  such that we have*

$$\mathcal{I}(T, \lambda) = \sqrt{\pi} \frac{\alpha^{\frac{1}{4}}}{\lambda^{\frac{3}{4}}} e^{-\frac{\alpha}{T}} e^{-2\sqrt{\alpha\lambda}} \left( 1 + \frac{R(T, \lambda)}{\lambda^{\frac{1}{4}}} \right), \quad |R(T, \lambda)| \leq C_T.$$

Next, this lemma allows us to link the operator  $\mathcal{I}(T, \mathcal{L})$  defined in 4.7 and the norms

$$\|u\|_{\delta, \theta, \sigma}^2 = \left\| (\mathcal{L} + 1)^{\frac{\sigma}{2}} e^{\theta \mathcal{L}^\delta} u \right\|_{L^2}^2 = \sum_{j \in \mathbb{N}} (\lambda_j + 1)^\sigma e^{2\theta \lambda_j^\delta} |a_j|^2, \quad \text{for } u = \sum_{j \in \mathbb{N}} a_j \varphi_j. \quad (4.19)$$

**Corollary 4.5.** *For all  $s \in \mathbb{R}$ , for all  $T, \alpha > 0$ , there exists  $C > 1$  such that we have*

$$C^{-1} \|u\|_{\frac{1}{2}, -2\sqrt{\alpha}, s - \frac{3}{2}} \leq \|\mathcal{I}(T, \mathcal{L})u\|_{\mathcal{H}_\mathcal{L}^s} \leq C \|u\|_{\frac{1}{2}, -2\sqrt{\alpha}, s - \frac{3}{2}},$$

The proof of the corollary only consists in remarking that, once  $\alpha, T$  are fixed, we have, according to Lemma 4.4, that

$$0 < \mathcal{I}(T, \lambda) \left( \frac{e^{-2\sqrt{\alpha\lambda}}}{(1+\lambda)^{\frac{3}{4}}} \right)^{-1} \rightarrow \sqrt{\pi} \alpha^{\frac{1}{4}} e^{-\frac{\alpha}{T}} > 0 \quad \text{as } \lambda \rightarrow +\infty,$$

and this quantity does not vanish on  $\mathbb{R}^+$  so that there is  $C > 1$  such that for all  $\lambda \geq 0$ , we have

$$C^{-1} \leq \mathcal{I}(T, \lambda) \left( \frac{e^{-2\sqrt{\alpha\lambda}}}{(1+\lambda)^{\frac{3}{4}}} \right)^{-1} \leq C,$$

which yields the result.

*Proof of Lemma 4.4.* Note first that, given  $\varepsilon \in (0, 1)$ , we may assume that  $\lambda_0$  is chosen such that  $T \geq \sqrt{\frac{\alpha}{\lambda}}(1 + \varepsilon)$  for  $\lambda \geq \lambda_0$ . We first change variables in  $\mathcal{I}(T, \lambda)$ , denoting  $\omega = \sqrt{\alpha\lambda}$  (new large parameter) and setting  $t = \sqrt{\frac{\alpha}{\lambda}}s = \frac{\alpha}{\omega}s$ , we have

$$\mathcal{I}(T, \lambda) = \frac{\alpha}{\omega} \int_0^{\frac{\omega T}{\alpha}} f_\omega(s) e^{-\omega(\frac{1}{s} + s)} ds, \quad f_\omega(s) = \exp\left(-\frac{\alpha}{T - \frac{\alpha s}{\omega}}\right).$$

The phase  $h(s) := \frac{1}{s} + s$  admits a single global strict (nondegenerate) minimum at the point  $s = 1$ , with  $h(1) = 2$ . Note also that  $0 \leq f_\omega \leq 1$ . Hence, using that  $\frac{\omega T}{\alpha} \geq 1 + \varepsilon$  by assumption, we have for  $\varepsilon \in (0, 1)$ , the estimate

$$\mathcal{I}(T, \lambda) = \frac{\alpha}{\omega} \int_{1-\varepsilon}^{1+\varepsilon} f_\omega(s) e^{-\omega(\frac{1}{s} + s)} ds + O_T(e^{-(h(1)+c_\varepsilon)\omega}), \quad c_\varepsilon > 0.$$

Let  $\phi : (1 - \varepsilon, 1 + \varepsilon) \rightarrow (-\varepsilon_1, \varepsilon_2)$  be a (Morse) diffeomorphism for some  $\varepsilon_1, \varepsilon_2 > 0$ , such that  $\phi(1) = 0$ , and with  $u = \phi(s)$ , we have

$$h(s) = h(1) + h''(1) \frac{u^2}{2} = 2 + u^2, \quad \text{sgn}(u) = \text{sgn}(s - 1).$$

Note that it is actually explicit, namely  $\phi(s) = (s - 1)/\sqrt{s}$ . We change variable, setting  $u = \phi(s)$ , and obtain

$$\mathcal{I}(T, \lambda) = \frac{\alpha}{\omega} \int_{-\varepsilon_1}^{\varepsilon_2} e^{-\omega(2+u^2)} f_\omega \circ \phi^{-1}(u) |(\phi^{-1})'(u)| du + O_T(e^{-(2+c_\varepsilon)\omega}), \quad c_\varepsilon > 0,$$

where  $(\phi^{-1})'(0) = 1$ .

Moreover, for  $s \in [1 - \varepsilon, 1 + \varepsilon] \subset [0, \frac{\omega T}{\alpha}]$ , we write  $f_\omega(s) = f_\omega(1) + R_\omega(s)$  with  $f_\omega(1) = \exp\left(-\frac{\alpha}{T - \frac{\alpha}{\omega}}\right)$  and

$$|R_\omega(s)| \leq |s - 1| \sup_{[1-\varepsilon, 1+\varepsilon]} |f'_\omega| \leq \frac{c}{\omega} |s - 1|, \quad (4.20)$$

where we used  $f'_\omega(s) = -\frac{1}{\omega} X^{-2} e^{-\frac{1}{X}}$  with  $X = (T - \frac{\alpha}{\omega}s)\alpha^{-1} \in [0, T/\alpha]$ .

As a consequence, we obtain

$$\begin{aligned} \mathcal{I}(T, \lambda) &= \frac{\alpha}{\omega} e^{-2\omega} \exp\left(-\frac{\alpha}{T - \frac{\alpha}{\omega}}\right) \int_{-\varepsilon_1}^{\varepsilon_2} e^{-\omega u^2} |(\phi^{-1})'(u)| du + \mathcal{R}(\omega) + O_T(e^{-(2+c_\varepsilon)\omega}) \\ &= \frac{\alpha}{\omega} e^{-2\omega} \exp\left(-\frac{\alpha}{T - \frac{\alpha}{\omega}}\right) \left\{ \int_{-\varepsilon_1}^{\varepsilon_2} e^{-\omega u^2} |(\phi^{-1})'(0)| du + \mathcal{R}'(\omega) \right\} + \mathcal{R}(\omega) + O_T(e^{-(2+c_\varepsilon)\omega}) \\ &= \frac{\alpha}{\omega} e^{-2\omega} \exp\left(-\frac{\alpha}{T - \frac{\alpha}{\omega}}\right) \left\{ \int_{\mathbb{R}} e^{-\omega u^2} |(\phi^{-1})'(0)| du + \mathcal{R}'(\omega) \right\} + \mathcal{R}(\omega) + O_T(e^{-(2+\tilde{c}_\varepsilon)\omega}), \quad \tilde{c}_\varepsilon > 0 \\ &= \frac{\alpha}{\omega} e^{-2\omega} \exp\left(-\frac{\alpha}{T - \frac{\alpha}{\omega}}\right) \left\{ \sqrt{\frac{\pi}{\omega}} + \mathcal{R}'(\omega) \right\} + \mathcal{R}(\omega) + O_T(e^{-(2+\tilde{c}_\varepsilon)\omega}), \quad \tilde{c}_\varepsilon > 0 \end{aligned}$$

with, using (4.20),

$$|\mathcal{R}(\omega)| = \left| \frac{\alpha}{\omega} \int_{-\varepsilon_1}^{\varepsilon_2} e^{-\omega(2+u^2)} R_\omega \circ \phi^{-1}(u) |(\phi^{-1})'(u)| du \right| \leq C \frac{\alpha}{\omega^2} e^{-2\omega} \int_{-\varepsilon_1}^{\varepsilon_2} |u| e^{-\omega u^2} du \leq \frac{\alpha}{\omega^2} e^{-2\omega} \frac{C}{\omega},$$

and

$$|\mathcal{R}'(\omega)| = \left| \int_{-\varepsilon_1}^{\varepsilon_2} e^{-\omega u^2} ((\phi^{-1})'(u) - (\phi^{-1})'(0)) du \right| \leq C \int_{-\varepsilon_1}^{\varepsilon_2} |u| e^{-\omega u^2} du \leq \frac{C}{\omega}.$$

Using finally that  $\exp\left(-\frac{\alpha}{T - \frac{\alpha}{\omega}}\right) = e^{-\frac{\alpha}{T}} (1 + O_T\left(\frac{\alpha}{\omega}\right))$ , we finally obtain

$$\mathcal{I}(T, \lambda) = \frac{\alpha}{\omega} e^{-2\omega} \exp\left(-\frac{\alpha}{T - \frac{\alpha}{\omega}}\right) \sqrt{\frac{\pi}{\omega}} \left(1 + O_T\left(\frac{1}{\sqrt{\omega}}\right)\right),$$

which, recalling that  $\omega = \sqrt{\alpha\lambda}$ , concludes the proof of the lemma.  $\square$

**Lemma 4.6.** *Let  $F : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  be any function and let  $G : F(\mathbb{R}^+) \rightarrow \mathbb{R}^+$  be a function such that  $G^2$  is convex (it is for instance the case if  $G$  is). Then, for all  $u \in \mathcal{H}_{\mathcal{L}}^F \cap D(G \circ F(\mathcal{L}))$ , we have*

$$G \left( \frac{\|u\|_{\mathcal{H}_{\mathcal{L}}^F}^2}{\|u\|_{L^2}^2} \right) \|u\|_{L^2} \leq \|G \circ F(\mathcal{L})u\|_{L^2}, \quad (4.21)$$

where the seminorm  $\mathcal{H}_{\mathcal{L}}^F$  is defined by  $\|u\|_{\mathcal{H}_{\mathcal{L}}^F}^2 = \sum_{j \in \mathbb{N}} F(\lambda_j) |a_j|^2$  if  $u = \sum_{j \in \mathbb{N}} a_j \varphi_j$ .

**Remark 4.7.** Using the previous lemma with  $F(s) = s + 1$  and  $G(s) = \frac{1}{\sqrt{s}}$ , we obtain the interpolation inequality

$$\frac{\|u\|_{L^2}}{\|u\|_{\mathcal{H}_{\mathcal{L}}^{-1}}} \leq \frac{\|u\|_{\mathcal{H}_{\mathcal{L}}^1}}{\|u\|_{L^2}},$$

comparing two types of “frequency functions”, the first of which being used e.g. in [Phu04] for the classical heat equation.

*Proof.* Dividing by  $\|u\|_{L^2}$  (if non zero, otherwise the inequality is clear), it is enough to prove (4.21) assuming  $\|u\|_{L^2} = 1$ . If so, we write  $u = \sum_{j \in \mathbb{N}} a_j \varphi_j$  with  $\sum_{j \in \mathbb{N}} |a_j|^2 = 1$ . Using the Jensen inequality with the convex function  $G^2$ , we have

$$\|G \circ F(\mathcal{L})u\|_{L^2}^2 = \sum_{j \in \mathbb{N}} G(F(\lambda_j))^2 |a_j|^2 \geq G^2 \left( \sum_{j \in \mathbb{N}} F(\lambda_j) |a_j|^2 \right) = G^2(\|u\|_{\mathcal{H}_{\mathcal{L}}^F}^2),$$

which concludes the proof of the lemma. □

**Lemma 4.8.** *Let  $F, G : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  be two nondecreasing continuous functions such that  $F(s)G(s) \rightarrow +\infty$  as  $s \rightarrow +\infty$ . Then, for all  $u \in D(F(\mathcal{L})G(\mathcal{L}))$ , we have*

$$\|F(\mathcal{L})u\|_{L^2} \|G(\mathcal{L})u\|_{L^2} \leq 2 \|F(\mathcal{L})G(\mathcal{L})u\|_{L^2} \|u\|_{L^2}.$$

Note that, replacing  $F$  and  $G$  by  $1/F$  and  $1/G$ , the same statement is true as well if  $F, G$  are nonvanishing, nonincreasing and such that  $F(s)G(s) \rightarrow 0$  as  $s \rightarrow +\infty$ .

**Corollary 4.9.** *Denoting, for  $\sigma > 0$  by  $\Lambda_\sigma(u) = \frac{\|u\|_{\mathcal{H}_{\mathcal{L}}^\sigma}}{\|u\|_{L^2}}$  the modified frequency functions, we have that for any  $H : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  nonvanishing, nonincreasing such that  $H(s) \rightarrow 0$  as  $s \rightarrow +\infty$ ,*

$$\Lambda_\sigma(H(\mathcal{L})u) \leq 2\Lambda_\sigma(u). \quad (4.22)$$

The corollary is obtained by taking  $F(s) = (s+1)^{\frac{\sigma}{2}}$  in Lemma 4.8,  $G = 1/H$  and  $u = H(\mathcal{L})v$ . Remark that the frequency function  $\Lambda$  used in the main part of the article is  $\Lambda = \Lambda_1$ .

**Remark 4.10.** The interpretation of the corollary is clearer. Indeed, in this context,  $H(\mathcal{L})$  is a compact operator of  $L^2$  and (4.22) only translates that the “average frequency” of  $H(\mathcal{L})u$  is smaller than the “average frequency” of  $u$ .



**Remark 4.11.** It is very likely that the previous Lemma (or at least Corollary 4.9) is still true with the constant 2 replaced by 1. Indeed, when taking  $H(s) = (s + 1)^{-\beta}$ , or  $H(s) = e^{-ts^\beta}$  with  $\beta, t > 0$ , Corollary 4.9 is true with a constant 1; in the first case, it is proved using Sobolev interpolation and in the second one using the monotonicity of the frequency function for solutions of the heat equation, see Phung [Phu04].

*Proof of Lemma 4.8.* First, since  $F$  and  $G$  are nondecreasing, by decomposing  $u = \sum a_j \varphi_j$  with frequency less than  $\lambda$  and greater than  $\lambda$ , we notice that for any  $\lambda > 0$ , we have

$$\begin{aligned} \|F(\mathcal{L})u\|_{L^2}^2 &= \sum_j F(\lambda_j)^2 |a_j|^2 = \sum_{\lambda_j \leq \lambda} F(\lambda_j)^2 |a_j|^2 + \sum_{\lambda_j > \lambda} F(\lambda_j)^2 |a_j|^2 \\ &\leq F(\lambda)^2 \sum_{\lambda_j \leq \lambda} |a_j|^2 + \sum_{\lambda_j > \lambda} F(\lambda_j)^2 \frac{G(\lambda_j)^2}{G(\lambda)^2} |a_j|^2 \\ &\leq F(\lambda)^2 \|u\|_{L^2}^2 + \frac{1}{G(\lambda)^2} \|F(\mathcal{L})G(\mathcal{L})u\|_{L^2}^2. \end{aligned}$$

Similarly, we have

$$\|G(\mathcal{L})u\|_{L^2}^2 \leq G(\lambda)^2 \|u\|_{L^2}^2 + \frac{1}{F(\lambda)^2} \|F(\mathcal{L})G(\mathcal{L})u\|_{L^2}^2.$$

Multiplying these two estimates, we obtain, for all  $\lambda > 0$

$$\|F(\mathcal{L})u\|_{L^2}^2 \|G(\mathcal{L})u\|_{L^2}^2 \leq F(\lambda)^2 G(\lambda)^2 \|u\|_{L^2}^4 + 2 \|u\|_{L^2}^2 \|F(\mathcal{L})G(\mathcal{L})u\|_{L^2}^2 + \frac{1}{F(\lambda)^2 G(\lambda)^2} \|F(\mathcal{L})G(\mathcal{L})u\|_{L^2}^4,$$

which is

$$\|F(\mathcal{L})u\|_{L^2} \|G(\mathcal{L})u\|_{L^2} \leq F(\lambda)G(\lambda) \|u\|_{L^2}^2 + \frac{1}{F(\lambda)G(\lambda)} \|F(\mathcal{L})G(\mathcal{L})u\|_{L^2}^2, \quad \text{for all } \lambda \geq 0.$$

Now, using that  $s \mapsto F(s)G(s)$  tends to  $+\infty$  as  $s \rightarrow +\infty$  and is nondecreasing, it is onto  $\mathbb{R}^+ \rightarrow [F(0)G(0), +\infty[$ . Since  $F(0)G(0) = \min FG \leq \frac{\|F(\mathcal{L})G(\mathcal{L})u\|_{L^2}}{\|u\|_{L^2}}$ , there is  $\lambda \geq 0$  such that  $F(\lambda)G(\lambda) = \frac{\|F(\mathcal{L})G(\mathcal{L})u\|_{L^2}}{\|u\|_{L^2}}$ , which together with the last estimate yields the sought result.  $\square$

## 5 A partially analytic example: Grushin type operators

In this section, we are concerned with the setting of Example 1.22 and give a proof of Theorem 1.23. As explained in Section 1.4, it only suffices to prove the analogue of Theorem 1.13 (with estimate (1.24) instead of (1.13)), that is for the hypoelliptic wave equation; all other results are then deduced as in Section 4.

The setting of Example 1.22 differs from the general setting of the paper (compact manifolds, analytic context) with two respects: (i) we do not suppose analyticity in all variables; (ii) the manifold  $\mathcal{M} = [-1, 1] \times \mathbb{T}$  has a boundary. Hence, there are four main differences in the proofs, the first of which being of geometric nature, the next two being linked to the analysis of [LL15], and the last one to hypoelliptic estimates:

1. the presence of the boundary makes it complicated to apply directly Theorem 3.5 coming from [RT05].
2. the partial analyticity assumption does not allow to make changes of variables to define the relation  $\triangleleft$ . For any couple of points  $x^0, x^1$ , we thus have to find some global set  $\Omega$  containing one (short) path linking them.
3. the application of the results in [LL15] yields an observation term of the form  $\|u\|_{H_b^1(\bar{\omega})}$  and we would expect it to be in  $L^2$ .
4. The available hypoelliptic estimates, similar to those of Theorem 1.4, do not apply directly in the presence of boundary.

The problem imposed by Item 1 is that because of the boundary, the shortest path between two points in  $\text{Int}(\mathcal{M})$  does not necessarily exists inside of  $\text{Int}(\mathcal{M})$ . To understand this issue, it may help to think about the flat metric in  $\mathbb{R}^n \setminus O$  where  $O$  a convex obstacle. The boundary  $\partial\mathcal{M} = \{x_1 = \pm 1\}$  can look like a convex obstacle for the Grushin case for instance (see Figure 3.1 of [BL13] for some drawing of geodesics in Grushin). The solution we propose is to apply the result of Rifford and Trélat [RT05] only locally away from the boundary. The drawback is then that our path is only piecewise normal geodesic. But this will be sufficient thanks to the variant Proposition 3.14 of Proposition 3.12.

The solution to the issue of Item 2 is the very simple geometry of  $[-1, 1]_{x_1} \times \mathbb{T}_{x_2}$ , so we almost do not perform any change of variable.

The solution to the issue of Item 3 is to use the fact that the operator  $P$  is elliptic in  $\{\zeta_a = 0\}$  where  $\zeta_a$  is the dual to the analytic variable  $z_a = (t, x_2)$ , see Section 5.3.

Concerning the issue of Item 4, we prove the necessary estimates in Section B.2. Recall that the operator is elliptic close to the boundary. So, we are left to patch the usual elliptic estimates close to the boundary with internal hypoelliptic estimates.

All in all, the proof of Estimate (1.24) is as above in two steps: first, proving (1.28) (hard part), and then performing energy estimates (soft part). The latter are done the same way as in Section 3.3.1, except that the hypoelliptic estimates of Corollary B.1 have to be replaced by those of Theorem B.3 (with boundary).

We now focus on the first part of the proof, that is, proving (1.28) in the context of Example 1.22. This corresponds to the above Step 1 (Section 3.1) and Step 2 (Section 3.2).

## 5.1 The geometric context

Denote  $\pi : [-1, 1]_{x_1} \times \mathbb{R}_{x_2} \rightarrow [-1, 1]_{x_1} \times \mathbb{T}_{x_2}$  the natural covering map,  $\pi(x_1, x_2) = (x_1, x_2 + \mathbb{Z})$ . The vector fields  $X_1$  and  $X_2$  can be lifted to  $[-1, 1]_{x_1} \times \mathbb{R}_{x_2}$ , which allows to define the natural sub-Riemannian distance on  $]-1, 1[_{x_1} \times \mathbb{R}_{x_2}$ . As for the case of  $[-1, 1]_{x_1} \times \mathbb{T}_{x_2}$ , the latter can be extended up to the boundary as well as all the other notions naturally inherited. We keep the same notations without leading to any confusion.

We will need the following Geometric Lemma, the proof of which relies on an iterative use of a slight variant of the result of Rifford-Trélat [RT05], see Theorem 3.5.

**Lemma 5.1.** *Let  $x^0 = (x_1^0, x_2^0)$  and  $x^1 = (x_1^1, x_2^1)$  in  $[-1, 1]_{x_1} \times \mathbb{T}_{x_2}$ . Then, for any  $\varepsilon > 0$ , there exists a continuous path  $\gamma : [0, 1] \mapsto [-1, 1]_{x_1} \times \mathbb{R}_{x_2}$  so that with  $\gamma(s) = (x_1(s), x_2(s))$  we have*

1.  $\pi(\gamma(0)) = x^0$  and  $\text{dist}(\pi(\gamma(1)), x^1) < \varepsilon$ ;

2.  $x_1(s) \notin \{-1, 1\}$  for  $s \in ]0, 1[$ ;
3.  $\gamma$  is piecewise normal geodesic in  $] -1, 1[_{x_1} \times \mathbb{R}_{x_2}$ ;
4. if  $x_1^1 = -1$  (resp.  $1$ ) then there is  $\delta > 0$  so that  $\gamma(s) = (-s, x_2^1)$  (resp.  $\gamma(s) = (s, x_2^1)$ ) for  $s \in [1 - \delta, 1]$ .  
Similarly, if  $x_1^0 = -1$  (resp.  $1$ ) then there is  $\delta > 0$  so that  $\gamma(s) = (-1 + s, x_2^0)$  (resp.  $\gamma(s) = (1 - s, x_2^0)$ ) for  $s \in [0, \delta]$ ;
5.  $\text{length}(\gamma) \leq \text{dist}(x^0, x^1) + \varepsilon$ .

*Proof.* Note first that the paths defined in Item 4 are normal geodesic paths (i.e. geodesics, since the metric is Riemannian near the boundary) corresponding to  $(\xi_1, \xi_2) = (\pm 1/2, 0)$  since  $f$  does not depend on  $x_2$  near the boundary. Therefore, by defining  $\gamma$  like this for  $s \in [1 - \delta, 1]$ , we have  $\gamma(1 - \delta) \in ] -1, 1[_{x_1} \times \mathbb{T}_{x_2}$  and  $\text{length}(\gamma(s), s \in [1 - \delta, 1]) = \delta$ ; hence up to changing the length of  $\gamma$  by  $\delta$ , we are left to the case where  $x^1$  does not belong to the boundary. The argument shows that we may assume as well that  $x^0$  does not belong to the boundary.

Let now  $\tilde{\gamma}$  be a smooth path on  $[-1, 1]_{x_1} \times \mathbb{T}_{x_2}$  so that  $\tilde{\gamma}(0) = x^0$ ,  $\tilde{\gamma}(1) = x^1$  and  $\text{length}(\tilde{\gamma}) \leq \text{dist}(x^0, x^1) + \varepsilon$ . We select one continuous lifting of  $\tilde{\gamma}$  on  $[-1, 1]_{x_1} \times \mathbb{R}_{x_2}$ , denoted by  $\gamma_1$ , so that  $\pi(\gamma_1(s)) = \tilde{\gamma}(s)$  for  $s \in [0, 1]$ . Moreover, we have  $\text{length}(\gamma_1) = \text{length}(\tilde{\gamma})$  that we denote by  $L$ . Since  $x^0, x^1 \notin \{\pm 1\} \times \mathbb{T}$ , then up to deforming a bit  $\gamma_1$  without changing  $\gamma_1(0)$  (and still denoting it with the same name), we can assume that  $\text{dist}(\gamma_1(s), \{\pm 1\} \times \mathbb{R}) > \eta > 0$  for all  $s \in [0, 1]$ , up to having only the estimate  $\text{length}(\gamma_1) \leq \text{dist}(x^0, x^1) + 2\varepsilon$ . Now, we choose  $N \in \mathbb{N}$  large enough so that  $(L + \varepsilon)/N < \eta$ . Up to reparametrization, we can also assume that  $\gamma_1 : [0, L] \rightarrow ] -1, 1[_{\times} \mathbb{R}$  has unit speed.

Denote  $t_i = \gamma_1(iL/N)$  for  $i = 0, \dots, N$ . In particular, we have  $\text{dist}(t_i, t_{i+1}) \leq L/N$  since  $\gamma_1$  has unit speed.

We now define  $m_i$  for  $i = 0, \dots, N$  by induction, so that the following properties are satisfied:

- (P1)  $\text{dist}(m_i, t_i) < \varepsilon/2N$ ;
- (P2) there is a minimizing normal geodesic between  $m_i$  and  $m_{i+1}$ .

Note that these properties imply in particular

$$\text{dist}(m_i, \{\pm 1\} \times \mathbb{R}) \geq \text{dist}(t_i, \{\pm 1\} \times \mathbb{R}) - \text{dist}(m_i, t_i) > \eta - \varepsilon/2N > (L + \varepsilon/2)/N. \quad (5.1)$$

Let us now construct the points  $m_i$  by induction as follows:

- $m_0 = \tilde{\gamma}(0)$ ;
- $m_i \rightarrow m_{i+1}$ : by iteration hypothesis,  $\text{dist}(m_i, t_i) \leq \varepsilon/N$ . Note that the ball  $B(m_i, (L + \varepsilon/2)/N)$  does not intersect the boundary thanks to (5.1) and that  $\text{dist}(m_i, t_{i+1}) \leq \text{dist}(m_i, t_i) + \text{dist}(t_i, t_{i+1}) < (L + \varepsilon/2)/N$ . In particular,  $t_{i+1} \in B(m_i, (L + \varepsilon/2)/N)$ , and there exists one ball of radius  $r < \varepsilon/2N$  so that  $B(t_{i+1}, r) \subset B(m_i, (L + \varepsilon/2)/N)$ .

A slight variant of Theorem 3.5 of Rifford-Trélat [RT05] implies that the image of the exponential map (given by  $T^*(\mathbb{R}^2) \rightarrow \mathbb{R}^2$ ,  $(m_0, \xi_0) \mapsto m(1)$  where  $(m(t), \xi(t))$  is the Hamiltonian curve issued from  $(m_0, \xi_0)$ , see Definition 3.2) from the point  $m_i$  is dense in  $B(m_i, (L + \varepsilon/2)/N)$ . In particular, there exists one point, which we choose as  $m_{i+1} \in B(t_{i+1}, r) \subset B(m_i, (L + \varepsilon/2)/N)$  so that there exists a minimizing normal geodesic between  $m_i$  and  $m_{i+1}$ , and (P2) is satisfied. Then, we have by construction  $\text{dist}(m_{i+1}, t_{i+1}) < r < \varepsilon/2N$ , so the first induction assumption (P1) is also fulfilled.

Once the process is finished, we have by construction,  $\text{dist}(m_i, m_{i+1}) \leq \text{dist}(m_i, t_i) + \text{dist}(t_i, t_{i+1}) + \text{dist}(m_{i+1}, t_{i+1}) \leq (L + \varepsilon)/N$ . Now, denote by  $\gamma$ , the path defined by concatenation of the above defined normal geodesic path linking  $m_i$  and  $m_{i+1}$ . We have  $\gamma(0) = m_0 = \tilde{\gamma}(0) = \gamma_1(0)$  by construction. Also, we have  $\text{dist}(\gamma(L), \gamma_1(L)) = \text{dist}(m_{N-1}, t_{N-1}) < \varepsilon/2N < \varepsilon$ . Moreover, since the geodesic linking  $m_i$  and  $m_{i+1}$  are minimizing, we have  $\text{length}(\gamma) = \sum_{i=0}^{N-1} \text{dist}(m_i, m_{i+1}) \leq L + \varepsilon$ . This concludes the proof of the lemma.  $\square$

## 5.2 A proof of Estimate (1.24)

Let us now sketch the proof of Estimate (1.24). It is very similar to that of Theorem 1.13, so that we only stress the main differences.

For this, we define  $\Omega = I \times U$  with  $I$  a bounded neighborhood of  $(-T, T)$  (where  $T$  is that of the statement of Theorem 1.23) and  $U$  a bounded neighborhood of  $\gamma$  in  $[-1, 1]_{x_1} \times \mathbb{R}_{x_2}$ . We consider the operator  $P = \partial_t^2 + \mathcal{L}$  in this set, and use the splitting of variables in  $\mathbb{R}^n = \mathbb{R}^{n_a} \times \mathbb{R}^{n_b} = \mathbb{R}^{d+1}$  with  $n = 3, n_a = 2, n_b = 1, d = 2$ , as

$$z = (z_a, z_b), \quad z_a = (t, x_2), \quad z_b = x_1,$$

with  $t$  being the time variable, and  $x = (x_1, x_2)$  the space variable.

Now, we follow the general proof. The geometrical context being made precise in Lemma 5.1, it only remains to check that we can apply the equivalent of Proposition 3.14, with the scheme of proof described in Remark 3.13. Since the appropriate piecewise geodesic path is constructed in Lemma 5.1, we only need to check that the local results (the equivalent of Lemma 3.11) can be applied in this setting.

There are two differences:

- We are in a situation where the only analytic variables are  $z_a = (t, x_2)$ . So, all Fourier multipliers defined in Section 2.2 (and therefore the associated relation  $\triangleleft$ ) are taken with respect to these variables, see Remark 2.4. The symbol of the wave operator  $P$  is

$$p(t, x_1, x_2, \tau, \xi_1, \xi_2) = -\tau^2 + \xi_1^2 + f(x_1, x_2)^2 \xi_2^2.$$

But we check that we are still in the situation of Remark 1.10 of [LL15] with  $z_a = (t, x_2)$  and  $z_b = x_1$ . Indeed,  $p(t, x_1, x_2, 0, \xi_1, 0) = \xi_1^2$  that is positive definite on  $\mathbb{R}_{\xi_1}$ .

- The equivalent of Lemma 3.11 should be obtained in the presence of boundary. We have to check that we can apply [LL15, Theorem 5.12], namely “propagation up to the boundary”  $\{x_1 = \pm 1\}$ . Let us only explain the construction near the boundary  $\{x_1 = 1\}$ , the other case being similar. One important thing is that we are in the geometric situation described in Lemma 5.1:  $P$  is already under the form of (3.5) and the choice of the geodesic close to the boundary of Item 4 of Lemma 5.1 is already the same straight line. Indeed, we almost do not need to perform any change of coordinates, but only a translation. We can directly construct the noncharacteristic hypersurfaces of Lemma 3.10 with  $l_0 = \varepsilon_0$ ,  $(t, \tilde{x}) = (t, x_2)$  (tangential) and  $x_d = x_1 + 1 - \varepsilon_0$  (normal). Everything works then as in the interior case precised before, except for the last hypersurface  $S_1 = \{\phi_1 = 0\}$ , which touches the the boundary  $\{x_1 = 1\}$  tangentially. For this last step, we apply the local propagation result up to the boundary [LL15, Theorem 5.12]. We only need to check that the additional assumptions of this result are fulfilled:

- The analytic variable  $z_a = (t, x_2)$  are tangential with respect to the boundary  $\{x_1 = 1\}$ ;

- Assumption 5.1 in [LL15] is fulfilled for  $P$  because close to the boundary,  $p = q_x((\tau, \xi_2)) + \tilde{q}_x(\xi_1)$  where  $q_x((\tau, \xi_2)) = -\tau^2 + f(x)^2 \xi_2^2$  and  $\tilde{q}_x(\xi_1) = \xi_1^2$  are both quadratic forms independent on  $t$  and  $x_2$ ;
- The boundary  $\{x_1 = 1\}$  is non characteristic for  $P$ ;
- To apply Theorem 5.12 of [LL15] in this context, calling  $(x', x_n)$  the variables in that reference (the domain is locally  $\{x_n > 0\}$  in [LL15], it is  $\{x_1 < 1\}$  here), one needs to set  $x_n = 1 - x_1$  and  $x' = (t, x_2)$  so that  $\{x_1 \leq 1\}$  is transformed into  $\mathbb{R}_+^n$ . The defining function of the last hypersurface  $\phi_1(t, \tilde{x}, x_d) = G((t, \tilde{x}), 1) - x_d$  is changed, for the application of [LL15, Theorem 5.12] into  $\tilde{\phi}_1(x', x_n) = G(x', 1) - (1 - x_n)$ . The assumption  $\partial_{x_n} \tilde{\phi}_1 = -\partial_{x_d} \phi_1 = 1 > 0$  of [LL15, Theorem 5.12] is hence satisfied.

This variant of Proposition 3.14, with an application of the boundary Theorem 5.12 of [LL15] as a last step, leads to the relation

$$\|M_\mu^{\beta\mu} \sigma_{r,\mu} u\|_1 \leq C e^{\kappa\mu} \left( \|M_\mu^{\alpha\mu} \vartheta_\mu u\|_1 + \|Pu\|_{L^2(\Omega)} \right) + C e^{-\kappa'\mu} \|u\|_1$$

where  $M_\mu^{\beta\mu}$  is defined with the analytic variables  $z_a = (t, x_2)$  and naturally extended to the boundary case since  $z_a$  is tangential. The function  $\vartheta \in C_0^\infty(\mathcal{M})$  is chosen supported close to  $x^0 = (x_1^0, x_2^0)$  and can therefore be taken in  $C_0^\infty(\cdot - T, T[\times\omega])$ . The function  $\sigma_r \in C_0^\infty(\mathcal{M})$  is equal to 1 in a small ball centered at  $x^1 = (x_1^1, x_2^1)$  of size  $r > 0$ .

Lemma 5.2 below allows to obtain with different constants

$$\|M_\mu^{\beta\mu} \sigma_{r,\mu} u\|_1 \leq C e^{\kappa\mu} \left( \|u\|_{L^2(\cdot - T, T[\times\omega])} + \|Pu\|_{L^2(\Omega)} \right) + C e^{-\kappa'\mu} \|u\|_1.$$

Again, as in [LL15, Section 4.2], this leads, after a rough estimate of the high frequency, to

$$\|u\|_{L^2(\cdot - \varepsilon, \varepsilon[\times B(x^1, r)])} \leq C e^{\kappa\mu} \left( \|u\|_{L^2(\cdot - T, T[\times\omega])} + \|Pu\|_{L^2(\Omega)} \right) + \frac{1}{\mu} \|u\|_{H^1(\cdot - T, T[\times\mathcal{M}])}.$$

This is the equivalent to Proposition 3.7 which leads to a result similar to Corollary 3.8 by a compactness argument.

As explained above, the last step to get estimates (1.24) corresponds to the energy estimates of Step 3 Section 3.3.1. There, they relied on the hypoelliptic estimates of Corollary B.2. The equivalent in the present situation with boundary is provided by Theorem B.3.

### 5.3 An observation term in $L^2$ in quantitative unique continuation estimates

In this section, we explain how the observation term  $\|u\|_{H_b^1(\tilde{\omega})} = \sum_{|\beta| \leq 1} \|D_b^\beta u\|_{L^2(\tilde{\omega})}$  in unique continuation estimates as (2.1) can actually be replaced by the weaker norm  $\|u\|_{L^2(\tilde{\omega})}$  under suitable assumptions.

**Lemma 5.2.** *Let  $\Omega$  be a bounded open set of  $\mathbb{R}^n$  with  $n = n_a + n_b$ . Let  $P$  be a differential operator of order 2, defined in a neighborhood of  $\Omega$ , with real principal symbol and coefficients independent on the variable  $z_a$ , and being elliptic in  $\{\zeta_a = 0\}$ . Let  $\omega \Subset \Omega$  and  $\vartheta \in C_0^\infty(\omega)$ . Then, for all  $\alpha > 0$ , there exists  $C > 0$  such that for every  $u \in C_0^\infty(\mathbb{R}^n)$  and  $\mu \geq 1$ , we have*

$$\|M_{\alpha\mu}^\mu \vartheta_\mu u\|_{H^1} \leq C \langle \mu \rangle \|u\|_{L^2(\omega)} + C \|Pu\|_{L^2(\Omega)} + C e^{-c\mu} \|u\|_{H^1}.$$

Recall that the regularization process  $\vartheta \rightarrow \vartheta_\mu$  and the Fourier multiplier  $M_{\alpha\mu}^\mu$  are defined at the beginning of Section 2.2.

*Proof.* Since  $P$  is elliptic (say positive to fix the ideas) in  $\zeta_a = 0$  and  $\overline{\Omega}$  is compact, we can find  $A > 0$  (fixed for the rest of the proof) so that  $A|\zeta_a|^2 + p(z_b, \zeta_a, \zeta_b)$  is elliptic on  $\overline{\Omega} \times \mathbb{R}^{n_a+n_b}$  (where  $p$  is the principal symbol of  $P$ ), see for instance [LL15, Lemma A.1]. Using then the Gårding inequality, there exists  $C > 0$  so that

$$\begin{aligned} \|v\|_{H^1}^2 &\leq C \operatorname{Re}((A|D_a|^2 + P)v, v)_{L^2} + C \|v\|_{L^2}^2 \\ &\leq C \| |D_a|v \|_{L^2}^2 + C \operatorname{Re}(Pv, v)_{L^2} + C \|v\|_{L^2}^2 \end{aligned}$$

for every  $v \in C_0^\infty(\Omega)$ . Let  $\varphi, \chi \in C_0^\infty(\Omega)$  being real valued and such that  $\varphi = 1$  on a neighborhood of  $\operatorname{supp}(\vartheta)$  and  $\chi = 1$  on a neighborhood of  $\operatorname{supp}\varphi$ . Applying this estimate to  $v = \varphi w$  for  $w \in C_0^\infty(\mathbb{R}^n)$ , we obtain that

$$\begin{aligned} \|\varphi w\|_{H^1}^2 &\leq C \|\nabla_a(\varphi w)\|_{L^2}^2 + C \operatorname{Re}(P(\varphi w), \varphi w)_{L^2} + C \|\varphi w\|_{L^2}^2 \\ &\leq C \|\nabla_a w\|_{L^2}^2 + C \operatorname{Re}(P(\varphi w), (\varphi w))_{L^2} + C \|w\|_{L^2}^2 \\ &\leq C \|\nabla_a w\|_{L^2}^2 + C \|P(\varphi w)\|_{H^{-1}} \|\varphi w\|_{H^1} + C \|w\|_{L^2}^2. \end{aligned}$$

Writing  $P(\varphi w) = \varphi Pw + [P, \varphi]w$ , where  $[P, \varphi]$  is of order 1, we have

$$\|P(\varphi w)\|_{H^{-1}} \leq \|\varphi Pw\|_{H^{-1}} + C \|w\|_{L^2},$$

so that, after absorption, we have proved the existence of  $C > 0$  such that for every  $w \in C_0^\infty(\mathbb{R}^n)$  (and so for  $w \in \mathcal{S}(\mathbb{R}^n)$ ), we have

$$\|\varphi w\|_{H^1}^2 \leq C \|\nabla_a w\|_{L^2}^2 + C \|\varphi Pw\|_{H^{-1}}^2 + C \|w\|_{L^2}^2. \quad (5.2)$$

We apply the previous estimate to  $w = M_{\alpha\mu}^\mu \vartheta_\mu u$ . For the first term in the right handside of (5.2), we have  $\|\zeta_a m_\mu(\zeta_a/\mu)\|_{L^\infty(\mathbb{R}^{n_a})} \leq \mu \|\zeta_a m_\mu(\zeta_a)\|_{L^\infty(\mathbb{R}^{n_a})} \leq C\mu$ , so that

$$\|\nabla_a(M_{\alpha\mu}^\mu \vartheta_\mu u)\|_{L^2} \leq C \langle \mu \rangle \|\vartheta_\mu u\|_{L^2}.$$

For this term, we further use that  $\vartheta \in C_0^\infty(\omega)$ , which according to Lemma 3.24 Item 2, gives

$$\|\vartheta_\mu u\|_{L^2} \leq \|u\|_{L^2(\omega)} + Ce^{-c\mu} \|u\|_{L^2}.$$

Note that this previous inequality also rules the term  $\|w\|_{L^2}$  in the right handside of (5.2).

It only remains to estimate the term  $\|\varphi Pw\|_{H^{-1}}$  in the right handside of (5.2). Using that  $P$  is invariant on  $z_a$  (and hence commutes with  $M_{\alpha\mu}^\mu$ ), it is

$$\begin{aligned} \|\varphi P M_{\alpha\mu}^\mu \vartheta_\mu u\|_{H^{-1}} &\leq \|[P, \vartheta_\mu]u\|_{H^{-1}} + \|\varphi M_{\alpha\mu}^\mu \vartheta_\mu P u\|_{H^{-1}} \\ &\leq \|[P, \vartheta_\mu]u\|_{H^{-1}} + \|\varphi M_{\alpha\mu}^\mu \vartheta_\mu \chi P u\|_{H^{-1}} + \|\varphi M_{\alpha\mu}^\mu (1 - \chi) \vartheta_\mu P u\|_{H^{-1}}. \end{aligned}$$

Note that, *a priori*, since  $P$  is not defined on the whole  $\mathbb{R}^n$  but only in a neighborhood of  $\Omega$ , the term  $Pu$  does not have any meaning. Yet, since  $P$  is invariant in  $z_a$ , the differential operator  $\vartheta_\mu P$  is a well defined

operator on  $\mathbb{R}^n$ , so as  $\chi P$  and  $[P, \vartheta_\mu]$ . In the end, all terms involved are well defined for all  $u \in C_0^\infty(\mathbb{R}^n)$ , even if not supported inside of  $\Omega$ .

Now, using  $\vartheta \in C_0^\infty(\omega)$ , we have by (a dual version of) Lemma 3.24 that

$$\|[P, \vartheta_\mu]u\|_{H^{-1}} \leq \|u\|_{L^2(\omega)} + Ce^{-c\mu} \|u\|_{L^2}.$$

Finally, since  $\text{supp}(\varphi) \cap \text{supp}(1 - \chi) = \emptyset$ , [LL15, Lemma 2.10] also yields

$$\|\varphi M_{\alpha\mu}^\mu(1 - \chi)\|_{H^{-1} \rightarrow H^{-1}} \leq Ce^{-c\mu}.$$

Combining the last five inequalities together with (5.2), we are led, after absorption, to the estimate

$$\|\varphi M_{\alpha\mu}^\mu \vartheta_\mu u\|_{H^1} \leq C \|Pu\|_{L^2(\Omega)} + C \langle \mu \rangle \|u\|_{L^2(\omega)} + Ce^{-c\mu} \|u\|_{H^1}.$$

The property  $\text{supp}(1 - \varphi) \cap \text{supp}(\vartheta) = \emptyset$  with [LL15, Lemma 2.10] gives the similar estimate

$$\|(1 - \varphi)M_{\alpha\mu}^\mu \vartheta_\mu u\|_{H^1} \leq Ce^{-c\mu} \|u\|_{H^1},$$

which allows to conclude the proof of the lemma.  $\square$

## A On the optimality: Proof of Proposition 1.12

In this appendix, we discuss the optimality of the results presented in the main part of the paper, in the situation of Example 1.11, i.e. we give a proof of Proposition 1.12. The estimates we use are mainly extracted from the article [BCG14] by Beauchard, Cannarsa and Guglielmi. They are slightly spread out in this reference so that the proof below mainly explains where in [BCG14] to pick the results. This is mainly the proof of Theorem 5, Section 3.2 and 3.3 in this reference.

*Proof of Proposition 1.12.* First, Fourier transforming the operator  $\mathcal{L}_\gamma = -(\partial_{x_1}^2 - x_1^{2\gamma} \partial_{x_2}^2)$  in the  $x_2$  variable, we obtain a family  $A_{n,\gamma}$  of 1-dimensional operators defined for  $n \in \mathbb{Z}$  by  $(A_{n,\gamma} f)(x_1) = -f''(x_1) + (n\pi)^2 x_1^{2\gamma} f(x_1)$  on  $] -1, 1[$ , with Dirichlet boundary conditions.

The sequence of eigenfunctions  $\varphi_n$  is then taken of the form  $\varphi_n(x_1, x_2) = \sqrt{2} v_n(x_1) \sin(n\pi x_2)$  where  $v_n$  is the first normalized eigenvector of  $A_{n,\gamma}$  (see Lemma 2 of [BCG14]). We have  $A_{n,\gamma} v_n = \lambda_{n,\gamma} v_n$ , with  $\lambda_{n,\gamma}$  the lowest eigenfunction of  $A_{n,\gamma}$ , and hence  $\mathcal{L}_\gamma \varphi_n = \lambda_{n,\gamma} \varphi_n$ . Moreover,  $v_n$  is even.

The following estimates hold:

1.  $\frac{1}{C} n^{\frac{2}{1+\gamma}} \leq \lambda_{n,\gamma} \leq C n^{\frac{2}{1+\gamma}}$ , see Proposition 4 of [BCG14].
2. for  $0 < a < b < 1$ , we have  $\int_a^b v_n(x)^2 dx \leq Ce^{-C_1(\gamma)na^{\gamma+1}} c_n$  with  $c_n \leq n^\beta$  for some appropriate  $\beta$ : this is inequality (35) of [BCG14] once we have checked that for  $n$  large enough  $\mu_n = C(\gamma)n$  (written in (33)) and the definition of  $x_n$  in (26). For  $\gamma = 1$ , a more precise result is stated [BCG14, Lemma 4], where the constant is computed, namely  $\lambda_{n,\gamma} \approx n\pi$  and  $\int_a^b v_n(x)^2 dx \approx \frac{e^{-a^2 n\pi}}{2a\pi\sqrt{n}}$  for  $a > 0$ .

So, in any cases, if  $a > 0$ , there are  $C, c > 0$ , so that  $\int_a^b v_n(x)^2 dx \leq Ce^{-cn} \leq Ce^{-c\lambda_{n,\gamma}^{\frac{1+\gamma}{2}}}$  where we have used Item 1. Then, since  $\bar{\omega} \cap \{x_1 = 0\} = \emptyset$  and  $v_n$  is even, there exists  $a, b, C > 0$  so that  $\|\varphi_n\|_{L^2(\omega)}^2 \leq C \int_a^b v_n(x)^2 dx$ . To finish the proof, it is enough to notice that  $v_n$  was chosen normalized in  $L^2(]-1, 1[)$  so that  $\varphi_n$  is normalized in  $L^2(\mathcal{M})$ .  $\square$

**Remark A.1.** It is very interesting to compare the estimates obtained by [BCG14] with respect to those obtained in the present paper, even if the techniques are quite different. The scheme of proof we followed may be summarized in two steps:

1. We prove an observability estimate where the cost is, more or less, exponential of the usual Sobolev frequency. This step is performed in [BCG14] by using only the analyticity of the coefficients in the  $x_2$  variable. In their Proposition 5, they prove a 1D Carleman estimate and the cost is of the order of  $e^{cn}$  where  $n$  is the frequency in the  $x_2$  (the analytic frequency).
2. Then, we use the decay of the heat flow using hypoelliptic estimates. For this,  $\frac{1}{C}n^{\frac{2}{1+\gamma}} \leq \lambda_{n,\gamma} \leq Cn^{\frac{2}{1+\gamma}}$  may be seen as a counterpart of the hypoelliptic estimates of Theorem 1.4. Indeed, these estimates roughly say that in the worst case, the operator  $\mathcal{L}$  counts (when we want estimates from below) as  $\frac{2}{1+\gamma} = \frac{2}{k}$  when compared to the usual derivatives (that is to the usual Sobolev norms).

## B Subelliptic estimates

### B.1 $H^s$ subelliptic estimates on compact manifolds

In this appendix, we draw classical consequences of the subelliptic estimate (1.4) of Rothschild-Stein [RS76] and Fefferman-Phong [FP83], that are used in the main part of the paper. The following corollary of the subelliptic estimate (1.4) might be written elsewhere, but we did not find any reference. The short proof below stresses that the sole subelliptic estimate we rely on in the paper is (1.4).

**Corollary B.1.** *Under Assumption 1.1, for any  $s \geq 0$  there is  $C > 0$  such that we have*

$$\|u\|_{H^{s+\frac{1}{k}}(\mathcal{M})}^2 \leq C \sum_{i=1}^m \|X_i u\|_{H^s(\mathcal{M})}^2 + C \|u\|_{L^2(\mathcal{M})}^2, \quad (\text{B.1})$$

$$\|u\|_{H^{s+\frac{2}{k}}(\mathcal{M})}^2 \leq C \|\mathcal{L}u\|_{H^s(\mathcal{M})}^2 + C \|u\|_{L^2(\mathcal{M})}^2, \quad (\text{B.2})$$

for any  $u \in C^\infty(\mathcal{M})$ .

The proof we give is inspired by [FP83] (see the beginning of the proof of Theorem 1). For this, we let  $\Lambda$  be an elliptic invertible pseudodifferential operator of order one in  $\mathcal{M}$ , being selfadjoint in  $L^2(\mathcal{M})$  (see e.g. [LL16, Remark 2.11] after having endowed  $\mathcal{M}$  with a Riemannian metric). Recall that the power operator  $\Lambda^s$  is an elliptic invertible pseudodifferential operator of order  $s$  in  $\mathcal{M}$ , being also selfadjoint in  $L^2(\mathcal{M})$ . All  $H^s$  norms are equivalent to  $\|\cdot\|_{H^s(\mathcal{M})} = \|\Lambda^s \cdot\|_{L^2(\mathcal{M})}$ .

*Proof.* We start proving Estimate (B.1), which is simpler due to the fact that  $X_i$  is only of order 1 and therefore  $[X_i, \Lambda^s]$  is of order  $s$  and hence an admissible remainder term (compared to the estimated norm



$H^{s+\frac{1}{k}}$ ). Using the  $L^2$  estimate (1.4), we have

$$\begin{aligned}
\|u\|_{H^{s+\frac{1}{k}}(\mathcal{M})}^2 &\leq C \|\Lambda^s u\|_{H^{\frac{1}{k}}(\mathcal{M})}^2 \leq C \sum_{i=1}^m \|X_i \Lambda^s u\|_{L^2(\mathcal{M})}^2 + C \|\Lambda^s u\|_{L^2(\mathcal{M})}^2 \\
&\leq C \sum_{i=1}^m \|\Lambda^s X_i u\|_{L^2(\mathcal{M})}^2 + C \sum_{i=1}^m \|[\Lambda^s, X_i] u\|_{L^2(\mathcal{M})}^2 + C \|\Lambda^s u\|_{L^2(\mathcal{M})}^2 \\
&\leq C \sum_{i=1}^m \|X_i u\|_{H^s(\mathcal{M})}^2 + C \|u\|_{H^s(\mathcal{M})}^2.
\end{aligned}$$

An interpolation estimate gives  $\|u\|_{H^s(\mathcal{M})} \leq \varepsilon \|u\|_{H^{s+\frac{1}{k}}(\mathcal{M})} + C_\varepsilon \|u\|_{L^2(\mathcal{M})}$  for any  $\varepsilon > 0$ , which yields (B.1) after having taken  $\varepsilon$  sufficiently small.

Concerning Estimate (B.2), we have to be more careful since the commutator  $[\mathcal{L}, \Lambda^s]$  is of order  $s+1$  and hence not an admissible remainder term (compared to the estimated norm  $H^{s+\frac{2}{k}}$ ). Following [FP83], we apply the  $L^2$ -based estimate (1.5) to  $\Lambda^{s+\frac{1}{k}} u$ , yielding

$$\begin{aligned}
\|u\|_{H^{s+\frac{2}{k}}(\mathcal{M})}^2 &\leq C \operatorname{Re}(\mathcal{L} \Lambda^{s+\frac{1}{k}} u, \Lambda^{s+\frac{1}{k}} u)_{L^2(\mathcal{M})} + C \left\| \Lambda^{s+\frac{1}{k}} u \right\|_{L^2(\mathcal{M})}^2 \\
&\leq C \operatorname{Re}(\Lambda^{s+\frac{1}{k}} \mathcal{L} u, \Lambda^{s+\frac{1}{k}} u)_{L^2(\mathcal{M})} + C \operatorname{Re}([\mathcal{L}, \Lambda^{s+\frac{1}{k}}] u, \Lambda^{s+\frac{1}{k}} u)_{L^2(\mathcal{M})} + C \|u\|_{H^{s+\frac{1}{k}}(\mathcal{M})}^2 \\
&\leq C \operatorname{Re}(\Lambda^s \mathcal{L} u, \Lambda^{s+\frac{2}{k}} u)_{L^2(\mathcal{M})} + C \operatorname{Re}([\mathcal{L}, \Lambda^{s+\frac{1}{k}}] u, \Lambda^{s+\frac{1}{k}} u)_{L^2(\mathcal{M})} + C \|u\|_{H^{s+\frac{1}{k}}(\mathcal{M})}^2 \\
&\leq \frac{1}{2} \|u\|_{H^{s+\frac{2}{k}}(\mathcal{M})}^2 + C \|\mathcal{L} u\|_{H^s(\mathcal{M})}^2 + C \operatorname{Re}(\Lambda^{s+\frac{1}{k}} [\mathcal{L}, \Lambda^{s+\frac{1}{k}}] u, u)_{L^2(\mathcal{M})} + C \|u\|_{H^{s+\frac{1}{k}}(\mathcal{M})}^2 \tag{B.3}
\end{aligned}$$

where we have used Cauchy-Schwarz inequality in the last step. The term with the commutator has to be taken carefully since it is a priori of order  $2s + \frac{2}{k} + 1$ . But the following simple remark is in order: this pseudodifferential operator has purely imaginary principal symbol. Hence, according to pseudodifferential calculus, it can be written as  $\Lambda^{s+\frac{1}{k}} [\mathcal{L}, \Lambda^{s+\frac{1}{k}}] = T_1 + T_2$  where  $T_1$  is a *skew-adjoint* pseudodifferential operator of order  $2s + \frac{2}{k} + 1$ , and  $T_2$  is a pseudodifferential operator of order  $2s + \frac{2}{k}$ . In particular, we have

$$\left| \operatorname{Re}(\Lambda^{s+\frac{1}{k}} [\mathcal{L}, \Lambda^{s+\frac{1}{k}}] u, u)_{L^2(\mathcal{M})} \right| = \left| \operatorname{Re}(T_2 u, u)_{L^2(\mathcal{M})} \right| \leq \|u\|_{H^{s+\frac{1}{k}}(\mathcal{M})}^2.$$

So, at this stage, we have proved

$$\|u\|_{H^{s+\frac{2}{k}}(\mathcal{M})}^2 \leq C \|\mathcal{L} u\|_{H^s(\mathcal{M})}^2 + C \|u\|_{H^{s+\frac{1}{k}}(\mathcal{M})}^2.$$

This concludes the proof of (B.2), after an interpolation argument as above.  $\square$

To conclude this section, we prove the continuous injection  $\mathcal{H}_{\mathcal{L}}^s \subset H^{s/k}(\mathcal{M})$  for all  $s \geq 0$ . We shall use the following classical operator theoretic (interpolation) result for which we refer e.g. to [SSV12, Corollary 12.15]. Given two selfadjoint nonnegative operators  $(A, D(A))$  and  $(B, D(B))$  on a Hilbert space, we have

$$\|Au\| \leq \|Bu\| \text{ for all } u \in D(B) \implies \|A^\alpha u\| \leq \|B^\alpha u\| \text{ for all } \alpha \in [0, 1] \text{ and } u \in D(B^\alpha). \tag{B.4}$$

Note that this result already yields the simple inequality: for  $s \geq 0$ , there is  $C > 0$  such that for all  $u \in H^s(\mathcal{M})$ ,

$$\|u\|_{\mathcal{H}_{\mathcal{L}}^s} \leq C \|u\|_{H^s(\mathcal{M})},$$

consequence of that obtained for  $s \in 2\mathbb{N}$ . It also yields by duality for all  $s \geq 0$  the existence of  $C > 0$  such that for all  $u \in \mathcal{H}_{\mathcal{L}}^{-s}$ ,

$$\|u\|_{H^{-s}(\mathcal{M})} \leq C \|u\|_{\mathcal{H}_{\mathcal{L}}^{-s}}. \quad (\text{B.5})$$

**Corollary B.2.** *For all  $s \geq 0$ , there exists  $C > 0$  such that for all  $u \in \mathcal{H}_{\mathcal{L}}^s$ , we have*

$$\|u\|_{H^{\frac{s}{k}}(\mathcal{M})} \leq C \|u\|_{\mathcal{H}_{\mathcal{L}}^s}.$$

Note that it also yields by duality for all  $s \geq 0$  the existence of  $C > 0$  such that for all  $u \in H^{-\frac{s}{k}}(\mathcal{M})$ ,

$$\|u\|_{\mathcal{H}_{\mathcal{L}}^{-s}} \leq C \|u\|_{H^{-\frac{s}{k}}(\mathcal{M})}.$$

*Proof.* We first prove the result for  $s = 2p$ ,  $p \in \mathbb{N}$ , and then conclude by interpolation. We prove by induction that for all  $p \in \mathbb{N}$ , we have

$$\|u\|_{H^{\frac{2p}{k}}(\mathcal{M})} \leq C \|u\|_{\mathcal{H}_{\mathcal{L}}^{2p}} = C \|(\mathcal{L} + 1)^p u\|_{L^2(\mathcal{M})}, \quad \text{for all } u \in C^\infty(\mathcal{M}). \quad (\text{B.6})$$

The case  $p = 0$  is clear. Assume now that this is satisfied for  $p$ , and estimate  $\|u\|_{H^{\frac{2(p+1)}{k}}(\mathcal{M})} = \|u\|_{H^{\frac{2p}{k} + \frac{2}{k}}(\mathcal{M})}$ . After having used (B.2), we have

$$\|u\|_{H^{\frac{2(p+1)}{k}}(\mathcal{M})}^2 \leq C \|\mathcal{L}u\|_{H^{\frac{2p}{k}}(\mathcal{M})}^2 + C \|u\|_{L^2(\mathcal{M})}^2,$$

which, using the induction assumption (B.6) to  $\mathcal{L}u$ , yields

$$\|u\|_{H^{\frac{2(p+1)}{k}}(\mathcal{M})}^2 \leq C \|(\mathcal{L} + 1)^p \mathcal{L}u\|_{L^2(\mathcal{M})}^2 + C \|u\|_{L^2(\mathcal{M})}^2.$$

Using the functional calculus (1.10), this implies

$$\|u\|_{H^{\frac{2(p+1)}{k}}(\mathcal{M})}^2 \leq C \|(\mathcal{L} + 1)^{p+1} u\|_{L^2(\mathcal{M})}^2 = C \|u\|_{\mathcal{H}_{\mathcal{L}}^{2(p+1)}}^2, \quad \text{for all } u \in C^\infty(\mathcal{M}),$$

which is the sought estimate.

Now for  $s \geq 0$ ,  $s \notin \mathbb{N}$ , pick  $p \in \mathbb{N}$  such that  $s \in [0, p]$ , write (B.6) as  $\|\Lambda^{\frac{2p}{k}} u\|_{L^2(\mathcal{M})} \leq C \|(\mathcal{L} + 1)^p u\|_{L^2(\mathcal{M})}$  and apply (B.4) to  $A = \Lambda^{\frac{2p}{k}}$ ,  $B = (\mathcal{L} + 1)^p$ , and  $\alpha = \frac{s}{p} \in [0, 1]$  to obtain the result.  $\square$

## B.2 Subelliptic estimates for manifolds with boundaries

In this section, we assume that  $\mathcal{M}$  is a compact manifold with a nonempty boundary  $\partial\mathcal{M}$ , and write  $\mathcal{M} = \text{Int}(\mathcal{M}) \cup \partial\mathcal{M}$ , with a disjoint union. We assume that the coefficients of  $X_j$ 's are smooth up to the boundary, and that  $\text{span}(X_1, \dots, X_m)(x) = T_x\mathcal{M}$  for  $x \in \mathcal{M} \setminus K$ , where  $K$  is a compact subset of  $\text{Int}(\mathcal{M})$  (i.e. the operator  $\mathcal{L}$  is elliptic in the neighborhood of the boundary) and that Assumption 1.1 is satisfied on  $K$ .

**Theorem B.3.** Denote by  $\Delta_D$  the Laplace-Dirichlet operator associated to some/any Riemannian metric equal to that issued from the vector fields  $(X_1, \dots, X_m)$  in  $\mathcal{M} \setminus K$ . Then, for all  $s \geq 0$ , we have  $\mathcal{H}_{\mathcal{L}}^s := D(\mathcal{L}^{\frac{s}{2}}) \subset D((-\Delta_D)^{\frac{s}{2k}})$ . Moreover, there exists  $C > 0$  such that for all  $u \in \mathcal{H}_{\mathcal{L}}^s$ , we have

$$\left\| (-\Delta_D)^{\frac{s}{2k}} u \right\|_{L^2(\mathcal{M})} \leq C \|u\|_{\mathcal{H}_{\mathcal{L}}^s}.$$

Note that the space  $D((-\Delta_D)^s)$  does not depend on the metric chosen to define  $\Delta$  inside  $\text{Int}(\mathcal{M})$  but only on its values in a neighborhood of  $\partial\mathcal{M}$ . Remark that we also have the converse simple inclusion  $D((-\Delta_D)^s) \subset D(\mathcal{L}^s)$ .

We now explain how the estimates in the previous section have to be modified to yield the statement of Theorem B.3.

We first let  $\Lambda^{\frac{1}{k}}$  (with a slight abuse of notation:  $\Lambda^{\frac{1}{k}}$  is not the  $\frac{1}{k}$  power of an operator) be

- a pseudodifferential operator of order  $\frac{1}{k}$  on  $\mathcal{M}$ , with kernel compactly supported in  $\mathcal{M} \times \mathcal{M}$ ,
- formally selfadjoint on  $L^2(\mathcal{M})$ ,
- locally elliptic on a neighborhood  $\mathcal{N}$  of  $K$  (i.e.  $K \Subset \text{Int}(\mathcal{N}) \Subset \text{Int}(\mathcal{M})$ ),
- with kernel compactly supported in an  $\varepsilon$ -neighborhood of  $\text{diag}(\mathcal{M} \times \mathcal{M})$ , with  $\varepsilon \ll \text{dist}(K, \mathcal{N}^c)$ .

This implies that for  $n \in \mathbb{N}$ ,  $(\Lambda^{\frac{1}{k}})^n$  has a kernel compactly supported in an  $n\varepsilon$ -neighborhood of  $\text{diag}(\mathcal{M} \times \mathcal{M})$ . Here, we will use the abuse of notation  $\Lambda^{\frac{n}{k}}$  instead of  $(\Lambda^{\frac{1}{k}})^n$ . This will not lead to any confusion since we will only use  $\Lambda^s$  for  $s$  of the form  $\frac{n}{k}$  with  $n \in \mathbb{N}$ .

When a maximal Sobolev exponent  $s_0$  is fixed, we will need to use such operators for  $n \leq n_0$  with  $n_0 = s_0 k$  and make proofs by induction using several cutoff functions. At each step, we shall need to make some estimates on some  $C\varepsilon$  neighborhood of the zone where we get the information, with  $C$  depending on  $s_0$  and some geometric properties of the cutoff functions. At the end, once the number of steps is fixed, we can select  $\varepsilon$  small enough (and the associated  $\Lambda^{\frac{1}{k}}$ ) so that all the reasoning is valid. To make the presentation more readable, we have chosen not to keep track of all the constants and the geometrical conditions involved. Yet, the proof will make it clear that there is  $\varepsilon_0 > 0$  depending on  $s_0, \mathcal{N}, K$  and  $\mathcal{M}$  so that all the support conditions of the following proof are fulfilled if  $0 < \varepsilon < \varepsilon_0$ .

That Assumption 1.1 is satisfied on  $K$  (and hence on  $\mathcal{N}$ , since  $\mathcal{L}$  is elliptic on  $\mathcal{M} \setminus K$ ) yields, for all  $\chi \in C_0^\infty(\mathcal{M})$  such that  $\chi = 1$  in a neighborhood of  $K$ , the existence of  $C > 0$  such that for all  $u \in C_0^\infty(\mathcal{M})$ , we have (see again [RS76] Theorem 17 and estimate (17.20) p311)

$$\left\| \Lambda^{\frac{1}{k}} \chi u \right\|_{L^2(\mathcal{M})}^2 \leq C \sum_{i=1}^m \|X_i u\|_{L^2(\mathcal{M})}^2 + C \|u\|_{L^2(\mathcal{M})}^2,$$

and hence, still for  $u \in C_0^\infty(\mathcal{M})$ ,

$$\left\| \Lambda^{\frac{1}{k}} \chi u \right\|_{L^2(\mathcal{M})}^2 \leq C(\mathcal{L}u, u)_{L^2(\mathcal{M})} + C \|u\|_{L^2(\mathcal{M})}^2. \quad (\text{B.7})$$

We now decompose the proof in several lemmata.

Several times in the proof, we shall use the following fact of pseudodifferential calculus. Given  $n \leq n_0$  and a function  $\varphi \in C_0^\infty(\mathcal{N})$ , we remark that  $\Lambda^{\frac{n}{k}}$  is elliptic of order  $\frac{n}{k}$  in a neighborhood of  $\text{supp}(\varphi)$ . As

a consequence, the classical parametrix construction (see for instance [Hör85, Proof of Theorem 18.1.9]) allows, for any  $N \in \mathbb{N}$ , to construct a pseudodifferential operator  $\tilde{\Lambda}_N^{-\frac{n}{k}}$  of order  $-\frac{n}{k}$ , elliptic on a neighborhood of  $\text{supp}(\varphi)$ , such that  $\tilde{\Lambda}_N^{-\frac{n}{k}} \Lambda^{\frac{n}{k}} = \varphi(x) + R_N^1$  and  $\Lambda^{\frac{n}{k}} \tilde{\Lambda}_N^{-\frac{n}{k}} = \varphi(x) + R_N^2$  with  $R_N^i$ ,  $i = 1, 2$ , pseudodifferential operators of order  $-N$  with kernel compactly supported in  $\mathcal{M} \times \mathcal{M}$ . In the applications,  $N$  will always be fixed, sufficiently large.

**Lemma B.4.** *For all  $\chi_0 \in C_0^\infty(\mathcal{N})$  such that  $\chi_0 = 1$  in a neighborhood of  $K$ , and all  $\chi_1 \in C_0^\infty(\mathcal{N})$  such that  $\chi_1 = 1$  in a neighborhood of  $\text{supp}(\chi_0)$ , for all  $s \in \mathbb{N}/k$  with  $s + 2/k \leq s_0$ , for  $\varepsilon$  small enough, there is  $C > 0$  such that for all  $u \in C_0^\infty(\mathcal{M})$ , we have*

$$\left\| \Lambda^{s+\frac{2}{k}} \chi_0 u \right\|_{L^2(\mathcal{M})}^2 \leq C \|\Lambda^s \chi_0 \mathcal{L}u\|_{L^2(\mathcal{M})}^2 + C \left\| \Lambda^{s+\frac{1}{k}} \chi_1 u \right\|_{L^2(\mathcal{M})}^2 + \|\chi_1 u\|_{L^2(\mathcal{M})}^2$$

*Proof.* The proof is almost the same as that of Estimate (B.2) in Corollary B.1, and only relies on the application of (B.7) (instead of (1.5)). First, estimate (B.7) applies for  $\chi u$  replaced by  $\Lambda^{s+\frac{1}{k}} \chi_0 u$  since  $\Lambda^{s+\frac{1}{k}} \chi_0 u = \tilde{\chi}_0 \Lambda^{s+\frac{1}{k}} \chi_0 u$  for  $\tilde{\chi}_0 = 1$  in an  $n\varepsilon$ -neighborhood of  $\text{supp}(\chi_0)$ , with  $n = sk + 1$ . This yields

$$\left\| \Lambda^{s+\frac{2}{k}} \chi_0 u \right\|_{L^2(\mathcal{M})}^2 \leq C \left( \mathcal{L} \Lambda^{s+\frac{1}{k}} \chi_0 u, \Lambda^{s+\frac{1}{k}} \chi_0 u \right)_{L^2(\mathcal{M})} + C \left\| \Lambda^{s+\frac{1}{k}} \chi_0 u \right\|_{L^2(\mathcal{M})}^2.$$

Then, a computation similar to (B.3), and the only difference comes from the estimate of the remainder term

$$\text{Re}(\chi_0 \Lambda^{s+\frac{1}{k}} [\mathcal{L}, \Lambda^{s+\frac{1}{k}} \chi_0] u, u)_{L^2(\mathcal{M})} = \text{Re}((T_1 + T_2)u, u)_{L^2(\mathcal{M})} = \text{Re}(T_2 u, u)_{L^2(\mathcal{M})},$$

where  $T_2$  is a pseudodifferential operator of order  $2s + \frac{2}{k}$ , with kernel supported in  $\text{supp}(\chi_0) \times \text{supp}(\chi_0)$ . Given  $\varphi \in C_0^\infty(\mathcal{N})$  such that  $\varphi = 1$  on  $\text{supp}(\chi_1)$ , we may define the associated parametrix  $\tilde{\Lambda}^{-(s+\frac{1}{k})}$  of  $\Lambda^{s+\frac{1}{k}}$  as above. Writing  $T_2 = \chi_1 \varphi T_2 \varphi \chi_1$ , we now have  $T_2 = \chi_1 \Lambda^{s+\frac{1}{k}} \tilde{\Lambda}^{-(s+\frac{1}{k})} T_2 \tilde{\Lambda}^{-(s+\frac{1}{k})} \Lambda^{s+\frac{1}{k}} \chi_1 + R$ , where  $R$  is a smoothing operator with kernel compactly supported in  $\mathcal{M} \times \mathcal{M}$ . The boundedness of  $\tilde{\Lambda}^{-(s+\frac{1}{k})} T_2 \tilde{\Lambda}^{-(s+\frac{1}{k})}$  (as a pseudodifferential operator of order zero) and  $R$  on  $L^2(\mathcal{M})$  then implies

$$|\text{Re}(T_2 u, u)_{L^2(\mathcal{M})}| = |\text{Re}(T_2 \chi_1 u, \chi_1 u)_{L^2(\mathcal{M})}| \leq C \left\| \Lambda^{s+\frac{1}{k}} \chi_1 u \right\|_{L^2(\mathcal{M})}^2 + C \|\chi_1 u\|_{L^2(\mathcal{M})}^2,$$

which concludes the proof. Note also that the term  $\left\| \Lambda^{s+\frac{1}{k}} \chi_0 u \right\|_{L^2(\mathcal{M})}^2$  has been bounded by  $\left\| \Lambda^{s+\frac{1}{k}} \chi_1 u \right\|_{L^2(\mathcal{M})}^2$  the same way using that  $\Lambda^{s+\frac{1}{k}} \chi_0 = \Lambda^{s+\frac{1}{k}} \chi_0 \varphi \chi_1 = \Lambda^{s+\frac{1}{k}} \chi_0 \tilde{\Lambda}_1^{-(s+\frac{1}{k})} \Lambda^{s+\frac{1}{k}} \chi_1 + R \chi_1$  with  $R$  smoothing.  $\square$

Before going further, recall that we can localize (B.7) under the following form.

**Lemma B.5.** *For all  $\chi_0 \in C_0^\infty(\mathcal{N})$  such that  $\chi_0 = 1$  in a neighborhood of  $K$ , and all  $\chi_1 \in C_0^\infty(\mathcal{N})$  such that  $\chi_1 = 1$  in a neighborhood of  $\text{supp}(\chi_0)$ , there is  $C > 0$  such that for all  $u \in C_0^\infty(\mathcal{M})$ , we have*

$$\left\| \Lambda^{\frac{1}{k}} \chi_0 u \right\|_{L^2(\mathcal{M})}^2 \leq C (\chi_0 \mathcal{L}u, \chi_0 u)_{L^2(\mathcal{M})} + C \|\chi_1 u\|_{L^2(\mathcal{M})}^2$$

*Proof.* We apply Estimate (B.7) with  $\chi$  such that  $\chi = 1$  on a neighborhood of  $\text{supp}(\chi_0)$ :

$$\begin{aligned} \left\| \Lambda^{\frac{1}{k}} \chi_0 u \right\|_{L^2(\mathcal{M})}^2 &= \left\| \Lambda^{\frac{1}{k}} \chi \chi_0 u \right\|_{L^2(\mathcal{M})}^2 \leq C (\mathcal{L} \chi_0 u, \chi_0 u)_{L^2(\mathcal{M})} + C \|\chi_0 u\|_{L^2(\mathcal{M})}^2 \\ &\leq C (\chi_0 \mathcal{L}u, \chi_0 u)_{L^2(\mathcal{M})} + \text{Re}([\mathcal{L}, \chi_0]u, \chi_0 u)_{L^2(\mathcal{M})} + C \|\chi_0 u\|_{L^2(\mathcal{M})}^2, \end{aligned}$$

where  $[\mathcal{L}, \chi_0]$  is a skew-adjoint first order differential operator: the principal part of  $\chi_0[\mathcal{L}, \chi_0]$  is thus skew-adjoint, and hence

$$\operatorname{Re}([\mathcal{L}, \chi_0]u, \chi_0 u)_{L^2(\mathcal{M})} = \operatorname{Re}(\chi_0[\mathcal{L}, \chi_0]\chi_1 u, \chi_1 u)_{L^2(\mathcal{M})} \leq C \|\chi_1 u\|_{L^2(\mathcal{M})}^2,$$

which, together with the preceding estimate, proves the lemma.  $\square$

**Lemma B.6.** *For all  $\chi_0, \chi_1, \chi_2 \in C_0^\infty(\mathcal{N})$  such that  $\chi_0 = 1$  in a neighborhood of  $K$ ,  $\chi_1 = 1$  in a neighborhood of  $\operatorname{supp}(\chi_0)$ , and  $\chi_2 = 1$  in a neighborhood of  $\operatorname{supp}(\chi_1)$ , for all  $p \in \mathbb{N}$  with  $\frac{p}{k} + \frac{2}{k} \leq s_0$  and  $\varepsilon$  small enough, there is  $C > 0$  such that for all  $u \in C_0^\infty(\mathcal{M})$ , we have*

$$\left\| \Lambda^{\frac{p}{k} + \frac{2}{k}} \chi_0 u \right\|_{L^2(\mathcal{M})}^2 \leq C \left\| \Lambda^{\frac{p}{k}} \chi_1 \mathcal{L}u \right\|_{L^2(\mathcal{M})}^2 + C \|\chi_2 u\|_{L^2(\mathcal{M})}^2.$$

*Proof.* We prove the statement by induction. The case  $p = 0$  follows directly from the estimate of Lemma B.4 with  $s = 0$ , combined with Lemma B.5 (using an additional cutoff function as done below). Assume now this is true for  $p$ , then, the estimate of Lemma B.4 with  $s = \frac{p+1}{k}$  yields, for some  $\tilde{\chi}_1$  such that  $\tilde{\chi}_1 = 1$  on a neighborhood of  $\operatorname{supp}(\chi_0)$  and  $\chi_1 = 1$  in a neighborhood of  $\operatorname{supp}(\tilde{\chi}_1)$ ,

$$\left\| \Lambda^{\frac{p+1}{k} + \frac{2}{k}} \chi_0 u \right\|_{L^2(\mathcal{M})}^2 \leq C \|\Lambda^{\frac{p+1}{k}} \chi_0 \mathcal{L}u\|_{L^2(\mathcal{M})}^2 + C \left\| \Lambda^{\frac{p+1}{k} + \frac{1}{k}} \tilde{\chi}_1 u \right\|_{L^2(\mathcal{M})}^2 + C \|\tilde{\chi}_1 u\|_{L^2(\mathcal{M})}^2.$$

Using then the induction assumption for  $p$  for the term  $\Lambda^{\frac{p}{k} + \frac{2}{k}} \tilde{\chi}_1 u$  gives, since  $\chi_1 = 1$  in a neighborhood of  $\operatorname{supp}(\tilde{\chi}_1)$ ,

$$\left\| \Lambda^{\frac{p+1}{k} + \frac{2}{k}} \chi_0 u \right\|_{L^2(\mathcal{M})}^2 \leq C \|\Lambda^{\frac{p+1}{k}} \chi_0 \mathcal{L}u\|_{L^2(\mathcal{M})}^2 + C \|\Lambda^{\frac{p}{k}} \chi_1 \mathcal{L}u\|_{L^2(\mathcal{M})}^2 + C \|\chi_1 u\|_{L^2(\mathcal{M})}^2. \quad (\text{B.8})$$

We now use pseudodifferential calculus and the parametrices of  $\Lambda^{\frac{p+1}{k}}$  and  $\Lambda^{\frac{p}{k}}$  to write, for  $\varphi = 1$  on  $\operatorname{supp}(\chi_1)$ ,

$$\begin{aligned} \Lambda^{\frac{p+1}{k}} \chi_0 &= \Lambda^{\frac{p+1}{k}} \chi_0 \varphi \chi_1 = \Lambda^{\frac{p+1}{k}} \chi_0 \tilde{\Lambda}^{-\frac{p+1}{k}} \Lambda^{\frac{p+1}{k}} \chi_1 + R\chi_1, \\ \Lambda^{\frac{p}{k}} \chi_1 &= \Lambda^{\frac{p}{k}} \varphi \chi_1 = \Lambda^{\frac{p}{k}} \tilde{\Lambda}^{-\frac{p+1}{k}} \Lambda^{\frac{p+1}{k}} \chi_1 + R\chi_1, \end{aligned}$$

and hence

$$\begin{aligned} \|\Lambda^{\frac{p+1}{k}} \chi_0 \mathcal{L}u\|_{L^2(\mathcal{M})} &\leq C \|\Lambda^{\frac{p+1}{k}} \chi_1 \mathcal{L}u\|_{L^2(\mathcal{M})} + C \|R\chi_1 \mathcal{L}u\|_{L^2(\mathcal{M})} \leq C \|\Lambda^{\frac{p+1}{k}} \chi_1 \mathcal{L}u\|_{L^2(\mathcal{M})} + C \|\chi_2 u\|_{L^2(\mathcal{M})}, \\ \|\Lambda^{\frac{p}{k}} \chi_1 \mathcal{L}u\|_{L^2(\mathcal{M})} &\leq C \|\Lambda^{\frac{p+1}{k}} \chi_1 \mathcal{L}u\|_{L^2(\mathcal{M})} + C \|\chi_2 u\|_{L^2(\mathcal{M})}, \end{aligned}$$

which, combined with (B.8) concludes the proof of the statement for  $p + 1$ , and hence of the lemma.  $\square$

**Lemma B.7.** *For all  $\chi_0 \in C_0^\infty(\mathcal{N})$  such that  $\chi_0 = 1$  in a neighborhood of  $K$ , and all  $\chi_1 \in C_0^\infty(\mathcal{M})$  such that  $\chi_1 = 1$  in a neighborhood of  $\operatorname{supp}(\chi_0)$ , for all  $p \in \mathbb{N}$  and  $\varepsilon$  small enough, there is  $C > 0$  such that for all  $u \in C_0^\infty(\mathcal{M})$ , we have*

$$\left\| \Lambda^{\frac{2p}{k}} \chi_0 u \right\|_{L^2(\mathcal{M})}^2 \leq C \sum_{j=0}^p \|\chi_1 \mathcal{L}^j u\|_{L^2(\mathcal{M})}^2.$$

*Proof.* Again, we prove this by an induction argument. For  $p = 1$ , this is the estimate of Lemma B.6 with  $p = 0$ .

Assume the result for  $p$ . Using Lemma B.6, we obtain, for some  $\tilde{\chi}_1, \tilde{\chi}_2$  such that  $\tilde{\chi}_1 = 1$  on a neighborhood of  $\text{supp}(\chi_0)$ ,  $\tilde{\chi}_2 = 1$  on a neighborhood of  $\text{supp}(\tilde{\chi}_1)$  and  $\chi_1 = 1$  in a neighborhood of  $\text{supp}(\tilde{\chi}_2)$ ,

$$\left\| \Lambda^{\frac{2p+2}{k}} \chi_0 u \right\|_{L^2(\mathcal{M})}^2 \leq C \left\| \Lambda^{\frac{2p}{k}} \tilde{\chi}_1 \mathcal{L} u \right\|_{L^2(\mathcal{M})}^2 + C \|\tilde{\chi}_2 u\|_{L^2(\mathcal{M})}^2.$$

Applying the induction assumption with  $u$  replaced by  $\mathcal{L} u$  (and  $\chi_0$  replaced by  $\tilde{\chi}_1$ ) yields (since  $\chi_1 = 1$  in a neighborhood of  $\text{supp}(\tilde{\chi}_1)$ ),

$$\left\| \Lambda^{\frac{2p+2}{k}} \chi_0 u \right\|_{L^2(\mathcal{M})}^2 \leq C \sum_{j=0}^p \|\chi_1 \mathcal{L}^j(\mathcal{L} u)\|_{L^2(\mathcal{M})}^2 + C \|\tilde{\chi}_2 u\|_{L^2(\mathcal{M})}^2,$$

which concludes the proof of the lemma since  $\chi_1 = 1$  in a neighborhood of  $\text{supp}(\tilde{\chi}_2)$ .  $\square$

Combining Lemma B.7 together with classical ellipticity at the boundary, we are now ready for proving the following results.

**Proposition B.8.** *For all  $m \in \mathbb{N}$ , there is  $C > 0$  such that for all  $u \in D(\mathcal{L}^{mk})$ , we have*

$$\|u\|_{H^{2m}(\mathcal{M})}^2 \leq C \sum_{j=0}^{mk} \|\mathcal{L}^j u\|_{L^2(\mathcal{M})}^2 \leq C \|(\mathcal{L} + 1)^{mk} u\|_{L^2(\mathcal{M})}.$$

From this proposition, we directly obtain by interpolation the statement of Theorem B.3 (see e.g. the proof of Corollary B.2) using that  $\mathcal{L}^j u|_{\partial\mathcal{M}} = \Delta u|_{\partial\mathcal{M}}$  for all  $u \in C^\infty(\mathcal{M})$ .

*Proof of Proposition B.8.* First write Lemma B.7 with  $p = mk$ ,  $\chi_0 \in C_0^\infty(\mathcal{N})$  with  $\chi_0 = 1$  in a neighborhood  $\tilde{\mathcal{N}} \subset \mathcal{N}$  of  $K$ , yielding

$$\|u\|_{H^{2m}(\tilde{\mathcal{N}})}^2 \leq \|\Lambda^{2m} \chi_0 u\|_{L^2(\mathcal{M})}^2 \leq C \sum_{j=0}^{mk} \|\chi_1 \mathcal{L}^j u\|_{L^2(\mathcal{M})}^2. \quad (\text{B.9})$$

Then, concerning estimates near the boundary, we first have the following statement. For all  $\theta_0 \in C^\infty(\mathcal{M})$  such that  $\theta_0 = 1$  in a neighborhood of  $\partial\mathcal{M}$ ,  $\text{supp}(\theta_0) \cap K = \emptyset$ , and all  $\theta_1 \in C^\infty(\mathcal{M})$  such that  $\theta_1 = 1$  in a neighborhood of  $\text{supp}(\theta_0)$ , since  $\mathcal{L}$  elliptic in  $\text{supp}(\theta_1)$ , there is  $C > 0$  such that for all  $u \in C^\infty(\mathcal{M})$  with  $u|_{\partial\mathcal{M}} = 0$ , we have

$$\|\theta_0 u\|_{H^{m+2}(\mathcal{M})}^2 \leq C \|\theta_1 \mathcal{L} u\|_{H^m(\mathcal{M})}^2 + C \|\theta_1 u\|_{L^2(\mathcal{M})}^2. \quad (\text{B.10})$$

This is actually a corollary of the usual proof of elliptic regularity up to the boundary. Yet, to check it directly, we can apply the global elliptic regularity result (see [Eva98, Theorem 5 p323]) to  $\theta_0 u$  with a global elliptic operator  $\tilde{\mathcal{L}}$  equal to  $\mathcal{L}$  on  $\text{supp}(\theta_1)$ . Let  $V$  be an open subset with  $\text{supp}(\nabla\theta_1) \Subset V \Subset U \Subset \mathcal{M}$ , where  $U$  in an open set so that  $\theta_1 = 1$  on  $U$ . This yields

$$\begin{aligned} \|\theta_0 u\|_{H^{m+2}(\mathcal{M})} &\leq C \|\theta_0 \tilde{\mathcal{L}} u\|_{H^m(\mathcal{M})} + C \|u\|_{H^{m+1}(V)} + C \|\theta_0 u\|_{L^2(\mathcal{M})} \\ &\leq C \|\theta_1 \mathcal{L} u\|_{H^m(\mathcal{M})} + C \|\theta_1 u\|_{L^2(\mathcal{M})}. \end{aligned}$$

where we have used interior regularity (see [Eva98, Theorem 2 p314]) that gives  $\|u\|_{H^{m+1}(V)} \leq C\|\tilde{\mathcal{L}}u\|_{H^m(U)} + C\|u\|_{L^2(U)} \leq C\|\theta_1\mathcal{L}u\|_{H^m(\mathcal{M})} + C\|\theta_1u\|_{L^2(\mathcal{M})}$ . This proves (B.10).

Then, from (B.10), another induction argument as in the proof of Lemma B.7 gives, for all  $m \in \mathbb{N}$  and  $\theta_0, \theta_1$  as above, the existence of  $C > 0$  such that for all  $u \in C^\infty(\mathcal{M})$  with  $u|_{\partial\mathcal{M}} = \mathcal{L}u|_{\partial\mathcal{M}} = \dots = \mathcal{L}^{m-1}u|_{\partial\mathcal{M}} = 0$ ,

$$\|\theta_0u\|_{H^{2m}(\mathcal{M})}^2 \leq C \sum_{j=0}^m \|\theta_1\mathcal{L}^j u\|_{L^2(\mathcal{M})}^2.$$

Combining this for a function  $\theta_0$  equal to 1 on a neighborhood of  $\mathcal{M} \setminus \tilde{\mathcal{N}}$  together with (B.9) now implies for all  $u \in C^\infty(\mathcal{M})$  such that  $\mathcal{L}^j u|_{\partial\mathcal{M}} = 0$ , for  $0 \leq j \leq m-1$ ,

$$\|u\|_{H^{2m}(\mathcal{M})}^2 \leq C \sum_{j=0}^{mk} \|\chi_1\mathcal{L}^j u\|_{L^2(\mathcal{M})}^2 + C \sum_{j=0}^m \|\theta_1\mathcal{L}^j u\|_{L^2(\mathcal{M})}^2.$$

Since the set of such functions  $u$  is dense in  $D(\mathcal{L}^{mk})$ , this yields the sought result.  $\square$

## C Sub-Riemannian norm of normal vectors

**Lemma C.1.** *Let  $X_i \in \mathbb{R}^d$  for  $i = 1, \dots, m$  and, for  $v \in \text{span}(X_i, i = 1, \dots, m)$ , set*

$$g(v) = \inf \left\{ \sum_{i=1}^m u_i^2, \quad (u_1, \dots, u_m) \in \mathbb{R}^m, \sum_{i=1}^m u_i X_i = v \right\}.$$

*Then, for any  $\xi \in (\mathbb{R}^d)^*$ , and for  $v_0 = \sum_{i=1}^m 2\langle \xi, X_i \rangle X_i$ , we have  $g(v_0) = 4\ell(\xi)$  where  $\ell(\xi) = \sum_{i=1}^m \langle \xi, X_i \rangle^2$ .*

Note that this is clear if the family  $(X_i)_{i=1, \dots, m}$  is linearly independent.

*Proof.* We want to compute the minimum

$$g(v_0) = \inf \left\{ \sum_{i=1}^m u_i^2, \quad (u_1, \dots, u_m) \in \mathbb{R}^m, \sum_{i=1}^m u_i X_i = \sum_{i=1}^m 2\langle \xi, X_i \rangle X_i \right\}.$$

First that taking  $u_i = 2\langle \xi, X_i \rangle$  in this definition directly yields that

$$g(v_0) \leq \sum_{i=1}^m (2\langle \xi, X_i \rangle)^2 = 4\ell(\xi). \tag{C.1}$$

Then it only remains to prove the converse inequality. To this aim, remark that

$$\ell(\xi) = \max \left\{ \sum_{i=1}^m u_i \langle \xi, X_i \rangle - \frac{1}{4} \sum_{i=1}^m u_i^2, \quad (u_1, \dots, u_m) \in \mathbb{R}^m \right\}.$$

As a consequence, we have

$$\ell(\xi) \geq \left\langle \xi, \sum_{i=1}^m u_i X_i \right\rangle - \frac{1}{4} \sum_{i=1}^m u_i^2, \quad \text{for all } (u_1, \dots, u_m) \in \mathbb{R}^m.$$

Hence, for all  $(u_1, \dots, u_m) \in \mathbb{R}^m$  such that  $\sum_{i=1}^m u_i X_i = \sum_{i=1}^m 2\langle \xi, X_i \rangle X_i$ , we obtain

$$\ell(\xi) \geq \left\langle \xi, \sum_{i=1}^m 2\langle \xi, X_i \rangle X_i \right\rangle - \frac{1}{4} \sum_{i=1}^m u_i^2 = \sum_{i=1}^m 2\langle \xi, X_i \rangle^2 - \frac{1}{4} \sum_{i=1}^m u_i^2 = 2\ell(\xi) - \frac{1}{4} \sum_{i=1}^m u_i^2,$$

that is  $4\ell(\xi) \leq \sum_{i=1}^m u_i^2$ , and hence  $4\ell(\xi) \leq g(v_0)$ . This, together with (C.1) concludes the proof of the lemma.  $\square$

## References

- [ABB16a] Andrei Agrachev, Davide Barilari, and Ugo Boscain. Introduction to geodesics in sub-Riemannian geometry. In *Geometry, Analysis and Dynamics in SR manifolds*, volume 2 of *EMS Lecture Notes series*, page 85pp. EMS, 2016.
- [ABB16b] Andrei Agrachev, Davide Barilari, and Ugo Boscain. *Introduction to Riemannian and sub-Riemannian geometry*. Lecture notes. unpublished, <https://webusers.imj-prg.fr/~d Davide.Barilari/2016-11-21-ABB.pdf>, 2016.
- [Bah86] Hajer Bahouri. Non prolongement unique des solutions d’opérateurs “somme de carrés”. *Ann. Inst. Fourier (Grenoble)*, 36(4):137–155, 1986.
- [BC17] Karine Beauchard and Piermarco Cannarsa. Heat equation on the Heisenberg group: Observability and applications. *J. Differential Equations*, 262(8):4475–4521, 2017.
- [BCG14] Karine Beauchard, Piermarco Cannarsa, and Roberto Guglielmi. Null controllability of Grushin-type operators in dimension two. *J. Eur. Math. Soc. (JEMS)*, 16(1):67–101, 2014.
- [BCN82] Pierre Bolley, Jacques Camus, and Jean-François Nourrigat. La condition de Hörmander-Kohn pour les opérateurs pseudo-différentiels. *Comm. Partial Differential Equations*, 7(2):197–221, 1982.
- [Bea14] Karine Beauchard. Null controllability of Kolmogorov-type equations. *Math. Control Signals Systems*, 26(1):145–176, 2014.
- [Bel96] André Bellaïche. The tangent space in sub-Riemannian geometry. In *Sub-Riemannian geometry*, volume 144 of *Progr. Math.*, pages 1–78. Birkhäuser, Basel, 1996.
- [BFKG12] Hajer Bahouri, Clotilde Fermanian-Kammerer, and Isabelle Gallagher. Phase-space analysis and pseudodifferential calculus on the Heisenberg group. *Astérisque*, (342):vi+127, 2012.
- [BHHR15] Karine Beauchard, Bernard Helffer, Raphael Henry, and Luc Robbiano. Degenerate parabolic operators of Kolmogorov type with a geometric control condition. *ESAIM Control Optim. Calc. Var.*, 21(2):487–512, 2015.
- [BL13] Ugo Boscain and Camille Laurent. The Laplace-Beltrami operator in almost-Riemannian geometry. *Ann. Inst. Fourier (Grenoble)*, 63(5):1739–1770, 2013.
- [BMM15] Karine Beauchard, Luc Miller, and Morgan Morancey. 2D Grushin-type equations: minimal time and null controllable data. *J. Differential Equations*, 259(11):5813–5845, 2015.
- [Bon69] Jean-Michel Bony. Principe du maximum, inégalité de Harnack et unicité du problème de Cauchy pour les opérateurs elliptiques dégénérés. *Ann. Inst. Fourier (Grenoble)*, 19(fasc. 1):277–304 xii, 1969.
- [BPS16a] Karine Beauchard and Karel Pravda-Starov. Null-controllability of hypoelliptic quadratic differential equations. *preprint*, <http://arxiv.org/abs/1603.05367>, 2016.
- [BPS16b] Karine Beauchard and Karel Pravda-Starov. Null-controllability of non-autonomous Ornstein-Uhlenbeck equations. *preprint*, <http://arxiv.org/abs/1607.05433>, 2016.



- [Bra14] Marco Bramanti. *An invitation to hypoelliptic operators and Hörmander's vector fields*. SpringerBriefs in Mathematics. Springer, Cham, 2014.
- [BZ09] Karine Beauchard and Enrique Zuazua. Some controllability results for the 2D Kolmogorov equation. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 26(5):1793–1815, 2009.
- [CDSZ93] Ferruccio Colombini, Daniele Del Santo, and Claude Zuily. Uniqueness and nonuniqueness in the Cauchy problem for degenerate elliptic operators. *Amer. J. Math.*, 115(6):1281–1297, 1993.
- [Cho39] Wei-Liang Chow. Über Systeme von linearen partiellen Differentialgleichungen erster Ordnung. *Math. Ann.*, 117:98–105, 1939.
- [CMV16] Piermarco Cannarsa, Patrick Martinez, and Judith Vancostenoble. Global Carleman estimates for degenerate parabolic operators with applications. *Mem. Amer. Math. Soc.*, 239(1133):ix+209, 2016.
- [Cor07] Jean-Michel Coron. *Control and nonlinearity*, volume 136 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2007.
- [CSL16] Felipe W. Chaves-Silva and Gilles Lebeau. Spectral inequality and optimal cost of controllability for the Stokes system. *ESAIM Control Optim. Calc. Var.*, 22(4):1137–1162, 2016.
- [Der71] Makhlof Derridj. Un problème aux limites pour une classe d'opérateurs du second ordre hypoelliptiques. *Ann. Inst. Fourier (Grenoble)*, 21(4):99–148, 1971.
- [DF88] Harold Donnelly and Charles Fefferman. Nodal sets of eigenfunctions on riemannian manifolds. *Invent. Math.*, 93:161–183, 1988.
- [Eva98] Lawrence C. Evans. *Partial differential equations*. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 1998.
- [EZ11a] Sylvain Ervedoza and Enrique Zuazua. Observability of heat processes by transmutation without geometric restrictions. *Math. Control Relat. Fields*, 1(2):177–187, 2011.
- [EZ11b] Sylvain Ervedoza and Enrique Zuazua. Sharp observability estimates for heat equations. *Arch. Ration. Mech. Anal.*, 202(3):975–1017, 2011.
- [FCZ00] Enrique Fernández-Cara and Enrique Zuazua. The cost of approximate controllability for heat equations: the linear case. *Adv. Differential Equations*, 5(4-6):465–514, 2000.
- [FI96] Andrei V. Fursikov and Oleg Yu. Imanuvilov. *Controllability of evolution equations*, volume 34 of *Lecture Notes Series*. Seoul National University Research Institute of Mathematics Global Analysis Research Center, Seoul, 1996.
- [FP81] Charles Fefferman and Duong Hong Phong. The uncertainty principle and sharp Gårding inequalities. *Comm. Pure Appl. Math.*, 34(3):285–331, 1981.
- [FP83] Charles Fefferman and Duong Hong Phong. Subelliptic eigenvalue problems. In *Conference on harmonic analysis in honor of Antoni Zygmund, Vol. I, II (Chicago, Ill., 1981)*, Wadsworth Math. Ser., pages 590–606. Wadsworth, Belmont, CA, 1983.
- [Gar93] Nicola Garofalo. Unique continuation for a class of elliptic operators which degenerate on a manifold of arbitrary codimension. *J. Differential Equations*, 104(1):117–146, 1993.
- [God82] Roger Godement. *Introduction à la théorie des groupes de Lie. Tome 2*, volume 12 of *Publications Mathématiques de l'Université Paris VII [Mathematical Publications of the University of Paris VII]*. Université de Paris VII, U.E.R. de Mathématiques, Paris, 1982.
- [Her63] Robert Hermann. On the accessibility problem in control theory. In *Internat. Sympos. Nonlinear Differential Equations and Nonlinear Mechanics*, pages 325–332. Academic Press, New York, 1963.
- [HN05] Bernard Helffer and Francis Nier. *Hypoelliptic estimates and spectral theory for Fokker-Planck operators and Witten Laplacians*, volume 1862 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2005.

- [Hör67] Lars Hörmander. Hypoelliptic second order differential equations. *Acta Math.*, 119:147–171, 1967.
- [Hör85] Lars Hörmander. *The Analysis of Linear Partial Differential Operators*, volume III. Springer-Verlag, 1985. Second printing 1994.
- [Hör92] Lars Hörmander. A uniqueness theorem for second order hyperbolic differential equations. *Comm. Partial Differential Equations*, 17(5-6):699–714, 1992.
- [Hör97] Lars Hörmander. On the uniqueness of the Cauchy problem under partial analyticity assumptions. In *Geometrical optics and related topics (Cortona, 1996)*, volume 32 of *Progr. Nonlinear Differential Equations Appl.*, pages 179–219. Birkhäuser Boston, Boston, MA, 1997.
- [Jea14] Frédéric Jean. *Control of nonholonomic systems: from sub-Riemannian geometry to motion planning*. Springer Briefs in Mathematics. Springer, Cham, 2014.
- [JSC87] David Jerison and Antonio Sánchez-Calle. Subelliptic, second order differential operators. In *Complex analysis, III (College Park, Md., 1985–86)*, volume 1277 of *Lecture Notes in Math.*, pages 46–77. Springer, Berlin, 1987.
- [Koe17] Armand Koenig. Non null controllability of the grushin operator in 2D. *preprint*, <https://arxiv.org/abs/1701.06467>, 2017.
- [Koh78] Joseph J. Kohn. Lectures on degenerate elliptic problems. In *Pseudodifferential operator with applications (Bressanone, 1977)*, pages 89–151. Liguori, Naples, 1978.
- [Koh05] Joseph J. Kohn. Hypoellipticity and loss of derivatives. *Ann. of Math. (2)*, 162(2):943–986, 2005. With an appendix by Makhlof Derridj and David S. Tartakoff.
- [Léa10] Matthieu Léautaud. Spectral inequalities for non-selfadjoint elliptic operators and application to the null-controllability of parabolic systems. *J. Funct. Anal.*, 258:2739–2778, 2010.
- [Leb92] Gilles Lebeau. Contrôle analytique. I. Estimations a priori. *Duke Math. J.*, 68(1):1–30, 1992.
- [Lee13] John M. Lee. *Introduction to smooth manifolds*, volume 218 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2013.
- [Ler88] Nicolas Lerner. Uniqueness for an ill-posed problem. *J. Differential Equations*, 71(2):255–260, 1988.
- [LL15] Camille Laurent and Matthieu Léautaud. Quantitative unique continuation for operators with partially analytic coefficients. Application to approximate control for waves. *accepted for publication in Journal of the European Mathematical Society*, <http://arxiv.org/abs/1506.04254>, 2015.
- [LL16] Camille Laurent and Matthieu Léautaud. Uniform observability estimates for linear waves. *ESAIM Control Optim. Calc. Var.*, 22(4):1097–1136, 2016.
- [LL17] Camille Laurent and Matthieu Léautaud. The cost function for the approximate control of waves. *work in progress*, 2017.
- [Lob70] Claude Lobry. Contrôlabilité des systèmes non linéaires. *SIAM J. Control*, 8:573–605, 1970.
- [LR95] Gilles Lebeau and Luc Robbiano. Contrôle exact de l’équation de la chaleur. *Comm. Partial Differential Equations*, 20:335–356, 1995.
- [LR97] Gilles Lebeau and Luc Robbiano. Stabilisation de l’équation des ondes par le bord. *Duke Math. J.*, 86:465–491, 1997.
- [LR05] Jérôme Le Rousseau and Luc Robbiano. A spectral inequality for the bi-laplace operator. *preprint*, <https://hal.archives-ouvertes.fr/hal-01194748v4/document>, 2005.
- [LRL12] Jérôme Le Rousseau and Gilles Lebeau. On Carleman estimates for elliptic and parabolic operators. Applications to unique continuation and control of parabolic equations. *ESAIM Control Optim. Calc. Var.*, 18(3):712–747, 2012.

- [LRM16] Jérôme Le Rousseau and Iván Moyano. Null-controllability of the Kolmogorov equation in the whole phase space. *J. Differential Equations*, 260(4):3193–3233, 2016.
- [LZ82] Richard Lascar and Claude Zuily. Unicité et non unicité du problème de Cauchy pour une classe d’opérateurs différentiels à caractéristiques doubles. *Duke Math. J.*, 49(1):137–162, 1982.
- [Mel86] Richard Melrose. Propagation for the wave group of a positive subelliptic second-order differential operator. In *Hyperbolic equations and related topics (Katata/Kyoto, 1984)*, pages 181–192. Academic Press, Boston, MA, 1986.
- [Mil06a] Luc Miller. The control transmutation method and the cost of fast controls. *SIAM J. Control Optim.*, 45(2):762–772, 2006.
- [Mil06b] Luc Miller. On the controllability of anomalous diffusions generated by the fractional laplacian. *Mathematics of Control, Signals, and Systems*, 3:260–271, 2006.
- [Mil10] Luc Miller. A direct Lebeau-Robbiano strategy for the observability of heat-like semigroups. *Discrete Contin. Dyn. Syst. Ser. B*, 14(4):1465–1485, 2010.
- [Mon02] Richard Montgomery. *A Tour of Subriemannian Geometries, Their Geodesics and Applications*, volume 91 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2002.
- [Mor15] Morgan Morancey. Approximate controllability for a 2D Grushin equation with potential having an internal singularity. *Ann. Inst. Fourier (Grenoble)*, 65(4):1525–1556, 2015.
- [Nag66] Tadashi Nagano. Linear differential systems with singularities and an application to transitive Lie algebras. *J. Math. Soc. Japan*, 18:398–404, 1966.
- [Phu04] Kim Dang Phung. Note on the cost of the approximate controllability for the heat equation with potential. *J. Math. Anal. Appl.*, 295(2):527–538, 2004.
- [Phu10] Kim Dang Phung. Waves, damped wave and observation. In *Some problems on nonlinear hyperbolic equations and applications*, volume 15 of *Ser. Contemp. Appl. Math. CAM*, pages 386–412. Higher Ed. Press, 2010.
- [Ras38] Petr K. Rashevski. About connecting two points of complete nonholonomic space by admissible curve. *Uch Zapiski ped. inst. Libknekhta*, 2:83–94, 1938.
- [Rif14] Ludovic Rifford. *Sub-Riemannian geometry and optimal transport*. Springer Briefs in Mathematics. Springer, Cham, 2014.
- [Rob91] Luc Robbiano. Théorème d’unicité adapté au contrôle des solutions des problèmes hyperboliques. *Comm. Partial Differential Equations*, 16(4-5):789–800, 1991.
- [Rob95] Luc Robbiano. Fonction de coût et contrôle des solutions des équations hyperboliques. *Asymptotic Anal.*, 10:95–115, 1995.
- [RS76] Linda Preiss Rothschild and Elias M. Stein. Hypoelliptic differential operators and nilpotent groups. *Acta Math.*, 137(3-4):247–320, 1976.
- [RS80] Michael Reed and Barry Simon. *Methods of modern mathematical physics*. Academic press, 1980.
- [RT05] Ludovic Rifford and Emmanuel Trélat. Morse-Sard type results in sub-Riemannian geometry. *Math. Ann.*, 332(1):145–159, 2005.
- [RT73] Jeffrey Rauch and Michael Taylor. Penetrations into shadow regions and unique continuation properties in hyperbolic mixed problems. *Indiana Univ. Math. J.*, 22:277–285, 1972/73.
- [Rus73] David L. Russell. A unified boundary controllability theory for hyperbolic and parabolic partial differential equations. *Studies in Appl. Math.*, 52:189–221, 1973.
- [RZ98] Luc Robbiano and Claude Zuily. Uniqueness in the Cauchy problem for operators with partially holomorphic coefficients. *Invent. Math.*, 131(3):493–539, 1998.

- [SSV12] René L. Schilling, Renming Song, and Zoran Vondraček. *Bernstein functions*, volume 37 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, second edition, 2012. Theory and applications.
- [Sus73] Héctor J. Sussmann. Orbits of families of vector fields and integrability of distributions. *Trans. Amer. Math. Soc.*, 180:171–188, 1973.
- [Tar07] Luc Tartar. *An introduction to Sobolev spaces and interpolation spaces*, volume 3 of *Lecture Notes of the Unione Matematica Italiana*. Springer, Berlin; UMI, Bologna, 2007.
- [Tat95] Daniel Tataru. Unique continuation for solutions to PDE’s; between Hörmander’s theorem and Holmgren’s theorem. *Comm. Partial Differential Equations*, 20(5-6):855–884, 1995.
- [Tat99a] Daniel Tataru. *Carleman estimates, unique continuation and applications*. Lecture notes. unpublished, <https://math.berkeley.edu/~tataru/papers/ucpnotes.ps>, 1999.
- [Tat99b] Daniel Tataru. Unique continuation for operators with partially analytic coefficients. *J. Math. Pures Appl. (9)*, 78(5):505–521, 1999.
- [Wat82] Kinji Watanabe. Sur l’unicité du prolongement des solutions des équations elliptiques dégénérées. *Tôhoku Math. J. (2)*, 34(2):239–249, 1982.