ACADÉMIE
DES SCIENCES
INSTITUT DE FRANCE

Comptes Rendus
Mathématique

Roberta Bianchini, Vincent Laheurte and Frank Sueur

A new look at the controllability cost of linear evolution systems with a long gaze at localized data

Volume 363 (2025), p. 241-266

Online since: 8 April 2025

https://doi.org/10.5802/crmath.725

CENTRE MERSENNE
The Comptes Rendus. Mathématique *are a member of the Mersenne Center for open scientific publishing*
www.centre-mersenne.org — e-ISSN : 1778-3569

ACADÉMIE
DES SCIENCES
INSTITUT DE FRANCE

Comptes Rendus. Mathématique
2025, Vol. 363, p. 241-266
https://doi.org/10.5802/crmath.725

Research article / *Article de recherche*
Control theory / *Théorie du contrôle*

A new look at the controllability cost of linear evolution systems with a long gaze at localized data

Un autre regard sur le coût de contrôle des systèmes d'évolutions linéaires, avec une attention particulière sur les données localisées

Roberta Bianchini [a], **Vincent Laheurte** [b] and **Frank Sueur** [*, c]

[a] Consiglio Nazionale delle Ricerche, Istituto per le Applicazioni del Calcolo, 00185 Rome, Italy

[b] IMB, Université de Bordeaux, France

[c] Department of Mathematics, Maison du nombre, 6 avenue de la Fonte, University of Luxembourg, L-4364 Esch-sur-Alzette, Luxembourg

E-mails: roberta.bianchini@cnr.it, vincent.laheurte@math.u-bordeaux.fr, franck.sueur@uni.lu

Abstract. We revisit the classical issue of the controllability/observability cost of linear first order evolution systems, starting with ODEs, before turning to some linear first order evolution PDEs in several space dimensions, including hyperbolic systems and pseudo-differential systems obtained by linearization in fluid mechanics. In particular we investigate the cost for localized initial data and, in the dispersive case, for initial data which are semiclassically microlocalized.

Résumé. Nous revisitons la question classique du coût de contrôle/d'observation des systèmes d'évolutions linéaires du premier ordre, en commençant par le cas des EDOs, avant de d'étendre l'analyse aux EDPs d'évolution, du premier ordre, en plusieurs dimensions d'espace. Notre analyse couvre le cas des systèmes hyperboliques et des systèmes pseudo-différentiels, notamment ceux obtenus par linéarisation en mécanique des fluides. En particulier, nous examinons le coût de contrôle de données initiales localisées et, dans le cas dispersif, de données initiales semi-classiquement micro-localisées.

Funding. F.S. was partially supported by the Agence Nationale de la Recherche, Project SINGFLOWS, grant ANR-18-CE40-0027-01, Project BOURGEONS, grant ANR-23-CE40-0014-01, and Institut Universitaire de France. R.B. acknowledges financial support by the Italian Ministry for Universities and Research (MUR), Project PRIN 2022 TESEO, grant 2022HSSYPN.

Manuscript received 14 January 2024, revised 23 January 2025, accepted 27 January 2025.

* Corresponding author

ISSN (electronic): 1778-3569 https://comptes-rendus.academie-sciences.fr/mathematique/

1. Reminder on the cost of controllability of linear ODEs

In the setting of linear ODEs in finite dimensions, with variable coefficients, the issue of exact controllability can be set as follows.

Question 1. *Given some time-dependent fields of matrices $A(t)$ and $B(t)$, a positive time T, some initial and final states b_0 and b_T, is there a time-dependent field $f(t)$ and a corresponding time-dependent field $b(t)$ such that*

$$b'(t)+A(t)b(t)=B(t)f(t) \text{ for } t\in[0,T] \text{ and } \big(b(0),b(T)\big)=(b_0,b_T)? \tag{1}$$

Above, it is understood that the dependence on the time t is smooth enough for the system to make sense and that the matrices and vectors are compatible in terms of their size. A positive answer to Question 1 means by definition that the differential system $b'(t)=A(t)b(t)$ is *exactly controllable* in time T.

Let us observe that such a result is only non-trivial whenever the system is under-actuated, that is in the cases where the matrices $B(t)$ are not invertible. Otherwise, it is sufficient to consider any smooth field $b(t)$ satisfying $\big(b(0),b(T)\big)=(b_0,b_T)$, and then to define the control force field $f(t)$ by

$$f(t):=B(t)^{-1}\big(b'(t)+A(t)b(t)\big) \text{ for } t\in[0,T],$$

and then, the answer to Question 1 is positive.

In the case of a positive answer to Question 1, one may investigate the minimal quadratic cost of such a control result.

Question 2. *What is the infimum $\mathscr{C}=\mathscr{C}(T,A,B,b_0,b_T)$ of*

$$\int_0^T \left|f(t)\right|^2 \mathrm{d}t,$$

among the control force fields $f(t)$ which verify Question 1?

The Hilbert Uniqueness Method (HUM), introduced by Jacques-Louis Lions, relates the notions of controllability and observability by means of a duality argument, including the computation of the cost, see for example [10, Chapter 1]. In the context above, in the particular case where $b(T)=0$, the so-called case of *null-controllability*, this method yields that $\mathscr{C}$ is the minimal constant for which, for any b_0, up to renaming the adjoints $A^*(t)$ and $B^*(t)$ respectively by $A(t)$ and $B(t)$, the unique solution to the uncontrolled equation

$$b'(t)+A(t)b(t)=0 \text{ for } t\in[0,T] \text{ and } b(0)=b_0, \tag{2}$$

satisfies

$$\left|b(T)\right|^2\leqslant \mathscr{C}\int_0^T\left|B(t)b(t)\right|^2\mathrm{d}t. \tag{3}$$

The interpretation of the inequality above is that the energy $|b(T)|^2$ of the final state reached by the uncontrolled system at time T can be bounded by the energy $\int_0^T|B(t)b(t)|^2\,\mathrm{d}t$ which is observed through the field of matrices $B(t)$ during the interval $(0,T)$, up to a constant factor $\mathscr{C}$. The minimal possible constant $\mathscr{C}$ in (3) is also called the *observability cost* of the differential system: $b'(t)+A(t)b(t)=0$ on $(0,T)$.

Before giving a characterization of $\mathscr{C}$, let us recall the following elementary result on differential Lyapunov equations.

Lemma 1. *Let A and B be some continuous time-dependent fields of matrices. Then there is a unique global solution G of the differential Lyapunov equation:*

$$\mathrm{Lyap}(A,B):\quad G'(t)-G(t)A(t)-A^*(t)G(t)=B^*(t)B(t) \text{ for } t\in[0,T] \text{ and } G(0)=0.$$

Moreover at any time t, the matrix $G(t)$ is symmetric and nonnegative. Finally, for any time T,

$$G(T) = \int_0^T R(t,T)^* B^*(t) B(t) R(t,T)\, \mathrm{d}t, \tag{4}$$

where, for any time T, the matrix-valued function $R(\cdot,T)$ is the solution to the linear system

$$\frac{\mathrm{d}}{\mathrm{d}t} R(t,T) = -A(t)R(t,T), \quad R(T,T) = \mathrm{Id}.$$

Proof. The existence and uniqueness of G and of $R(\cdot,T)$ for any time T, as solutions of their respective problems, follow from the linear finite-dimensional version of the Cauchy–Lipschitz theorem. Then the symmetry follows from the uniqueness by observing that G and G^* satisfy the same equation. Similarly, the formula (4) follows from the uniqueness by observing that the r.h.s. satisfies the differential Lyapunov equation. Finally, the nonnegativity at any time $t \in [0,T]$ of the matrix $G(t)$ follows from this representation. □

Let us now recall a characterization of $\mathscr{C}$ as the inverse of the minimal eigenvalue $\lambda_{\min}\big(G(T)\big)$ of the so-called *observability Gramian matrix* $G(T)$.

Theorem 2. *The observability cost $\mathscr{C}$ satisfies $\mathscr{C} = \lambda_{\min}\big(G(T)\big)^{-1}$, where $G(T)$ is the value at time T of the solution G to the differential Lyapunov equation* $\mathrm{Lyap}(A,B)$.

A byproduct of Theorem 2 is that a necessary and sufficient condition for a linear time-varying finite-dimensional control system to be observable is that the observability Gramian $G(T)$ is invertible.

Let us also mention that for time-invariant finite-dimensional control systems, this condition reduces to the famous Kalman rank condition, an algebraic condition which does not require any integration of linear differential equations.

Theorem 2 is well-known, see for example [10, Chapter 1]. However, we give below a proof setting up a method which will be repeatedly used in some more difficult settings in the sequel.

Proof of Theorem 2. The starting idea is that there is an extra integral over time in (3) so that $\mathscr{C}^{-1}$ can be interpreted as a time. To restore this time in the left-hand side of (3), we introduce the following scalar field:

$$E_R(t) := G(t)b(t)\cdot b(t), \tag{5}$$

where the dot in (5) is the usual euclidean inner product. The quantity E_R will be called the *recorded energy*.

In particular the recorded energy starts with zero initial value, that is

$$E_R(0) = 0. \tag{6}$$

Moreover, by Leibniz' rule and by using the equations of G and b, its time derivative is given by

$$E_R' = B^* B b \cdot b = |Bb|^2,$$

and then,

$$E_R(T) = \int_0^T \big|B(t)b(t)\big|^2\, \mathrm{d}t. \tag{7}$$

Thus, at the final time T, the recorded energy is equal to the observed energy.

On the other hand, from Definition 5, one may compare $E_R(T)$ with the final energy $|b(T)|^2$ by

$$\inf_{b_0\neq 0} \frac{G(T)b(T)\cdot b(T)}{|b(T)|^2} = \inf_{b(T)\neq 0} \frac{G(T)b(T)\cdot b(T)}{|b(T)|^2} = \lambda_{\min}\big(G(T)\big).$$

This concludes the proof of Theorem 2. □

While the proof above recalls the multiplier method, see for instance [21], it is interesting to reinterpret in a slightly different way these computations, more precisely the identity (7), for the sequel. In this direction, first, from (2), we deduce that the rank-one matrix $b(t)\otimes b(t) = b(t)b(t)^*$, which has to be interpreted as the energy tensor, satisfies the following ODE:

$$(b\otimes b)' + A(b\otimes b) + (b\otimes b)A^* = 0. \tag{8}$$

Moreover, the recorded energy reads

$$E_R = G:(b\otimes b), \tag{9}$$

where the notation : stands for the Frobenius product, that is for two matrices A and B of the same size,

$$A:B = \operatorname{tr}(A^*B) = B:A.$$

Recalling the following interplay between the Frobenius product and the matrix product:

$$A:(BC) = (AC^*):B \quad \text{and} \quad A:(CB) = (C^*A):B, \tag{10}$$

we observe that the differential Lyapunov equation $\mathrm{Lyap}(A,B)$ is the dual equation of (8) with source term B^*B, meaning that if one tests the differential Lyapunov equation $\mathrm{Lyap}(A,B)$ with a time-dependent matrix $C(t)$ by means of the Frobenius product and integrates by parts in time, then

$$\int_0^T G:(C' + AC + CA^*) = (G:C)(T) - \int_0^T B^*B:C. \tag{11}$$

This way, the identity (7) follows from choosing $C = b\otimes b$, recalling (6), (8) and (9).

2. Case of a transport PDE

Let us now turn to the case of a simple PDE, the following linear transport equation:

$$\partial_t \mathbf{u} + (v\cdot\nabla)\mathbf{u} + A\mathbf{u} = 0, \ \text{ for } \ t\in[0,T], \tag{12}$$

where the space variable is $x\in\mathbb{R}^d$ with $d\in\mathbb{N}^*$, $v\colon \mathbb{R}^d\to\mathbb{R}^d$ is a given regular vector field and A is a given regular field of $N\times N$ matrices depending on $x\in\mathbb{R}^d$, with $N\in\mathbb{N}^*$. These matrices encode some damping or amplification along the transportation by v. The final time $T>0$ is also given and the unknown is the vector field $\mathbf{u}\colon [0,T]\times\mathbb{R}^d\to\mathbb{R}^N$. In this case, the cost is defined as follows.

Definition 3. *The* cost of observability *by a regular field of $n\times N$ matrices B depending on $x\in\mathbb{R}^d$ is the minimal constant $\mathscr{C} = \mathscr{C}(T,v,A,B)$ such that for any initial condition $\mathbf{u}_0\in L^2(\mathbb{R}^d)$, the associated solution $\mathbf{u}(t)$ to* (12) *verifies the observability inequality:*

$$\|\mathbf{u}(T,\cdot)\|^2_{L^2(\mathbb{R}^d;\mathbb{R}^N)} \leqslant \mathscr{C}\int_0^T \|B\mathbf{u}(t)\|^2_{L^2(\mathbb{R}^d;\mathbb{R}^n)}\,\mathrm{d}t. \tag{13}$$

Let us highlight that, in this setting, the HUM method, which was mentioned in the previous section, still holds and the identification of $\mathscr{C}$ with the cost of controllability of the adjoint system as well, see [10].

The natural question that arises is whether one can give a condition for the observability of the system (12) and compute its cost, as done in Theorem 2 for ODE systems.

The following result, generalizing Theorem 2, shows that $\mathscr{C}$ can be computed in terms of the infimum over $x_0\in\mathbb{R}^d$ of the minimal eigenvalues $\lambda_{\min}\big(G_{x_0}(T)\big)$ of a family of Gramian matrices $G_{x_0}(T)$ associated with time-dependent Lyapunov ODEs involving the coefficients of the PDE (12) evaluated along the characteristic flow induced by the vector field v. This flow is denoted by $X_{x_0}(t)$ and satisfies, for any $x_0\in\mathbb{R}^d$,

$$X'_{x_0}(t) = v\big(X_{x_0}(t)\big) \ \text{ for } \ t\in[0,T] \ \text{ and } \ X_{x_0}(0) = x_0. \tag{14}$$

On the other hand, the Gramian matrices G_{x_0} are given as solutions to $\mathrm{Lyap}\big(A_{x_0}, B_{x_0}\big)$ where, for any $x_0 \in \mathbb{R}^d$ and for any $t \in [0,T]$,

$$A_{x_0}(t) := A\big(X_{x_0}(t)\big) - \frac{1}{2}(\operatorname{div} v)\big(X_{x_0}(t)\big)\,\mathrm{Id}_N \quad \text{and} \quad B_{x_0}(t) := B\big(X_{x_0}(t)\big). \tag{15}$$

Theorem 4. *The observability cost $\mathscr{C}$ satisfies*

$$\mathscr{C} = \left(\inf_{x_0\in\mathbb{R}^d} \lambda_{\min}\big(G_{x_0}(T)\big)\right)^{-1}. \tag{16}$$

A byproduct of Theorem 4 is that a necessary and sufficient condition for (12) to be observable by the field of matrices B is that the Gramian matrices $\big(G_{x_0}(T)\big)_{x_0\in\mathbb{R}^d}$ are uniformly invertible.

The exact controllability, the null-controllability, and the boundary stabilization problem of hyperbolic systems in *one space dimension* have been widely investigated in the literature for almost half a century. We refer here to [3,11] for a recent account on the literature on this subject. On the other hand less is known for hyperbolic systems in *several space dimensions*. Let us however mention here the related works [8,18] on Carleman estimates for first-order symmetric hyperbolic systems in several space dimensions and [20] for some results on the stabilization of general linear multi-dimensional hyperbolic systems by a boundary feedback control. While we are not aware of any mention of the precise statement of Theorem 4 in the literature, it can be considered as part of the folklore and is mainly preparatory to the next sections.

To prove Theorem 4 we will mimic the proof of Theorem 2 by introducing a recorded energy to which we will apply the following lemma, where we assume that M and F are regular enough.

Lemma 5. *Let $\mathbf{u}(t)$ be a solution to* (12) *associated with the initial condition $\mathbf{u}_0 \in L^2(\mathbb{R}^d)$. Let M satisfy the following PDE*

$$\partial_t M + (v\cdot\nabla)M + (\operatorname{div} v)M - A^*M - MA = F, \tag{17}$$

where $F(t,x)$ is a time-dependent family of matrices. Let

$$E[M,u] := \int_{\mathbb{R}^d} (\mathbf{u}\otimes\mathbf{u}) : M. \tag{18}$$

Then its time derivative satisfies, for all $t\in[0,T]$:

$$E[M,u]' = \int_{\mathbb{R}^d} (\mathbf{u}\otimes\mathbf{u}) : F.$$

Proof. By Leibniz' rule, applying (12) twice, an integration by parts and a transposition, we obtain that the time derivative $E[M,u]'$ of $E[M,u]$ is given by

$$E[M,u]' = \int_{\mathbb{R}^d} \Big(\partial_t M + (v\cdot\nabla)M + (\operatorname{div} v)M - A^*M - MA\Big)\mathbf{u}\cdot\mathbf{u}\,\mathrm{d}x. \tag{19}$$

Then it is sufficient to use (17) to conclude. □

Another elementary ingredient in the proof of Theorem 4 is the following result.

Lemma 6. *Let $m\colon \mathbb{R}^d \to \mathbb{R}^{N\times N}$ be a continuous field of symmetric nonnegative matrices. Then*

$$\inf_{\|u\|_{L^2}=1} \int_{\mathbb{R}^d} m(x)u(x)\cdot u(x)\,\mathrm{d}x = \inf_{x\in\mathbb{R}} \lambda_{\min}\big(m(x)\big),$$

where the infimum on the left-hand side is taken over all the vector fields u in $L^2(\mathbb{R}^d)$ with $\|u\|_{L^2} = 1$.

Proof. Let

$$C := \inf_{x\in\mathbb{R}^d} \lambda_{\min}\big(m(x)\big).$$

On the one hand, for any $u\in L^2(\mathbb{R}^d)$, for any $x\in\mathbb{R}^d$,

$$m(x)u(x)\cdot u(x) \geqslant C\big|u(x)\big|^2,$$

so that, by integration,

$$\int_{\mathbb{R}^d} m(x)u(x)\cdot u(x)\,\mathrm{d}x \geqslant C\|u\|_{L^2}^2. \tag{20}$$

On the other hand, for any $\varepsilon > 0$, we choose $x_\varepsilon \in \mathbb{R}^d$ such that $\lambda_{\min}\big(m(x_\varepsilon)\big) \le C+\varepsilon$, as well as b_ε in the sphere $\mathscr{S}^{N-1}$ an eigenvector of $m(x_\varepsilon)$ associated with $\lambda_{\min}\big(m(x_\varepsilon)\big)$, so that

$$m(x_\varepsilon)b_\varepsilon = \lambda_{\min}\big(m(x_\varepsilon)\big)b_\varepsilon.$$

Finally, we set, for any $x \in \mathbb{R}^d$,

$$u_\varepsilon(x) := (\pi\varepsilon)^{-\frac{d}{4}} b_\varepsilon e^{-\frac{|x-x_\varepsilon|^2}{2\varepsilon}}.$$

Then, for any $\varepsilon > 0$, $\|u_\varepsilon\|_{L^2} = 1$ and, since m is continuous,

$$\lim_{\varepsilon\to 0}\int_{\mathbb{R}^d} m(x)u_\varepsilon(x)\cdot u_\varepsilon(x)\,\mathrm{d}x = C.$$

Combining this equality with (20), we conclude the proof of the lemma. □

We are now equipped to start the proof of Theorem 4.

Proof of Theorem 4. Let M be given by setting for any t and x,

$$G_x(t) = M\big(t, X_x(t)\big). \tag{21}$$

We highlight that, for any t, the mapping $x \mapsto X_x(t)$ is invertible, since its Jacobian satisfies a linear first order homogeneous ODE with 1 as initial condition. Then we observe that it follows from the fact that the Gramian matrices G_x are solutions to $\mathrm{Lyap}\big(A_x, B_x\big)$ and from the chain rule that $M(t,x)$ is a nonnegative symmetric matrix for any t and x, that M satisfies an equation of the type (17) with $F = B^*B$ and that $M(0,\cdot) = 0$. Therefore, by Lemma 5 and an integration on $[0,T]$, the recorded energy defined by $E_R := E[M,u]$ satisfies

$$E_R(T) = \int_0^T \int_{\mathbb{R}^d} \big|B\mathbf{u}(t,x)\big|^2\,\mathrm{d}x\mathrm{d}t. \tag{22}$$

On the other hand, since the map $\mathbf{u}_0 \mapsto \mathbf{u}(T,\cdot)$ is a bijection from $L^2(\mathbb{R}^d)$ to itself, by Lemma 6, we obtain

$$\inf_{\mathbf{u}_0\neq 0} \frac{E_R(T)}{\|\mathbf{u}(T,\cdot)\|_{L^2}^2} = \inf_{x\in\mathbb{R}^d} \lambda_{\min}\big(M(T,x)\big) = \inf_{x\in\mathbb{R}^d} \lambda_{\min}\big(G_x(T)\big), \tag{23}$$

since $x \mapsto X_x(T)$ is a bijection from $\mathbb{R}^d$ to itself. We conclude the proof of Theorem 4 by gathering (22) and (23). □

3. A localization in space and amplitude of the cost

While the definition of observability cost is rather *global* in space, we consider below the observability cost

$$\mathscr{C}[\mathbf{u}_0] := \frac{\|\mathbf{u}(T,\cdot)\|_{L^2}^2}{\int_0^T \|B\mathbf{u}(t)\|_{L^2}^2\,\mathrm{d}t}, \tag{24}$$

of the initial data $\mathbf{u}_0$ whose local energy tensor $\mathbf{u}_0 \otimes \mathbf{u}_0$ is *localized*, for the weak-* topology of the space $\mathscr{M}\big(\mathbb{R}^d;\mathbb{R}^{N\times N}\big)$ of the matrix-valued measures, at $(b_0\otimes b_0)\delta_{x_0}$, for some $x_0 \in \mathbb{R}^d$ and $b_0 \in \mathbb{R}^N$. Above $\mathbf{u}$ is the solution of the transport PDE (12) associated with $\mathbf{u}_0$. Since it is a linear equation, the cost $\mathscr{C}[\mathbf{u}_0]$ is homogeneous of degree 0 with respect to the initial data $\mathbf{u}_0$.

Theorem 7. *Let $x_0 \in \mathbb{R}^d$ and set $G_T := G_{x_0}(T)$, where G_{x_0} is the solution to the differential Lyapunov equation* $\mathrm{Lyap}(A_{x_0}, B_{x_0})$, *where we recall the notation* (15). *Let $b_0 \in \mathbb{R}^N \setminus \{0\}$ and denote by*

$$b_T^{\mathcal{N}} := \frac{b_{x_0,b_0}(T)}{|b_{x_0,b_0}(T)|},$$

where $b_{x_0,b_0} \in \mathbb{R}^N \setminus \{0\}$ is the solution of the linear ordinary differential equation:

$$b'_{x_0,b_0} = -A_{x_0} b_{x_0,b_0}, \quad \text{with } b_{x_0,b_0}(0) = b_0. \tag{25}$$

Let $(\mathbf{u}_0^n)_n$ be a sequence in $C^\infty(\mathbb{R}^d) \cap L^2(\mathbb{R}^d;\mathbb{R}^N)$ such that

$$\mathbf{u}_0^n \otimes \mathbf{u}_0^n \xrightarrow{w*} (b_{x_0,b_0} \otimes b_{x_0,b_0})\delta_{x_0} \quad \text{in } \mathscr{M}(\mathbb{R}^d;\mathbb{R}^{N\times N}), \quad \text{as } n \to +\infty. \tag{26}$$

Then,

$$\mathscr{C}[\mathbf{u}_0^n]^{-1} \longrightarrow (b_T^{\mathcal{N}})^* G_T b_T^{\mathcal{N}}, \quad \text{as } n \to +\infty.$$

To prove Theorem 7, it is interesting to follow the viewpoint introduced at the end of Section 1 and to consider the local energy tensor $u \otimes u$. By elementary operations, we deduce from (12) that $u \otimes u$ satisfies the transport PDE:

$$(\partial_t + v \cdot \nabla)(u \otimes u) + A(u \otimes u) + (u \otimes u)A^* = 0. \tag{27}$$

While we consider some vector field u with $L^2(\mathbb{R}^d;\mathbb{R}^N)$-regularity in space, the equation (27) makes sense and the Cauchy problem is well-posed, even if $u \otimes u$ is a matrix-valued measure μ (or even more generally for a distribution) rather than a $L^1(\mathbb{R}^d;\mathbb{R}^{N\times N})$-function. As we are interested in the case where the local energy tensor $u \otimes u$ concentrates in a Dirac mass, it is useful to recall the following elementary result on the Cauchy problem for transport equations with measure solutions.

Lemma 8. *Let $n, N \in \mathbb{N}^*$ and $T > 0$. Let V be a regular field of vector fields from $\mathbb{R}^n$ to $\mathbb{R}^n$ and A, $\widetilde{A}$ some smooth vector fields of matrices in $\mathbb{R}^{N\times N}$. Let $\mathscr{M}(\mathbb{R}^n;\mathbb{R}^{N\times N})$ be the space of matrix-valued measures and*

$$\mathfrak{C} := C\Big([0,T]; \mathscr{M}(\mathbb{R}^n;\mathbb{R}^{N\times N}) - w*\Big)$$

the space of the measures depending on time in a continuous way when $\mathscr{M}(\mathbb{R}^n;\mathbb{R}^{N\times N})$ is endowed by its weak- topology. Then:*

(1) *For any μ_0 in $\mathscr{M}(\mathbb{R}^n;\mathbb{R}^{N\times N})$, there exists a unique solution $\mu \in \mathfrak{C}$, to the PDE:*

$$\partial_t \mu + \operatorname{div}(\mu V) + A\mu + \mu \widetilde{A} = 0, \tag{28}$$

with initial value μ_0.

(2) *If the initial datum μ_0 is of the form $\mu_0^{y_0,P_0} = P_0 \delta_{y_0}$ with $y_0 \in \mathbb{R}^n$ and $P_0 \in \mathbb{R}^{N\times N}$ then the unique corresponding solution is given by*

$$\mu^{y_0,P_0}(t,\cdot) = P(t)\delta_{y(t)}, \tag{29}$$

where y solves the nonlinear ODE:

$$y' = V(y) \quad \text{and} \quad y(0) = y_0, \tag{30}$$

and $P \in C^1([0,T];\mathbb{R}^{N\times N})$ solves the linear ODE:

$$P' + (A \circ y)P + P(\widetilde{A} \circ y) = 0 \quad \text{and} \quad P(0) = P_0. \tag{31}$$

(3) *Finally, if $(\mu_0^n)_n$ is a sequence in $\mathscr{M}(\mathbb{R}^n;\mathbb{R}^{N\times N})$ which converges weakly-* to $\mu_0^{y_0,P_0}$, then the corresponding solution μ^n converges to μ^{y_0,P_0} in $\mathfrak{C}$ as $n \to +\infty$.*

Proof. The proof is very standard and we only give here a few words of comments. The existence follows from the characteristics method. Indeed, solutions can be represented as superpositions of the atomic case considered in the second item. The uniqueness and stability parts of the statement can be proved by a duality argument, in the spirit of [15], or by a quantitative approach, thanks to matricial Wasserstein distances, see [6,26]. □

Equipped with this result, we are now ready to prove Theorem 7.

Proof of Theorem 7. Let X_{x_0} be given by (14) and b_{x_0,b_0} by (25). Then, by Leibniz' rule,

$$\big(b_{x_0,b_0}\otimes b_{x_0,b_0}\big)' = -A_{x_0}\big(b_{x_0,b_0}\otimes b_{x_0,b_0}\big) - \big(b_{x_0,b_0}\otimes b_{x_0,b_0}\big)A^*_{x_0},$$

where we recall that A_{x_0} is given by (15). Let

$$\mu(t,\cdot) := \big(b_{x_0,b_0}(t)\otimes b_{x_0,b_0}(t)\big)\delta_{X_{x_0}(t)}.$$

Therefore, by the second item of Lemma 8, μ is the unique solution to

$$\partial_t\mu + \operatorname{div}(\mu V) + \left(A - \frac{1}{2}(\operatorname{div} v)\,\mathrm{Id}\right)\mu + \mu\left(A^* - \frac{1}{2}(\operatorname{div} v)\,\mathrm{Id}\right) = 0, \tag{32}$$

with the initial data $\mu_0 = (b_0\otimes b_0)\delta_{x_0}$.

Let $\mathbf{u}^n$ be the solution of the transport PDE (12) associated with the initial data $\mathbf{u}^n_0$, and set $\mu^n := \mathbf{u}^n\otimes\mathbf{u}^n$ the corresponding local energy tensor. Then by (27),

$$\partial_t\mu^n + \operatorname{div}(\mu^n V) + \left(A - \frac{1}{2}(\operatorname{div} v)\,\mathrm{Id}\right)\mu^n + \mu^n\left(A^* - \frac{1}{2}(\operatorname{div} v)\,\mathrm{Id}\right) = 0. \tag{33}$$

Then, by (26) and the third item of Lemma 8, we obtain that μ^n converges to μ in $\mathfrak{C}$ as $n\to+\infty$.

Now, considering again M given by (21), one may re-use (22), which can actually be re-obtained as a consequence of the fact that (17) is the dual equation of (27) with B^*B as a source term, that is:

$$\int_0^T \big\|B\mathbf{u}^n(t)\big\|^2_{L^2}\,\mathrm{d}t = \int_{\mathbb{R}^d} M(T,x) : \mu^n(T,x)\,\mathrm{d}x, \tag{34}$$

while

$$\big\|\mathbf{u}^n(T,\cdot)\big\|^2_{L^2} = \int_{\mathbb{R}^d}\mathrm{Id} : \mu^n(T,x)\,\mathrm{d}x. \tag{35}$$

Finally, the proof of Theorem 7 follows by passing to the limit in these two expressions, recalling (21) for the first one. □

4. Microlocalized cost

As we have seen before, the concept of energy is at the center of the observability issue. In the previous sections, in the case of some transport equations, we considered the energy tensor $\mathbf{u}\otimes\mathbf{u}$, which encodes the localization of the energy in space, with respect to the components of $\mathbf{u}$. However, for more general evolution linear first order PDEs, in several space dimensions, dispersive effects occur and the propagation of energy depends not only on the space variable x, but also on the frequency variable ξ. Hence we need a tool to encode the distribution of energy in the phase space, generalizing the energy tensor $\mathbf{u}\otimes\mathbf{u}$. To this end, the natural object to consider is the matrix Wigner transform, which we consider in its semiclassical version, focusing on high frequencies.

The Wigner transform and the so-called defect measures were first used in a controllability context in [7,24] in the context of the control of wave equations, it is therefore all the more natural to investigate the use of this tool in our context.

4.1. *Semiclassical Wigner matrix transform*

Below $\varepsilon \in (0,1]$. However the regime of interest later on corresponds to this parameter ε converging to 0, its inverse standing for the typical size of the frequencies under consideration.

Definition 9. *For u and v in $L^2(\mathbb{R}^d;\mathbb{C}^N)$, we define the* Wigner matrix $W^\varepsilon[u,v](x,y)$, *for any $x,\xi \in \mathbb{R}^d \times \mathbb{R}^d$, by*

$$W^\varepsilon[u,v](x,\xi) := \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} u\left(x-\varepsilon\frac{y}{2}\right)\otimes \bar{v}\left(x+\varepsilon\frac{y}{2}\right)e^{iy\cdot\xi}\,\mathrm{d}y.$$

The mapping $(u,v)\mapsto W^\varepsilon[u,v]$ is bilinear and verifies

$$W^\varepsilon[u,v] = \left(W^\varepsilon[v,u]\right)^*. \tag{36}$$

In the case where $u=v$ we simply denote by $W^\varepsilon[u]$ the Wigner transform $W^\varepsilon[u,u]$, which is a field of symmetric matrices. The Wigner matrix encodes the energy in the phase space in the following sense: the marginals of $W^\varepsilon[u,v]$ respectively satisfy, formally, for any $x\in\mathbb{R}^d$, for any $\xi\in\mathbb{R}^d$,

$$\int_{\mathbb{R}^d} W^\varepsilon[u,v]\,\mathrm{d}\xi = u(x)\otimes\bar{v}(x),$$
$$\int_{\mathbb{R}^d} W^\varepsilon[u,v]\,\mathrm{d}x = \frac{1}{(2\pi\varepsilon)^d}\,\widehat{u}\left(\frac{\xi}{\varepsilon}\right)\otimes(\widehat{v})\left(\frac{\xi}{\varepsilon}\right),$$

where $\widehat{u},\widehat{v}$ denote the Fourier transforms of u and v. In particular,

$$\iint_{\mathbb{R}^d\times\mathbb{R}^d}\left(W^\varepsilon[u]:\mathrm{Id}\right)\mathrm{d}x\mathrm{d}\xi = \|u\|^2_{L^2(\mathbb{R}^d;\mathbb{R}^N)}, \tag{37}$$

where we recall that the notation : stands for the Frobenius product.

We also have, for any $1\le j,k\le N$,

$$\left|W^\varepsilon_{j,k}[u,v](x,\xi)\right| \le \frac{1}{(\pi\varepsilon)^d}\|u_j\|_{L^2(\mathbb{R}^d;\mathbb{C})}\|v_k\|_{L^2(\mathbb{R}^d;\mathbb{C})},$$

where $W^\varepsilon_{j,k}[u,v]$ denotes the coefficient (j,k) of the Wigner matrix, and u_j and v_k respectively denote the j-th component of u and the k-th component of v. This implies that the Wigner matrix $W[u,v]$ is in L^∞, from $\mathbb{R}^d\times\mathbb{R}^d$ to $\mathbb{R}^{N\times N}$ equipped with the Frobenius norm, and that

$$\left\|W^\varepsilon[u,v]\right\|_{L^\infty} \le \frac{1}{(\pi\varepsilon)^d}\|u\|_{L^2}\|v\|_{L^2}.$$

The Wigner matrix $W^\varepsilon[u,v]$ is also in $L^2(\mathbb{R}^d\times\mathbb{R}^d;\mathbb{R}^{N\times N})$, with

$$\left\|W^\varepsilon[u,v]\right\|_{L^2} = \|u\|_{L^2}\|v\|_{L^2}.$$

For more details about those properties, we refer, for instance, to [25].

Remark 10. The Wigner matrix $W^\varepsilon[u]$ carries information about the energy tensor $u\otimes u$, but is not necessarily nonnegative. However, for positivity, one may consider the Husimi transform instead, or Wigner measures, which are weak limits of Wigner matrices as ε goes to 0, in some sense. For our purpose, a sign is not required so we will not use those tools.

Below we recall some classical examples.

Example 11. An important example is given by the coherent states defined, for any $x_0,\xi_0\in\mathbb{R}^d$, for any $\mathbf{b}_0\in\mathbb{R}^N$, as follows:

$$\mathbf{u}^\varepsilon_{x_0,\xi_0,\mathbf{b}_0}(x) := (\pi\varepsilon)^{-\frac{d}{4}}\mathbf{b}_0 e^{i\frac{(x-x_0)\cdot\xi_0}{\varepsilon}}e^{-\frac{|x-x_0|^2}{2\varepsilon}}.$$

Then one may compute

$$W^\varepsilon\left[\mathbf{u}^\varepsilon_{x_0,\xi_0,\mathbf{b}_0}\right](x,\xi) = (\pi\varepsilon)^{-d}(\mathbf{b}_0\otimes\mathbf{b}_0)e^{-\frac{|x-x_0|^2+|\xi-\xi_0|^2}{\varepsilon}}.$$

In particular, the matrix Wigner transform $\mathscr{W}^\varepsilon[\mathbf{u}^\varepsilon_{x_0,\xi_0,\mathbf{b}_0}]$ weakly converges, as $\varepsilon \to 0$, in the sense of distributions, to $(\mathbf{b}_0 \otimes \mathbf{b}_0)\delta_{x_0,\xi_0}$.

Example 12. Another classical example is given by the WKB states defined, for any $\alpha \in (0,1]$, any $\xi_0 \in \mathbb{R}^d$ and any amplitude function $a(x) \in L^2(\mathbb{R}^d)$, by

$$u^\varepsilon_{a,\alpha}(x) := a(x) e^{i\frac{x\cdot\xi_0}{\varepsilon^\alpha}}.$$

Then, as $\varepsilon \to 0$, for any $\alpha \in (0,1)$,

$$W^\varepsilon[u^\varepsilon_{a,\alpha}] \xrightarrow{w*} a(x)\otimes a(x)\delta_0(\xi) \text{ in } \mathscr{S}'\big(\mathbb{R}^d \times \mathbb{R}^d\big),$$

while in the case where $\alpha = 1$,

$$W^\varepsilon[u^\varepsilon_{a,\alpha}] \xrightarrow{w*} a(x)\otimes a(x)\delta_{\xi_0}(\xi) \text{ in } \mathscr{S}'\big(\mathbb{R}^d \times \mathbb{R}^d\big).$$

4.2. *Pseudo-differential operators*

Our analysis allows to consider not only PDEs but also pseudo-differential equations, which is interesting to tackle equations from incompressible fluid mechanics. In microlocal analysis, different kinds of quantizations are available, see for instance to [29, Chapter 4]. Here we consider the Weyl quantization, which associates to a symbol $a = a(x,\xi)$ in a suitable class the semiclassical pseudo-differential operator $\mathrm{op}^{\mathrm{w}}_\varepsilon[a]$ defined through its action on functions $u \in L^2(\mathbb{R}^d)$, by

$$\big(\mathrm{op}^{\mathrm{w}}_\varepsilon[a]u\big)(x) := (2\pi\varepsilon)^{-d}\int e^{\frac{i(x-y)\cdot\xi}{\varepsilon}} a\left(\frac{x+y}{2},\xi\right) u(y)\,\mathrm{d}\xi\,\mathrm{d}y.$$

Throughout this paper, we will simply assume every symbol to be in the class $S(1)$ of functions bounded as well as all their derivatives, defined by:

$$S(1) := \Big\{a \in C^\infty(\mathbb{R}^{2d}) : \forall\, \alpha \in \mathbb{N}^{2d},\ \exists\, C_\alpha > 0,\ \forall\,(x,\xi)\in\mathbb{R}^{2d},\ \|\partial^\alpha a\|_\infty \le C_\alpha\Big\}. \tag{38}$$

Wider classes of symbols may be studied, but they are not required for the analysis which we perform in this paper. With that choice of class, the pseudo-differential operator given above defines a linear continuous operator from $L^2\big(\mathbb{R}^d;\mathbb{C}^N\big)$ to itself, and there exists a constant $C > 0$, uniform in ε, as well as a universal constant γ, such that, for all $a \in S(1)$, it holds

$$\big\|\mathrm{op}^{\mathrm{w}}_\varepsilon[a]\big\|_{\mathscr{L}(L^2;L^2)} \le C \sup_{|\alpha|+|\beta|\le\gamma d}\big\|\partial^\alpha_x\partial^\beta_\xi a\big\|_\infty, \tag{39}$$

where $\mathscr{L}(L^2;L^2)$ denotes the space of linear bounded operators from L^2 to itself, endowed with the usual operator norm. From now on, we denote for k in $\mathbb{N}$ by $\mathscr{N}_k$ the norm on $S(1)$:

$$\mathscr{N}_k(\cdot) := \sup_{|\alpha|+|\beta|\le k}\big\|\partial^\alpha_x\partial^\beta_\xi\cdot\big\|_\infty. \tag{40}$$

We list below some useful properties of calculus in the Weyl quantization. More details, other choices of quantizations and other classes of symbols can be found in [25,29].

Lemma 13 (Adjoint operator). *Let $a \in S(1)$ be a symbol. Then the adjoint operator of* $\mathrm{op}^{\mathrm{w}}_\varepsilon[a]$ *is given by*

$$\big(\mathrm{op}^{\mathrm{w}}_\varepsilon[a]\big)^* = \mathrm{op}^{\mathrm{w}}_\varepsilon[a^*]. \tag{41}$$

Lemma 14 (Composition of operators). *Let $a, b \in S(1)$ be two symbols. Then the following composition law holds:*

$$\mathrm{op}^{\mathrm{w}}_\varepsilon[a]\,\mathrm{op}^{\mathrm{w}}_\varepsilon[b] = \mathrm{op}^{\mathrm{w}}_\varepsilon[ab] + \frac{\varepsilon}{2i}\mathrm{op}^{\mathrm{w}}_\varepsilon\big[\{a,b\}\big] + O(\varepsilon^2)_{\mathscr{L}(L^2;L^2)}, \tag{42}$$

where $\{\cdot,\cdot\}$ *denotes the Poisson bracket*

$$\{a,b\} = (\partial_\xi a)(\partial_x b) - (\partial_x a)(\partial_\xi b).$$

The Wigner transform is linked to the semiclassical Weyl quantization by the relation:

$$\iint_{\mathbb{R}^d\times\mathbb{R}^d}\big(W^\varepsilon[u,v]:a(x,\xi)\big)\,\mathrm{d}x\mathrm{d}\xi=\big\langle \mathrm{op}_\varepsilon^{\mathrm{w}}[a^*]u,v\big\rangle_{L^2}, \tag{43}$$

for any $a(x,\xi)$ in $S(1)$, and any $u,v\in L^2(\mathbb{R}^d;\mathbb{C}^N)$. This property allows us to view Wigner matrices as tempered distributions in $S(1)'$, the dual space of $S(1)$. Indeed, for any $u\in L^2(\mathbb{R}^d;\mathbb{C}^N)$, combining the property (43) with the estimate (39) yields the existence of a constant $C=C(u)>0$, uniform in ε, such that for any $a\in S(1)$ it holds:

$$\left|\iint_{\mathbb{R}^d\times\mathbb{R}^d}W^\varepsilon[u](x,\xi):a(x,\xi)\,\mathrm{d}x\,\mathrm{d}\xi\right|\le C\mathcal{N}_{\gamma d}(a).$$

This allows us to consider $W^\varepsilon[u]$ as a distribution of order at most γd, where $W^\varepsilon[u]=W^\varepsilon[u,u]$ as in Definition 9. More precisely, for any $u,v\in L^2(\mathbb{R}^d;\mathbb{C}^N)$, we have $W^\varepsilon[u,v]\in S(1)'_{\gamma d}(\mathbb{R}^{2d})$, where

$$S(1)'_k(\mathbb{R}^{2d})=\Big\{\psi(x,\xi)\in S(1)' : \exists\, C>0,\ \forall\,\varphi\in S(1),\ \big|\langle\psi,\varphi\rangle_{\mathcal{S}',\mathcal{S}}\big|\le C\mathcal{N}_k(\varphi)\Big\}.$$

We endow $S(1)'_k(\mathbb{R}^{2d})$ with the norm

$$\|\psi\|_{\mathcal{S}'_k}:=\sup_{\phi\in S(1),\mathcal{N}_k(\phi)\le1}\langle\psi,\phi\rangle_{S(1)',S(1)}.$$

For any $u,v\in L^2(\mathbb{R}^d;\mathbb{C}^N)$, and $a\in S(1)$,

$$W^\varepsilon\big[\mathrm{op}_\varepsilon^{\mathrm{w}}[a]u,v\big]=aW^\varepsilon[u,v]+\frac{\varepsilon}{2i}\big\{a,W^\varepsilon[u,v]\big\}+\varepsilon^2\mathcal{R}_\varepsilon, \tag{44}$$

where the remainder term $\mathcal{R}_\varepsilon\in S(1)'_{\gamma d+2}(\mathbb{R}^{2d})$ verifies

$$\|\mathcal{R}_\varepsilon\|_{S(1)'_{\gamma d+2}}=O(1).$$

4.3. *Statement of the result*

With the previous material, it is natural to extend our setting to hyperbolic pseudo-differential equations of the form

$$\varepsilon\partial_t\mathbf{u}+\mathrm{op}_\varepsilon^{\mathrm{w}}\big[iH(x,\xi)\,\mathrm{Id}+\varepsilon A(x,\xi)\big]\mathbf{u}=0,\quad \text{ on } [0,T], \tag{45}$$

where the principal part H is a real scalar symbol in $S(1)$ and the sub-principal part A is a matrix-valued symbol in $S(1)$.

Our setting also allows to consider a wider class of observability operators given as well by pseudo-differential operators associated with a symbol matrix B in $S(1)$. Thus we are led to the following extension of the notion of observability cost $\mathcal{C}[\mathbf{u}_0]$ of an initial data $\mathbf{u}_0$ at time T for the equation (45) by the pseudo-differential observability operator of symbol B in $S(1)$.

Definition 15. *Let $T>0$. Let $(x_0,\xi_0,b_0)\in\mathbb{R}^d\times\big(\mathbb{R}^d\setminus\{0\}\big)\times\mathbb{R}^N$ and $\mathbf{u}_0$ in $L^2(\mathbb{R}^d;\mathbb{R}^N)$. The observability cost $\mathcal{C}[\mathbf{u}_0]$ of the initial data $\mathbf{u}_0$ at time T for the equation* (45) *by the observability operator of symbol B in $S(1)$ is defined by*

$$\mathcal{C}[\mathbf{u}_0]:=\frac{\|\mathbf{u}(T,\cdot)\|_{L^2}^2}{\int_0^T\|\mathrm{op}_\varepsilon^{\mathrm{w}}[B]\mathbf{u}(t)\|_{L^2}^2\,\mathrm{d}t}, \tag{46}$$

where $\mathbf{u}$ is the solution of (45) *associated with $\mathbf{u}_0$.*

The purpose of this section is to study the observability cost of initial data which are localized at $(x_0,\xi_0,\mathbf{b})$ in the following sense.

Definition 16. *A family $(\mathbf{u}_0^\varepsilon)_{\varepsilon\in(0,1]}$ in $L^2(\mathbb{R}^d;\mathbb{R}^N)$ is said to be localized at $(x_0,\xi_0,b_0)\in\mathbb{R}^d\times\big(\mathbb{R}^d\setminus\{0\}\big)\times\mathcal{S}^{N-1}$ if its* Wigner matrix *satisfies*

$$\mathcal{W}^\varepsilon[\mathbf{u}_0^\varepsilon]\xrightarrow{w*}(b_0\otimes b_0)\delta_{x_0,\xi_0}\quad \textit{in } \big(\mathcal{S}(1)\big)',\quad \textit{as } \varepsilon\to0.$$

The question at stake is then the following.

Question 3. *Let $T>0$. Let $(x_0,\xi_0,b_0)\in\mathbb{R}^d\times(\mathbb{R}^d\setminus\{0\})\times\mathbb{R}^N$ and a family $(\mathbf{u}_0^\varepsilon)_{\varepsilon\in(0,1]}$ in $L^2(\mathbb{R}^d;\mathbb{R}^N)$ localized at (x_0,ξ_0,b_0). Can we determine the limit of the observability cost $\mathscr{C}[\mathbf{u}_0^\varepsilon]$ as $\varepsilon\to0$?*

The result below provides an answer in terms of the Hamiltonian flow induced by the principal symbol H, the microlocalized observability Gramian matrices, the bicharacteristic equations and the bicharacteristic amplitude equation whose definitions are now given.

Definition 17. *Let H be a real scalar symbol in $S(1)$ and v_H the divergence-free vector field from $\mathbb{R}^d\times\mathbb{R}^d$ to $\mathbb{R}^d\times\mathbb{R}^d$ defined by*

$$v_H:=\big(\nabla_\xi H,-\nabla_x H\big). \tag{47}$$

For any $x_0,\xi_0\in\mathbb{R}^d$, we define the Hamiltonian flow $\big(X_{x_0,\xi_0}(t),\Xi_{x_0,\xi_0}(t)\big)$ induced by H as the unique solution to the nonlinear ODE:

$$\big(X_{x_0,\xi_0},\Xi_{x_0,\xi_0}\big)'=v_H\big((X_{x,\xi},\Xi_{x_0,\xi_0})\big),\quad \big(X_{x_0,\xi_0},\Xi_{x_0,\xi_0}\big)(0)=(x_0,\xi_0). \tag{48}$$

Remark 18. It is worth noticing that the determinant $J(t):=\det\big(D\big(X_{x_0,\xi_0}(t),\Xi_{x_0,\xi_0}(t)\big)\big)$ of the Jacobian matrix associated with the Hamiltonian flow satisfies $J(t)=1$ for all $t\in[0,T]$ since $J'(t)=\mathbf{div}\big(\nabla_\xi H,-\nabla_x H\big)J(t)=0$ and $J(0)=1$. Here and below we use the notation **div** to denote the divergence operator in $\mathbb{R}^d\times\mathbb{R}^d$.

We now define the microlocalized observability Gramian matrices, as solutions of some differential Lyapunov equations involving the coefficients of (45) along the Hamiltonian flow.

Definition 19. *For any $x_0,\xi_0\in\mathbb{R}^d$, we define the Gramian matrix microlocalized at (x_0,ξ_0) as the unique solution G_{x_0,ξ_0} to the differential Lyapunov equation* $\mathrm{Lyap}\big(A_{x_0,\xi_0},B_{x_0,\xi_0}\big)$ *where the time-dependent matrices A_{x_0,ξ_0} and B_{x_0,ξ_0} are given by*

$$A_{x_0,\xi_0}(t):=A\big(X_{x_0,\xi_0}(t),\Xi_{x_0,\xi_0}(t)\big)\quad\text{and}\quad B_{x_0,\xi_0}(t):=B\big(X_{x_0,\xi_0}(t),\Xi_{x_0,\xi_0}(t)\big).$$

Finally we define the bicharacteristic amplitude.

Definition 20. *For any $x_0,\xi_0\in\mathbb{R}^d$, for any $\mathbf{b}_0$ in $\mathbb{R}^N$, we define the bicharacteristic amplitude $\mathbf{b}_{x_0,\xi_0,\mathbf{b}_0}(t)$ associated with H and A as the unique solution to the linear ODE:*

$$\mathbf{b}'_{x_0,\xi_0,\mathbf{b}_0}=-A_{x_0,\xi_0}\,\mathbf{b}_{x_0,\xi_0,\mathbf{b}_0},\quad \mathbf{b}_{x_0,\xi_0,\mathbf{b}_0}(0)=\mathbf{b}_0. \tag{49}$$

The result below answers Question 3 and identifies the limit of the observability cost $\mathscr{C}[\mathbf{u}_0^\varepsilon]$ of a family $(\mathbf{u}_0^\varepsilon)_{\varepsilon\in(0,1]}$ in $L^2(\mathbb{R}^d;\mathbb{R}^N)$ localized at (x_0,ξ_0,b_0) as $\varepsilon\to0$ in terms of the Rayleigh quotient of the microlocalized observability Gramian matrix along the bicharacteristic flow with the bicharacteristic amplitude, at the final time.

Theorem 21. *Let $T>0$. Let $x_0\in\mathbb{R}^d$, $\xi_0\in\mathbb{R}^d\setminus\{0\}$ and set $G_T:=G_{x_0,\xi_0}(T)$, where G_{x_0,ξ_0} is the Gramian matrix microlocalized at (x_0,ξ_0) as given by Definition 19. Let $\mathbf{b}_0\in\mathscr{S}^{N-1}$ and set*

$$\mathbf{b}_T^{\mathscr{N}}:=\frac{\mathbf{b}_{x_0,\xi_0,\mathbf{b}_0}(T)}{|\mathbf{b}_{x_0,\xi_0,\mathbf{b}_0}(T)|},$$

where $\mathbf{b}_{x_0,\xi_0,\mathbf{b}_0}$ is the transported amplitude associated with H and A given by Definition 20. Let $(\mathbf{u}_0^\varepsilon)_{\varepsilon\in(0,1]}\in C^\infty(\mathbb{R}^d)\cap L^2(\mathbb{R}^d;\mathbb{R}^N)$ be localized at (x_0,ξ_0,b_0) in the sense of Definition 16. Then the observability cost $\mathscr{C}[\mathbf{u}_0^\varepsilon]$ satisfies

$$\mathscr{C}[\mathbf{u}_0^\varepsilon]^{-1}\to(\mathbf{b}_T^{\mathscr{N}})^*\,G_T\,\mathbf{b}_T^{\mathscr{N}},\quad \text{as } \varepsilon\to0.$$

Let us highlight that the interest of the result above is that the cost of observability of initial data which are localized at (x_0,ξ_0,b_0) converges to a limit which can be determined by ODEs. It can be considered as a broad extension of the result obtained in the case of transport equations in Theorem 7 to general hyperbolic pseudo-differential equations of the form (45), in the high frequency limit, despite the extra dispersive effects. As a matter of fact, the equation (12) can be recast as an equation of the form (45) with

$$H(x,\xi) = v(x)\cdot\xi \quad \text{and} \quad A(x,\xi) := A(x) - \frac{1}{2}\big(\operatorname{div} v(x)\big)\,\mathrm{Id}. \tag{50}$$

Then the bicharacteristic equations (48) reduce to

$$X'_{x_0,\xi_0}(t) = v\big(X_{x_0,\xi_0}(t)\big) \quad \text{and} \quad \Xi'_{x_0,\xi_0}(t) = -\big(\nabla_x v\big(X_{x_0,\xi_0}(t)\big)\big)^*\,\Xi_{x_0,\xi_0}(t) \quad \text{for } t\in(0,T). \tag{51}$$

In particular, the first equation does not depend on ξ, translating the lack of dispersion in the case of transport equations. Moreover the matrix $A(x,\xi)$ defined in (50) is the one which is involved in the Lyapunov equation satisfied by the Gramian matrix G_{x_0}, see (15).

The result stated in Theorem 21 will be extended to a more general setting below, see Theorem 28, and we refer to the comment after this statement, in particular for these results to be placed in perspective.

4.4. *Proof of Theorem 21*

For any (x,ξ), let $G_{x,\xi}$ be the Gramian matrix microlocalized at (x,ξ) as given by Definition 19. We define the symbolic multiplier $M(t,x,\xi)$ by setting, for any (t,x,ξ),

$$G_{x,\xi}(t) =: M\big(t,\big(X_{x,\xi}(t),\Xi_{x,\xi}(t)\big)\big). \tag{52}$$

As the Hamiltonian flow is invertible, see Remark 18, the above M is uniquely defined. Similarly to the reformulation in (34) and (35) performed in the non-dispersive case in Section 3, we are going to use a reformulation of the observability cost in terms of a recorded energy by the use of a multiplier. The latter is precisely chosen as the symbolic multiplier $M(t,x,\xi)$, and the corresponding recorded energy is then set as

$$\iint_{\mathbb{R}^d\times\mathbb{R}^d} \mathscr{W}^\varepsilon(t,x,\xi) : M(t,x,\xi)\,\mathrm{d}x\mathrm{d}\xi = \Big\langle \mathrm{op}^{\mathrm{w}}_\varepsilon\big[M(t,x,\xi)\big]\mathbf{u},\mathbf{u}\Big\rangle_{L^2(\mathbb{R}^d)}, \tag{53}$$

where we use the shorthand notation $\mathscr{W}^\varepsilon$ for the Wigner matrix transform $\mathscr{W}^\varepsilon[\mathbf{u}^\varepsilon]$, recalling the property (43).

Let $T>0$ and $(\mathbf{u}_0^\varepsilon)_{\varepsilon\in(0,1]}$ be bounded in $L^2\big(\mathbb{R}^d;\mathbb{R}^N\big)$. A key point of the proof of Theorem 21 is to establish that, as $\varepsilon\to 0$, the observability cost $\mathscr{C}[\mathbf{u}_0^\varepsilon]$ satisfies

$$\mathscr{C}[\mathbf{u}_0^\varepsilon] = \frac{\int_{\mathbb{R}^d\times\mathbb{R}^d} \mathscr{W}^\varepsilon(T,\cdot) : \mathrm{Id}\,\mathrm{d}x\,\mathrm{d}\xi}{\int_{\mathbb{R}^d\times\mathbb{R}^d} \mathscr{W}^\varepsilon(T,\cdot) : M(T,\cdot)\,\mathrm{d}x\,\mathrm{d}\xi} + o(1). \tag{54}$$

Recalling (37), we already have:

$$\big\|\mathbf{u}(T,\cdot)\big\|^2_{L^2} = \iint_{\mathbb{R}^d\times\mathbb{R}^d} \mathscr{W}^\varepsilon(T,\cdot) : \mathrm{Id}\,\mathrm{d}x\mathrm{d}\xi. \tag{55}$$

The reformulation of the denominator is more subtle. First, by (43), the properties of the Frobenius product and (44), we observe that, at any time,

$$\big\|\mathrm{op}^{\mathrm{w}}_\varepsilon[B]\mathbf{u}\big\|^2_{L^2} = \iint_{\mathbb{R}^d\times\mathbb{R}^d} \big(\mathscr{W}^\varepsilon : B^*B\big)\,\mathrm{d}x\mathrm{d}\xi + O(\varepsilon). \tag{56}$$

Then, using the chain rule and (48), we observe that

$$G'_{x,\xi}(t) = \big(\partial_t M + \mathbf{div}(M v_H)\big)\big(t, X_{x,\xi}(t),\Xi_{x,\xi}(t)\big),$$

so that, since $G_{x,\xi}$ verifies the differential Lyapunov equation $\mathrm{Lyap}\big(A_{x,\xi},B_{x,\xi}\big)$, we infer that the symbolic multiplier M satisfies the following PDE:

$$\begin{cases} \partial_t M + \mathbf{div}(M v_H) - MA - A^*M = B^*B, \\ \qquad\qquad\qquad\qquad\qquad\quad M(0) = 0, \end{cases} \tag{57}$$

where B^*B appears as a source term. On the other hand, the dynamics of the Wigner matrix $\mathscr{W}^\varepsilon$ is given by the following result.

Lemma 22. *There exists a remainder term $\mathscr{R}_\varepsilon$ in $(S(1))'_{\gamma d+2}$ with $\|\mathscr{R}_\varepsilon\|_{\mathscr{S}'_{\gamma d+2}} = O(\varepsilon)$ such that*

$$\partial_t \mathscr{W}^\varepsilon + \mathbf{div}(\mathscr{W}^\varepsilon v_H) + A\mathscr{W}^\varepsilon + \mathscr{W}^\varepsilon A^* = \mathscr{R}_\varepsilon. \tag{58}$$

Proof. By bilinearity and the symmetry property (36), the Wigner matrix transform $\mathscr{W}^\varepsilon$ satisfies

$$\partial_t \mathscr{W}^\varepsilon = W^\varepsilon[\partial_t \mathbf{u}, \mathbf{u}] + W^\varepsilon[\partial_t \mathbf{u}, \mathbf{u}]^*. \tag{59}$$

Moreover, by (45), it follows that

$$W^\varepsilon[\partial_t \mathbf{u}, \mathbf{u}] = -W^\varepsilon\left[\mathrm{op}_\varepsilon^{\mathrm{w}}\left[\frac{i}{\varepsilon} H(x,\xi)\,\mathrm{Id} + A(x,\xi)\right]\mathbf{u}, \mathbf{u}\right],$$

so that, using the property (44), we arrive at

$$W^\varepsilon[\partial_t \mathbf{u}, \mathbf{u}] = -\frac{i}{\varepsilon} H\mathscr{W}^\varepsilon - \frac{1}{2}\{H, \mathscr{W}^\varepsilon\} - A\mathscr{W}^\varepsilon + O_{(S(1))'_{\gamma d+2}}(\varepsilon).$$

Substituting into (59) and using that H is a real scalar symbol, we obtain

$$\partial_t \mathscr{W}^\varepsilon + \{H\,\mathrm{Id}, \mathscr{W}^\varepsilon\} + A\mathscr{W}^\varepsilon + \mathscr{W}^\varepsilon A^* = O_{(S(1))'_{\gamma d+2}}(\varepsilon). \tag{60}$$

Recalling that v_H is the divergence-free vector field defined by (47), the conclusion follows. □

As in the proof of Theorem 7 we rely on a duality argument given by the following result.

Lemma 23. *Let $W(t,x,\xi)$ and $M(t,x,\xi)$ be two matrix-valued time-dependent symbols such that*

$$\begin{cases} \partial_t W + \mathbf{div}(W v_H) + AW + WA^* = \mathscr{R}, \\ \partial_t M + \mathbf{div}(M v_H) - MA - A^* M = \mathscr{F}. \end{cases}$$

Then

$$\left(\iint_{\mathbb{R}^d\times\mathbb{R}^d} \big(W : M\big)\,\mathrm{d}x\mathrm{d}\xi\right)(T) = \left(\iint_{\mathbb{R}^d\times\mathbb{R}^d} \big(W : M\big)\,\mathrm{d}x\mathrm{d}\xi\right)(0) + \int_0^T \left(\iint_{\mathbb{R}^d\times\mathbb{R}^d} \big(\mathscr{R} : M + W : \mathscr{F}\big)\,\mathrm{d}x\mathrm{d}\xi\right)(t)\,\mathrm{d}t. \tag{61}$$

Proof. We observe that the time derivative E' of

$$E(t) := \iint_{\mathbb{R}^d\times\mathbb{R}^d} M : W\,\mathrm{d}x\,\mathrm{d}\xi,$$

satisfies

$$E'(t) = \iint_{\mathbb{R}^d\times\mathbb{R}^d} \big(-\mathbf{div}(M v_H) + MA + A^* M + \mathscr{F}\big) : W\,\mathrm{d}x\,\mathrm{d}\xi + \iint_{\mathbb{R}^d\times\mathbb{R}^d} M : \big(-\mathbf{div}(W v_H) - AW - WA^* + \mathscr{R}\big)\,\mathrm{d}x\,\mathrm{d}\xi. \tag{62}$$

Using that $\mathbf{div}(v_H) = 0$, an integration by parts yields

$$\iint_{\mathbb{R}^d\times\mathbb{R}^d} \mathbf{div}(M v_H) : W\,\mathrm{d}x\,\mathrm{d}\xi = -\iint_{\mathbb{R}^d\times\mathbb{R}^d} M : \mathbf{div}(W v_H)\,\mathrm{d}x\,\mathrm{d}\xi.$$

For the other terms, simple computations give

$$(MA + A^* M) : W = M : (AW + WA^*).$$

Using those two equalities in (62) leads to

$$E'(t) = \iint_{\mathbb{R}^d\times\mathbb{R}^d} \big(\mathscr{F} : W + M : \mathscr{R}\big)\,\mathrm{d}x\,\mathrm{d}\xi. \tag{63}$$

Integrating this equality in time, from 0 to T, we conclude the proof of Lemma 23. □

We now apply Lemma 23 with $W = \mathscr{W}^\varepsilon$ and M, which yields, thanks to (57) and Lemma 22,

$$\left(\iint_{\mathbb{R}^d\times\mathbb{R}^d} \big(\mathscr{W}^\varepsilon : M\big)\,\mathrm{d}x\mathrm{d}\xi\right)(T) = \int_0^T \left(\iint_{\mathbb{R}^d\times\mathbb{R}^d} \big(\mathscr{W}^\varepsilon : B^*B\big)\,\mathrm{d}x\mathrm{d}\xi\right)(t)\,\mathrm{d}t + O(\varepsilon). \tag{64}$$

From (55) and (64), we deduce (54).

Now let $x_0 \in \mathbb{R}^d, \xi_0 \in \mathbb{R}^d\setminus\{0\}$ and $\mathbf{b}_0 \in \mathscr{S}^{N-1}$ and assume that $(\mathbf{u}_0^\varepsilon)_{\varepsilon\in(0,1]}$ is localized at (x_0,ξ_0,b_0) in the sense of Definition 16. To pass to the limit in (54), we rely on the following intermediate result.

Lemma 24. *Let $n, N \in \mathbb{N}^*$ and $T > 0$. Let V be a regular field of vector fields from $\mathbb{R}^d\times\mathbb{R}^d$ to itself, $A, \widetilde{A} \in S(1)$ be two smooth vector fields of matrices in $\mathbb{R}^{N\times N}$, and $(\mathscr{R}_\varepsilon)_{\varepsilon\in(0,1]} \subset (S(1))'_{\gamma d+2}$ such that*

$$\lim_{\varepsilon\to 0}\left\|\mathscr{R}_\varepsilon(t)\right\|_{\mathscr{S}'_{\gamma d+2}} = 0, \ \textit{uniformly in time.} \tag{65}$$

Let

$$\mathfrak{C} := C\big([0,T];(S(1))' - w*\big),$$

the space of the distributions in the dual space to $S(1)$, depending on time in a continuous way when $(S(1))'$ is endowed by its weak- topology. Then*

(1) *For any μ_0 in $(S(1))'$, ε in $(0,1]$, there exists a unique solution $\mu \in \mathfrak{C}$, to the PDE:*

$$\partial_t\mu + \mathbf{div}(\mu V) + A\mu + \mu\widetilde{A} = \mathscr{R}_\varepsilon, \tag{66}$$

with initial value μ_0.

(2) *If the initial data μ_0 is of the form $\mu_0^{x_0,\xi_0,P_0} = P_0\delta_{x_0,\xi_0}$ with $x_0,\xi_0 \in \mathbb{R}^d$ and $P_0 \in \mathbb{R}^{N\times N}$ then the unique corresponding solution to the homogeneous equation is given by*

$$\mu^{x_0,\xi_0,P_0}(t,\cdot) = P(t)\delta_{x(t),\xi(t)}, \tag{67}$$

where (x,ξ) solves the nonlinear ODE:

$$(x,\xi)' = V(x,\xi) \quad \textit{and} \quad (x,\xi)(0) = (x_0,\xi_0), \tag{68}$$

and P solves the linear ODE:

$$P' + A\big(x(\cdot),\xi(\cdot)\big)P + P\widetilde{A}\big(x(\cdot),\xi(\cdot)\big) = 0 \quad \textit{and} \quad P(0) = P_0. \tag{69}$$

(3) *Finally, if $(\mu_0^\varepsilon)_{\varepsilon\in(0,1]}$ converges weakly-* to $\mu_0^{x_0,\xi_0,P_0}$ in $(S(1))'$ as ε goes to 0, then the corresponding solution $(\mu^\varepsilon)_{(0,1]}$ converges to μ^{x_0,ξ_0,M_0} in $\mathfrak{C}$.*

Proof. The existence of the solution, as well as the form of the solution associated with the initial datum $M_0\delta_{x_0,\xi_0}$, follows from the characteristics method. The uniqueness and stability can be proved by a standard duality argument. □

By the second item of Lemma 24, for any $(x_0,\xi_0,\mathbf{b}_0) \in \mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^N$, the unique solution μ^{x_0,ξ_0,b_0} to

$$\partial_t\mu^{x_0,\xi_0,b_0} + \mathbf{div}\big(\mu^{x_0,\xi_0,b_0}v_H\big) + A\mu^{x_0,\xi_0,b_0} + \mu^{x_0,\xi_0,b_0}A^* = 0, \tag{70}$$

associated with the initial data $\mu_0^{x_0,\xi_0,b_0} := (\mathbf{b}_0\otimes\mathbf{b}_0)\delta_{x_0,\xi_0}$ is

$$\mu^{x_0,\xi_0,b_0} = \big(\mathbf{b}_{x_0,\xi_0,b_0}(t)\otimes\mathbf{b}_{x_0,\xi_0,b_0}(t)\big)\delta_{X_{x_0,\xi_0}(t),\Xi_{x_0,\xi_0}(t)},$$

where we recall that the Hamiltonian flow $\big(X_{x_0,\xi_0}(t),\Xi_{x_0,\xi_0}(t)\big)$ is defined by the equations (48) and (49). Moreover, by the third item of Lemma 24, since, as $\varepsilon\to 0$, $\mathscr{W}^\varepsilon \xrightarrow{w*} (b_0\otimes b_0)\delta_{x_0,\xi_0}$ in $\big(\mathscr{S}(1)\big)'$, we have that, as $\varepsilon\to 0$,

$$\mathscr{W}^\varepsilon(T) \xrightarrow{w*} \big(\mathbf{b}_{x_0,\xi_0,\mathbf{b}_0}(t)\otimes\mathbf{b}_{x_0,\xi_0,\mathbf{b}_0}(T)\big)\delta_{X_{x_0,\xi_0}(T),\Xi_{x_0,\xi_0}(T)}. \tag{71}$$

Passing to the limit in (54) and recalling (52) concludes the proof of Theorem 21.

5. Application to the incompressible Euler equations linearized at a Couette flow in velocity form for a constant non-invertible observability matrix

An interesting example of application of our method is represented by the incompressible Euler equations in $\mathbb{R}^2$, in velocity form, which we linearize at a steady state $\bar{\mathbf{u}}$ satisfying $\mathrm{div}(\bar{\mathbf{u}}) = 0$, so that the velocity field $\mathbf{u} = \mathbf{u}(t,x) \in \mathbb{R}^N$ and the pressure $p = p(t,x) \in \mathbb{R}$ satisfy

$$\begin{cases} \partial_t \mathbf{u} + (\bar{\mathbf{u}} \cdot \nabla)\mathbf{u} + (\mathbf{u} \cdot \nabla)\bar{\mathbf{u}} + \nabla p = 0, \\ \mathrm{div}(\mathbf{u}) = 0. \end{cases} \tag{72}$$

Using the Leray projection onto divergence-free vector fields, the semiclassical approximation of (72) then reads

$$\varepsilon \partial_t \mathbf{u} + \mathrm{op}_\varepsilon^{\mathrm{w}}\left[iH(x,\xi)\,\mathrm{Id} + \varepsilon A(x,\xi)\right]\mathbf{u} = 0, \tag{73}$$

where

$$H(x,\xi) = \bar{\mathbf{u}}(x) \cdot \xi, \quad A(x,\xi) = \left(\mathrm{Id} - 2\frac{\xi \otimes \xi}{|\xi|^2}\right) D\bar{\mathbf{u}}. \tag{74}$$

The formulation (73) is not strictly equivalent to (72), as it should also feature a remainder source term of lower order, which we may neglect here as we wish to study the high-frequency case.

Since we want to study the behavior of solutions that are localized at high frequencies, we may multiply the symbols A and H by a cut-off function that is equal to 0 near $\xi = 0$ to avoid the singularity of A at $\xi = 0$, so that one may consider a symbol A in $S(1)$ which is given by (74) for ξ outside of a ball of radius say $1/2$.

We consider the case where the steady state $\bar{\mathbf{u}}$ is the Couette flow $\bar{\mathbf{u}} = (x_2, 0)$, and where the observation matrix is

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

Then for any x_0, ξ_0 in $\mathbb{R}^2$, the Gramian matrix G_{x_0,ξ_0} microlocalized at (x_0,ξ_0) is given by

$$G_{x_0,\xi_0}(T) = \int_0^T \left(R_{x_0,\xi_0}(t,T)^* B^* B R_{x_0,\xi_0}(t,T)\right) \mathrm{d}t,$$

where $R_{x_0,\xi_0}(t,T)$ is the resolvent matrix, given by solving the linear system

$$\frac{\mathrm{d}}{\mathrm{d}t} R_{x_0,\xi_0}(t,T) = -A\big(X_{x_0,\xi_0}(t), \Xi_{x_0,\xi_0}(t)\big) R_{x_0,\xi_0}(t,T), \quad R(T,T) = \mathrm{Id}.$$

A tedious computation yields, with the notation $\xi_0 = \big(\xi_{1,0}, \xi_{2,0}\big)$,

$$R_{x_0,\xi_0}(t,T) = \begin{pmatrix} 1 & f(t,\xi_0) \\ 0 & g(t,\xi_0) \end{pmatrix},$$

where

$$f(t,\xi_0) = \begin{cases} T - t & \text{if } \xi_{1,0} = 0, \\ -\dfrac{\big(\xi_{2,0} - t\xi_{1,0}\big)\big(\xi_{1,0}^2 + \big(\xi_{2,0} - T\xi_{1,0}\big)^2\big)}{\xi_{1,0}\big(\xi_{1,0}^2 + \big(\xi_{2,0} - t\xi_{1,0}\big)^2\big)} + \dfrac{\xi_{2,0} - T\xi_{1,0}}{\xi_{1,0}} & \text{otherwise,} \end{cases}$$

and

$$g(t,\xi_0) = \begin{cases} 1 & \text{if } \xi_{1,0} = 0, \\ \dfrac{\xi_{1,0}^2 + \big(\xi_{2,0} - T\xi_{1,0}\big)^2}{\xi_{1,0}^2 + \big(\xi_{2,0} - t\xi_{1,0}\big)^2} & \text{otherwise.} \end{cases}$$

Moreover

$$G_{x_0,\xi_0}(T) = \int_0^T \begin{pmatrix} 1 & f(t,\xi_0) \\ f(t,\xi_0) & f^2(t,\xi_0) \end{pmatrix} \mathrm{d}t.$$

We now consider $(x_0,\xi_0,\mathbf{b}_0)\in\mathbb{R}^2\times\left(\mathbb{R}^2\setminus\{0\}\right)\times\mathscr{S}^1$ and $(\mathbf{u}_0^\varepsilon)_{\varepsilon\in(0,1]}\subset L^2\left(\mathbb{R}^2;\mathbb{R}^2\right)$ a family of initial data localized at $(x_0,\xi_0,\mathbf{b}_0)$. By Theorem 21, the observability cost $\mathscr{C}[\mathbf{u}_0^\varepsilon]$ satisfies, as ε goes to 0,

$$\lim_{\varepsilon\to 0}\left(\mathscr{C}[\mathbf{u}_0^\varepsilon]\right)^{-1}=(\mathbf{b}_T^{\mathscr{N}})^*G_{x_0,\xi_0}(T)\mathbf{b}_T^{\mathscr{N}},$$

with

$$\mathbf{b}_T^{\mathscr{N}}=\frac{\mathbf{b}_T}{|\mathbf{b}_T|},\quad \mathbf{b}_T=R_{x_0,\xi_0}(T,0)\mathbf{b}_0=\begin{pmatrix}1 & -\frac{f(0,\xi_0)}{g(0,\xi_0)}\\ 0 & \frac{1}{g(0,\xi_0)}\end{pmatrix}\mathbf{b}_0.$$

We can notice that, for any value of $\xi_0\neq 0$, the function $f(\cdot,\xi_0)$ is continuous and not constant, so that, by the Cauchy–Schwarz inequality,

$$\det G_{x_0,\xi_0}(T)=T\int_0^T f^2(t,\xi_0)\,\mathrm{d}t-\left(\int_0^T f(t,\xi_0)\,\mathrm{d}t\right)^2>0.$$

Thus the system is observable in high frequencies for initial data localized at any $(x_0,\xi_0,\mathbf{b}_0)$, and for any positive time $T>0$. The limit of the observability cost as ε goes to 0 is finite, and can be explicitly computed in terms of ξ_0 and $\mathbf{b}_0$. The short-time and long-time behaviors of the observability cost can be given by further computations. First, in the case where $\xi_{1,0}=0$, the Gramian matrix has the simple expression:

$$G_{x_0,\xi_0}(T)=\begin{pmatrix}T & \frac{T^2}{2}\\ \frac{T^2}{2} & \frac{T^3}{3}\end{pmatrix}.$$

Immediate computations then yield

$$\lambda_{\min}\left(G_{x_0,\xi_0}(T)\right)=\frac{T}{2}+\frac{1}{6}\left(T^3-\sqrt{T^6+3T^4+9T^2}\right),$$

so that in particular $\lambda_{\min}\left(G_{x_0,\xi_0}(T)\right)^{-1}\sim\frac{12}{T^3}$ for short times and $\lambda_{\min}\left(G_{x_0,\xi_0}(T)\right)^{-1}\sim\frac{4}{T}$ for long times. Note that $\lambda_{\min}\left(G_{x_0,\xi_0}(T)\right)^{-1}$ is the "worst-case" cost for initial data localized at (x_0,ξ_0), namely

$$\lambda_{\min}\left(G_{x_0,\xi_0}(T)\right)^{-1}=\sup_{\mathbf{b}_0\in\mathscr{S}^{N-1}}\left\{\lim_{\varepsilon\to 0}\mathscr{C}(\mathbf{u}_0^\varepsilon)\,:\,(\mathbf{u}_0^\varepsilon)\text{ is localized at }(x_0,\xi_0,\mathbf{b}_0)\right\}.$$

The asymptotic behaviour of the cost at short times can be extended to the other values of ξ_0. If $\xi_{1,0}\neq 0$ and $\xi_{1,0}^2\neq\xi_{2,0}^2$, then by using the Taylor expansion of f at first order, we obtain

$$\lambda_{\min}\left(G_{x_0,\xi_0}(T)\right)\sim\frac{T^3\left(\xi_{1,0}^2-\xi_{2,0}^2\right)^2}{12|\xi_0|^4}.\tag{75}$$

Otherwise, that is if $\xi_{1,0}\neq 0$ and $\xi_{1,0}^2=\xi_{2,0}^2$, then by using the Taylor expansion of f at second order, we obtain

$$\lambda_{\min}\left(G_{x_0,\xi_0}(T)\right)\sim\frac{T^5}{45}.\tag{76}$$

Those estimates are consistent with the results of [27]. Indeed, in the case where $\xi_{1,0}^2\neq\xi_{2,0}^2$, the rank condition is fulfilled on the second bracket, namely

$$\operatorname{rank}\left[B\;\middle|\;A_{x_0,\xi_0}(0)B\right]=2,\quad\text{ while }\operatorname{rank}(B)<2.$$

Then the analysis of [27] gives a cost of order T^{-3}, as given by (75). When $\xi_{1,0}^2=\xi_{2,0}^2$, we need to take one more bracket to verify the rank condition, namely

$$\operatorname{rank}\left[B\;\middle|\;A_{x_0,\xi_0}(0)\;\middle|\;\left(A_{x_0,\xi_0}^2(0)+A'_{x_0,\xi_0}(0)\right)B\right]=2,\quad\text{ while }\operatorname{rank}\left[B\;\middle|\;A_{x_0,\xi_0}(0)B\right]<2.$$

The results in [27] then yield a cost of order T^{-5}, as given by (76). Our approach furthermore gives an explicit constant factor.

6. Case of systems without crossing modes

This section is devoted to an extension of the results of Section 4 to more general systems.

6.1. *Statement of the result*

In this section we extend the previous analysis to the case of pseudo-differential hyperbolic systems of the form:

$$\partial_t \mathbf{u} + \mathrm{op}_\varepsilon^{\mathrm{w}}\left[\frac{i}{\varepsilon}H + A\right]\mathbf{u} = 0, \quad (x,\xi) \in \mathbb{R}^d \times \mathbb{R}^d. \tag{77}$$

Above, A and H are two matrices of symbols which are supposed to fulfill the following assumption.

Assumption. *We assume that A and H satisfy the following.*

(H1) *$H = H(x,\xi)$ is a $N \times N$ field of Hermitian matrices in $S(1)$, and there exist $1 \le m \le N$ and $C > 0$ such that for all $(x,\xi) \in \mathbb{R}^d \times \mathbb{R}^d$, the matrix $H(x,\xi)$ has m real semi-simple eigenvalues $\lambda_k(x,\xi)$, for $1 \leqslant k \leqslant m$, with $\left|\lambda_j(x,\xi) - \lambda_k(x,\xi)\right| > C$ for all (j,k) with $1 \le j \ne k \le m$.*

(H2) *$A = A(x,\xi)$ is a $N \times N$ field of matrices in $S(1)$.*

As a consequence of (H1) the matrix $\mathscr{U}(x,\xi) := (\mathbf{r}_\ell)_{\ell=1,\dots,N}$ of the right normalized eigenvectors $(\mathbf{r}_\ell)_{\ell=1,\dots,N}$ of $H(x,\xi)$ is in $S(1)$, takes values which are unitary matrices, and

$$\mathscr{U}^* H = \mathscr{D}\mathscr{U}^* = (\lambda_\ell \mathbf{r}_\ell)^*_{\ell=1,\dots,N}, \tag{78}$$

and

$$\mathscr{D}(x,\xi) := (\mathscr{U}^* H \mathscr{U})(x,\xi) = \sum_{k=1}^{m} \lambda_k(x,\xi)\Pi_k \in S(1), \tag{79}$$

where the $(\Pi_k)_{1\le k\le m}$ are the constant orthogonal projectors onto $\ker\big(\mathscr{D}(x,\xi) - \lambda_k(x,\xi)\,\mathrm{Id}\big)$, which are independent of $(x,\xi) \in \mathbb{R}^d$.

Our aim is to answer Question 3 in the case of the system (77), that is to determine, as $\varepsilon \to 0$, the limit observability cost $\mathscr{C}[\mathbf{u}_0^\varepsilon]$, as given by Definition 15, for a family of initial data $(\mathbf{u}_0^\varepsilon)_{\varepsilon\in(0,1]}$ in $L^2\big(\mathbb{R}^d;\mathbb{R}^N\big)$ which is localized at $(x_0,\xi_0,b_0) \in \mathbb{R}^d \times \big(\mathbb{R}^d \setminus \{0\}\big) \times \mathbb{R}^N$, for the corresponding solution $\mathbf{u}$ of the system (77) and an observability operator $B \in S(1)$. The main result of this section provides an answer to this question which generalizes the earlier case where a single characteristic was considered. Again the result involves a few ingredients, which we now recall in this more general setting. First it makes use of the Hamiltonian flows associated with each eigenvalue λ_k, with $1 \le k \le m$, as given by Definition 20. For the sake of clarity we explicitly recall this definition below, which involves nonlinear ODEs.

Definition 25. *For $1 \le k \le m$, for any (x,ξ), we define the Hamiltonian flow $\big(X_{k,x,\xi}, \Xi_{k,x,\xi}\big)$ induced by the k-th eigenvalue λ_k as the unique smooth solution of the following Cauchy problem:*

$$\big(X_{k,x,\xi}, \Xi_{k,x,\xi}\big)' = v_{\lambda_k}\big(X_{k,x,\xi}, \Xi_{k,x,\xi}\big), \quad \big(X_{k,x,\xi}, \Xi_{k,x,\xi}\big)(0) = (x,\xi), \tag{80}$$

where v_{λ_k} is the divergence-free vector field from $\mathbb{R}^d \times \mathbb{R}^d$ to itself defined by

$$v_{\lambda_k} = \big(\nabla_\xi \lambda_k, -\nabla_x \lambda_k\big). \tag{81}$$

We also rely on some Gramian matrices associated with each Hamiltonian flow and with H and A, similarly to Definition 19, but with some surprising additional terms, which encode the interactions between the various modes of the system. More precisely, for $1 \le k \le m$, let

$$A_k^{\mathrm{eff}} := \Pi_k \mathscr{U}^* A \mathscr{U} \Pi_k + \frac{1}{2}\Pi_k\Big(\{\mathscr{U}^*, H\}\mathscr{U} - \{\mathscr{D}, \mathscr{U}^*\}\mathscr{U}\Big)\Pi_k, \tag{82}$$

$$B_k^{\mathrm{eff}} := B\mathscr{U}\Pi_k, \tag{83}$$

which both belong to $S(1)$. We may now give the definition of the Gramian matrices, which involves linear ODEs.

Definition 26. *Let $x_0, \xi_0 \in \mathbb{R}^d$ and $1 \le k \le m$. We define the k-th Gramian matrix microlocalized at (x_0, ξ_0) as the unique solution G_{k,x_0,ξ_0} to the differential Lyapunov equation*

$$\mathrm{Lyap}\big(A_{k,x_0,\xi_0}, B_{k,x_0,\xi_0}\big),$$

where

$$A_{k,x_0,\xi_0}(t) := A_k^{\mathrm{eff}}\big(X_{k,x_0,\xi_0}(t), \Xi_{k,x_0,\xi_0}(t)\big) \quad \textit{and} \quad B_{k,x_0,\xi_0}(t) := B_k^{\mathrm{eff}}\big(X_{k,x_0,\xi_0}(t), \Xi_{k,x_0,\xi_0}(t)\big).$$

We also define the bicharacteristic amplitudes, which are given as solutions to linear ODEs.

Definition 27. *For any $x_0, \xi_0 \in \mathbb{R}^d$, for any $\mathbf{b}_0$ in $\mathbb{R}^N$ and $1 \le k \le m$, we define the k-th bicharacteristic amplitude $\mathbf{b}_{k,x_0,\xi_0,\mathbf{b}_0}$ as the unique solution to the linear ODE:*

$$\mathbf{b}'_{k,x_0,\xi_0,\mathbf{b}_0} = -A_{k,x_0,\xi_0}\,\mathbf{b}_{k,x_0,\xi_0,\mathbf{b}_0}, \quad \mathbf{b}_{k,x_0,\xi_0,\mathbf{b}_0}(0) = \Pi_k \mathbf{b}_0.$$

The main result of this section is that the inverse of the observability cost $\mathscr{C}[\mathbf{u}_0^\varepsilon]$ for a family of initial data $(\mathbf{u}_0^\varepsilon)_{\varepsilon\in(0,1]}$ in $L^2\big(\mathbb{R}^d; \mathbb{C}^N\big)$ which is localized at $(x_0, \xi_0, \mathbf{b}_0)$, in the sense of Definition 16, converges to a weighted mean of Rayleigh quotients of the m Gramian matrices with the bicharacteristic amplitudes.

Theorem 28. *Let $(x_0, \xi_0, b_0) \in \mathbb{R}^d \times \big(\mathbb{R}^d \setminus \{0\}\big) \times \mathbb{R}^N$. For $1 \le k \le m$, let $G_{T,k}$ be the k-th Gramian matrix given at time T, that is $G_{T,k} := G_{k,x_0,\xi_0}(T)$, and $\mathbf{b}_{T,k}$ the characteristic amplitude at time T, that is $\mathbf{b}_{T,k} := \mathbf{b}_{k,x_0,\xi_0,\mathbf{b}_0}(T)$. Let $(\mathbf{u}_0^\varepsilon)_{\varepsilon\in(0,1]} \in L^2\big(\mathbb{R}^d; \mathbb{R}^N\big)$ be localized at (x_0, ξ_0, b_0) in the sense of Definition 16.*

Then the observability cost $\mathscr{C}[\mathbf{u}_0^\varepsilon]$ verifies, as $\varepsilon \to 0$,

$$\mathscr{C}[\mathbf{u}_0^\varepsilon]^{-1} \longrightarrow \frac{\sum_{k=1}^m \mathbf{b}_{T,k}^* G_{T,k} \mathbf{b}_{T,k}}{\sum_{k=1}^m |\mathbf{b}_{T,k}|^2}. \tag{84}$$

A few comments are in order. First, the use of microlocal techniques in controllability theory is rather old; one may date it back to the pioneering papers [2,7,23], where the controllability of some hyperbolic equations, the wave equation and the Lamé system in elasticity in particular is related to a geometric optics condition, by using Wigner scalar or matrix transforms. Later on, an alternative approach has been proposed in [14] and pursued in [12,13,22], based on the Egorov theorem, leading to quantitative estimates of the controllability cost. There, the analysis covers the case of the coupled wave equations and of some Klein–Gordon type equations, by using a preliminary diagonalization which then allows to apply the scalar Egorov theorem to the resulting decoupled equations. Theorem 28 extends these results to the general case of hyperbolic systems, as long as there is no eigenvalues crossing. In the proof below, we extend the approach performed in the previous sections in terms of recorded energy and multiplier. Actually, some counterparts of the Egorov theorem for systems have been established in [1,5,9] in the context of quantum mechanics, and could be perhaps implemented in an alternative proof. This illustrates a quite ubiquitous dichotomy between an approach where one represents the solution in a relevant way to subsequently extract the desired information and another approach where one estimates an energy-type quantity to track the desired information along the time evolution. In the latter, some preliminary reductions are also necessary, see Lemma 29 and Lemma 30 below where we respectively diagonalize the principal part, which encodes propagation, and the subprincipal part which encodes the amplification or attenuation along the propagation. To do so, we follow a method, initially used by Poincaré in the setting of ODEs, to further diagonalize the system into a pseudo-differential normal form where the term of order $1/\varepsilon$ is diagonal and the term of order 1 is block-diagonal. However the diagonalization of the principal part creates some extra sub-principal terms, see the last term in (82), which encodes the interaction between modes and influences the cost. This will be illustrated in Section 7 in the case of compressible gas dynamics in a rapidly rotating frame. Let us point out that the apparition of a similar term can be found in [19,28] in different contexts.

6.2. *Proof of Theorem 28*

The first step of the proof is to rewrite the system in normal form. To this end, we need the two following preliminary results, that we will prove later. Recall that $\mathscr{D}$ is defined in (79), and set

$$2A^\flat(x,\xi) := \{\mathscr{U}^*, H\}\mathscr{U} - \{\mathscr{D}, \mathscr{U}^*\}\mathscr{U}, \tag{85}$$

which is in $S(1)$. Let us also recall the notation (40).

Lemma 29. *Let* $\mathbf{u}$ *be the solution to* (77) *with initial datum* $\mathbf{u}_0$*, and let* $\widetilde{\mathbf{u}} := \mathrm{op}_\varepsilon^{\mathrm{w}}[\mathscr{U}^*]\mathbf{u}$*. Then* $\widetilde{\mathbf{u}}$ *solves the system*

$$\begin{cases} \partial_t \widetilde{\mathbf{u}} + \dfrac{i}{\varepsilon}\mathrm{op}_\varepsilon^{\mathrm{w}}[\mathscr{D}]\widetilde{\mathbf{u}} + \mathrm{op}_\varepsilon^{\mathrm{w}}[\Gamma]\widetilde{\mathbf{u}} = \varepsilon\, \mathrm{op}_\varepsilon^{\mathrm{w}}[r_\varepsilon]\widetilde{\mathbf{u}}, \\ \widetilde{\mathbf{u}}(0) = \mathrm{op}_\varepsilon^{\mathrm{w}}[\mathscr{U}^*]\mathbf{u}_0, \end{cases} \tag{86}$$

where

$$\Gamma := \mathscr{U}^* A \mathscr{U} + A^\flat,$$

and the remainder $r_\varepsilon \in S(1)$ *is such that* $\mathscr{N}_{\gamma d}(r_\varepsilon) = O(1)$.

For all $(x,\xi) \in \mathbb{R}^{2d}$, we set

$$Q(x,\xi) := -\sum_{(j,k)/\, j\neq k} \frac{i\Pi_j \Gamma(x,\xi)\Pi_k}{\lambda_j(x,\xi) - \lambda_k(x,\xi)}. \tag{87}$$

It holds that $Q \in S(1)$. Recall that A_k^{eff} is given by (82).

Lemma 30. *Let* $\widetilde{\mathbf{u}}$ *be the solution to* (86)*. We set*

$$\mathbf{v} := \big(\mathrm{Id} + \varepsilon\, \mathrm{op}_\varepsilon^{\mathrm{w}}[Q]\big)\widetilde{\mathbf{u}} \quad \textit{and} \quad \mathbf{v}_k := \Pi_k \mathbf{v} \quad \textit{for } 1 \le k \le m.$$

Then

$$\begin{cases} \partial_t \mathbf{v}_k + \dfrac{i}{\varepsilon}\mathrm{op}_\varepsilon^{\mathrm{w}}[\lambda_k \mathrm{Id}]\mathbf{v}_k + \mathrm{op}_\varepsilon^{\mathrm{w}}[A_k^{\mathrm{eff}}]\mathbf{v}_k = \varepsilon\, \mathrm{op}_\varepsilon^{\mathrm{w}}[r_\varepsilon]\mathbf{v}, \\ \mathbf{v}_k(0) = \big(\mathrm{Id} + \varepsilon\, \mathrm{op}_\varepsilon^{\mathrm{w}}[Q]\big)\mathrm{op}_\varepsilon^{\mathrm{w}}[\mathscr{U}^*]\mathbf{u}_0, \end{cases} \tag{88}$$

where r_ε *is again a remainder in* $S(1)$ *such that* $\mathscr{N}_{\gamma d}(r_\varepsilon) = O(1)$.

Let us take these two results for granted and finish the proof of Theorem 28. Using that $\mathscr{U}\mathscr{U}^* = \mathscr{U}^*\mathscr{U} = \mathrm{Id}$, we may rewrite both energy terms involved in the observability cost in terms of $\mathbf{v}$:

$$\|\mathbf{u}(T)\|_{L^2}^2 = \sum_{k=1}^m \|\mathbf{v}_k(T)\|_{L^2}^2 \big(1 + O(\varepsilon)\big), \tag{89}$$

$$\|\mathrm{op}_\varepsilon^{\mathrm{w}}[B]\mathbf{u}\|_{L^2}^2 = \sum_{k=1}^m \|\mathrm{op}_\varepsilon^{\mathrm{w}}[B_k^{\mathrm{eff}}]\mathbf{v}_k\|_{L^2}^2 + \sum_{j\neq k}\Big\langle \mathrm{op}_\varepsilon^{\mathrm{w}}[B\mathscr{U}]\mathbf{v}_j, \mathrm{op}_\varepsilon^{\mathrm{w}}[B\mathscr{U}]\mathbf{v}_k\Big\rangle_{L^2} + O(\varepsilon)\|\mathbf{v}\|_{L^2}^2. \tag{90}$$

Note that we split the diagonal and off-diagonal terms in the observed energy because they exhibit very distinct behaviors.

For the off-diagonal terms, we rely on the following lemma.

Lemma 31. *Let* $\Lambda(x,\xi) \in S(1)$ *be a matrix-valued symbol. Let* $1 \le j, k \le m$ *such that* $j \neq k$*. Then it holds*

$$\int_0^T \Big\langle \mathrm{op}_\varepsilon^{\mathrm{w}}[\Pi_k \Lambda \Pi_j]\mathbf{v}(t), \mathbf{v}(t)\Big\rangle_{L^2} \mathrm{d}t = O(\varepsilon)\|\mathbf{v}(T)\|_{L^2}^2. \tag{91}$$

Applying this lemma with $\Lambda = \mathscr{U}^* B^* B \mathscr{U}$ yields, for $1 \le j \ne k \le m$,

$$\left\langle \mathrm{op}_\varepsilon^{\mathrm{w}}[B\mathscr{U}]\mathbf{v}_j, \mathrm{op}_\varepsilon^{\mathrm{w}}[B\mathscr{U}]\mathbf{v}_k \right\rangle_{L^2} = O(\varepsilon)\left\|\mathbf{v}(T)\right\|_{L^2}^2. \tag{92}$$

For the diagonal terms, for each $1 \le k \le m$, we apply Theorem 21 on the system (88), for $1 \le k \le m$, to obtain that

$$\left\|\mathbf{v}_k(T)\right\|_{L^2}^2 = |\mathbf{b}_{T,k}|^2 + o(1), \tag{93}$$

$$\int_0^T \left\|\mathrm{op}_\varepsilon^{\mathrm{w}}[B\mathscr{U}]\mathbf{v}_k\right\|_{L^2}^2 = \mathbf{b}_{T,k}^* G_{T,k}(T)\mathbf{b}_{T,k} + o(1). \tag{94}$$

Combining (92), (93) and (94), and substituting them into (89) and (90), we arrive at

$$\left\|\mathbf{u}(T)\right\|_{L^2}^2 = \sum_{k=1}^m \left|\mathbf{b}_{k,x_0,\xi_0,\mathbf{b}_0}(T)\right|^2 + o(1),$$

$$\int_0^T \left\|\mathrm{op}_\varepsilon^{\mathrm{w}}[B]\mathbf{u}\right\|_{L^2}^2 = \sum_{k=1}^m \mathbf{b}_{k,x_0,\xi_0,\mathbf{b}_0}(T)^* G_{k,x_0,\xi_0}(T)\mathbf{b}_{k,x_0,\xi_0,\mathbf{b}_0}(T) + o(1).$$

Plugging those estimates into the definition of the cost as given by (46) concludes the proof.

6.3. *Diagonalization of the propagation term. Proof of Lemma 29*

We consider the system (77), and apply to it the operator $\mathrm{op}_\varepsilon^{\mathrm{w}}[\mathscr{U}^*]$, giving

$$\partial_t \widetilde{\mathbf{u}} + \frac{i}{\varepsilon}\,\mathrm{op}_\varepsilon^{\mathrm{w}}[\mathscr{U}^*]\,\mathrm{op}_\varepsilon^{\mathrm{w}}[H]\mathbf{u} + \mathrm{op}_\varepsilon^{\mathrm{w}}[\mathscr{U}^*]\,\mathrm{op}_\varepsilon^{\mathrm{w}}[A]\mathbf{u} = 0. \tag{95}$$

Applying Lemma 14 for the identity (78), we obtain

$$\mathrm{op}_\varepsilon^{\mathrm{w}}[\mathscr{U}^*]\,\mathrm{op}_\varepsilon^{\mathrm{w}}[H] + \frac{\varepsilon}{2i}\,\mathrm{op}_\varepsilon^{\mathrm{w}}\big[\{\mathscr{U}^*, H\}\big] = \mathrm{op}_\varepsilon^{\mathrm{w}}[\mathscr{D}]\,\mathrm{op}_\varepsilon^{\mathrm{w}}[\mathscr{U}^*] + \frac{\varepsilon}{2i}\,\mathrm{op}_\varepsilon^{\mathrm{w}}\big[\{\mathscr{D}, \mathscr{U}^*\}\big] + \varepsilon^2\,\mathrm{op}_\varepsilon^{\mathrm{w}}[r_\varepsilon], \tag{96}$$

where $r_\varepsilon \in S(1)$ is a remainder symbol such that $\mathscr{N}_{\gamma d}(r_\varepsilon)$ is bounded uniformly in ε. By a slight abuse of notation, we will denote all such remainders r_ε, despite them not being the same in all the equalities. Rearranging, using that $\mathscr{U}\mathscr{U}^* = \mathrm{Id}$ and recalling the definition of $A^\flat$ in (85), it follows that

$$\mathrm{op}_\varepsilon^{\mathrm{w}}[\mathscr{U}^*]\,\mathrm{op}_\varepsilon^{\mathrm{w}}[H] = \mathrm{op}_\varepsilon^{\mathrm{w}}[\mathscr{D}]\,\mathrm{op}_\varepsilon^{\mathrm{w}}[\mathscr{U}^*] + \frac{\varepsilon}{i}\,\mathrm{op}_\varepsilon^{\mathrm{w}}[A^\flat]\,\mathrm{op}_\varepsilon^{\mathrm{w}}[\mathscr{U}^*] + \varepsilon^2\,\mathrm{op}_\varepsilon^{\mathrm{w}}[r_\varepsilon]. \tag{97}$$

Meanwhile, for the term of order 0, we have

$$\mathrm{op}_\varepsilon^{\mathrm{w}}[\mathscr{U}^*]\,\mathrm{op}_\varepsilon^{\mathrm{w}}[A] = \mathrm{op}_\varepsilon^{\mathrm{w}}[\mathscr{U}^* A \mathscr{U}]\,\mathrm{op}_\varepsilon^{\mathrm{w}}[\mathscr{U}^*] + \varepsilon\,\mathrm{op}_\varepsilon^{\mathrm{w}}[r_\varepsilon]. \tag{98}$$

Substituting (97) and (98) into (95), it finally yields

$$\partial_t \widetilde{\mathbf{u}} + \left(\frac{i}{\varepsilon}\,\mathrm{op}_\varepsilon^{\mathrm{w}}[\mathscr{D}] + \mathrm{op}_\varepsilon^{\mathrm{w}}[A^\flat] + \mathrm{op}_\varepsilon^{\mathrm{w}}[\mathscr{U}^* A \mathscr{U}]\right)\mathrm{op}_\varepsilon^{\mathrm{w}}[\mathscr{U}^*]\mathbf{u} = \varepsilon\,\mathrm{op}_\varepsilon^{\mathrm{w}}[r_\varepsilon]\widetilde{\mathbf{u}}, \tag{99}$$

and the claim immediately follows.

Remark 32. For any $1 \le k \le m$, multiplying equality (97) on the left by Π_k and on the right by $\mathrm{op}_\varepsilon^{\mathrm{w}}[\mathscr{U}]\Pi_k$ gives, denoting by σ_p the principal symbol of a pseudo-differential operator,

$$\sigma_p\Big(\Pi_k\Big(\mathrm{op}_\varepsilon^{\mathrm{w}}[\mathscr{U}^*]\,\mathrm{op}_\varepsilon^{\mathrm{w}}[H]\,\mathrm{op}_\varepsilon^{\mathrm{w}}[\mathscr{U}] - \mathrm{op}_\varepsilon^{\mathrm{w}}[\mathscr{D}]\,\mathrm{op}_\varepsilon^{\mathrm{w}}[\mathscr{U}^*]\,\mathrm{op}_\varepsilon^{\mathrm{w}}[\mathscr{U}]\Big)\Pi_k\Big) = -i\varepsilon\Pi_k A^\flat \Pi_k. \tag{100}$$

By the property (41) of the Weyl quantization, the first operator in the left-hand side of (100) is self-adjoint, and the second one is also self-adjoint, as a consequence of the symbolic identity $(\Pi_k\mathscr{D})(x,\xi) = (\lambda_k\Pi_k)(x,\xi)$. Therefore the symbolic matrix $i(\Pi_k A^\flat \Pi_k)(x,\xi)$ is Hermitian for $1 \le k \le m$.

6.4. *Block-diagonalization of the amplification term. Proof of Lemma 30*

Since

$$\left(\mathrm{Id} + \varepsilon \operatorname{op}_\varepsilon^{\mathrm{w}}[Q]\right)^{-1} = \mathrm{Id} - \varepsilon \operatorname{op}_\varepsilon^{\mathrm{w}}[Q] + \varepsilon^2 \operatorname{op}_\varepsilon^{\mathrm{w}}[r_\varepsilon],$$

we get

$$\left(\mathrm{Id} + \varepsilon \operatorname{op}_\varepsilon^{\mathrm{w}}[Q]\right)\left(\operatorname{op}_\varepsilon^{\mathrm{w}}\left[\frac{i}{\varepsilon}\mathscr{D} + \Gamma\right]\right)\left(\mathrm{Id} + \varepsilon \operatorname{op}_\varepsilon^{\mathrm{w}}[Q]\right)^{-1} = \operatorname{op}_\varepsilon^{\mathrm{w}}\left[\frac{i}{\varepsilon}\mathscr{D} - i[\mathscr{D}, Q] + \Gamma\right] + \varepsilon \operatorname{op}_\varepsilon^{\mathrm{w}}[r_\varepsilon]. \tag{101}$$

Now, we observe from the definition of Q in (87) that

$$-i[\mathscr{D}, Q] + \Gamma = \sum_{k=1}^{m} \Pi_k \Gamma \Pi_k. \tag{102}$$

Substituting (102) into this equality then yields

$$\left(\mathrm{Id} + \varepsilon \operatorname{op}_\varepsilon^{\mathrm{w}}[Q]\right)\left(\operatorname{op}_\varepsilon^{\mathrm{w}}\left[\frac{i}{\varepsilon}\mathscr{D} + \Gamma\right]\right)\left(\mathrm{Id} + \varepsilon \operatorname{op}_\varepsilon^{\mathrm{w}}[Q]\right)^{-1} = \operatorname{op}_\varepsilon^{\mathrm{w}}\left[\frac{i}{\varepsilon}\mathscr{D} + \sum_{k=1}^{m} \Pi_k \Gamma \Pi_k\right] + \varepsilon \operatorname{op}_\varepsilon^{\mathrm{w}}[r_\varepsilon]. \tag{103}$$

Performing the change of variable $\mathbf{v} = \left(\mathrm{Id} + \varepsilon \operatorname{op}_\varepsilon^{\mathrm{w}}[Q]\right)\tilde{\mathbf{u}}$ in the system (86) then gives

$$\partial_t \mathbf{v} + \operatorname{op}_\varepsilon^{\mathrm{w}}\left[\frac{i}{\varepsilon}\mathscr{D} + \sum_{k=1}^{m} \Pi_k \Gamma \Pi_k\right]\mathbf{v} = \varepsilon \operatorname{op}_\varepsilon^{\mathrm{w}}[r_\varepsilon]\mathbf{v}. \tag{104}$$

For $1 \le k \le m$, multiplying this equation on the left by Π_k finally leads to the system (88) and the claim is proven.

6.5. *Estimate of the off-diagonal energy terms. Proof of Lemma 31*

We follow the approach of [17, Chapter 6] to prove the smallness of the extra-diagonal energy terms.

By (104) and standard pseudo-differential calculations, for any $1 \le j \ne k \le m$, recalling that λ_j, λ_k denote the eigenvalues of $\mathscr{D}$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\langle \operatorname{op}_\varepsilon^{\mathrm{w}}\left[\frac{\Pi_k \Lambda \Pi_j}{\lambda_j - \lambda_k}\right]\mathbf{v}, \mathbf{v}\right\rangle = -\frac{i}{\varepsilon}\left\langle \operatorname{op}_\varepsilon^{\mathrm{w}}\left[\left[\frac{\Pi_k \Lambda \Pi_j}{\lambda_j - \lambda_k}, \mathscr{D}\right]\right]\mathbf{v}, \mathbf{v}\right\rangle + O(1)\|\mathbf{v}\|_{L^2}^2.$$

Now, we observe that, for any $1 \le j \ne k \le m$,

$$\left[\frac{\Pi_k \Lambda \Pi_j}{\lambda_j - \lambda_k}, \mathscr{D}\right] = \Pi_k \Lambda \Pi_j,$$

and we set

$$E_{j,k}(t) := \left\langle \operatorname{op}_\varepsilon^{\mathrm{w}}[\Pi_k \Lambda \Pi_j]\mathbf{v}, \mathbf{v}\right\rangle_{L^2}.$$

This immediately yields

$$E_{j,k}(t) = i\varepsilon \frac{\mathrm{d}}{\mathrm{d}t}\left\langle \operatorname{op}_\varepsilon^{\mathrm{w}}\left[\frac{\Pi_k \Lambda \Pi_j}{\lambda_j - \lambda_k}\right]\mathbf{v}, \mathbf{v}\right\rangle + O(\varepsilon)\left\|\mathbf{v}(T)\right\|_{L^2}^2. \tag{105}$$

Then, integrating in time gives

$$\begin{aligned}\int_0^T E_{j,k}(t)\,\mathrm{d}t = i\varepsilon &\left\langle \operatorname{op}_\varepsilon^{\mathrm{w}}\left[\frac{\Pi_k \mathscr{U}^* B^* B \mathscr{U} \Pi_j}{\lambda_j - \lambda_k}\right]\mathbf{v}(T), \mathbf{v}(T)\right\rangle_{L^2} \\ &- i\varepsilon \left\langle \operatorname{op}_\varepsilon^{\mathrm{w}}\left[\frac{\Pi_k \mathscr{U}^* B^* B \mathscr{U} \Pi_j}{\lambda_j - \lambda_k}\right]\mathbf{v}(0), \mathbf{v}(0)\right\rangle_{L^2} + O(\varepsilon)\left\|\mathbf{v}(T)\right\|_{L^2}^2.\end{aligned}$$

By the non-crossing condition, $\operatorname{op}_\varepsilon^{\mathrm{w}}\left[\frac{\Pi_k \mathscr{U}^* B^* B \mathscr{U} \Pi_j}{\lambda_j - \lambda_k}\right]$ is a bounded operator on L^2, and the claim immediately follows.

7. Compressible gas dynamics in a rapidly rotating frame

As an illustration, we consider the non-isentropic compressible Euler equations for perfect gases in a rotating frame (see for instance [4,16]) in 2D:

$$\begin{cases} \partial_t \mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} + \rho^{-1}\nabla p = 0, \\ \partial_t p + \mathbf{u}\cdot\nabla p + \rho c^2 \nabla\cdot\mathbf{u} + \dfrac{b(x_2)}{\mathrm{Ro}}\mathbf{u}^{\perp} = 0, \\ \partial_t s + \mathbf{u}\cdot\nabla s = 0, \end{cases} \tag{106}$$

where ρ is the density, $\mathbf{u}$ the velocity, p the pressure, s the entropy, $\mathrm{Ro}\in\mathbb{R}^+$ the Rossby number, $\mathbf{u}^{\perp} = (-u_2, u_1)$, b is a smooth function of x_2 and c is the sound speed. We consider the case of polytropic gases, for which the pressure law is $p = (\gamma-1)\rho^{\gamma} e^{\frac{s}{c}}$, for $\gamma > 1$, and the sound speed c is given by $c = \sqrt{\gamma p/\rho}$. Then, setting

$$\pi := \left(\frac{p}{\gamma-1}\right)^{2\gamma},$$

the system (106) is recast as

$$\begin{cases} \partial_t \mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} + 2\gamma e^{\frac{s}{\gamma c}}\pi\nabla\pi + \dfrac{b(x_2)}{\mathrm{Ro}}\mathbf{u}^{\perp} = 0, \\ \partial_t \pi + \mathbf{u}\cdot\nabla\pi + \dfrac{\gamma-1}{2}\pi\nabla\cdot\mathbf{u} = 0, \\ \partial_t s + \mathbf{u}\cdot\nabla s = 0. \end{cases}$$

The linearization of the system at a steady solution $(\bar{\mathbf{u}}, \bar\pi, \bar s)$, where $\bar{\mathbf{u}} = \big(\bar u_1(x_2), 0\big)$ is a shear flow, $\bar s\in\mathbb{R}^+$ is a constant value and $\bar\pi = \bar\pi(x_2)$ satisfies the relation

$$2\gamma\mathrm{Ro}\exp\left(\frac{\bar s}{c\gamma}\right)(\bar\pi\bar\pi')(x_2) = -b(x_2)\bar u_1(x_2).$$

In the case of small Rossby number, that is when $\mathrm{Ro} = \varepsilon$, the semiclassical approximation yields the following system:

$$\partial_t \mathbf{U} + \mathrm{op}^{\mathrm{w}}_{\varepsilon}\left[\frac{i}{\varepsilon}H + A\right]\mathbf{U} = 0, \tag{107}$$

where

$$\mathbf{U} = \begin{pmatrix} u_1 \\ u_2 \\ pi \\ s \end{pmatrix}, \quad H = \begin{pmatrix} \bar u_1(x_2)\xi_1 & -ib(x_2) & 2\gamma e^{\frac{\bar s}{\gamma c}}\bar\pi\xi_1 & 0 \\ ib(x_2) & \bar u_1(x_2)\xi_1 & 2\gamma e^{\frac{\bar s}{\gamma c}}\bar\pi\xi_2 & 0 \\ \frac{\gamma-1}{2}\bar\pi\xi_1 & \frac{\gamma-1}{2}\bar\pi\xi_2 & \bar u_1(x_2)\xi_1 & 0 \\ 0 & 0 & 0 & \bar u_1(x_2)\xi_1 \end{pmatrix},$$

and

$$A = \begin{pmatrix} 0 & \bar u_1'(x_2) & 0 & 0 \\ 0 & 0 & 2\gamma e^{\frac{\bar s}{\gamma c}}\bar\pi'(x_2) & \frac{2}{c}e^{\frac{\bar s}{\gamma c}}\bar\pi\bar\pi'(x_2) \\ 0 & \bar\pi'(x_2) & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The principal symbol H is not symmetric; however, setting

$$S := \mathrm{diag}\left(1, 1, 4\frac{\gamma}{\gamma-1}\exp\left(\frac{\bar s}{c\gamma}\right), 1\right),$$

which only depends on x, and computing

$$\widetilde H = \sqrt{S}H\sqrt{S}^{-1} = \begin{pmatrix} \bar u_1\xi_1 & -ib & \eta\xi_1 & 0 \\ ib & \bar u_1\xi_1 & \eta\xi_2 & 0 \\ \eta\xi_1 & \eta\xi_2 & \bar u_1\xi_1 & 0 \\ 0 & 0 & 0 & \bar u_1\xi_1 \end{pmatrix},$$

where $\eta := \sqrt{\gamma(\gamma-1)}\exp\left(\frac{\bar{s}}{2c\gamma}\right)\bar{\pi}$ and $\sqrt{S}$ is the positive square root of S, we observe that $\widetilde{H}$ is symmetric. Performing the change of unknowns $\mathbf{W} := \sqrt{S}\mathbf{U}$, we are then led to a system of the form:

$$\partial_t \mathbf{W} + \mathrm{op}_\varepsilon^{\mathrm{w}}\left[\frac{i}{\varepsilon}\widetilde{H} + \widetilde{A}\right]\mathbf{W} = \varepsilon\,\mathrm{op}_\varepsilon^{\mathrm{w}}[r_\varepsilon]\mathbf{W}, \tag{108}$$

with $\widetilde{A} = \sqrt{S}A\sqrt{S}^{-1}$. The eigenvalues of $\widetilde{H}$ are $\lambda_1 = \bar{u}_1(x_2)\xi_1$ with multiplicity 2,

$$\lambda_2 = \bar{u}_1\xi_1 + \sqrt{b^2 + \eta^2|\xi|^2},$$

with multiplicity 1, and

$$\lambda_3 = \bar{u}_1\xi_1 - \sqrt{b^2 + \eta^2|\xi|^2},$$

with multiplicity 1. The non-crossing assumption is verified for (x,ξ) such that

$$b^2(x_2) + \eta^2|\xi|^2 \neq 0.$$

Two orthogonal eigenvectors of λ_1 are given by

$$(0,0,0,1)^T, \quad \frac{1}{\langle\xi\rangle_b}\left(-i\eta\xi_2, i\eta\xi_1, b, 0\right)^T, \quad \text{where } \langle\xi\rangle_b = \sqrt{\eta^2|\xi^2| + b^2}.$$

Taking again $\mathscr{U}$ the matrix of the right normalized eigenvectors of $\widetilde{H}$, we can then compute the block of the matrix $(\mathscr{U}^*\widetilde{A}\mathscr{U})(x,\xi)$ and of the (block-diagonal) matrix $A^\flat(x,\xi)$ in (85) associated with λ_1:

$$A_{\lambda_1}(x,\xi) = \frac{1}{\langle\xi\rangle_b^2}\begin{pmatrix} 0 & 0 \\ i\alpha(x_2,\xi) & i\beta(x_2,\xi)\end{pmatrix} \quad \text{and} \quad A_{\lambda_1}^\flat(x,\xi) = \frac{1}{\langle\xi\rangle_b^2}\begin{pmatrix} 0 & 0 \\ 0 & -i\delta(x_2,\xi)\end{pmatrix},$$

where

$$\alpha(x_2,\xi) = -\frac{2}{c}\eta\xi_1\bar{\pi}\bar{\pi}'\exp\left(\frac{\bar{s}}{\gamma c}\right), \quad \beta(x_2,\xi) = 2b\eta\xi_1\sqrt{\frac{\gamma}{\gamma-1}}\exp\left(\frac{\bar{s}}{2\gamma c}\right)\bar{\pi}' - \eta^2\xi_1\bar{\pi}'b, \quad \delta(x_2,\xi) = \eta\xi_1 b'.$$

Then, if one wants to compute the short-time controllability cost for initial data micro-localized at (x,ξ) and living in the eigenspace associated with $\lambda_1(x,\xi)$, distinct behaviors may occur depending on the choice of the observation symbol $B(x,\xi)$.

For example, if $B(x,\xi) \in \mathbb{C}^{4\times 1}$ is such that the block corresponding to λ_1 in $B\mathscr{U}$ is given by

$$B_{\lambda_1} = \begin{pmatrix} -\frac{\alpha}{\beta} \\ 1 \end{pmatrix},$$

it holds

$$\mathrm{rank}\left[B_{\lambda_1}(x,\xi) \;\middle|\; A_{\lambda_1(x,\xi)}B_{\lambda_1}(x,\xi)\right] = 1, \quad \text{whereas } \mathrm{rank}\left[B_{\lambda_1}(x,\xi) \;\middle|\; \left(A_{\lambda_1} + A_{\lambda_1}^\flat\right)(x,\xi)B_{\lambda_1}(x,\xi)\right] = 2.$$

In this case, the presence of the bracket term $A^\flat$ therefore improves the short-time controllability cost, lowering it to $O(t^{-3})$ instead of $\Omega(t^{-5})$, with the Hardy and Littlewood notation.

On the contrary, if

$$B_{\lambda_1} = \begin{pmatrix} \frac{\delta-\beta}{\alpha} \\ 1 \end{pmatrix},$$

the ranks of the Kalman matrices rather are

$$\mathrm{rank}\left[B_{\lambda_1}(x,\xi) \;\middle|\; A_{\lambda_1(x,\xi)}B_{\lambda_1}(x,\xi)\right] = 2, \quad \text{and} \quad \mathrm{rank}\left[B_{\lambda_1}(x,\xi) \;\middle|\; \left(A_{\lambda_1} + A_{\lambda_1}^\flat\right)(x,\xi)B_{\lambda_1}(x,\xi)\right] = 1.$$

In this case, the bracket term $A^\flat(x,\xi)$ worsens the short-time controllability cost, making it $\Omega(t^{-5})$ instead of $O(t^{-3})$.

In conclusion, despite its skew-hermitian structure, the symbol $A^\flat(x,\xi)$ cannot be neglected in general as it may reduce or amplify the controllability cost.

Declaration of interests

The authors do not work for, advise, own shares in, or receive funds from any organization that could benefit from this article, and have declared no affiliations other than their research organizations.

References

[1] M. Assal, "Long time semiclassical Egorov theorem for $\hbar$-pseudodifferential systems", *Asymptotic Anal.* **101** (2017), no. 1-2, pp. 17–67.

[2] C. Bardos, G. Lebeau and J. Rauch, "Sharp sufficient conditions for the observation, control, and stabilization of waves from the boundary", *SIAM J. Control Optim.* **30** (1992), no. 5, pp. 1024–1065.

[3] G. Bastin and J.-M. Coron, *Stability and boundary stabilization of 1-D hyperbolic systems*, Progress in Nonlinear Differential Equations and their Applications, Birkhäuser, 2016, pp. xiv+307.

[4] S. Benzoni-Gavage and D. Serre, *Multidimensional hyperbolic partial differential equations. First-order systems and applications*, Oxford Mathematical Monographs, Clarendon Press, 2007, pp. xxvi+508.

[5] J. Bolte and R. Glaser, "A semiclassical Egorov theorem and quantum ergodicity for matrix valued operators", *Commun. Math. Phys.* **247** (2004), no. 2, pp. 391–419.

[6] Y. Brenier and D. Vorotnikov, "On optimal transport of matrix-valued measures", *SIAM J. Math. Anal.* **52** (2020), no. 3, pp. 2849–2873.

[7] N. Burq and G. Lebeau, "Mesures de défaut de compacité, application au système de Lamé", *Ann. Sci. Éc. Norm. Supér. (4)* **34** (2001), no. 6, pp. 817–870.

[8] P. Cannarsa, G. Floridia and M. Yamamoto, "Observability inequalities for transport equations through Carleman estimates", in *Trends in control theory and partial differential equations*, Springer INdAM Series, Springer, 2019, pp. 69–87.

[9] H. O. Cordes, "A version of Egorov's theorem for systems of hyperbolic pseudodifferential equations", *J. Funct. Anal.* **48** (1982), no. 3, pp. 285–300.

[10] J.-M. Coron, *Control and nonlinearity*, Mathematical Surveys and Monographs, American Mathematical Society, 2007, pp. xiv+426.

[11] J.-M. Coron and H.-M. Nguyen, "On the optimal controllability time for linear hyperbolic systems with time-dependent coefficients", 2021. Online at https://arxiv.org/abs/2103.02653.

[12] Y. Cui, C. Laurent and Z. Wang, "On the observability inequality of coupled wave equations: the case without boundary", *ESAIM, Control Optim. Calc. Var.* **26** (2020), article no. 14 (37 pages).

[13] B. Dehman, J. Le Rousseau and M. Léautaud, "Controllability of two coupled wave equations on a compact manifold", *Arch. Ration. Mech. Anal.* **211** (2014), no. 1, pp. 113–187.

[14] B. Dehman and G. Lebeau, "Analysis of the HUM control operator and exact controllability for semilinear waves in uniform time", *SIAM J. Control Optim.* **48** (2009), no. 2, pp. 521–550.

[15] R. J. DiPerna and P.-L. Lions, "Ordinary differential equations, transport theory and Sobolev spaces", *Invent. Math.* **98** (1989), no. 3, pp. 511–547.

[16] E. Feireisl, I. Gallagher and A. Novotný, "A singular limit for compressible rotating fluids", *SIAM J. Math. Anal.* **44** (2012), no. 1, pp. 192–205.

[17] C. Fermanian-Kammerer, "Multi-scale analysis and Schrödinger operators", 2021. Online at https://perso.math.u-pem.fr/fermanian.clotilde/M2_v3.pdf.

[18] G. Floridia, H. Takase and M. Yamamoto, "A Carleman estimate and an energy method for a first-order symmetric hyperbolic system", *Inverse Probl. Imaging* **16** (2022), no. 5, pp. 1163–1178.

[19] I. Gallagher, T. Paul and L. Saint-Raymond, "On the propagation of oceanic waves driven by a strong macroscopic flow", in *Nonlinear partial differential equations*, Abel Symposia, Springer, 2012, pp. 231–254.

[20] M. Herty and F. Thein, "Boundary feedback control for hyperbolic systems", *ESAIM, Control Optim. Calc. Var.* **30** (2024), article no. 71 (16 pages).

[21] V. Komornik, *Exact controllability and stabilization. The multiplier method*, Research in Applied Mathematics, Masson, 1994, pp. viii+156.

[22] C. Laurent and M. Léautaud, "Uniform observability estimates for linear waves", *ESAIM, Control Optim. Calc. Var.* **22** (2016), no. 4, pp. 1097–1136.

[23] G. Lebeau, "Control for hyperbolic equations", in *Analysis and optimization of systems: state and frequency domain approaches for infinite-dimensional systems (Sophia-Antipolis, 1992)*, Lecture Notes in Control and Information Sciences, Springer, 1993, pp. 160–183.

[24] G. Lebeau, "Équation des ondes amorties", in *Algebraic and geometric methods in mathematical physics (Kaciveli, 1993)*, Mathematical Physics Studies, Kluwer Academic Publishers, 1996, pp. 73–109.

[25] N. Lerner, *Integrating the Wigner distribution on subsets of the phase space, a survey*, Memoirs of the European Mathematical Society, European Mathematical Society, 2024, pp. viii+216.

[26] L. Ning and T. Georgiou, "Metrics for matrix-valued measures via test functions", in *53rd IEEE Conference on Decision and Control*, IEEE, 2014, pp. 2642–2647.

[27] T. I. Seidman, "How violent are fast controls?", *Math. Control Signals Syst.* **1** (1988), no. 1, pp. 89–95.

[28] H. Spohn, "Semiclassical limit of the Dirac equation and spin precession", *Ann. Phys.* **282** (2000), no. 2, pp. 420–431.

[29] M. Zworski, *Semiclassical analysis*, Graduate Studies in Mathematics, American Mathematical Society, 2012, pp. xii+431.